THE UNIVERSITY OF CHICAGO

COSMOLOGY THROUGH LARGE-SCALE STRUCTURE: BARYON ACOUSTIC OSCILLATIONS IN THE BISPECTRUM AND EVOLUTION OF HALO PROFILES

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS

BY
HILLARY LYN CHILD

CHICAGO, ILLINOIS
MARCH 2019

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my advisor, Salman Habib; all past and present members of the Cosmological Physics and Advanced Computing group at Argonne National Laboratory; my collaborators, Takahiro Nishimichi, Zachary Slepian, Tomomi Sunayama, and Masahiro Takada; and the members of my thesis committee, Professors Michael Gladders, Stephan Meyer, Robert Wald, and Paul Wiegmann, for their support and contributions to this work.

This work was supported by the National Science Foundation Graduate Research Fellowship Program under Grants No. DGE-1144082 and DGE-1746045, an international travel allowance through Graduate Research Opportunities Worldwide, a Japan Society for the Promotion of Science International Research Fellowship [Postdoctoral Fellowships for Research in Japan (Strategic)], and the United States Department of Energy under Contract No. DE-AC02-06CH11357.

ABSTRACT

The history of the Universe is a story of structure formation: fluctuations in the post-inflation density field collapse to form a "cosmic web" of dark matter halos, and observations of this large-scale structure offer constraints on cosmological parameters. This thesis presents two studies of large-scale structure probes of cosmology.

We first study the distribution of structure as a probe of expansion history. The two-point matter correlation function (2PCF), which counts pairs of dark matter halos separated by a given distance, and its Fourier transform the power spectrum both show the signature of baryons in the early Universe. The baryon acoustic oscillation (BAO) feature represents a fixed distance scale, or standard ruler, which expands with the scale factor and thus constrains expansion history and dark energy. The 2PCF constrains the BAO scale to 1% precision in modern galaxy surveys, but higher-point statistics like the three-point correlation function and its Fourier transform the bispectrum also carry BAO information. We present a new technique to strategically select the bispectrum triangles most sensitive to BAO. A small number of bispectrum measurements can improve BAO precision by around 15%, equivalent to lengthening a survey by 30%. To understand why some bispectrum triangles are more sensitive to BAO than others, we study the structure of BAO in the bispectrum in detail.

We then turn to the internal structure of dark matter halos and the evolution of their profiles. The spherically-averaged density of a halo is well described by the Navarro-Frenk-White (NFW) profile, a function of two parameters: concentration, which describes the density of the inner regions of the halo, and halo mass. These two parameters are correlated: the concentration-mass ($c - M$) relation, which predicts concentration as a function of halo mass, has negative slope at low redshifts. We measure the concentrations of tens of millions of simulated halos to compute a robust $c - M$ relation across a wide range of redshifts and masses. Our $c - M$ relation can be used in preparation for and analysis of next-generation surveys. We then study the relationship between halo formation time and concentration by tracking the evolution of individual halos' masses and concentrations.

CHAPTER 1

INTRODUCTION

In the hierarchical model of structure formation, initial overdensities in the dark matter field attract more dark matter and collapse to form sheets, filaments, and finally halos. These halos are roughly spherical structures of dense dark matter—hundreds of times the average density of the Universe. Both their distribution and internal structure can be observed through the galaxies that populate them, and both depend on the evolution of the Universe, so both are powerful probes of cosmology.

Linear perturbation theory describes the evolution of small overdensities and is a good description of the evolution of the Universe at large scales and early times. As density fluctuations grow, however, the perturbations are no longer small and linear theory can no longer be used to predict the evolution of the density field. Cosmological N-body simulations, which model the motions of particles under full Newtonian gravity, are needed. This thesis uses N-body simulations to address topics in two regimes of structure formation: the mildly nonlinear regime, at cosmological scales where nonlinear effects contaminate a signal used to constrain dark energy; and the deeply nonlinear regime, in the interior regions of dark matter halos where galaxies form. In the remainder of the Introduction, we discuss structure formation in further detail, introduce cosmological N-body simulation methods, provide background on bispectrum baryon acoustic oscillations (BAOs) and halo profiles, then outline the body of the thesis.

1.1 Structure Formation

The cosmic web of large-scale structure we observe today grew from small initial density fluctuations after inflation. Through the effects of gravity in an expanding universe, these small fluctuations grew into the large overdensities observable today. The theory of large scale structure offers a connection between cosmological observables, like the number density

of objects with a given mass or the probability of two objects to be separated by any given distance scale, and cosmological parameters like the properties of dark matter and dark energy. The process of structure formation is outlined briefly in this section, but described in detail in e.g. Ref. [240].

Most models of inflation predict that after inflation ends, the mass density of the universe is relatively, but not entirely, uniform. The density field $\rho(\mathbf{x})$ can be expressed in terms of a homogeneous mean density $\bar{\rho}$ and a perturbation $\delta(\mathbf{x})$,

$$\rho(\mathbf{x}) = \bar{\rho}\left(1 + \delta(\mathbf{x})\right). \tag{1.1}$$

The perturbation $\delta(\mathbf{x})$ arises from random quantum mechanical fluctuations in the inflaton field that drives inflation, so it is a Gaussian random field, statistically fully described by the power spectrum $P(k)$. The initial perturbations are very small, so at least initially, the evolution of the density field is well described by linear theory: the Fourier modes evolve independently and the field remains Gaussian.

After inflation, the universe expands (or contracts) at a rate determined by its contents (radiation, matter, and dark energy) according to the Friedmann equations

$$\begin{aligned} H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 &= \frac{8\pi G}{3c^2}\rho \\ \frac{\ddot{a}}{a} &= -\frac{4\pi G}{3}\left(\rho + \frac{3P}{c^2}\right), \end{aligned} \tag{1.2}$$

where H is the Hubble parameter, a is the scale factor describing the size of the universe, $G = 6.67 \times 10^{-11} Nm^2/kg^2$ is the gravitational constant, c is the speed of light, ρ is density, and P is pressure. These equations describe the expansion of space, but the evolution of the (pressureless and nonrelativistic) matter field is governed by three further equations: the

continuity, Euler, and Poisson equations

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) &= 0 \\ \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} &= -\nabla \phi \\ \nabla^2 \phi &= 4\pi G \rho, \end{aligned} \tag{1.3}$$

where $\mathbf{v}$ is the fluid velocity and ϕ is the gravitational potential.

Combining these three equations in an expanding universe, to first order in δ and $\mathbf{v}$ the small density perturbations obey the second-order differential equation

$$\frac{\partial^2 \delta}{\partial t^2} + 2H(z)\frac{\partial \delta}{\partial t} = 4\pi G \bar{\rho} \delta. \tag{1.4}$$

The solutions to this wave equation are a growing mode and a decaying mode:

$$\delta(\mathbf{x}, t) = \delta_+(\mathbf{x}) D_+(t) + \delta_-(\mathbf{x}) D_-(t). \tag{1.5}$$

Time dependence enters only through the functions D_+ and D_-, while δ_+ and δ_- depend only on the initial conditions. The growing solution, D_+, is known as the growth factor. So long as the density perturbations are small relative to the background density, the growth of structure at all scales is described by the growth factor.

As time proceeds, the density fluctuations continue to grow. When δ is no longer much smaller than unity, Fourier modes are no longer independent, so linear theory can no longer describe the evolution of structure. This happens first at small scales like the size of an individual halo, where δ can be much larger than unity, indicating a deeply nonlinear regime. On larger scales ($k \sim 0.2\ h/\mathrm{Mpc}$), nonlinear effects are less dramatic, but can interfere with our ability to observe features of the linear power spectrum. At all scales, nonlinear structure formation moves power into higher order statistics, and the overdensity field is no longer fully described by two-point statistics. Even scales as large as $k \sim 0.1\ h/\mathrm{Mpc}$ are mildly nonlinear,

and N-body simulations are needed to fully model the evolution of structure.

1.2 N-body Simulations

N-body simulations have wide applications in astrophysics and cosmology. These simulations, which evolve large numbers of particles under the influence of gravity, fully account for the effects of nonlinearity and model the evolution of systems ranging from stars, to galaxies, to cosmological scales. For example, N-body simulations can be used to understand the formation of the earth and moon, or to track the evolution of structure in the universe.

The motion of every particle in an N-body simulation depends on the forces acting on it due to the other particles, as well as any external potential field. The force acting on a particle of mass m_i at location $\mathbf{r_i}$ is, by Newton's law, a sum of the forces due to all other particles:

$$\mathbf{F} = -\sum_{j \neq i} \frac{G m_i m_j \left(\mathbf{r}_i - \mathbf{r}_j\right)}{\left|\mathbf{r}_i - \mathbf{r}_j\right|}, \tag{1.6}$$

where $G = 6.67 \times 10^{-11} N m^2/kg^2$ is the gravitational constant. Given an initial set of particles, positions, and velocities, the differential equations governing the particles' motions can be solved numerically. At each timestep, the total force acting on each particle are calculated and their positions and velocities are updated accordingly. Naively, this calculation is $\mathcal{O}(N^2)$ for N particles and the calculation becomes very computationally intensive in the context of problems relevant to cosmology. The diameter of the observable universe is about 30 gigaparsecs; modeling a large fraction of it, for example a cubic gigaparsec, requires a trillion particles to obtain a mass resolution of 10^8 solar masses. Many techniques have been developed to optimize code performance and enable simulations of cosmological volumes at high mass resolution. The simulations analyzed in this work were performed using the HACC (Hardware/Hybrid Accelerated Cosmology Codes) framework, which combines a particle-mesh solver at large scales with a particle-based solver at short ranges [119].

The results of these cosmological simulations enable many studies of cosmology. For

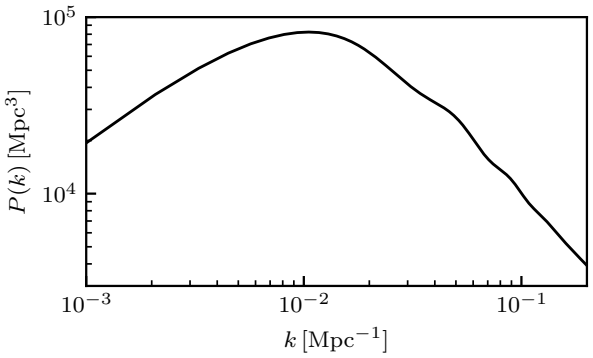

Figure 1.1: Emulated nonlinear matter power spectrum at $z = 0$, $\omega_m = 0.133$, showing baryon acoustic oscillations at $\log(k) \sim -1.5$.

example, simulations are used in analysis of survey data to improve our knowledge of the dark contents of the universe, as in Chapter 2. The evolution of structure in simulations can also be tracked and compared to observations, both to constrain cosmological parameters and to better understand structure formation, as in Chapters 4 and 5.

1.3 Mildly Nonlinear Regime: Baryon Acoustic Oscillations

The distribution of galaxies—and the dark matter haloes that surround them—has been measured by multiple surveys and used to constrain cosmological parameters. The two-point correlation function (2PCF) is one such statistic. The 2PCF counts the number of observations separated by a given distance, then compares this count to that expected from a random distribution. A large value for a given scale indicates that structure is strongly correlated on that scale. The Fourier transform of the 2PCF is the power spectrum $P(k)$, shown in Fig. 1.1.

The concepts of the 2PCF and power spectrum can be extended to higher order statistics. While the 2PCF counts the number of pairs of objects separated by a given distance, the three-point correlation function (3PCF) measures the correlation between triplets of halos. It is therefore a function of three triangle side lengths. The Fourier transform of the 3PCF, the bispectrum, similarly depends on three wavenumbers.

Perturbation theory provides a highly successful description of the observed power spectrum, particularly at the largest scales. At smaller scales, however, Fourier modes are no longer independent and perturbation theory begins to break down. One process in the early Universe leaves a signature at these scales known as baryon acoustic oscillations.

In the early Universe, the opposing forces of gravity and pressure produce oscillations in the baryon-photon plasma. An overdense shell of coupled baryons and photons expands until recombination, when the photons free-stream away. Now that there is no pressure force, the shell's radius freezes at a fixed radius, the sound horizon. As structure forms, this overdense region continues to attract more matter. This process happens at many points throughout the Universe, so everywhere there is more structure separated by the sound horizon ($\sim 150\,\mathrm{Mpc}$) than by other similar scales. The 2PCF, measured by wide-area galaxy surveys like the Baryon Oscillation Spectroscopic Survey (BOSS)[1] or Dark Energy Spectroscopic Instrument (DESI)[2], therefore shows a peak at the sound horizon: statistically, more galaxies are separated by this distance than by distances slightly larger or smaller. The peak in the correlation function corresponds to oscillations in the matter power spectrum $P(k)$, shown in Fig. 1.1. The location of the peak (and frequency of the power spectrum oscillation) gives the baryon acoustic oscillation scale, a distance measurement that has been used to calculate the expansion history of the Universe and constrain dark energy [30, 83, 141, 180, 278]. Applied to Baryon Oscillation Spectroscopic Survey (BOSS) data, the BAO method has offered 1% distance constraints [2, 262], while the upcoming DESI collaboration forecasts distance errors of 0.2% [174].

In the time between recombination and the low redshifts of particular importance for dark energy studies, nonlinear effects, galaxy bias, and redshift space distortions change the shape and position of the BAO peak. BAO reconstruction, introduced in Ref. [84] and discussed further in Refs. [226, 234, 235], is a technique used to correct for nonlinear

1. `http://www.sdss3.org/surveys/boss.php`

2. `http://desi.lbl.gov`

effects and recover the unbiased BAO signal. Existing methods to reconstruct the BAO peak improve the error in distance measurement by a factor of greater than 1.5 (e.g. Ref. [236]), but efforts continue to further increase the quality of reconstruction through new algorithms [270] to be applied to current and future survey data.

Higher-order statistics may also be used to constrain the BAO distance scale. The advantage of higher-order statistics is that they can be measured on data that is already collected, without incurring the costs associated with extending a survey. Additionally, they provide a cross-check on reconstruction that is subject to different systematics, does not require any theoretical assumptions, and does not need to be rerun. The matter bispectrum, for example, provides further information about the matter distribution, and is very sensitive to baryon acoustic oscillations [275]. The BAO feature has recently been detected in the three-point correlation function and the bispectrum $B(k)$ [102, 239, 292, 293].

These works all find that three-point information can be used to improve constraints on the BAO scale over $P(k)$ or the 2PCF alone, but challenges remain. Unlike two-point statistics, where each measurement probes only one scale, each bispectrum measurement carries information from three different wavenumbers. Measurement of the bispectrum is therefore more computationally expensive than the power spectrum, but this problem has largely been solved using sophisticated algorithms. However, bispectrum constraining power remains limited by the number of mock catalogs available to calculate covariances. For example, in the case of Ref. [239], a covariance matrix estimated with more mocks would improve their 1.1% precision joint power spectrum+bispectrum constraints substantially to 0.7%, a gain of more than 30%. Measuring all bispectrum triangles would provide the optimal constraint, if a perfect covariance matrix were available. But estimation of an accurate and invertible covariance matrix requires a large number of mock catalogs proportional to the number of measurements, as discussed in Ref. [241]. In order to be convinced of the accuracy of the bispectrum, each mock should be fully N-body to fully capture nonlinear dynamics. Since N-body mocks are computationally expensive, the number of mocks available is restricted.

Current resources therefore limit the number of triangles that can reasonably be used to constrain BAO in the bispectrum. Given a limited number of bispectra measurements, which ones should we choose to most efficiently constrain BAO?

Some triangles should be more sensitive to the BAO scale than others, and if bispectrum BAO studies can be restricted to the set of triangles most sensitive to BAO, fewer mock catalogs will be needed to obtain improved constraints. However, it is not obvious how to select the best subset of triangles. We propose a method, informed by perturbation theory, to select triangle configurations that maximize or minimize the BAO signal. We then measure bispectra for selected configurations on simulated halo catalogs to estimate the improvement in constraints on the BAO scale over those obtained from $P(k)$ alone. Having shown the value of our method to constrain BAO, we detail the physical reasons why some triangle configurations carry a much stronger BAO signal than others.

1.4 Deeply Nonlinear Regime: Internal Halo Structure

The overdensities associated with halos are hundreds of times the background overdensity, so the formation of halos is a highly nonlinear process. Theoretical models of spherical collapse using the extended Press-Schechter [248] formalism (e.g. Refs. [14, 196, 218, 250, 265, 328]) and ellipsoidal collapse (e.g. Refs. [179, 186]) predict the formation of dense halos. In the simplest picture of halo formation, top-hat spherical collapse, a single spherical top-hat perturbation collapses into a dense dark matter halo and relaxes to virial equilibrium while the background Universe continues to expand. The average overdensity of the virialized halo is about 200 times the critical density of the Universe, motivating one common definition of a halo boundary: the radius of a halo encloses the region within which the average density is $\Delta_{\rm vir} \sim 200$ times the critical density, as shown in Figure 1.2.

This simple picture breaks down as the halo virializes, however, and simulations remain the only way to fully model halo formation. The first simulations of halo evolution (e.g. Ref. [217]) included only a small number of halos, but found that the spherically-averaged

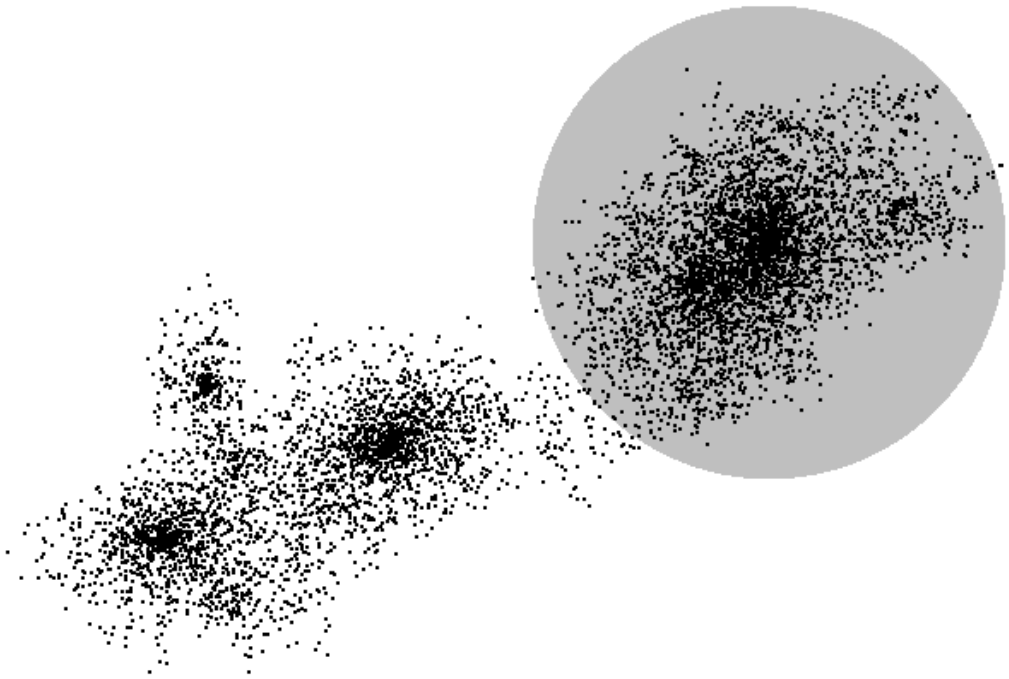

Figure 1.2: An example of a simulated halo. Points represent a subsample of the particles identified as belonging to the halo; the radius of the shaded region is r_{200}, enclosing the sphere within which the average density is 200 times the critical density. The points on the left belong to smaller halos in the process of merging into the main halo.

radial mass density of all dark matter halos is well described by a universal profile, known as the Navarro-Frenk-White (NFW) profile. This result has been confirmed in modern simulations that model tens of millions of halos, or smaller numbers of halos at extremely high resolution.

The NFW profile predicts that the density of dark matter in a halo depends on distance from the center of the halo and two parameters: concentration (a dimensionless quantity related to the density of the inner regions of the halo) and halo mass. These two parameters are correlated, and the concentration-mass (c–M) relation describes how the concentration of a halo depends on its mass. Generally, more massive halos are found to have lower concentration, but the shape of the c–M relation depends on redshift—some simulations have even shown that at sufficiently high redshift, it is the most massive halos that have the highest concentration. Analytical models of structure formation predict that the interior density of a halo is related to the background density of the Universe at the time the halo formed. Background density falls with time, and the most massive halos have formed only recently, so the low-mass result that concentration falls with mass matches the theoretical expectation. Concentration that rises with mass, however, is counterintuitive.

Halo masses and concentrations can be measured in both simulations and observations, so the c–M relation provides a comparison between theory and observation. In preparation for future surveys with much smaller uncertainty in concentration measurements, robust predictions of the c–M relation are needed to correspond to the different types of observations. Strong lensing can be used to measure the profiles of individual clusters, while the weak lensing signal is only strong enough for stacks of multiple halo observations.

We use new simulations with state-of-the-art mass resolution and statistics to measure concentrations for tens of millions of individual and stacked simulated halos, compute a new c–M relation, and better characterize its redshift dependence. We find that concentration falls with mass at low redshift, but the c–M relation flattens as redshift increases—at high redshifts, halos of all masses have concentration $c \sim 3 - 4$. When masses are scaled by the

nonlinear mass scale $M_\star$ (a function of redshift and cosmology, falling rapidly as redshift increases), the c–M relation is well fit by a broken power-law across eight decades in $M/M_\star$. This universal form enables better comparison with observations and other simulations, and agrees well with weak- and strong-lensing observations. In addition, we fit halo profiles to the Einasto profile, a function with three parameters to allow for deviations from the NFW shape. We also develop multiple methods to measure concentration and test how well halos are fit by NFW, and test the effects of several potential sources of bias in concentration measurements.

The internal structure of halos depends on their evolution. As a halo evolves and its mass changes, both via slow accretion of small amounts of matter and more violent mergers like the one shown in Figure 1.2, its concentration does as well. Concentration is understood to be related to the density of the Universe at the time the halo forms; older halos formed in a higher-density Universe, so their inner regions are denser and their concentration is higher. However, while halo age is broadly understood to influence concentration, the literature contains many different definitions of halo age, each capturing a slightly different feature of the halo's mass accretion history. All of these definitions rely on halo merger trees, which track the evolution of individual halos through time. We use these merger trees to identify different halo formation times and study the co-evolution of mass and concentration. Some halos accrete mass slowly and steadily, but other halos undergo major merger events, where mass rapidly as much as doubles and concentration may fall. Halo concentration and all definitions of formation time both depend on the relative importance of slow accretion and major mergers in the individual halo's mass accretion history, but some definitions of formation time are much more strongly correlated with concentration than are others.

Both the c–M relation and further understanding of the relationship between assembly history and halo structure are crucial for analysis of future surveys. Current measurements of the c–M relation have large error bars, but these are expected to shrink in coming years. Precision predictions of the c–M relation can be compared to these improved measurements

to constrain cosmological parameters like σ_8 that can change the shape of the c–M relation. Understanding of halo evolution is also intimately tied to assembly bias and the galaxy halo connection, both of which are crucial for the development of accurate mock catalogs for survey analysis.

1.5 Thesis Outline

We first address structure formation in the mildly nonlinear regime: Chapters 2 and 3 present a new technique to constrain the baryon acoustic oscillation distance scale using a tailored set of bispectrum measurements. We then turn to the deeply nonlinear regime and the internal structure of halos. Chapter 4 presents measurements of the concentration-mass relation in simulations, its evolution with redshift, and a new fitting formula for the mass- and redshift-dependence of concentration. In Chapter 5, we study the relationship between halo concentration and formation history. Finally, in Chapter 6 we conclude by summarizing the results and discussing promising directions for further work on these topics.

CHAPTER 2

BISPECTRUM AS BARYON ACOUSTIC OSCILLATION INTERFEROMETER

The galaxy bispectrum, measuring excess clustering of galaxy triplets, offers a probe of dark energy via baryon acoustic oscillations (BAOs). However up to now it has been severely underused due to the combinatorically explosive number of triangles. Here we exploit interference in the bispectrum to identify triangles that amplify BAOs. This approach reduces the computational cost of estimating covariance matrices, offers an improvement in BAO constraints equivalent to lengthening BOSS by 30% and simplifies adding bispectrum BAO information to future large-scale redshift survey analyses.

This chapter appears in Physical Review D as "Bispectrum as Baryon Acoustic Oscillation Interferometer" [45].

2.1 Introduction

The baryon acoustic oscillation (BAO) method exploits the imprint of sound waves in the prerecombination plasma on the late-time clustering of galaxies to measure the expansion history of the Universe and constrain the dark energy equation of state [30, 83, 141, 180, 278]. Applied to Baryon Oscillation Spectroscopic Survey (BOSS) data, the BAO method has offered 1% distance constraints [2, 262]; future surveys such as DESI [1] and Subaru PFS [305] promise to tighten these to subpercent precision. The BAO precision from the power spectrum $P(k)$ and two-point correlation function is further improved by reconstruction [84, 226, 235, 236, 270], which uses the density field as sampled by galaxies to partially reverse the smearing effects of nonlinear structure formation on the BAO peak.

Direct measurements of higher-point functions may yet provide an additional improvement on BAO constraints, and recent algorithms [272, 288, 289, 291, 299] have reduced the

1. `http://desi.lbl.gov`

computational complexity of calculating three-point statistics. The BAO feature has recently been detected in the three-point correlation function [102, 292, 293] and bispectrum $B(k_1, k_2, k_3)$ [239]. Like reconstruction, three-point information can improve constraints on the BAO scale: for example, Ref. [288] finds a 6% improvement [292] using CMASS data. However the high number of triangles necessitates a large number of mock catalogs to directly compute covariance matrices. One approach to this challenge is an analytic covariance template [288], which improves the covariance matrix calculated from a smaller number of mocks. Alternatively, using only a diagonal covariance matrix [105] or measuring bispectra on only a subset of all possible triangles dramatically reduces the size of the covariance matrix. However, simple rules for selecting triangles (e.g. isosceles, or one side an integer multiple of another [105, 106, 202]) may be far from optimal for probing BAOs.

Here we identify triangle configurations that maximize or minimize the BAO signal in the bispectrum, enabling precise BAO constraints with relatively few bispectrum measurements. Our approach allows intuitive visualization of bispectra as functions of a single variable, because we set two of the three wave numbers to depend on the first.

2.2 Perturbation Theory Model

We first explore the perturbation theory (PT) bispectrum $B^{\rm PT}$ to study the BAO feature's dependence on triangle configuration. In contrast to the power spectrum, the bispectrum depends on a closed triangle formed by the three wave vectors $(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$. We consider the isotropic bispectrum, where six degrees of freedom are redundant, so we can specify a triangle by e.g. three wave numbers (k_1, k_2, k_3). The tree-level matter bispectrum in real space (i.e., without redshift-space distortions) [273] is

$$B^{\rm PT}(k_1, k_2, k_3) = 2P^{\rm L}(k_1)P^{\rm L}(k_2)F_2(\mathbf{k}_1, \mathbf{k}_2) + \text{cyc.}, \tag{2.1}$$

with

$$F_2(\mathbf{k}_i, \mathbf{k}_j) = \frac{5}{7} + \frac{1}{2}\left(\frac{k_i}{k_j} + \frac{k_j}{k_i}\right)(\hat{\mathbf{k}}_i \cdot \hat{\mathbf{k}}_j) + \frac{2}{7}(\hat{\mathbf{k}}_i \cdot \hat{\mathbf{k}}_j)^2, \tag{2.2}$$

where $P^{\rm L}(k)$ is the *linear* matter power spectrum. We refer to $2P^{\rm L}(k_1)P^{\rm L}(k_2)F_2(\mathbf{k}_1, \mathbf{k}_2)$ as the precyclic term and to the terms denoted by cyc as the postcyclic terms.

The linear power spectrum involves the square of the matter transfer function:

$$P^{\rm L}(k) = P_{\rm pri}(k)T_{\rm m}^2(k) \tag{2.3}$$

where $P_{\rm pri}(k)$ is the primordial power spectrum.

We split the transfer function into smooth and oscillatory pieces as $T_{\rm m}(k) = T_{\rm sm}(k) + \epsilon(k)j_0(k\tilde{s})$, where $T_{\rm sm}(k)$ and $\epsilon(k)$ are smooth functions of k [82], the oscillations come from BAOs, and ϵ is small (as $\Omega_{\rm b}/\Omega_{\rm m} \ll 1$). The effective sound horizon $\tilde{s}(k)$ vanishes at low k and is within 1% of the true sound horizon for $k \gtrsim 0.1\,h/{\rm Mpc}$ [82]; we use $\tilde{s}_f \equiv \tilde{s}(k_f)$ at the fiducial wave number $k_f = 0.2\,h/{\rm Mpc}$. The spherical Bessel function $j_0(k\tilde{s}_f)$ has wavelength $2\pi/\tilde{s}_f$ with a decaying envelope $1/(k\tilde{s}_f)$. In each term of $B^{\rm PT}$ [Eq. (2.1)], BAOs enter through the product of transfer functions $T_{\rm m}^2(k_i)T_{\rm m}^2(k_j)$. To leading order in ϵ, the oscillatory part scales as

$$\frac{\left[T_{\rm m}^2(k_i)T_{\rm m}^2(k_j)\right]_{\rm osc}}{T_{\rm sm}^2(k_i)T_{\rm sm}^2(k_j)} \propto \frac{\epsilon(k_i)}{T_{\rm sm}(k_i)}\frac{\sin(k_i\tilde{s}_f)}{k_i\tilde{s}_f} + \frac{\epsilon(k_j)}{T_{\rm sm}(k_j)}\frac{\sin(k_j\tilde{s}_f)}{k_j\tilde{s}_f}. \tag{2.4}$$

In the precyclic term, defined below Eq. (2.1), interference will therefore depend on the phase difference between k_1 and k_2 in units of $2\pi/\tilde{s}_f$ (i.e. the BAO wavelength in Fourier space), motivating the triangle parametrization

$$k_1, \quad k_2 - k_1 = \delta\left(\frac{\pi}{\tilde{s}_f}\right), \quad \cos\theta = \hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2. \tag{2.5}$$

When δ is an odd integer, interference in the precyclic term is perfectly destructive, suppressing BAOs; for even integer δ, interference is perfectly constructive, amplifying BAOs. At

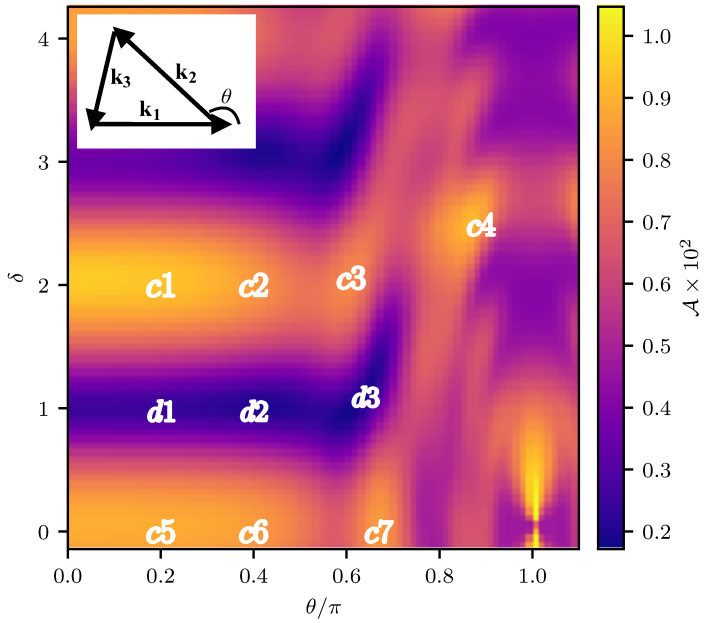

Figure 2.1: The standard deviation $\mathcal{A}$ of the bispectrum BAO feature in triangle configurations parametrized by (δ, θ); the inset shows the definition of θ. We measure labeled configurations in simulations to improve constraints on the BAO scale.

fixed δ, θ determines both the phase structure of the postcyclic terms and, through F_2, the relative contributions of all three terms to $B^{\rm PT}$. We compute $B^{\rm PT}$ for a range of $k_2(k_1)$ and all possible triangular $k_3(k_1, k_2)$ and select the configurations that maximize or minimize the BAO features. We refer to these as "constructive" and "destructive" configurations, where "configuration" means a set of triangles with fixed δ and θ over a range of k_1.

The amplitude of the bispectrum does include BAOs and redshift-space distortion information, but here we focus on the oscillatory behavior, isolated in the ratio

$$R(k_1, \delta, \theta) = \frac{B^{\rm PT}(k_1, \delta, \theta)}{B^{\rm PT}_{\rm nw}(k_1, \delta, \theta)}. \tag{2.6}$$

$B^{\rm PT}$ involves a power spectrum from CAMB [175]; we use a flat ΛCDM cosmology with $\Omega_{\rm m} = 0.2648$, $\Omega_{\rm b}h^2 = 0.02258$, $n_s = 0.963$, $\sigma_8 = 0.8$, and $h = 0.71$. In this cosmology, $\tilde{s}_f \approx 109.5\,{\rm Mpc}/h$ via the fit of Ref. [82]. $B^{\rm PT}_{\rm nw}$ is the bispectrum computed using the no-wiggle power spectrum [82].

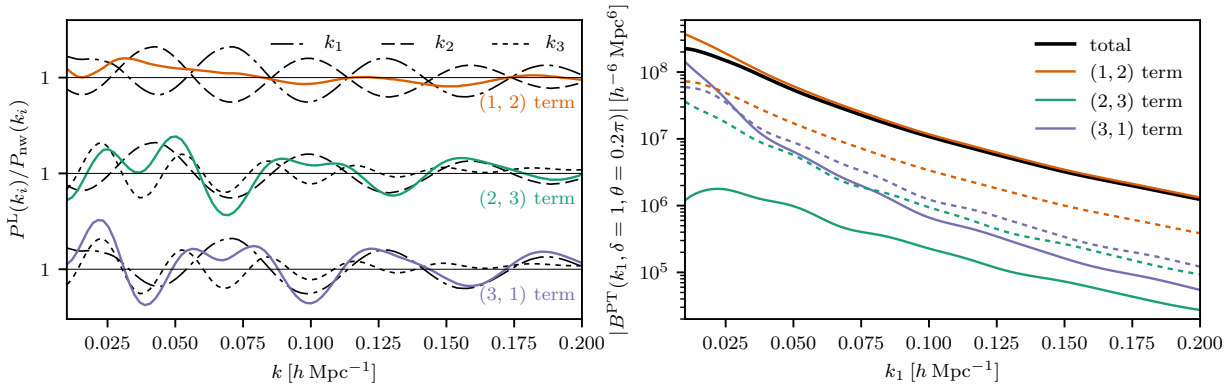

Figure 2.2: For a destructive configuration such as d_1 (with $\delta = 1$ and $\theta = 0.2\pi$), BAOs are out of phase in the dominant term or terms of $B^{\rm PT}$. Left: Interference of the BAO feature for each pair of wave numbers. Black curves show the ratio of the linear to the no-wiggle power spectrum, $P^{\rm L}(k_i)/P_{\rm nw}(k_i)$, for a single wave number; the product of each pair of ratios is shown in color. For example, the oscillations in the k_1 and k_2 terms are out of phase, so the product $P^{\rm L}(k_1)P^{\rm L}(k_2)/[P_{\rm nw}(k_1)P_{\rm nw}(k_2)]$ (orange) shows nearly no oscillation. Right: $B^{\rm PT}$ and its three terms: the products of power spectra weighted by the F_2 kernel (solid, color curves) that sum to the total $B^{\rm PT}$ (solid, black curves). Dashed curves show products of power spectra before weighting by the F_2 kernel. The $(1, 2)$ term (orange) is the primary contribution to $B^{\rm PT}$; because the BAOs destructively interfere in this term, $B^{\rm PT}$ shows little oscillation and so the configuration is destructive.

Constructive interference enhances "wiggles" and increases the standard deviation, denoted $\mathcal{A}$, of Eq. (2.6). We thus quantify BAO interference by

$$\mathcal{A}^2(\delta, \theta) \equiv \int_{0.01}^{0.2} \left[R(k_1, \delta, \theta) - \bar{R}(\delta, \theta)\right]^2 \frac{{\rm d}k_1}{[h/{\rm Mpc}]}, \tag{2.7}$$

where $\bar{R}(\delta, \theta)$ is the mean of $R(k_1, \delta, \theta)$ on the same range, $0.01 \leq k_1/[h/{\rm Mpc}] \leq 0.2$. Beyond the lower edge of this range, cosmic variance will limit the usefulness of bispectrum measurements, while beyond the upper end, Silk damping as well as late-time nonlinear structure formation damp BAOs. We display $\mathcal{A}$ in Figure 2.1. This figure is a guide to identify constructive bispectrum configurations, where we expect the strongest BAO signal.

Three effects combine to determine the oscillations' amplitude in each configuration. First, the broadband behavior of $P^{\rm L}(k)$ sets the magnitude of each $P^{\rm L}(k_i)P^{\rm L}(k_j)$ permutation. Second, the multiplication by $F_2(k_i, k_j)$ modulates this overall magnitude. Thus in any

configuration the dominant term will be the one with large $P^{\rm L}(k_i)P^{\rm L}(k_j)$, further enhanced by a large $F_2(k_i, k_j)$. Third, the amplitude of oscillation in each $P^{\rm L}(k_i)P^{\rm L}(k_j)$ depends on whether the oscillations in $P^{\rm L}(k_j)$ are in or out of phase with the oscillations in $P^{\rm L}(k_i)$. Therefore, the phase shift in the dominant term (or terms) sets whether the configuration is constructive (amplifying BAOs) or destructive (suppressing BAOs). As an example, the contributions to $B^{\rm PT}$ in a destructive configuration, d_1, are shown in Figure 2.2.

2.3 Simulations

Motivated by the above perturbative analysis, we now explore the power of our interferometric approach to constrain the BAO scale in full N-body simulations, which accurately solve for the nonlinear structure formation giving rise to the bispectrum. We measure halo power spectra and bispectra from four $z = 0.55$ MockBOSS halo catalogs [301] with box size $L = 4000\,{\rm Mpc}/h$ and particle mass $m_p = 6.8 \times 10^{10}\,{\rm M}_\odot/h$ using a friends-of-friends (FOF) finder with linking length 0.168. Our threshold halo mass is $M_{\rm h} \geq 10^{13}\,{\rm M}_\odot/h$ (100 particles in the FOF group), giving number density $n = 3.8 \times 10^{-4} h^3/{\rm Mpc}^3$. Our halo sample is thus roughly matched to the number density ($3 \times 10^{-4} h^3/{\rm Mpc}^3$) and average redshift ($z = 0.57$) of the BOSS CMASS galaxy sample [62]. To measure $B(k_1, \delta, \theta)$ and $P(k)$ from the halo catalog, we use an FFT-based algorithm [15] with aliasing correction via interlacing [276].

Constraining power differs between constructive and destructive configurations, so we focus on ten triangle configurations shown in Figure 2.1: seven constructive ($c_1, \ldots, c_7$) and three destructive (d_1, d_2, d_3). We ignore the maximum-amplitude configurations with $\theta = \pi$; these correspond to three collinear points in Fourier space, suggesting they are less independent from the power spectrum [42]. Additionally, the small values of $k_3 \sim 10^{-3}\,h/{\rm Mpc}$ for these configurations make them subject to cosmic variance in practice, as there are few modes this large in typical surveys (e.g. BOSS or DESI).

Each configuration comprises 19 triangles because k_1 varies in 19 bins between $0.01\,h/{\rm Mpc}$ and $0.2\,h/{\rm Mpc}$. In each bin, the other two wave numbers k_2 and k_3 are computed from k_1

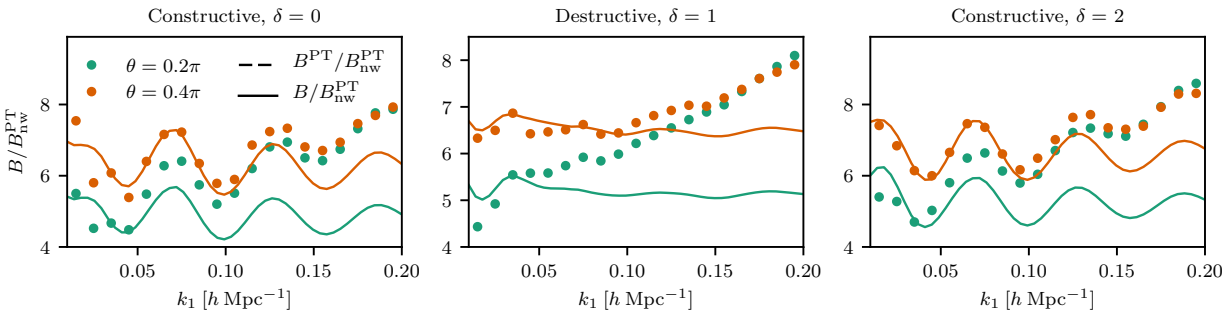

Figure 2.3: Bispectrum results for four constructive (left and right panels) and two destructive (middle panel) configurations measured from the MockBOSS simulations. Points show the ratio of the simulated bispectrum B to the PT no-wiggle bispectrum $B^{\rm PT}_{\rm nw}$. Curves show the ratio of the PT bispectrum $B^{\rm PT}$ to the no-wiggle $B^{\rm PT}_{\rm nw}$. Each PT curve $B^{\rm PT}/B^{\rm PT}_{\rm nw}$ oscillates about $B/B^{\rm PT}_{\rm nw} = 1$ but has been multiplied by a constant to account for linear bias and allow comparison with corresponding measurements.

according to the parameters δ and θ, held fixed for any one configuration. Thus across the ten configurations we measure 190 triangles in total. This number represents roughly one-fourth of the 805 triangles that can be formed from three sides in our range and binning. We note that some triangles we measure have one or two sides exceeding $0.2\,h$/Mpc; the maximum k_2 we measure is $0.3\,h$/Mpc, while k_3 reaches $0.45\,h$/Mpc. Little BAO information comes from these scales, as the BAO feature is heavily damped for $k \gtrsim 0.2\,h$/Mpc.

Figure 2.3 shows that simulated bispectra for destructive configurations have no BAO feature, while pronounced oscillations appear in the constructive configuration bispectra. Nonlinear effects are more important for the configurations with larger θ: as θ rises, so does k_3. As θ increases, we therefore expect the measured $B(k_1, \delta, \theta)$ to depart from $B^{\rm PT}(k_1, \delta, \theta)$ at smaller k_1, reflecting the increasingly nonlinear behavior of $P(k_3)$.

We perform a Fisher matrix analysis to estimate the improvement in BAO precision offered by our bispectrum approach. We use the 1σ uncertainty σ_α in the shift parameter α as a measure of constraining power. We introduce nuisance parameters A_P and $\mathbf{A_B}$ to model possible uncertainties (e.g. due to halo bias) in the amplitude of $P(k)/P_{\rm nw}(k)$ and each $B(k_1, \delta, \theta)/B^{\rm PT}_{\rm nw}(k_1, \delta, \theta)$, respectively. That is, our models for the power spectrum and

the bispectrum for the ith triangle configuration are

$$P(k)/P_{\rm nw}(k) = A_P P(\alpha k)/P_{\rm nw}(\alpha k), \tag{2.8}$$

$$R_i(k_1, \delta, \theta) = A_{B,i} R_i(\alpha k_1, \delta, \theta), \tag{2.9}$$

where R_i denotes the ratio of the bispectrum B_i to the corresponding no-wiggle bispectrum as in Eq. (2.6). We test an additional set of nuisance parameters such that $P(k)/P_{\rm nw}(k) = A_P\,[P(\alpha k)/P_{\rm nw}(\alpha k)] + C_p k$ and $R_i(k_1, \delta, \theta) = A_{B,i} R_i(\alpha k_1, \delta, \theta) + C_{B,i} k_1$. The parameters C_P and $\mathbf{C_B}$ are motivated by the behavior of the R_i shown in Figure 2.3, which rise linearly as k_1 increases. Marginalizing over all four nuisance parameters changes our results by $\lesssim 10\%$, similar to other sources of error in our rough Fisher analysis.

The Fisher matrix is

$$F_{\mu\nu} = \frac{\partial \mathbf{d}^T}{\partial p_\mu} \mathbf{C}^{-1} \frac{\partial \mathbf{d}}{\partial p_\nu}, \tag{2.10}$$

where $p_\mu \in \{\alpha, A_P, \mathbf{A}_B\}$, $\mathbf{d}$ is the data vector of power spectrum and bispectrum measurements divided by their no-wiggle analogs, and $\mathbf{C}^{-1}$ is the inverse covariance matrix. The dimension of the data vector is 209: 19 bins for $P(k)$ plus 190 triangles. To estimate the covariance matrix, we subdivide each of the four $(4\,{\rm Gpc}/h)^3$ simulations into subvolumes of $(500\,{\rm Mpc}/h)^3$. The resulting 2048 subvolumes exceed the 209 power spectrum and bispectrum bins by roughly an order of magnitude, and the derived covariance matrix is well conditioned. When computing the inverse of the covariance matrix, we include the correction factor in Ref. [125] to obtain its unbiased estimate.

To compute the partial derivatives of $P(k)/P_{\rm nw}(k)$ and $B(k_1)/B^{\rm PT}_{\rm nw}(k_1)$ with respect to α, we stretch and compress the simulation box by a factor $(1 \pm \epsilon)$ with $\epsilon = 0.05$, recompute the measured $P(k)$ and $B(k_1, \delta, \theta)$, and correct for the change in amplitude due to the change in box volume. This correction, the Jacobian of the integration measure between Fourier and real space, is $(1+\epsilon)^3$ for $P(k)$ and $(1+\epsilon)^6$ for the bispectrum [225, 278]. With these amplitude shifts corrected, the power spectrum oscillations are simply stretched and

compressed relative to the standard box. The oscillations as a function of k_1 are stretched in the bispectrum measurements as well. But even configurations where oscillations are suppressed contain BAO scale information: when $\alpha \neq 1$, configurations are selected based on an incorrect estimate of the sound horizon, so the measured configurations depart from the desired configurations. Consequently, the *amplitude* of oscillation in $B(k_1)/B^{\rm PT}_{\rm nw}(k_1)$ changes as well as the frequency. This behavior provides a small additional constraint on α—even destructive configurations contain BAO scale information.

The variance in α including marginalization over the amplitude parameters is

$$\sigma_\alpha^2 = (\mathbf{F}^{-1})_{\alpha\alpha}. \tag{2.11}$$

We compute σ_α for $P(k)/P_{\rm nw}(k)$ alone and for $P(k)/P_{\rm nw}(k)$ combined with $B(k_1)/B^{\rm PT}_{\rm nw}(k_1)$ for different sets of triangle configurations. We also compute $P(k)/P_{\rm nw}(k)$ for the postreconstruction $P_{\rm r}(k)$, where reconstruction uses the algorithm of Ref. [236] with bias measured from $P(k)$ and a smoothing scale of $R_s = 15\,{\rm Mpc}/h$. Our results are unchanged at the percent level with a smaller smoothing scale of $R_s = 5\,{\rm Mpc}/h$. Both reconstruction and the addition of bispectrum information decrease the uncertainty in α. Computing the improvement in σ_α relative to the prereconstruction, $P(k)$-only constraint, we find that the relative improvement depends on the number and choice of triangle configurations. Three destructive configurations (57 triangles) only improve σ_α by less than 3%, but a different 57 triangles in three constructive triangle configurations (c_1, c_2, and c_3) improve σ_α by roughly 8%. With an additional four constructive configurations (for a total of 133 triangles in seven constructive configurations), the improvement reaches 12%. The marginal improvement from each configuration depends as well on its covariance with previously selected configurations; future work will explore this dependence more fully. All ten triangle configurations (seven constructive and three destructive) provide an improvement of roughly 14% over the precision from $P(k)$ alone. This initial choice of ten configurations does not exhaust the information avail-

Table 2.1: At all redshifts tested, reconstruction gives roughly twice the improvement in α of our ten bispectrum configurations.

Redshift	n (h^3/Mpc^3)	Reconstruction	Bispectrum
0.15	3.8×10^{-4}	38%	17%
0.55	3.8×10^{-4}	32%	14%
0.8	8.5×10^{-4}	26%	10%

able in the bispectrum; the seventh and last constructive configuration, for example, still raises the improvement in constraints by about 20%. We thus anticipate that opportunities remain to further improve constraints on α by strategically selecting triangle configurations with minimal covariance. We will study the covariance structure of the bispectrum and its implications for triangle selection in detail in future work.

Fisher analysis also gives an estimate of how constraints from the bispectrum compare to reconstruction, currently the best available technique to improve BAO precision. Following the same Fisher analysis, we estimate that the reconstructed $P_{\rm r}(k)$ improves on the prereconstruction α constraint by roughly 30%. This is roughly twice the improvement we find using ten bispectrum triangle configurations. Both bispectrum and reconstruction perform better at lower redshift, as shown in Table 2.1. At all three redshifts, reconstruction still gives roughly twice the improvement of our ten bispectrum configurations. We note that additional configurations will likely narrow this gap [113], and the comparison between bispectrum and reconstruction improvements does depend on the range of wave numbers used in the analysis. Decreasing the maximum wave number $k_{\rm max}$ from the value of $0.2\,h$/Mpc used above, bispectrum information provides more improvement in σ_α while reconstruction provides less. The bispectrum starts to perform better than reconstruction below $k_{\rm max} \sim 0.15\,h$/Mpc; for example, at $k_{\rm max} = 0.14\,h$/Mpc and $z = 0.55$, ten bispectrum configurations provide a 20% improvement, compared to 18% from reconstruction.

2.4 Discussion

Measuring the bispectrum for only 190 triangles can improve constraints on the BAO scale by 14%, corresponding to an increase in survey time of roughly 30%. Our 14% improvement with tailored triangles is comparable to the roughly 10% improvement found by Ref. [239] using a more complete set of triangles, and the 6% improvement of Ref. [288]. More critically, our central finding is that the improvement depends not only on the *number* of bispectrum measurements, but on the *choice* of measurements. By selecting triangle configurations where interference effects amplify the BAO feature in the bispectrum, we obtain constraints with relatively few measurements, decreasing the number of mock catalogs needed to estimate the covariance matrix $\mathbf{C}$. Our method also opens a new avenue for numerically obtaining the cross covariance between P and B, which is less easily treated with an analytic template than is the autocovariance. Additionally, our (k_1, δ, θ) parametrization enables visualization of the bispectrum BAO feature in simple 1D plots (Figure 2.3).

Reconstruction is a dynamical method that, applied to a single realization of the Universe, partially removes nonlinear effects using the full density field information—including information that is not captured even by higher-order statistics. Bispectrum measurements of BAOs are operationally independent from reconstruction, so agreement between the two methods will demonstrate robust measurements of the BAO scale. For example, the bispectrum provides an additional check for sources of error in reconstruction, such as those described in Ref. [282]: incorrect assumptions of bias, redshift-space distortions, or distance parameters. With better understanding of the covariance between the postreconstruction power spectrum and the bispectrum triangles most relevant for BAOs, it may be possible to combine pre- and postreconstruction measurements to further improve constraints on the BAO scale. Furthermore, the bispectrum can easily be extended to constrain BAOs even in modified gravity models. Once measured, bispectra can simply be compared to any modified gravity model; reconstruction, in contrast, must be modified according to each specific model and reapplied to the data before repeating power spectrum measurements to constrain

BAOs.

Our technique highlights phase effects; it thus may constrain sources of phase shifts in the power spectrum. One source of a phase shift is an isocurvature perturbation, where the oscillation is proportional to $\cos k\tilde{s}$ instead of $\sin k\tilde{s}$ [140]. We expect that destructive configurations (as identified in the no-isocurvature model) may now show a BAO feature. In future work, we will explore this phase shift as well as that induced by relativistic species such as neutrinos [19, 20, 90].

Future work will also discuss the dependence of BAO amplitude on the triangle parameters θ and δ, explore whether additional triangle configurations offer any improvement in constraints, and study the independence of bispectrum information from that used in reconstruction. When combined with reconstruction, our result may represent a further improvement in BAO precision. Whether or not it does so on a statistical level, bispectrum measurements are operationally independent from reconstruction and therefore subject to different systematic effects. The bispectrum thus offers at the very minimum a cross-check that, added to those of Ref. [2], will be valuable for analysis of BAOs in future large-scale surveys.

CHAPTER 3

A PHYSICAL PICTURE OF BISPECTRUM BARYON ACOUSTIC OSCILLATIONS IN THE INTERFEROMETRIC BASIS

We present a picture of the matter bispectrum in a novel "interferometric" basis designed to highlight interference of the baryon acoustic oscillations (BAOs) in the power spectra composing it. Triangles where constructive interference amplifies BAOs provide stronger cosmic distance constraints than triangles with destructive interference. We show that the amplitude of the BAO feature in the full cyclically summed bispectrum can be decomposed into simpler contributions from single terms or pairs of terms in the perturbation theory bispectrum, and that across large swathes of our parameter space the full BAO amplitude is described well by the amplitude of BAO in a single term. The dominant term is determined largely by the $F^{(2)}$ kernel of Eulerian standard perturbation theory. We present a simple physical picture of the BAO amplitude in each term; the BAO signal is strongest in triangle configurations where two wavenumbers differ by a multiple of the BAO fundamental wavelength.

This chapter has been submitted for publication in Monthly Notices of the Royal Astronomical Society as "A Physical Picture of Bispectrum Baryon Acoustic Oscillations in the Interferometric Basis" [44].

3.1 Introduction

The Baryon Acoustic Oscillation (BAO) method [30, 83, 141, 180, 278] has become a central means of pursuing the essential nature of dark energy, a mysterious substance making up roughly 72% of the present-day Universe. The BAO method uses the imprint of sound waves in the early Universe on the late-time clustering of galaxies to probe the cosmic expansion history, which through general relativity is linked to dark energy.

The BAO method has been applied to the galaxy 2-point correlation function (2PCF) and power spectrum [2, 262], as well as more recently to the galaxy 3-point correlation function (3PCF) [102, 292, 293] and bispectrum [239], measuring respectively 2- and 3-point clustering over random in configuration space and Fourier space. While measurements of the bispectrum and 3PCF have improved BAO constraints over those of the power spectrum alone, optimal constraints are difficult to obtain given the large number of bispectrum triangles. Bispectrum covariance matrices are often estimated from mock catalogs ([102, 239]; though other approaches exist, see Refs. [292, 293]). In order to properly estimate covariance matrices, the number of mock catalogs must greatly exceed the number of triangles [241]; when many triangles are used to constrain BAO, the number of mocks needed is unrealistic with present computational resources. For example, Ref. [239] measured the bispectrum for a large number of triangles, but noted that their error bars were limited by the number of mocks available to estimate the covariance matrix; more mocks would improve their 1.1% precision joint constraints substantially to 0.7%, a gain of more than 30%.

With better understanding of which bispectrum triangles are most sensitive to BAO, future studies could obtain better BAO constraints with a smaller set of triangles and, therefore, a smaller covariance matrix. These triangles could be identified by measuring all bispectrum triangles and their covariances, but such an approach also faces the problem of limited mock catalogs. Because fully N-body mocks cannot presently provide a good estimate of the full covariance matrix of all bispectrum triangles, we must select a set of optimal triangles for BAO constraints—without knowledge of the full covariance matrix.

Chapter 2 proposed one technique to select bispectrum measurements that are sensitive to BAO. We highlighted that BAO in the bispectrum constructively interfere in certain triangle configurations, amplifying the BAO signal. In a short work, we measured bispectra only on triangles where the BAO signal is amplified. With this relatively small set of bispectrum measurements we found substantial improvements in BAO constraints over power spectrum measurements alone, equivalent to lengthening BOSS by roughly 30%. Our method for

triangle selection greatly reduced the number of bispectrum measurements necessary to obtain such an improvement.

In detail, at leading order the bispectrum involves products of two power spectra; each power spectrum introduces an oscillatory BAO feature. When the oscillations are in phase, they amplify the BAO signal. In Chapter 2, we showed that this constructive interference increases the amplitude of BAO in the bispectrum. We introduced a new parametrization of bispectrum triangles: instead of the three triangle sides k_1, k_2, and k_3, we use the length of one triangle side k_1, the difference in length of the second from the first in units of the BAO fundamental wavelength, and the angle between them θ. We computed the BAO amplitude for a selection of triangle configurations, producing a map of the root-mean-square (RMS) amplitude as a function of the length difference and the opening angle. Using this RMS map, we can surgically identify the configurations most suitable for studies of BAO in the bispectrum.

The "interferometric basis" proposed in our earlier work promises other applications beyond improvement in BAO constraints with relatively few bispectrum measurements. First, since our method identifies the triangles that are most sensitive to BAO, it offers an approach to more efficiently investigate the independence of bispectrum information from that obtained via reconstruction [84, 226, 235, 236] and the covariance between the bispectrum and power spectrum. Second, our parameterization allows intuitive visualization of BAO in the bispectrum. Third, our interferometric approach is sensitive to phase shifts associated with $N_{\rm eff}$ (such as that driven by relativistic neutrinos at high redshift [20]), spinning particles in the early Universe [210], or relative velocities between baryons and dark matter [61, 294, 311, 340].

In this chapter, we investigate more fully the physics of BAO in the interferometric basis. The bispectrum is a sum of three cyclic terms, but for many triangles, the cyclic sum is dominated by only one or two of the three terms. Which term dominates is determined primarily by the $F^{(2)}$ kernel of Eulerian standard perturbation theory (SPT). Products of

power spectra also enter the leading-order perturbation theory (PT) bispectrum, but their role in the dominance structure is secondary; instead, they introduce the oscillatory features whose interference is highlighted by our basis.

When the bispectrum is dominated by only one or two terms, the amplitude of BAO in the full bispectrum can be approximated by the BAO amplitude in the dominant term or terms. We show analytically that because BAO are a small feature in the power spectrum, the difference between the RMS amplitude computed under this approximation and the full RMS amplitude vanishes at leading order. We can therefore decompose the full RMS map into approximate "eigen-RMSes" computed from individual terms. Numerical work verifies that this decomposition successfully reproduces the primary features of the full RMS map.

To understand these features, we study the behavior of BAO in each term and pair of terms that can dominate the bispectrum. The structure of the power spectrum, in particular BAO and their envelope due to Silk damping, determines the RMS amplitude. For each triangle configuration, the BAO amplitude in each term is driven by one of four interactions between power spectra: interference, incoherence, feathering, or single power spectrum. The first, interference, can dramatically amplify BAO amplitude in certain configurations, like those used in Chapter 2 to constrain the BAO scale.

In general, then, the $F^{(2)}$ kernel determines which pairs of power spectra set the amplitude of BAO in the measured bispectrum. The remainder of the chapter details this broad picture as follows. In §3.2, we review the interferometric basis as presented in our earlier work, and define the RMS amplitude of the ratio of physical to "no-wiggle" bispectrum we use in our analysis. §3.3 introduces notation used throughout. §3.4 shows which triangle configurations are dominated by a single term of the bispectrum cyclic sum; in these regions, the BAO signal in the bispectrum simplifies to the BAO signal in a single term, as shown analytically in §3.5. In §3.6, we numerically calculate the BAO amplitude in each term of the cyclic sum. These "eigen-RMSes" approximate the BAO amplitude in the full bispectrum in the regions where the corresponding terms dominate, and they assemble into a picture

that matches the full map of BAO amplitude. §3.7 presents implications of our study for the reduced bispectrum, the 3PCF, and multipole expansions of the 3PCF and bispectrum. §3.8 concludes.

Throughout we adopt a spatially flat ΛCDM cosmology at $z = 0$ consistent with the WMAP-7 [163] parameters of the MockBOSS simulations [301] used in Chapter 2: $\Omega_{\rm m} = 0.2648$, $\Omega_{\rm b}h^2 = 0.02258$, $n_s = 0.963$, $\sigma_8 = 0.80$, and $h = 0.71$.

3.2 Interferometric Basis

Several sets of triangle parameters have been used in previous works for the isotropic bispectrum and three-point correlation function (3PCF). Refs. [274], [157], and [15] considered the bispectrum for equilateral or isosceles triangles. Refs. [273], [271], [26], [274], [16], [107], and [134] used one side k_1, the ratio of a second side to the first k_2/k_1, and the angle between the two θ_{12}; the third side can also be parameterized simply as k_3 [239] or by its ratio to k_1 [15, 149]. Refs. [271] and [275] allowed each triangle side to be any integer multiple of the bin width $\Delta k \simeq 0.015\ h/\mathrm{Mpc}$.

Like the bispectrum, the 3PCF can be parametrized using one side r_1, the ratio of a second side to the first r_2/r_1, and the opening angle θ [164, 198, 199]. Other parameterizations use two sides r_1 and r_2. The third parameter can be the opening angle θ between them [114, 115, 201, 202], the cosine μ of the opening angle [102], a shape parameter combining the three side lengths [150, 223, 240, 327], or a multipole expansion of the dependence on opening angle [237, 292, 293, 303]. Many studies of the anisotropic bispectrum and 3PCF, which retain information about the line of sight, have also employed a multipole basis with respect to the line of sight in redshift space (e.g. Refs. [41, 69, 94, 216, 338, 339]). Refs. [291] and [299] use a spherical harmonic expansion of the 3PCF and bispectrum, which includes information on both the internal angle and the angles to the line of sight.

In Chapter 2, we introduced a new basis for bispectrum work motivated by the physics of BAO in the bispectrum. The bispectrum B involves products of linear matter power spectra

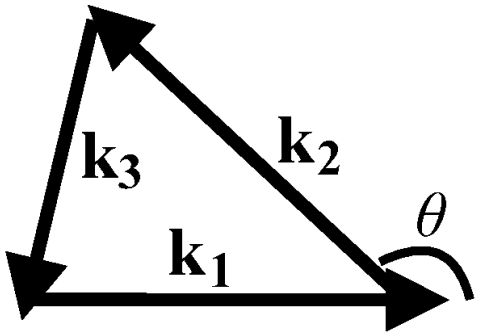

Figure 3.1: The triangle parameter θ is defined in equation (3.3) as the exterior angle between $\mathbf{k}_1$ and $\mathbf{k}_2$.

P (e.g. Ref. [273]), as

$$B(k_1, k_2, k_3) = 2P(k_1)P(k_2)F^{(2)}(k_1, k_2; \hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2) + \text{cyc.} \tag{3.1}$$

We refer to $2P(k_1)P(k_2)F^{(2)}(k_1, k_2; \hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2)$ as the pre-cyclic term, and to the terms denoted by cyc. as the post-cyclic terms. $F^{(2)}$ is the Eulerian standard perturbation theory kernel that generates a second-order density field when integrated against two linear density fields. The $F^{(2)}$ kernel depends only on two side lengths k_i and k_j and the angle between them (through the dot product $\hat{\mathbf{k}}_i \cdot \hat{\mathbf{k}}_j$):

$$F^{(2)}(k_i, k_j; \hat{\mathbf{k}}_i \cdot \hat{\mathbf{k}}_j) = \frac{5}{7} + \frac{1}{2}\left(\frac{k_i}{k_j} + \frac{k_j}{k_i}\right)(\hat{\mathbf{k}}_i \cdot \hat{\mathbf{k}}_j) + \frac{2}{7}(\hat{\mathbf{k}}_i \cdot \hat{\mathbf{k}}_j)^2. \tag{3.2}$$

We will refer to the middle term of equation (3.2), $(k_i/k_j + k_j/k_i)$, as the dipole contribution to $F^{(2)}$.

The power spectra in equation (3.1) have BAO features that oscillate sinusoidally. These features can thus interfere, motivating us to consider a parameterization of the bispectrum that transparently captures the phase structure. In particular, we set up our parametrization to capture the phase structure of the pre-cyclic term (first term in equation 3.1) in k_1 and

k_2, as follows:

$$k_1, \quad k_2 - k_1 = \delta \left(\frac{\lambda_f}{2} \right), \quad \cos\theta = \hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2. \tag{3.3}$$

Throughout this work we assume without loss of generality that δ is positive, so $k_2 > k_1$. The external angle θ is shown in Figure 3.1. The fundamental wavelength of the BAO in Fourier space λ_f is given by

$$\lambda_f = \frac{2\pi}{\tilde{s}_f} \approx 0.0574 \; h/\mathrm{Mpc}, \tag{3.4}$$

where $\tilde{s}_f = 109.5 \; \mathrm{Mpc}/h$ is the effective sound horizon evaluated at a fiducial wavenumber $k_f = 0.2 \; h/\mathrm{Mpc}$. As discussed in Ref. [82], the lowest-wavenumber nodes of the baryonic transfer function occur at higher wavenumber than the nodes of $\sin ks$. The effective sound horizon grows with k for $k \lesssim 0.05 \; h/\mathrm{Mpc}$ and asymptotes to the sound horizon s for $k \gtrsim 0.05 \; h/\mathrm{Mpc}$; we define λ_f according to its asymptotic value at $k_f = 0.2 \; h/\mathrm{Mpc}$.

We use "configuration" to describe a set of triangles with fixed δ and θ over the range $0.01 \leq k_1/[h/\mathrm{Mpc}] \leq 0.2$. For each configuration, we divide k_1 into 100 bins of width $1.9 \times 10^{-3} \; h/\mathrm{Mpc}$. The other two wavenumbers k_2 and k_3 are calculated from each k_1 according to equation (3.3). We study configurations with $0 \leq \delta \leq 4.25$ and $0 \leq \theta \leq 1.1$, with 80 points in θ and 80 in δ for a total of 6400 configurations. We sample many (δ, θ) points simply to produce well-resolved figures, but we note that configurations that are close to each other in (δ, θ) space are highly covariant.

We restrict k_1 to the range $0.01 \leq k_1/[h/\mathrm{Mpc}] \leq 0.2$ and δ to be less than about 4 to capture most of the effects of BAO. For most configurations, k_1 is smaller than the other two triangle sides, but for small δ and $\theta/\pi \sim 1$, k_3 can be smaller than k_1. In Chapter 2, these configurations were not used to constrain BAO because they are subject to cosmic variance and covariant with the power spectrum (as discussed in §2.3).

Below our minimum wavenumber of 0.01 h/Mpc, cosmic variance becomes significant;

given the mock catalogs we use one cannot make a sufficient number of subdivisions to estimate covariances on these large scales. Of course at large scales the covariance should be dominated by the Gaussian Random Field (GRF) contribution, so a template could be used to model the covariance (e.g., as Ref. [288] does for the isotropic 3PCF and Ref. [291] for the anisotropic 3PCF).

However, even with an adequate covariance, the contribution of low-wavenumber modes to BAO constraints should be small given the small number of large-scale modes in the volume of a survey such as DESI[1] [67]. Our minimum wavenumber corresponds to a physical scale of 628 Mpc/h; DESI will have volume of order 50 [Gpc/h]^3 equivalent to a box side length of roughly 3700 Mpc/h. Thus there are of order 200 modes of wavelength 628 Mpc/h in the box, enabling measurement to about 7% precision. The contribution of low-wavenumber bispectrum modes to BAO constraints will therefore be negligible compared to the 0.1% precision DESI will achieve using power spectrum BAO at higher wavenumber. It *is* the case that the other two triangle sides can probe higher wavenumbers—our maximum δ studied is 4, so $k_1 = 0.01$ h/Mpc corresponds to at most $k_2 = 2\lambda_f + 0.01$ h/Mpc $= 0.125$ h/Mpc and $k_3 = k_1 + k_2 = 0.135$ h/Mpc. These wavenumbers do access BAO scales, but nonetheless, the bispectrum error bar will not be competitive with DESI power spectrum precision as the total bispectrum error bars of such configurations will be dominated by the cosmic variance of the shortest side k_1.

At higher wavenumbers than our maximum, even at the level of linear theory Silk damping [285] degrades the BAO signal. The Silk damping scale $k_{\rm Silk}$ (equation 7 of Ref. [82]), is approximately 0.125 h/Mpc for our cosmology; at wavenumbers above $k_{\rm Silk}$, the BAO signal in the transfer function is increasingly suppressed as $\exp[-(k/k_{\rm Silk})^{1.4}]$ (equation 21 of Ref. [82]).

Wavenumbers above $k_{\rm NL} \sim 0.1$ h/Mpc are nonlinear, so perturbation theory no longer provides an accurate model of the bispectrum at these scales [251]. Effective field theory

1. `http://desi.lbl.gov`

(EFT) models perform reasonably well up to $k \sim 0.3\ h/\mathrm{Mpc}$; Ref. [40] describes the power spectrum to the percent level for $k \lesssim 0.3\ h/\mathrm{Mpc}$. In the case of the bispectrum (e.g. Refs. [27, 63, 215]), the maximum wavenumber at which EFT agrees with simulations depends on configuration and cosmology, but EFT models of the real-space matter bispectrum perform well up to $k \sim 0.2\ h/\mathrm{Mpc}$ [7, 15]. In redshift space, however, perturbation theory models break down at yet smaller wavenumbers, differing from bispectrum measurements at the 10% level by $k = 0.1\ h/\mathrm{Mpc}$ [295].

Baryonic effects, which are not as yet satisfactorily modeled, also become important at wavenumbers above our maximum. For $k \gtrsim 1\ h/\mathrm{Mpc}$, hydrodynamical simulations find a $5-15\%$ alteration in the power spectrum relative to dark-matter-only simulations [46]. As the bispectrum scales roughly as P^2 with P the power spectrum, this $\sim$10% uncertainty in the power spectrum likely translates to $\sim$20% in the bispectum, which at tree level is proportional to the square of the power spectrum (see also the hierarchical ansatz of Ref. [112]). In the absence of a theoretical model for baryonic effects, the uncertainty in high-wavenumber models of the bispectrum is much too large to measure BAO to sub-percent precision.

Overall, then, the range of wavenumbers we consider is a conservative cut to isolate the regime where BAO effects are most prominent and the bispectrum is best understood. Within this range of scales, our interferometric basis identifies the configurations where constructive interference of power spectra amplifies the BAO "wiggles." To quantify the presence of BAO in each configuration, we compute the RMS amplitude of the ratio R of the bispectrum $B(k_1, \delta, \theta)$ to its no-wiggle analog $B^{\mathrm{nw}}(k_1, \delta, \theta)$. We have

$$R(k_1, \delta, \theta) = \frac{B(k_1, \delta, \theta)}{B^{\mathrm{nw}}(k_1, \delta, \theta)}, \tag{3.5}$$

where the numerator is computed using power spectra $P(k)$ from CAMB [175] and the denominator using power spectra $P^{\mathrm{nw}}(k)$ from the fitting formula for the no-wiggle transfer

function of Ref. [82]. The variance is

$$\mathcal{A}^2(\delta,\theta) \equiv \int_{0.01}^{0.2} \left[R(k_1,\delta,\theta) - \bar{R}(\delta,\theta)\right]^2 \frac{\mathrm{d}k_1}{[h/\mathrm{Mpc}]}, \tag{3.6}$$

where $\bar{R}(\delta,\theta)$ is the mean of $R(k_1,\delta,\theta)$ on the same range, $0.01 \leq k_1/[h/\mathrm{Mpc}] \leq 0.2$. Figure 3.2 shows the root-mean-square amplitude $\mathcal{A}$ for a selection of configurations. Throughout this work we will refer to Figure 3.2 as the root-mean-square (RMS) map.

Our basis is a transformation of the triangle sides (k_1, k_2, k_3); the axes θ and δ of our RMS map correspond roughly to k_3 and k_2, respectively. In our basis, the wavenumbers k_2 and k_3 depend on k_1, δ, and θ as

$$\begin{aligned} k_2 &= k_1 + \delta\lambda_f/2, \\ k_3 &= \sqrt{k_1^2 + k_2^2 + 2k_1k_2\cos\theta} \\ &= \sqrt{k_1\left(1+\cos\theta\right)\left(2k_1+\delta\lambda_f\right) + \left(\delta\lambda_f/2\right)^2}, \end{aligned} \tag{3.7}$$

where the first equality for k_3 stems from the orientation of θ shown in Figure 3.1 and the law of cosines. Configurations with the same k_3 therefore lie along sloped curves in the (δ,θ) plane. As the power spectrum depends only on the magnitude of the wavenumber, these curves are also traces of constant $P(k_3)$.

The parameter δ was chosen to produce constructive interference in the precyclic term of the bispectrum (first term in equation 3.1). However, constructive interference is not limited only to this single term: we expect interference as well where k_2 and k_3, or k_3 and k_1, differ by integer multiples n of the BAO wavelength λ_f. We calculate the configurations for which these conditions are satisfied. Curves where $k_2 = k_1 + n\lambda_f$ are horizontal lines in the (θ,δ) plane, as shown in the left panel of Figure 3.3; that is, $k_2 = k_1 + n\lambda_f$ where

$$\delta/2 = n. \tag{3.8}$$

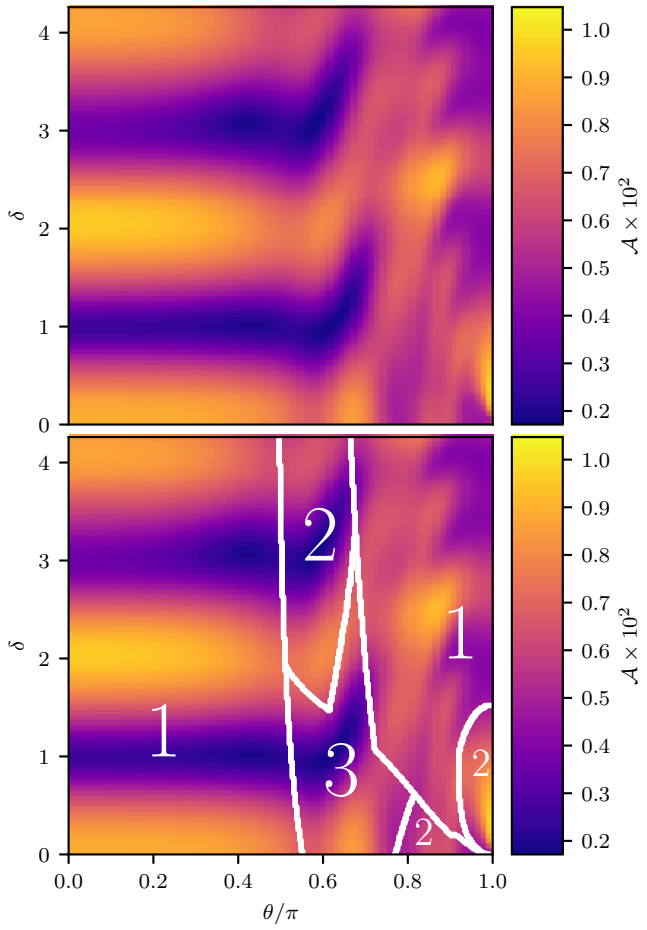

Figure 3.2: *Top*—The root-mean-square amplitude $\mathcal{A}$ (equation 3.6) of the bispectrum BAO feature in triangle configurations parameterized by (δ, θ). Maxima and minima are set by the constructive and destructive interference of BAO oscillations in the bispectrum. *Bottom*—In many regions of the RMS map, a single term or pair of terms in the bispectrum cyclic sum (3.1) dominates the sum. The boundaries between these regions (white lines) correspond to changes in the behavior of the RMS map. The numerals indicate the number of terms that must be considered to accurately approximate the bispectrum.

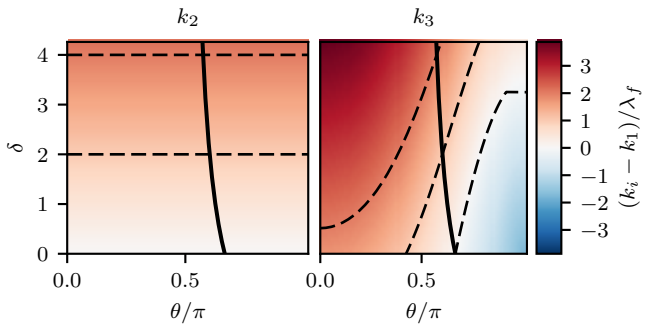

Figure 3.3: In our basis, k_2 depends only on δ, while k_3 varies with both δ and θ. The behavior of these two wavenumbers in the (δ, θ) basis is critical for understanding both the power spectrum and $F^{(2)}$ kernel. As $P(k) \sim k$, the structure of $P(k)$ in the δ-θ plane is similar to that of the individual wavenumbers, while $F^{(2)}$ is a more complicated function (as shown in Figure 3.7). k_2 and k_3 are calculated according to equation (3.7) with $k_1 = 0.1$ h/Mpc. Dashed lines in the left panel show configurations for which $k_2 = k_1 + n\lambda_f$ (equation 3.8) for $n = 1$ and 2; $n = 0$ coincides with the θ axis. In the right panel, dashed curves show configurations for which $k_3 = k_1 + n\lambda_f$ (equation 3.10) for $n = 0, 1$, and 2. Solid curves show configurations for which $k_2 = k_3$; the color is red where k_i is larger than k_1, blue where k_i is smaller than k_1, and white where $k_i = k_1$.

The curves where $k_2 = k_3 + n\lambda_f$, shown for $n = 0$ in Figure 3.3, are given by

$$\delta = \frac{-2k_1^2 \cos\theta - k_1^2 - 2n\lambda_f k_1 + n^2\lambda_f^2}{n\lambda_f^2 + \lambda_f k_1 \cos\theta}. \tag{3.9}$$

We only show the $n = 0$ case as higher harmonics of $k_2 = k_3 + n\lambda_f$ do not correspond to features in the RMS map, as discussed in §3.6.1 below. The curves where $k_3 = k_1 + n\lambda_f$, shown as dashed curves in the right panel of Figure 3.3, follow

$$\delta = \frac{2}{\lambda_f}\left[\sqrt{k_1\left(k_1\cos^2\theta + n\lambda_f\right) + n^2\lambda_f^2} - k_1(1 + \cos\theta)\right]. \tag{3.10}$$

In general, equations (3.9) and (3.10) depend on both k_1 and θ. In the special case of the equilateral configuration, the k_1 dependence cancels; that is, when $\theta/\pi = 2/3 = 0.67$ and $n = 0$, equation (3.9) gives $\delta = 0$. For these configurations, k_3 equals k_2 for all k_1. For all other configurations, however, equations (3.9) and (3.10) can only be satisfied for a single k_1. When necessary, we choose a representative value of $k_1 = 0.1$ h/Mpc to compute the

configurations for which equations (3.9) and (3.10) hold.

3.3 Notation

Here we define notation for several combinations of power spectra, $F^{(2)}$ kernels, and bispectra that will be used throughout.

BAO in the bispectrum come only from oscillations in the power spectrum, which we isolate as

$$P_i^{\rm BAO} = \frac{P(k_i)}{P^{\rm nw}(k_i)} \equiv 1 + w_i, \tag{3.11}$$

where w_i is defined through this equality and represents the BAO-only piece of the power spectrum. We note that $w_i \ll 1$; the baryon fraction $f_{\rm b}$ in our Universe is small ($f_{\rm b} \equiv \Omega_{\rm b}/\Omega_{\rm m} \sim 20\%$), so the BAO are a small feature in the power spectrum.

Each term in the cyclic sum (3.1) is denoted by

$$B_{ij} = 2P_{ij}F_{ij}^{(2)} \tag{3.12}$$

with

$$P_{ij} = P(k_i)P(k_j) \tag{3.13}$$

and

$$F_{ij}^{(2)} = F^{(2)}(k_i, k_j; \hat{\mathbf{k}}_i \cdot \hat{\mathbf{k}}_j). \tag{3.14}$$

The ratio of each term to its no-wiggle analog is

$$R_{ij} = \frac{B_{ij}}{B_{ij}^{\rm nw}} = \frac{P_{ij}}{P_{ij}^{\rm nw}} = P_i^{\rm BAO}P_j^{\rm BAO}, \tag{3.15}$$

where the second equality holds because the $F^{(2)}$ kernel is unaltered going from a physical to a "no-wiggle" cosmological model. The kernel stems from the Newtonian gravity solution of the equations of perturbation theory assuming an Einstein-De Sitter (matter-dominated)

cosmology, and is thus independent of the input linear power spectrum. We note that all cosmological parameters of the no-wiggle model, including the matter density, are identical to the physical model.

Using the definition (3.11) of w, we may rewrite

$$R_{ij} = 1 + w_i + w_j + w_i w_j \equiv 1 + w_{ij} \tag{3.16}$$

where the last equality defines w_{ij}, the oscillatory piece of one term of the bispectrum (3.1).

To refer to a ratio where one term in the sum is negligible, we use

$$R_{ij+jk} = \frac{B_{ij} + B_{jk}}{B_{ij}^{\rm nw} + B_{jk}^{\rm nw}}. \tag{3.17}$$

The sum $R_{12} + R_{23} + R_{31}$ is not equal to R of equation (3.5).

3.4 Regions of Dominance

In order to understand the behavior of the BAO amplitude shown in the RMS map (Figure 3.2), we seek to identify configurations where the cyclic sum of the perturbation theory bispectrum (3.1) simplifies. That is, we ask whether there are any regions where the behavior of the full bispectrum is determined by only one or two of the three terms in the cyclic sum.

Figure 3.4, our "dominance map," shows that many of the configurations are indeed dominated by a single term (red and blue regions), and others are dominated by two terms while the third is negligible (green and purple regions). The RMS map (Figure 3.2) reflects the dominance structure. The horizontal bands at $\theta/\pi \lesssim 0.4$ (red region, B_{12} dominant) transition to sloped bands in the purple and black regions. The blue region (B_{31} dominant) corresponds to a pattern of small maxima and minima in the RMS map that we call "feathering." Finally, in the green region at low δ and high θ, RMS amplitude is maximized for triangles where two wavevectors are nearly antiparallel and the third is small. The

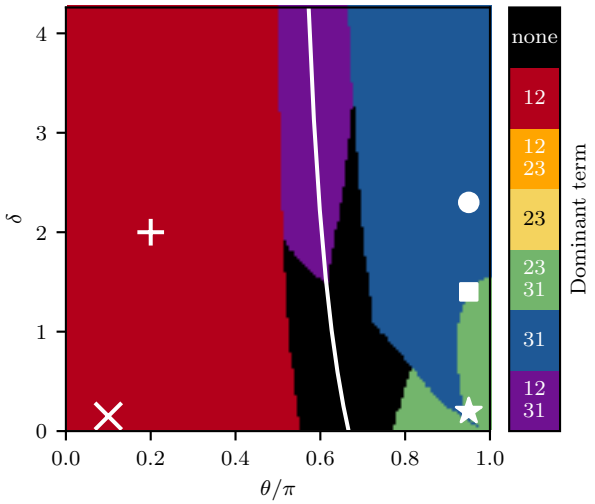

Figure 3.4: The dominance map (§3.4) shows regions where the bispectrum cyclic sum simplifies to a single term or pair of terms. A term dominates (§3.4.1) if the median of its ratio with each of the other two terms is at least 5. In our color scheme, primary colors (red, yellow, and blue) represent single terms, while the secondary colors (orange, green, and purple) represent pairs of terms. For an equilateral configuration ($\delta = 0$, $\theta/\pi = 0.67$), all three terms are identical so none can dominate; the black region surrounds this configuration. Symbols indicate representative configurations that are discussed in detail in §3.4.3. The white curve shows $k_2 = k_1$ (equation 3.9), where at least two terms must be of comparable magnitude (§3.4.3).

mechanisms that drive these different patterns in each region are described in detail in §3.6 below.

In this section, we detail the calculation and behavior of the dominance map (Figure 3.4). In §3.4.1, we present our definitions of dominance, which require the choice of a factor f. The specific ways in which triangle geometry determines which term dominates are discussed in detail in §3.4.3; the dominance map is driven primarily by the behavior of the $F^{(2)}$ kernel reinforced by the broadband behavior of the power spectrum, as we will further detail in §3.4.2. In most regions of the dominance map, the maximum and minimum terms in the $F^{(2)}$ kernel also maximize or minimize the power spectrum; the exceptions are discussed in §3.4.4. In general, in the squeezed limit where one side of the triangle can be much larger than the smallest, terms including the largest wavenumber are small. In the other limit, an equilateral triangle, all three sides are similar so all functions of them are similar as well and no sides dominate. The dominance plot shows the transitions between these two regimes.

We note that we assume positive δ, and for $\delta > 0$, no region is dominated by B_{23} (yellow) or the pair of terms $B_{12} + B_{23}$ (orange). When $\delta < 0$, k_1 and k_2 interchange. This would correspond to mirroring across the θ-axis; blue would become yellow.

3.4.1 Definition of "Dominance" and Choice of Dominance Ratio f

We identify dominant terms by comparing the magnitudes of terms $B_{ij}^{\rm nw}$ across k_1. The dominance structure is determined by the broadband behavior of the bispectrum terms, so we use the no-wiggle bispectrum $B_{ij}^{\rm nw}$ to fully isolate the broadband. Results are similar when the full bispectrum B_{ij} is used instead, as BAO are small relative to the broadband.

At each (δ, θ) configuration, we calculate the ratios between each pair of terms as a function of k_1. We then compare the medians, denoted med, of these ratios to a factor f. We use the median because it is a smooth function of our parameters δ and θ, unlike the mean, which can be skewed by large ratios between the terms at small k_1. The median is more representative of the typical ratio across all k_1 we consider.

Dominance criterion—If B_{ij} exceeds each other term by at least a factor of f, that is, if

$$\mathrm{med}\left|\frac{B_{ij}^{\mathrm{nw}}}{B_{jk}^{\mathrm{nw}}}\right| > f, \;\; \mathrm{med}\left|\frac{B_{ij}^{\mathrm{nw}}}{B_{ik}^{\mathrm{nw}}}\right| > f, \tag{3.18}$$

we consider B_{ij} dominant.

Double dominance criterion—If two of the terms that enter the bispectrum determine its behavior while the third is relatively small, we say that two terms are double dominant. Two ratios must be within a factor of f of each other but both exceed the third by at least a factor of f. Because terms can be either positive or negative, this comparison alone is not obviously sufficient; one term could be large and positive, and the other large and negative, such that their sum is smaller than the third term. Therefore we also require the sum of the two dominant terms to exceed the third term by a factor f. Our double dominance criterion is thus a set of four conditions: B_{ij} and B_{ik} are both dominant and only B_{jk} is negligible when

$$\mathrm{med}\left|\frac{B_{ij}^{\mathrm{nw}}}{B_{jk}^{\mathrm{nw}}}\right| > f, \;\; \mathrm{med}\left|\frac{B_{ik}^{\mathrm{nw}}}{B_{jk}^{\mathrm{nw}}}\right| > f, \;\; \mathrm{med}\left|\frac{B_{ij}^{\mathrm{nw}}}{B_{ik}^{\mathrm{nw}}}\right| < f, \;\; \mathrm{med}\left|\frac{B_{ij}^{\mathrm{nw}} + B_{ik}^{\mathrm{nw}}}{B_{jk}^{\mathrm{nw}}}\right| > f. \tag{3.19}$$

In practice, the final condition is not relevant for any configuration we test; the differences between large positive and negative terms remain much larger than the third term, for example in the region described in further detail in §3.4.3.

No term dominant—If the medians of all three ratios are within f of each other, then no term is dominant.

The dominance region plot is weakly dependent on the choice of the factor f, as shown in Figure 3.5. As the threshold for dominance rises, less of the plane is dominated by a single term; the $B_{12} + B_{31}$-dominant, $B_{23} + B_{31}$-dominant, and no-term-dominant regions encroach on the single-term-dominant regions. We choose $f = 5$ as our standard threshold for dominance, as it is sufficiently large to separate the term that dominates the RMS amplitude. With this choice of f, the non-dominant terms are typically less than 20% of

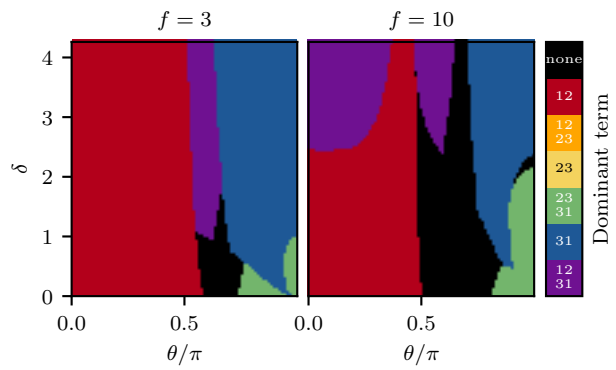

Figure 3.5: The shapes and locations of regions of dominance are not highly sensitive to f, the factor by which a term must exceed all others to be called "dominant" (see §3.4.1, equations 3.18 and 3.19). As f increases from the left panel to the right, the black and purple regions (where no single term exceeds all others by at least a factor of f) expand. That is, when the criterion for a single term to dominate is more strict, fewer configurations are dominated by a single term.

the dominant term, so the ratio of the bispectrum to the no-wiggle bispectrum (3.5) can be Taylor-expanded about the ratio R_{ij} of a single term in the cyclic sum to its no-wiggle analog (as we do in §3.5).

3.4.2 $F^{(2)}$ Kernel Drives Dominance Structure

As seen in Figure 3.6, the structure of the full dominance plot strongly resembles that of a dominance plot for $F^{(2)}$ alone, which itself reflects the behavior of $F^{(2)}$ in the δ-θ plane (Figure 3.7).

Including P_{ij} expands some regions (near their borders, P_{ij} can move the maximum term from just under to just over 5× the next-largest term). In Figure 3.6, we choose $f = \sqrt{5}$ for the P_{ij} and $F^{(2)}$ panels to agree with our choice of $f = 5$ for the product $P_{ij}F^{(2)}$. As discussed in §3.4.3 (equation 3.20), the P_{ij} and $F^{(2)}$ dominance criteria cannot simply be multiplied together to determine dominance in B_{ij}, but these two contributions independently illuminate the full bispectrum dominance map.

The dynamic range of $F^{(2)}$ is larger than that of P_{ij}, so the $F^{(2)}$ kernel determines most of the dominance map of Figure 3.4. The middle term of the $F^{(2)}$ kernel, $\left(k_i/k_j + k_j/k_i\right)\left(\hat{\mathbf{k}}_i \cdot \hat{\mathbf{k}}_j\right)$

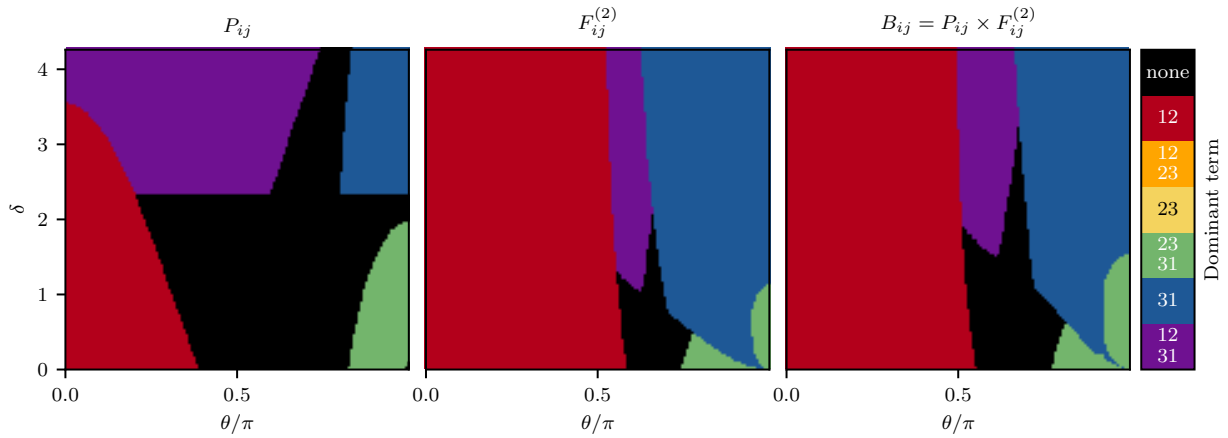

Figure 3.6: The regions of dominance are determined primarily by the $F^{(2)}$ kernel, as discussed in §3.4.2; the structure of the full dominance plot (right panel) is very similar to that of the $F_{ij}^{(2)}$ dominance plot (middle panel), with some modification from the products of power spectra P_{ij} (left panel). For the left and middle panels, P_{ij} and $F_{ij}^{(2)}$, a term is dominant if exceeds the other two by a factor of $\sqrt{5}$ (chosen for consistency with $f = 5$ for the product $P_{ij}F_{ij}^{(2)}$). For the third panel, a term is dominant if the median of its ratio with each of the other terms is at least 5 (as in Figure 3.4).

in equation (3.2), varies the most between configurations: it can be positive or negative, and can be very large when one side is much smaller than the other (for example, surrounding $\delta = 0, \theta/\pi = 1$ in Figure 3.7). Alternatively, $F^{(2)}$ can approach arbitrarily close to zero (black curves in Figure 3.7).

3.4.3 Regions of the Dominance Map

In this section, we step through each region of the dominance plot of Figure 3.4 from left to right to discuss the dominance behavior. In general, the relative magnitudes of the B_{12}, B_{23}, and B_{31} differ across configurations due to differences in the (δ, θ) dependence of the three wavenumbers k_1, k_2 and k_3 (given in equation (3.7) and Figure 3.3).

For each region, we discuss the behavior of the P_{ij} and $F_{ij}^{(2)}$ that enter the bispectrum. We build up understanding of each region by first analyzing their behavior separately, then considering the implications for the full dominance plot. We take this approach because

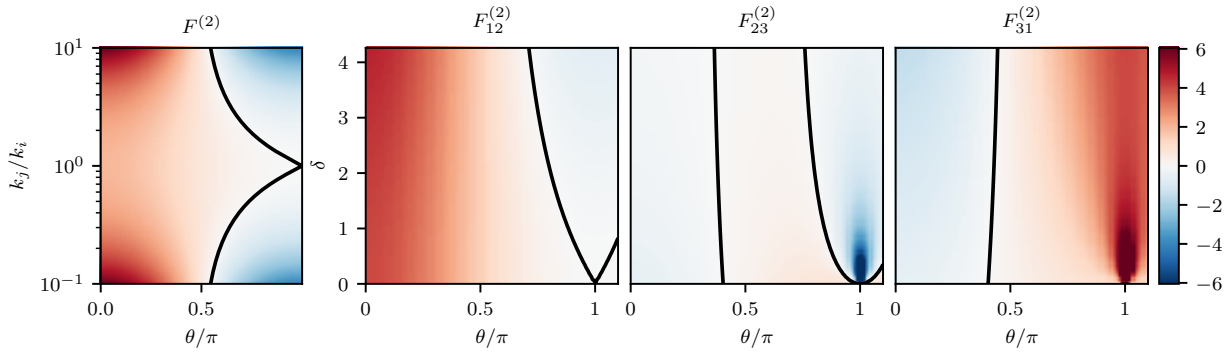

Figure 3.7: The $F^{(2)}$ kernel drives the structure of the dominance plot. As shown in the left panel, $F^{(2)}$ depends only on the angle between two sides through θ (Figure 3.1) and the ratio of their lengths k_j/k_i. The kernel can be positive or negative, and crosses zero (black curves). The dynamic range therefore exceeds that of the power spectrum product P_{ij}, which varies only by a factor of 500 across the triangles shown. The remaining three panels show the $F_{ij}^{(2)}$ that enter the bispectrum, evaluated at $k_1 = 0.1\ h/\text{Mpc}$. These three panels determine the behavior of the $F^{(2)}$ dominance plot (middle panel of Figure 3.6): $F_{12}^{(2)}$ (middle left panel) is the largest $F_{ij}^{(2)}$ in the red region of the middle panel of Figure 3.6, $F_{31}^{(2)}$ (right panel) dominates in the blue region of Figure 3.6, and in the green region of Figure 3.6, both $F_{23}^{(2)}$ (middle right panel) and $F_{31}^{(2)}$ are large while $F_{12}^{(2)}$ is small.

the power spectrum products behave very differently from the $F^{(2)}$ kernels, even though dominance is determined by the median ratios of terms $B_{ij}^{\rm nw}/B_{jk}^{\rm nw}$, which are medians of products and not products of medians:

$$\text{med}\left|\frac{B_{ij}^{\rm nw}}{B_{jk}^{\rm nw}}\right| = \text{med}\left|\frac{P_{ij}^{\rm nw}F_{ij}^{(2)}}{P_{jk}^{\rm nw}F_{jk}^{(2)}}\right| \neq \text{med}\left|\frac{P_{ij}^{\rm nw}}{P_{jk}^{\rm nw}}\right| \times \text{med}\left|\frac{F_{ij}^{(2)}}{F_{jk}^{(2)}}\right|. \tag{3.20}$$

We note that the power spectrum is maximal at $k_{\rm peak} \approx 0.015\ h/\text{Mpc}$; above $k_{\rm peak}$, $P^{\rm nw}(k)$ declines monotonically as $1/k$. Since our analysis covers $0.01 \leq k_1/[h/\text{Mpc}] \leq 0.2$, it is a good approximation that in our k-range of interest the broadband power spectrum falls as $P(k) \propto 1/k$. This approximation fails only in the low-δ, high-θ region where k_3 can be sufficiently small that $P(k_3)$ increases with k_3 (discussed in §3.4.3 below).

Red region, B_{12} dominant

When $\theta/\pi \lesssim 0.5$, Figure 3.4 shows that B_{12} is the dominant term in the bispectrum cyclic sum (3.1). In this red region, configurations are constructive where $P_1^{\rm BAO}$ and $P_2^{\rm BAO}$ (defined in equation 3.11) are in phase, and destructive where they are out of phase. B_{12} dominates because in both the $F^{(2)}$ kernel and the products of power spectra P_{ij}, the pre-cyclic terms are largest.

$F_{12}^{(2)}$ is much larger than $F_{23}^{(2)}$ and $F_{31}^{(2)}$ (as shown in Figure 3.7) because only in $F_{12}^{(2)}$ is the sign of the dipole contribution $(k_i/k_j + k_j/k_i)$ positive. The dot product of the unit vectors $\hat{\mathbf{k}}_1$ and $\hat{\mathbf{k}}_2$, which determines the sign of the dipole contribution, approaches $+1$ in $F_{12}^{(2)}$. For $F_{23}^{(2)}$ and $F_{31}^{(2)}$, the relevant dot product instead approaches -1. While $F_{23}^{(2)}$ and $F_{31}^{(2)}$ therefore contain both positive and negative contributions of similar magnitude, all contributions to $F_{12}^{(2)}$ are positive, so $F_{12}^{(2)}$ will be the largest $F_{ij}^{(2)}$. For example, where δ vanishes as well as θ (in the lower left corner, $\times$ symbol), two sides, k_1 and k_2, are equal, while $k_3 = 2k_1$. With these side lengths and dot products, $F_{12}^{(2)} = 2$ and $F_{31}^{(2)} = F_{23}^{(2)} = -0.25$. $F_{12}^{(2)}$ exceeds the other two $F_{ij}^{(2)}$ by a factor of eight.

The effect of the power spectrum products is to further separate the three cyclic terms. For $0 \leq \theta/\pi < 0.5$, the triangle is obtuse (see θ in Figure 3.1). As $\theta \to 0$ and the triangle fully opens, k_3 approaches $k_1 + k_2$. For these obtuse triangles $k_3 > k_2 > k_1$ (see Figure 3.3) because k_2 always exceeds k_1. The power spectrum is monotonically decreasing, so the k_i ordering implies $P(k_1) > P(k_2) > P(k_3)$. Thus $P_{12} > P_{31} > P_{23}$, reinforcing the order of the $F_{ij}^{(2)}$.

For nonzero δ (e.g., + symbol in Figure 3.4), B_{31} grows with δ, but B_{12} remains dominant by our dominance criterion of $f = 5$. For large δ, k_1 is small relative to $\delta\lambda_f/2$. The other two wavenumbers k_2 and k_3 are both larger than k_1 (Figure 3.3), so $P(k_3)$ approaches $P(k_2)$. As a result, P_{31} and P_{12} are of similar magnitude (as in the purple region at low θ and high δ in the leftmost panel of Figure 3.6). The magnitudes of the $F_{12}^{(2)}$ and $F_{31}^{(2)}$ kernels also grow as δ increases, with the $F_{31}^{(2)}$ kernel approaching but remaining smaller than $F_{12}^{(2)}$. As δ continues to increase, B_{31} comes within a factor of 10 of B_{12}, causing a purple region to appear in the upper left corner of the right panel of Figure 3.5. We note that this effect is too small to appear when the dominance criterion is $f = 5$, our choice in the main analysis of this work (as in Figure 3.4).

Middle region, no term dominant (black) or B_{12} and B_{31} dominant (purple)

Around $\theta/\pi = 0.6$, at most one term in the cyclic sum can be neglected. In the black region at low δ, all three terms are of comparable magnitude; in the purple region at larger δ, B_{23} shrinks, but B_{12} and B_{31} are still large and of similar magnitude. The black region contains triangles that are nearly equilateral; triangles with $\theta/\pi = 2/3$ and $\delta = 0$ are equilateral for all k_1. Since all three sides and angles are equal, all three terms in the bispectrum are identical, and no term can dominate any other.

In the purple region, B_{23} is negligible compared to B_{12} and B_{31}. As δ increases along the $k_2 = k_3$ line shown in Figures 3.3 and 3.4, k_2 and k_3 grow larger than k_1. Since $k_2 = k_3$, B_{12} and B_{31} remain equal; their ratio will deviate little from unity. But as δ increases, both $F_{23}^{(2)}$

and P_{23} shrink relative to the other terms. In particular, $F_{23}^{(2)}$ approaches zero. For large δ along $k_2 = k_3$, the unit vectors $\hat{\mathbf{k}}_2$ and $\hat{\mathbf{k}}_3$ are antiparallel as k_1 is relatively small, and their dot product $(\hat{\mathbf{k}}_2 \cdot \hat{\mathbf{k}}_3)$ becomes -1. Since $k_2 = k_3$, the dipole contribution $(k_2/k_3 + k_3/k_2) = 2$ in the $F_{23}^{(2)}$ kernel, and $F_{23}^{(2)}$ therefore vanishes. Furthermore, Figure 3.3 shows that at large δ, k_2 and k_3 are much larger than k_1. The product of power spectra P_{23} is therefore smaller than the other two products P_{12} and P_{31}, which both involve the much larger $P(k_1)$. Both $F_{23}^{(2)}$ and P_{23} shrink as δ increases, so B_{23} becomes smaller than the other two terms and can be neglected. Thus B_{12} and B_{31} dominate the bispectrum cyclic sum.

Blue region, B_{31} dominant

At the right side of Figure 3.4, where θ is large (circle symbol), B_{31} dominates the cyclic sum. Both $F_{31}^{(2)}$ and P_{31} are large relative to the other $F^{(2)}$ kernels and products of power spectra.

In the $F^{(2)}$ kernel, as shown in Figure 3.7, $F_{12}^{(2)}$ vanishes, and negative contributions to $F_{23}^{(2)}$ make it smaller than $F_{31}^{(2)}$. $F_{12}^{(2)}$ vanishes because in this region, triangles are in the squeezed limit, where $\mathbf{k}_2 \approx -\mathbf{k}_1$. The third wavenumber k_3 approaches $k_2 - k_1$, meaning

$$k_3 \to \frac{\delta\lambda_f}{2}, \tag{3.21}$$

so k_3 is small relative to the other two wavenumbers (see Figure 3.3). At the same time, the dot product $\hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2$ approaches -1, so $F_{12}^{(2)}$ behaves as

$$F_{12}^{(2)} \to 1 - \frac{1}{2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) \sim 0. \tag{3.22}$$

In equation (3.22), the difference $\delta\lambda_f/2$ between k_1 and k_2 (3.3) is much smaller than k_1, so $k_1 \sim k_2$ and $F_{12}^{(2)}$ vanishes. Meanwhile, the other two terms do not vanish; $\hat{\mathbf{k}}_2 \cdot \hat{\mathbf{k}}_3 = -1$ but

$\hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_3 = +1$, so $F_{23}^{(2)}$ and $F_{31}^{(2)}$ approach

$$F_{23}^{(2)} \to 1 - \frac{1}{2}\left(\frac{k_2}{k_3} + \frac{k_3}{k_2}\right) \tag{3.23}$$

$$F_{31}^{(2)} \to 1 + \frac{1}{2}\left(\frac{k_3}{k_1} + \frac{k_1}{k_3}\right). \tag{3.24}$$

The magnitudes of k_1 and k_2 are comparable, so the dipole contributions $(k_2/k_3 + k_3/k_2)$ in $F_{23}^{(2)}$ (equation 3.23) and $(k_3/k_1 + k_1/k_3)$ in $F_{31}^{(2)}$ (equation 3.24) are similar in magnitude. Only in $F_{31}^{(2)}$ do both contributions have the same sign, so $F_{31}^{(2)}$ is larger than $F_{23}^{(2)}$.

The P_{ij} reinforce the behavior of the $F^{(2)}$ kernel. k_3 is small, but still large enough that $P(k_3)$ is monotonically decreasing; for δ of a few, equation (3.21) is near $k_{\rm Peak}$ (§3.4.3). In this limit $k_2 = k_1 + k_3$, so k_2 must be the largest of the three wavenumbers. In the power spectrum, then, $P(k_2) < P(k_3), P(k_1)$. P_{12} and P_{23}, which include $P(k_2)$, are therefore smaller than P_{31}, which does not. The largest power spectrum term is therefore P_{31} as in the $F^{(2)}$ kernel, and B_{31} dominates the bispectrum.

Green region, B_{23} and B_{31} dominant

In the lower right corner of Figure 3.4 around $\delta = 0$, $\theta/\pi = 1$ (star symbol), the dominant terms are B_{23} and B_{31}. Dominance is once again driven by the $F^{(2)}$ kernel; $F_{23}^{(2)}$ and $F_{31}^{(2)}$ are very large (as shown in Figure 3.7) while $F_{12}^{(2)}$ vanishes.

$F_{12}^{(2)}$ vanishes for the same reason it does in the blue region, but the magnitude of $F_{23}^{(2)}$ is more comparable to that of $F_{31}^{(2)}$ than it is at higher δ. Along the line where $\theta/\pi = 1$, the blue region transitions to green at roughly $\delta \approx 1$ (square symbol in Figure 3.4). Here, $k_3 = \lambda_f/2 \approx 0.0285\ h/\mathrm{Mpc}$, which is nearing the lowest k_1 in our range ($k_1 = 0.01\ h/\mathrm{Mpc}$). Both k_2 and k_1 therefore exceed k_3 by up to a factor of 10. As in the blue region, triangles in the green region are squeezed, so the dot products between unit wavevectors are the same as in the blue region. The kernels then behave as equations (3.22–3.24), and $F_{12}^{(2)}$ (equation

3.22) vanishes. Unlike in the blue region, however, the dipole contribution $(k_i/k_j + k_j/k_i)$ is large enough that the constants in equations (3.23) and (3.24) become insignificant. $F_{23}^{(2)}$ and $F_{31}^{(2)}$ are thus both comparably large; $F_{31}^{(2)}$ is large and positive ($\hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_3 = +1$) while $F_{23}^{(2)}$ is large and negative ($\hat{\mathbf{k}}_2 \cdot \hat{\mathbf{k}}_3 = -1$). This region satisfies our double dominance criterion (3.19), as even the sum $B_{23} + B_{31}$ exceeds the very small third term B_{12}—for example, by a factor of 10^5 at $\delta = 0.01, \theta/\pi = 1$.

Our assumption that $P(k)$ falls monotonically with k breaks down in the green region. Because k_3 is proportional to δ (as in equation 3.21), k_3 becomes very small for small δ (see Figure 3.3). Our previous analysis assumed all wavenumbers were large enough that the power spectrum is monotonically decreasing, but in the green region, k_3 can be in the regime where $P(k)$ increases with k. Both k_1 and k_2 are still greater than k_{Peak}, in the regime where $P(k)$ falls with k. For very small k_3, then, $P(k_3)$ may be smaller than $P(k_1)$ and $P(k_2)$. As a result, P_{12} can be the largest product of power spectra, despite the fact that k_1 and k_2 are much larger than k_3. However, the power spectrum is overshadowed by the behavior of the $F^{(2)}$ kernel. Even for $\delta = 0.01$, where P_{12} can exceed the other two products of power spectra by a factor of 25 at small k_1, the median ratio of power spectra products, $\mathrm{med}\,[P_{12}/P_{23}]$, is only 5. But the median ratio of $F^{(2)}$ kernels, $\mathrm{med}\left[F_{12}^{(2)}/F_{23}^{(2)}\right]$, is of order 10^{-8} because k_1 so nearly equals k_2, driving $F_{12}^{(2)}$ to vanish. The dominance structure is thus driven primarily by the $F^{(2)}$ kernel.

3.4.4 Ordering of Subdominant Terms

The B_{ij} of equation (3.1) with maximum median P_{ij} is everywhere also the term with maximum median $|F_{ij}^{(2)}|$. The ordering of terms differs only in the regions shown in Figure 3.8, where the two subdominant terms are swapped. These regions arise because the $F^{(2)}$ kernel can be either positive or negative. In the range of wavenumbers of interest, the power spectrum is always positive but monotonically decreasing, so the products P_{ij} change smoothly. P_{23} and P_{31}, smaller than P_{12} for small θ (see the left panel of Figure 3.6 and the top panel

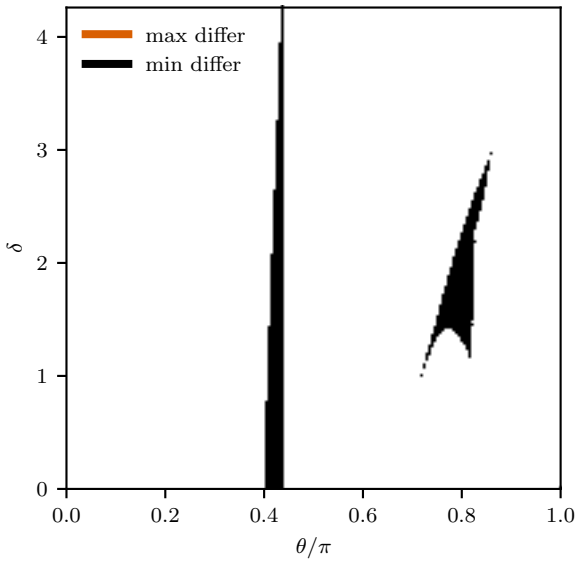

Figure 3.8: The dominant term in the cyclic sum composing the bispectrum (3.1) is determined by both P_{ij} and $F^{(2)}$ (as shown in Figure 3.6), and the two almost always act in the same direction. The maximum term never differs between P_{ij} and $F^{(2)}$ (orange regions, none shown), but the minimum terms are swapped in two regions (black). In these regions, discussed in §3.4.4, the term with minimum P_{ij} has the second-largest $F_{ij}^{(2)}$.

of Figure 3.9), cross above P_{12} around the equilateral configuration ($\theta/\pi = 0.67$). P_{31} is the first to cross above P_{12} because $k_2 \geq k_1$, so $P_{31} \geq P_{23}$. P_{23} lags behind (see the top panel of Figure 3.9).

The behavior of the $F^{(2)}$ kernel (equation 3.2) is not as simple, as shown in the lower panel of Figure 3.9. First, it can be either positive or negative (see Figure 3.7), explaining the $\theta/\pi \sim 0.4$ region in Figure 3.8 where the ordering of the subdominant terms differs between $F^{(2)}$ and P_{ij}. The difference arises because the absolute values of the two subdominant terms $F_{23}^{(2)}$ and $F_{31}^{(2)}$ spuriously cross when both are small. In detail, we take the absolute value of each $F_{ij}^{(2)}$, since the magnitudes of each term, not their signs, determine dominance. For $\theta \sim 0$, both $F_{23}^{(2)}$ and $F_{31}^{(2)}$ are negative, with $F_{31}^{(2)}$ more negative than $F_{23}^{(2)}$. For $\theta/\pi \gtrsim 0.4$, both terms are positive, with $F_{31}^{(2)}$ more positive than $F_{23}^{(2)}$. $F_{31}^{(2)}$ must therefore cross above $F_{23}^{(2)}$, and it does so in the same region around $\theta/\pi \sim 0.4$ where both terms cross zero—but the terms are not equal to zero where they cross each other. As shown in Figure 3.7, the

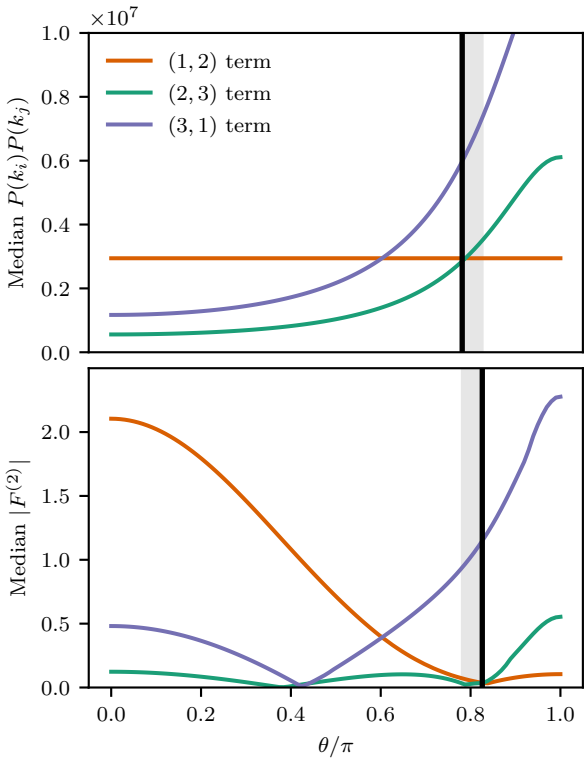

Figure 3.9: In the shaded region, the ordering of subdominant terms differs between P_{ij} and $F_{ij}^{(2)}$: P_{12} is the smallest P_{ij}, while the minimum $F_{ij}^{(2)}$ is $F_{23}^{(2)}$ (compare Figure 3.8). As discussed in §3.4.4, this region arises due to differences in the behavior of the median between the power spectrum and the $F^{(2)}$ kernel. Around $\theta/\pi \sim 0.6$, both $F_{12}^{(2)}$ and $F_{31}^{(2)}$ are positive for all k_1, so their medians cross at the same θ as the medians of P_{12} and P_{31}. At $\theta/\pi = 0.78$, however, P_{23} crosses above P_{12} (solid vertical line in top panel); $F_{23}^{(2)}$ lags behind, crossing above $F_{12}^{(2)}$ at $\theta/\pi = 0.83$ (solid vertical line in bottom panel). The ordering also differs around $\theta/\pi \sim 0.4$, as further discussed in §3.4.4.

value of θ at which $F_{ij}^{(2)}$ crosses zero depends on the ratio between the two sides k_i and k_j, so $F_{23}^{(2)}$ becomes positive at slightly lower θ than does $F_{31}^{(2)}$. As $F_{31}^{(2)}$ approaches zero, its absolute value falls below the small and positive $F_{23}^{(2)}$ at $\theta/\pi = 0.42$ (with $\delta = 2.1$, for example, as in Figure 3.9). After $F_{31}^{(2)}$ becomes positive, it crosses $F_{23}^{(2)}$ at $\theta/\pi = 0.44$. Meanwhile in the product of power spectra, the median P_{31} exceeds the median P_{23} for all θ. Therefore, in this narrow region where the median $F_{31}^{(2)}$ falls below the median $F_{23}^{(2)}$, the smallest P_{ij} term is not the smallest $F_{ij}^{(2)}$ term.

The order of the subdominant terms also differs around $\theta/\pi \sim 0.8$ (Figure 3.8). Figure 3.9 shows that this region arises due to differences in the behavior of the median between the power spectrum and the $F^{(2)}$ kernel. The power spectrum decreases monotonically, so the median P_{ij} occurs at the median k_1. Therefore P_{23} and P_{12} are equal when $k_3 = k_1$ is evaluated at the median k_1 (see equation 3.10). In Figure 3.9, P_{23} crosses above P_{12} at $\theta/\pi = 0.78$ (with $\delta = 2.1$). Unlike the power spectrum, the $F^{(2)}$ kernel can be positive or negative, and in some configurations it is positive for some values of k_1 and negative for others. For these configurations, the absolute value of $F^{(2)}$ is not a monotonic function of k_1, so its median does not necessarily occur at the median k_1. Therefore $F_{23}^{(2)}$ crosses above $F_{12}^{(2)}$ at θ such that the k_3 that corresponds to the median $F_{23}^{(2)}$ equals the k_1 that corresponds to the median $F_{12}^{(2)}$. In the example of Figure 3.9, the solution is $\theta/\pi = 0.83$. P_{12} becomes the minimum P_{ij} at $\theta/\pi = 0.78$ while $F_{12}^{(2)}$ does not become the minimum $F_{ij}^{(2)}$ until $\theta/\pi = 0.83$. Therefore in the shaded region of Figure 3.9 between these two crossings ($0.78 \leq \theta/\pi \leq 0.83$), the order of the subdominant terms differs.

In contrast, around $\theta/\pi \sim 0.6$, both $F_{12}^{(2)}$ and $F_{31}^{(2)}$ are positive for all k_1. Their medians are both found at the median value of k_1, so $F_{31}^{(2)}$ crosses above $F_{12}^{(2)}$ where $k_2 = k_3$ (equation 3.9, evaluated at the median k_1). As the median P_{ij} occurs at the median k_1, P_{31} also crosses above P_{12} where equation (3.9) is evaluated at the median k_1. $F_{31}^{(2)}$ therefore becomes the maximum $F_{ij}^{(2)}$ at the same value of θ where P_{31} becomes the maximum P_{ij}.

Though we set out to explain the regions where the minimum $F_{ij}^{(2)}$ differs from the

minimum P_{ij}, our analysis also explains why the largest P_{ij} is always also the largest $F_{ij}^{(2)}$ (see Figure 3.8). The median behaves most simply for $F_{12}^{(2)}$ and P_{12}, which are both the maximum term at low θ. As θ increases, P_{31} is always the first to cross P_{12}, and $F_{31}^{(2)}$ is always the first to cross $F_{12}^{(2)}$. $F_{23}^{(2)}$ and P_{23} are never the maximum term because $k_2 \geq k_1$. The complexity of the ordering of subdominant terms arises from the $F_{23}^{(2)}$ and P_{23} terms, but because these terms are never the maximum terms, the maximum term never differs between P_{ij} and $F_{ij}^{(2)}$.

3.5 Decomposition into Eigen-Root-Mean-Square Plots

We now show how the RMS amplitude $\mathcal{A}$ (equation 3.6) can frequently be approximated as a linear combination of three terms. We first require an expression for the ratio R (equation 3.5) of the full bispectrum (3.1) to its no-wiggle analog. We wish to leverage the fact that the BAO are a small fractional feature in the power spectrum, so we write

$$P_{ij} = P_{ij}^{\mathrm{nw}}[1 + w_{ij}] \tag{3.25}$$

with P_{ij}^{nw} the product of two no-wiggle power spectra and w_{ij} the BAO feature in the product P_{ij} (3.13) of linear power spectra. In particular we split the matter transfer function T_{m} into smooth and oscillatory pieces as

$$T_{\mathrm{m}}(k) = T_{\mathrm{sm}}(k) + \omega(k) j_0(k\tilde{s}), \tag{3.26}$$

where $T_{\mathrm{sm}}(k)$ and $\omega(k)$ are smooth functions of k [82], ω is small (because $\Omega_{\mathrm{b}}/\Omega_{\mathrm{m}} \ll 1$), and $j_0(x) = \sin(x)/x$ is the order zero spherical Bessel function.

The power spectrum is proportional to the primordial power spectrum $P_{\mathrm{pri}}(k)$ and the matter transfer function as

$$P(k) = P_{\mathrm{pri}}(k) T_{\mathrm{m}}^2(k). \tag{3.27}$$

We suppress the redshift dependence of the power spectrum for simplicity, as it does not affect our analysis. The products of power spectra are then

$$P_{ij} = P_{\rm pri}(k_i)P_{\rm pri}(k_j)T^2_{\rm sm}(k_i)T^2_{\rm sm}(k_j) \times \left[1 + \frac{\omega(k_i)j_0(k_i s)}{T_{\rm sm}(k_i)}\right]^2 \left[1 + \frac{\omega(k_j)j_0(k_j s)}{T_{\rm sm}(k_j)}\right]^2. \quad (3.28)$$

Taylor-expanding the fractions $\omega(k_i)j_0(k_i s)/T_{\rm sm}(k_i)$ to leading order in ω we have

$$w_{ij} = \frac{P_{ij}}{P^{\rm nw}_{ij}} - 1 \approx 2\left[\frac{\omega(k_i)j_0(k_i s)}{T_{\rm sm}(k_i)} + \frac{\omega(k_j)j_0(k_j s)}{T_{\rm sm}(k_j)}\right], \quad (3.29)$$

where we used the fact that the no-wiggle power spectrum is simply $P^{\rm nw}(k) = P_{\rm pri}(k)T^2_{\rm sm}(k)$.

In the remainder of this section, we show that in regions where only one or two terms dominate the bispectrum, the variance $\mathcal{A}^2$ of the full bispectrum is approximated to leading order by the variance of only the dominant term or terms. We first consider the case where one term is dominant and the other two negligible, and we then consider the case where one term is negligible and the other two must be retained.

3.5.1 Single Term Dominant

We first consider the case where one term in the bispectrum dominates over the other two; without loss of generality we take this to be the first term.

We calculate the RMS amplitude $\mathcal{A}$ (equation 3.6) from the variance of the ratio R, defined in equation (3.5) as

$$R = \frac{B_{12} + B_{23} + B_{31}}{B^{\rm nw}_{12} + B^{\rm nw}_{23} + B^{\rm nw}_{31}}. \quad (3.30)$$

Our goal is to show that the variance $\mathcal{A}^2_{12}$ of the approximate ratio

$$R_{12} = \frac{B_{12}}{B^{\rm nw}_{12}} = 1 + w_{12} \quad (3.31)$$

is the same as that of the full ratio, $\mathcal{A}^2$, to leading order in one or the other of two small parameters we define,

$$\epsilon_{ij} \equiv \frac{B_{ij}}{B_{12}}, \quad \epsilon_{ij}^{\rm nw} \equiv \frac{B_{ij}^{\rm nw}}{B_{12}^{\rm nw}}. \tag{3.32}$$

We first notice that the variance of R_{12} is

$$\mathcal{A}_{12}^2 = \left\langle w_{12}^2 \right\rangle - \langle w_{12} \rangle^2 \,; \tag{3.33}$$

the constant term in equation (3.31) of course contributes no variance. Since $\mathcal{A}_{12}^2$ is second-order in the small parameter w_{12}, we neglect all corrections at third order and higher. We will find that the difference between the full variance and the variance of R_{12} vanishes at second order.

Factoring out the dominant term in the numerator and denominator of equation (3.31) and using the definitions above, we have the ratio as

$$\begin{aligned} R =& \left(\frac{B_{12}}{B_{12}^{\rm nw}}\right) \frac{1+\epsilon_{23}+\epsilon_{31}}{1+\epsilon_{23}^{\rm nw}+\epsilon_{31}^{\rm nw}} \\ \approx& R_{12}\left(1+\epsilon_{23}+\epsilon_{31}\right)\left[1-\epsilon_{23}^{\rm nw}-\epsilon_{31}^{\rm nw}+(\epsilon_{23}^{\rm nw}+\epsilon_{31}^{\rm nw})^2\right] \\ =& R_{12}\left[1+(\epsilon_{23}-\epsilon_{23}^{\rm nw})+(\epsilon_{31}-\epsilon_{31}^{\rm nw})-(\epsilon_{23}+\epsilon_{31})(\epsilon_{23}^{\rm nw}+\epsilon_{31}^{\rm nw})+(\epsilon_{23}^{\rm nw}+\epsilon_{31}^{\rm nw})^2+\mathcal{O}(\epsilon^3)\right]. \end{aligned} \tag{3.34}$$

In the second, approximate equality, we have Taylor-expanded the denominator to second order in ϵ. We include the second-order term for the moment but see that it drops out of our end result.

We now seek to exploit the fact that the BAO feature itself is small, i.e. $w_{ij} \ll 1$ (equation 3.29 with small $\omega(k)$). Using w_{ij} as defined in equation (3.16) to simplify differences of ϵ_{ij}

and $\epsilon_{ij}^{\rm nw}$, we have

$$\begin{aligned}\epsilon_{23} - \epsilon_{23}^{\rm nw} &= \frac{B_{23}^{\rm nw}}{B_{12}^{\rm nw}}\left[\frac{1+w_{23}}{1+w_{12}} - 1\right] \\ &\approx \epsilon_{23}^{\rm nw}(w_{23} - w_{12} + w_{12}^2 - w_{12}w_{23}) \\ &= \epsilon_{23}^{\rm nw}\Delta_{23,12}^{w}(1-w_{12}),\end{aligned} \tag{3.35}$$

where to obtain the first equality we substituted the definitions (3.32) and to obtain the second we Taylor-expanded to leading order in w. In the third equality, we defined

$$\Delta_{23,12}^{w} = w_{23} - w_{12}, \tag{3.36}$$

which is $\mathcal{O}(w)$. The analog of equation (3.35) holds for the 31 term by switching 23 to 31 everywhere.

We know the variance at leading order is $\mathcal{O}(w^2)$ from equation (3.33), so we only retain terms that are second order in a combination of ϵ and w. Our approximate expression for the ratio R is now

$$R \approx R_{12}\left[1 + \epsilon_{23}^{\rm nw}\Delta_{23,12}^{w} + \epsilon_{31}^{\rm nw}\Delta_{31,12}^{w} - (\epsilon_{23}+\epsilon_{31})(\epsilon_{23}^{\rm nw}+\epsilon_{31}^{\rm nw}) + (\epsilon_{23}^{\rm nw}+\epsilon_{31}^{\rm nw})^2\right]. \tag{3.37}$$

The first term is $\mathcal{O}(1)$, the second and third $\mathcal{O}(\epsilon w)$, and the fourth and fifth $\mathcal{O}(\epsilon^2)$. These last two terms cancel each other to second order; $\epsilon_{ij} - \epsilon_{ij}^{\rm nw}$ (equation 3.35) is itself second order, so at leading order the second to last term is equal to the last. Equation (3.38) then simplifies:

$$R \approx R_{12}\left[1 + \epsilon_{23}^{\rm nw}\Delta_{23,12}^{w} + \epsilon_{31}^{\rm nw}\Delta_{31,12}^{w}\right]. \tag{3.38}$$

Now computing the expectation value of R and factoring out $\langle R_{12}\rangle$ to enable further Taylor

expansions, we find

$$\langle R\rangle \approx \langle R_{12}\rangle \left\{1 + \langle R_{12}\rangle^{-1} \left\langle R_{12}\epsilon_{23}^{\rm nw}\Delta_{23,12}^{w}\right\rangle + \langle R_{12}\rangle^{-1} \left\langle R_{12}\epsilon_{31}^{\rm nw}\Delta_{31,12}^{w}\right\rangle\right\}. \tag{3.39}$$

We now square the form above and multiply out to find

$$\langle R\rangle^2 \approx \langle R_{12}\rangle^2 + 2\langle R_{12}\rangle \times \left\{\left\langle R_{12}\epsilon_{23}^{\rm nw}\Delta_{23,12}^{w}\right\rangle + \left\langle R_{12}\epsilon_{31}^{\rm nw}\Delta_{31,12}^{w}\right\rangle\right\}. \tag{3.40}$$

Now using equation (3.38) to compute the expectation value of R^2, we obtain

$$\left\langle R^2\right\rangle \approx \left\langle R_{12}^2\right\rangle + 2\left\{\left\langle R_{12}^2\epsilon_{23}^{\rm nw}\Delta_{23,12}^{w}\right\rangle + \left\langle R_{12}^2\epsilon_{31}^{\rm nw}\Delta_{31,12}^{w}\right\rangle\right\}. \tag{3.41}$$

The variance of R is then

$$\begin{aligned}\left\langle R^2\right\rangle - \langle R\rangle^2 \approx &\left\langle R_{12}^2\right\rangle - \langle R_{12}\rangle^2 \\ &+2\Bigg\{\left\langle R_{12}^2\epsilon_{23}^{\rm nw}\Delta_{23,12}^{w}\right\rangle - \langle R_{12}\rangle\left\langle R_{12}\epsilon_{23}^{\rm nw}\Delta_{23,12}^{w}\right\rangle \\ &\quad + \left\langle R_{12}^2\epsilon_{31}^{\rm nw}\Delta_{31,12}^{w}\right\rangle - \langle R_{12}\rangle\left\langle R_{12}\epsilon_{31}^{\rm nw}\Delta_{31,12}^{w}\right\rangle\Bigg\}.\end{aligned} \tag{3.42}$$

Recalling that $R_{12} = 1 + w_{12}$ (equation 3.31) and denoting the variance of R_{12} as $\mathcal{A}_{12}^2$, we find

$$\begin{aligned}&\mathcal{A}^2 - \mathcal{A}_{12}^2 \approx \\ &2\times\left\{\left\langle(1+w_{12})^2\,\epsilon_{23}^{\rm nw}\Delta_{23,12}^{w}\right\rangle - \langle 1+w_{12}\rangle\left\langle(1+w_{12})\,\epsilon_{23}^{\rm nw}\Delta_{23,12}^{w}\right\rangle\right. \\ &\left.\quad + \left\langle(1+w_{12})^2\,\epsilon_{31}^{\rm nw}\Delta_{31,12}^{w}\right\rangle - \langle 1+w_{12}\rangle\left\langle(1+w_{12})\,\epsilon_{31}^{\rm nw}\Delta_{31,12}^{w}\right\rangle\right\}.\end{aligned} \tag{3.43}$$

To second order, the difference (3.43) cancels. Thus, the difference between $\mathcal{A}^2$ and $\mathcal{A}_{12}^2$ is suppressed by one order relative to $\mathcal{A}_{12}^2$. Therefore $\mathcal{A}_{12}^2$ is the leading contribution to the variance $\mathcal{A}^2$. When a single term dominates the bispectrum cyclic sum, the variance

of that single term is a good approximation of the variance of the full bispectrum. In §3.6, we use this fact to better understand the behavior of the RMS map (Figure 3.2) in regions dominated by a single term.

3.5.2 Double Dominance

Our goal is to show that if the sum of two terms dominates the third in R, then the variance of R, $\mathcal{A}^2$, is well approximated by that of the two dominant terms, and that the error in making this approximation is one order higher than the result itself. The requirement that the sum of two terms is much larger than the remaining term is only one condition of our double dominance criterion (3.19), but this condition is sufficient to show that $\mathcal{A}$ is well approximated by the contribution of only two terms. Our double dominance criterion is more strict in order to distinguish regions where two terms are both large from those where one term nearly dominates the full bispectrum, and its sum with either of the other two (comparably small) terms is much larger than the remaining term.

Without loss of generality we take $B_{31} \ll B_{12}+B_{23}$. We again begin from equation (3.30) and calculate $\mathcal{A}^2$ in terms of the variance of the two dominant terms, $\mathcal{A}^2_{12+23}$. We have

$$\begin{aligned} R &= \left(\frac{B_{12}+B_{23}}{B^{\rm nw}_{12}+B^{\rm nw}_{23}}\right)\frac{1+B_{31}/(B_{12}+B_{23})}{1+B^{\rm nw}_{31}/(B^{\rm nw}_{12}+B^{\rm nw}_{23})} \\ &\approx R_{12+23}\left(1+\epsilon_{31/(12+23)}\right)\left[1-\epsilon^{\rm nw}_{31/(12+23)}+\left(\epsilon^{\rm nw}_{31/(12+23)}\right)^2\right] \end{aligned} \tag{3.44}$$

where to obtain the second line we Taylor-expanded the denominator to second order in $\epsilon \ll 1$. We defined ϵ and its no-wiggle analog as

$$\epsilon_{31/(12+23)} = \frac{B_{31}}{B_{12}+B_{23}}, \quad \epsilon^{\rm nw}_{31/(12+23)} = \frac{B^{\rm nw}_{31}}{B^{\rm nw}_{12}+B^{\rm nw}_{23}}. \tag{3.45}$$

We also defined R_{12+23} as the first factor in the first line of equation (3.44), as in equation

(3.17). Multiplying out equation (3.44) and dropping $\mathcal{O}(\epsilon^3)$ terms, we obtain

$$\begin{aligned} R &\approx R_{12+23}\left[1+\Delta^{\epsilon}-\Delta^{\epsilon\times}\right], \\ \Delta^{\epsilon} &\equiv \epsilon_{31/(12+23)}-\epsilon^{\rm nw}_{31/(12+23)}, \\ \Delta^{\epsilon\times} &\equiv \epsilon^{\rm nw}_{31/(12+23)}\epsilon_{31/(12+23)}-\left(\epsilon^{\rm nw}_{31/(12+23)}\right)^2. \end{aligned} \tag{3.46}$$

The two small parameters $\epsilon_{31/(12+23)}$ and $\epsilon^{\rm nw}_{31/(12+23)}$ differ only in the BAO feature, which we again parameterize by w_{ij} (equation 3.16):

$$\epsilon_{31/(12+23)} = B^{\rm nw}_{31}\frac{1+w_{31}}{B^{\rm nw}_{12}+B^{\rm nw}_{23}}\left[1+\frac{B^{\rm nw}_{12}w_{12}+B^{\rm nw}_{23}w_{23}}{B^{\rm nw}_{12}+B^{\rm nw}_{23}}\right]^{-1} \tag{3.47}$$

motivating us to define

$$\bar{w}_{12,23} = \frac{B^{\rm nw}_{12}w_{12}+B^{\rm nw}_{23}w_{23}}{B^{\rm nw}_{12}+B^{\rm nw}_{23}} \equiv \bar{w}, \tag{3.48}$$

where in the second, identical equality we are simply noting that we will drop the subscripts on $\bar{w}$. Physically, $\bar{w}$ is the weighted average of the BAO features in the B_{12} and B_{23} terms in the bispectrum (3.1). Expanding equation (3.47) to second order in w and $\bar{w}$, we find

$$\epsilon_{31/(12+23)} \approx \epsilon^{\rm nw}_{31/(12+23)}\left[1+(w_{31}-\bar{w})+\bar{w}^2-w_{31}\bar{w}\right]. \tag{3.49}$$

So we see that

$$\begin{aligned} \Delta^{\epsilon} &\approx \epsilon^{\rm nw}_{31/(12+23)}(w_{31}-\bar{w})(1-\bar{w}), \\ \Delta^{\epsilon\times} &\approx \left(\epsilon^{\rm nw}_{31/(12+23)}\right)^2(w_{31}-\bar{w})(1-\bar{w}). \end{aligned} \tag{3.50}$$

Retaining only terms at second order and lower, we then have

$$\begin{aligned}\Delta^{\epsilon} &\approx \epsilon^{\mathrm{nw}}_{31/(12+23)}\left[w_{31}-\bar{w}\right],\\ \Delta^{\epsilon\times} &\sim \mathcal{O}(\epsilon^2 w)\end{aligned} \tag{3.51}$$

Now we have that to second order $R = R_{12+23}\left(1+\Delta^{\epsilon}\right)$, and we find that

$$\begin{aligned}\left\langle R^2\right\rangle &\approx \left\langle R^2_{12+23}\right\rangle + 2\left\langle R^2_{12+23}\Delta^{\epsilon}\right\rangle,\\ \langle R\rangle^2 &\approx \langle R_{12+23}\rangle^2 + 2\langle R_{12+23}\rangle\langle R_{12+23}\Delta^{\epsilon}\rangle\end{aligned} \tag{3.52}$$

including all terms at second order and below.

Thus to second order the variance $\mathcal{A}^2_{12+23}$ in the dominant terms differs from the variance $\mathcal{A}^2$ in the full bispectrum only by

$$\mathcal{A}^2 - \mathcal{A}^2_{12+23} \approx 2\left\{\left\langle R^2_{12+23}\Delta^{\epsilon}\right\rangle - \langle R_{12+23}\rangle\langle R_{12+23}\Delta^{\epsilon}\rangle\right\}. \tag{3.53}$$

As in §3.5.1 above, the leading contribution to the difference has $R_{12+23}\approx 1$, in which case the two terms on the right side of equation (3.53) cancel. The error is therefore $\mathcal{O}\left((\epsilon w)^{3/2}\right)$, where our notation $(\epsilon w)^{3/2}$ indicates that the error is third order in small quantities but can have any combination of ϵ and w reaching that order. In contrast, the result $\mathcal{A}^2_{12+23}$ is second-order in w. Thus we have shown that the error of approximating the variance of the full ratio R by that of the ratio of the first two terms, R_{12+23}, is one order smaller than the variance itself.

3.6 Numerical Eigen-Root-Mean-Square Calculations

As shown analytically in §3.5, in the regions of single and double dominance identified in §3.4, the RMS amplitude of BAO in the full bispectrum simplifies to the RMS amplitude

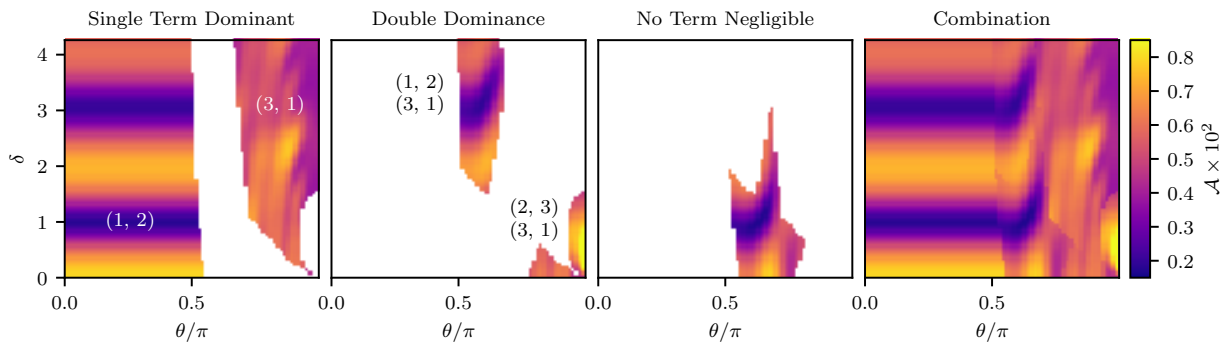

Figure 3.10: The detailed structure of the full RMS map can be understood by considering the RMS amplitude produced by only single terms or pairs of terms in the bispectrum cyclic sum. *Left*—The single-term-dominant contribution: BAO amplitude associated with R_{12} and R_{31} in regions where only B_{12} or B_{31} (indicated on the plot) dominates, detailed in §3.6.1. *Middle Left*—The double-term-dominant contribution: regions where one term is negligible, detailed in §3.6.2 (upper middle is $B_{12} + B_{31}$ dominant; lower right is $B_{23} + B_{31}$ dominant). *Middle Right*—The no-term-negligible contribution: regions where all terms are of comparable magnitude, detailed in §3.6.3. *Right*—By combining the other three panels, we reproduce the full RMS map of Figure 3.2.

of BAO in the dominant term or terms of the bispectrum cyclic sum (3.1). We calculate RMS maps for each single term B_{ij} and pair of terms $B_{ij} + B_{jk}$. In Figure 3.10, we combine the single- and double-term RMS maps in the corresponding single- and double-dominance regions; the result matches the full RMS map of Figure 3.2 reasonably well. The simplified maps of $\mathcal{A}$ therefore provide a good approximation to the full RMS map. In the remainder of this section, we fully detail the RMS maps for each single term (§3.6.1) and pair of terms (§3.6.2) over the full (δ, θ) plane.

3.6.1 Single Term Dominant

While the BAO amplitude in the full bispectrum (3.1) is a complicated function of triangle configuration, many configurations have only a single term dominant, as discussed in §3.4. In those regions, the behavior of the RMS amplitude can be understood through the interaction between pairs of oscillating power spectra. In the red ($\theta \lesssim 0.5\pi$) and blue ($\theta/\pi \sim 1$) regions of the dominance map (Figure 3.4) respectively, B_{12} and B_{31} dominate. We expect that in

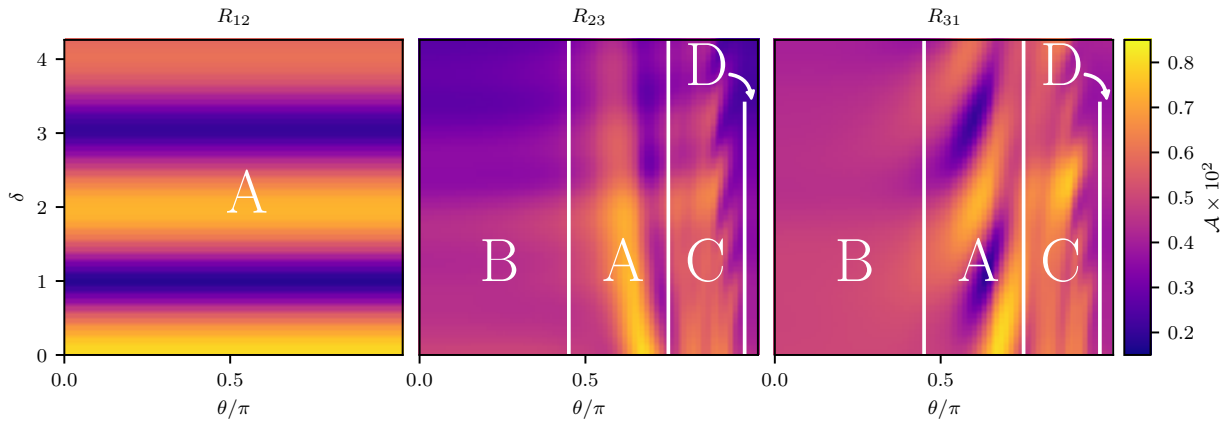

Figure 3.11: In each labeled region of the single-term-dominant RMS maps, the RMS amplitude $\mathcal{A}$ of the BAO feature is driven by a different mechanism. $\mathcal{A}$ is shown for R_{12} (left panel), R_{23} (middle panel), and R_{31} (right panel). The mechanisms are discussed in detail in §3.6.1: interference in region A (§3.6.1), incoherence in region B (§3.6.1), feathering in region C (§3.6.1), and single power spectrum in region D (§3.6.1). The labeled regions are identical for R_{23} and R_{31}, while interference is the only mechanism in R_{12}.

these regions, the RMS map (Figure 3.2) is well approximated by the RMS amplitude of BAO in the dominant term only (left panel of Figure 3.10).

Figure 3.11 zooms in on the RMS maps for each of the single terms, that is, the RMS amplitude (3.6) of the ratio of each term to its no-wiggle analog (3.15). No region with $\delta > 0$ has B_{23} dominant (see §3.4), but we discuss this term as well, both for completeness and because it nonetheless shares interesting physics with B_{31}.

Each labeled region of the single-term-dominant RMS maps (Figure 3.11) is driven by one of the following mechanisms: interference (region B, §3.6.1), incoherence (region A, §3.6.1), feathering (region C, §3.6.1), or single power spectrum (region D, §3.6.1). Only the first mechanism, interference, applies to B_{12}, while all four mechanisms occur in B_{23} and B_{31}.

The incoherence, feathering, and single power spectrum mechanisms arise from differences in the rate at which k_3 varies with k_1 across a configuration. At fixed δ and θ, the wavenumbers k_2 and k_3 vary with k_1 according to equation (3.7); their derivatives with

respect to k_1 at fixed δ and θ are

$$\begin{aligned}\frac{\mathrm{d}k_2}{\mathrm{d}k_1} &= 1, \\ \frac{\mathrm{d}k_3}{\mathrm{d}k_1} &= \frac{(k_1+k_2)(1+\cos\theta)}{k_3} \\ &= \frac{(2k_1+\delta\lambda_f/2)(1+\cos\theta)}{\sqrt{k_1\left(1+\cos\theta\right)\left(2k_1+\delta\lambda_f\right)+\left(\delta\lambda_f/2\right)^2}}.\end{aligned} \tag{3.54}$$

The behavior of dk_3/dk_1 differs across the three regions marked in the right two panels of Figure 3.11. First, at the left edge of the RMS map (region B) where $\theta = 0$, k_1 and k_2 are parallel, so $k_3 = k_1 + k_2$ and $\cos\theta = 1$. The derivative in equation (3.54) then simplifies to $dk_3/dk_1 = 2$. Second, in regions C and D where θ approaches π, $\mathbf{k}_2$ is antiparallel to $\mathbf{k}_1$, so k_3 is the difference between the other two sides: $k_3 = k_2 - k_1 = \delta$. In other words, as $\theta \to \pi$ in these configurations, $dk_3/dk_1 \to 0$, and k_3 is independent of k_1 for any fixed δ and θ configuration. Third, the only configuration where dk_3/dk_1 is unity for all k_1 is the equilateral triangle in region A, where $\theta/\pi = 2/3$ and $\delta = 0$, implying $k_1 = k_2 = k_3$. In general, the rate of change of k_3 with k_1 increases as θ decreases or δ increases.

In regions approaching $\theta = 0$ or $\theta/\pi = 1$, therefore, k_3 may vary across a configuration twice as quickly as k_1, or not at all. When shown as a function of k_1, the oscillations in P_3^{BAO} are consequently stretched (as $\theta \to \pi$) or compressed (as $\theta \to 0$) relative to the oscillations in P_1^{BAO}. While the wavelength of oscillations in P_1^{BAO} and P_2^{BAO} is the BAO fundamental wavelength λ_f (equation 3.4), the wavelength of oscillations in P_3^{BAO} can be infinitely large or as small as $\lambda_f/2$. We find that the interference picture is a good description of the interaction between two oscillations when the ratio of their wavelengths is less than roughly 1.4; when the wavelength of P_3^{BAO} differs from that of P_2^{BAO} and P_1^{BAO} by more than this factor, the concept of a phase shift between P_3^{BAO} and the other power spectrum in the product becomes meaningless because the wavelengths are simply too different. In the products $P_2^{\mathrm{BAO}} P_3^{\mathrm{BAO}}$ and $P_3^{\mathrm{BAO}} P_1^{\mathrm{BAO}}$, then, the BAO amplitude is no longer determined

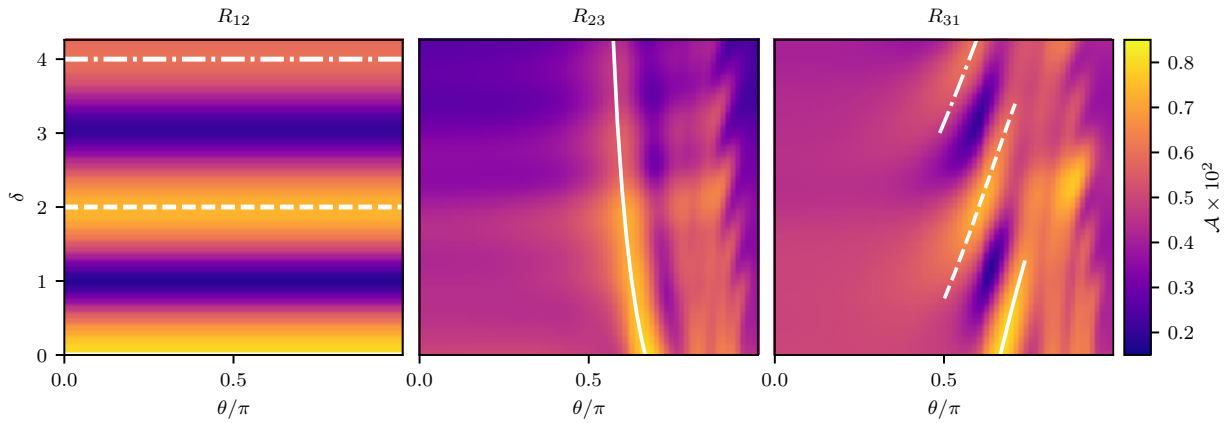

Figure 3.12: When two power spectra are in phase—that is, when k_i and k_j differ by a multiple n of the BAO fundamental wavelength λ_f—constructive interference increases $\mathcal{A}$, as discussed further in §3.6.1. The curves show $k_2 = k_1 + n\lambda_f$ (left panel, where $n = \delta/2$ as odd integer values of δ produce destructive interference), $k_3 = k_2 + n\lambda_f$ (middle panel), and $k_3 = k_1 + n\lambda_f$ (right panel). Solid curves have $n = 0$, dashed $n = 1$, and dot-dashed $n = 2$. For R_{23} and R_{31} (middle and right), the curves are calculated assuming $k_1 = 0.1\,\mathrm{Mpc}/h$, i.e., in the middle of the k_1 range used in this work. Curves are shown only where the wavelength of P_i^{BAO} as a function of k_1 differs from the wavelength of P_1^{BAO} by less than a factor of 1.4 (as explained in §3.6.1), which in R_{23} is the case only for $n = 0$. At higher and lower θ where the wavelengths differ more widely, the RMS amplitude is driven not by the phase difference between the two spectra, but instead by the alignment of individual peaks, as described in §3.6.1, §3.6.1, and §3.6.1.

by any phase shift between the oscillations, but instead by the alignment of the first (and highest, as subsequent peaks will be suppressed by Silk damping) peaks in each.

Interference

Our basis was designed to highlight interference effects between pairs of power spectra, which determine $\mathcal{A}$ in the regions marked "A" of Figure 3.11. As outlined in §3.2, when two wavenumbers k_i and k_j differ by a multiple of the fundamental BAO wavelength λ_f (equation 3.4), the two power spectra P_i^{BAO} and P_j^{BAO} interfere constructively and amplify BAO. In all three terms shown in Figure 3.12, interference produces bright ridges of amplified BAO corresponding to the configurations given by equations (3.8–3.10), where two wavenumbers differ by $n\lambda_f$.

The left panel of Figure 3.12 is clearly similar to the low-θ region of the full RMS map (Figure 3.2) where B_{12} dominates. RMS amplitude is maximized for $\delta = 0$, where $k_1 = k_2$ and the two BAO features are perfectly in phase. Constructive interference repeats wherever the phase difference between the two power spectra is a multiple of the BAO fundamental wavelength λ_f, that is, for even integer values of δ. The first two harmonics are marked in the left panel of Figure 3.12. As δ increases, the maximum $\mathcal{A}$ at even δ declines. This is a result of the declining amplitude of the BAO feature at small scales due to Silk damping. As δ rises, the k_1 dependence is unchanged, but k_2 becomes large enough that Silk damping reduces the amplitude of $P_2^{\rm BAO}$; as the BAO wiggle contribution is damped, it provides less enhancement. In practice, nonlinear structure formation would also degrade the BAO signal at large δ, similar to the effects at large k_1 discussed in §3.2.

Power spectra also interfere constructively and destructively to produce distinct ridges and troughs in the R_{23} and R_{31} RMS maps (right two panels of Figure 3.12). As in R_{12}, we expect the RMS amplitude in R_{23} to be maximized when $P_2^{\rm BAO}$ and $P_3^{\rm BAO}$ are in phase or differ by a multiple of the wavelength, and the RMS amplitude in R_{31} to be maximized when $P_3^{\rm BAO}$ and $P_1^{\rm BAO}$ are in phase or differ by a multiple of λ_f. The solutions (equations 3.9 and 3.10) to $k_2 = k_3 + n\lambda_f$ and $k_3 = k_1 + n\lambda_f$, however, depend not only on δ and θ, but also on k_1. As a result, for a single choice of (δ, θ), it is not possible for k_2 to equal k_3 (or k_3 to equal k_1) for all k_1 in a configuration. We therefore choose $k_1 = 0.1$ Mpc$/h$ (that is, in the middle of our k_1 range) as a representative value of k_1. We evaluate equation (3.9) at $k_1 = 0.1$ Mpc$/h$ to compute the curve of $k_3 = k_2$ shown in the middle panel of Figure 3.12. This curve does correspond to maximum $\mathcal{A}$, but the $k_3 = k_2 + \lambda_f$ curve does not; it falls in region B (labeled in Figure 3.11 and discussed in §3.6.1), where the wavelengths of $P_3^{\rm BAO}$ and $P_2^{\rm BAO}$ are widely different and the interference picture no longer applies. We also evaluate equation (3.10) at $k_1 = 0.1$ Mpc$/h$ to compute the curves of $k_1 = k_3 + n\lambda_f$ shown in the right panel of Figure 3.12.

Incoherence

In Region B (labeled in Figure 3.11) of the R_{23} and R_{31} RMS maps, the RMS amplitude is relatively uniform; $\mathcal{A}$ is neither maximized nor minimized for these configurations. Because the wavelength of $P_3^{\rm BAO}$ is much shorter than that of the other two power spectra in this region, the power spectra entering the products $P_2^{\rm BAO}P_3^{\rm BAO}$ and $P_3^{\rm BAO}P_1^{\rm BAO}$ are incoherent: they cannot interfere constructively or destructively, and patterns in $\mathcal{A}$ arise from the amplitudes of the largest-scale peaks in the power spectra. BAO amplitude can only be enhanced when two peaks—a single pair—in the two power spectra align with and amplify each other, and the greatest amplitude occurs where these peaks are at large scales and therefore minimally Silk-damped.

As $\theta \to 0$, each configuration spans a wider range of k_3 for a fixed range of k_1, meaning that any change in k_1 maps to a larger change in k_3 (equation 3.54). In the $\theta = 0$ limit, for example, $k_3 = k_1 + k_2$ spans at least twice the range of k_1. The wavelength of $P_3^{\rm BAO}$ as a function of k_1 is compressed relative to that of $P_2^{\rm BAO}$ as a function of k_1 (k_2 everywhere changes with k_1 at the same rate, since these two are related by addition of δ). In the small-θ region, therefore, all the products of power spectra in terms that include k_3—R_{23} and R_{31}—are products of oscillations with very different wavelengths (see Figure 3.13).

Although most of the pattern is washed out in the low-θ region of the R_{23} RMS map (middle panel of Figure 3.12), faint banding is still visible around integer values of δ. The maxima diminish with increasing δ as they do in R_{12}—the amplitude of the BAO oscillation in the power spectrum drops at higher wavenumbers. The banding is a result of the relative phase (controlled by δ) between $P_2^{\rm BAO}$ and $P_3^{\rm BAO}$ at low k_1. Because the wavelengths of $P_2^{\rm BAO}$ and $P_3^{\rm BAO}$ are very different, the peaks do not align more than once. The RMS amplitude is highest when the pair of aligned peaks are both large, but the amplitude of the oscillation in each $P^{\rm BAO}$ falls with increasing wavenumber. Therefore, BAO are maximized when the lowest-wavenumber peak (or trough) in $P_2^{\rm BAO}$ aligns with the lowest-wavenumber peak (or trough) in $P_3^{\rm BAO}$. When the lowest-wavenumber peak in $P_2^{\rm BAO}$ aligns with a trough

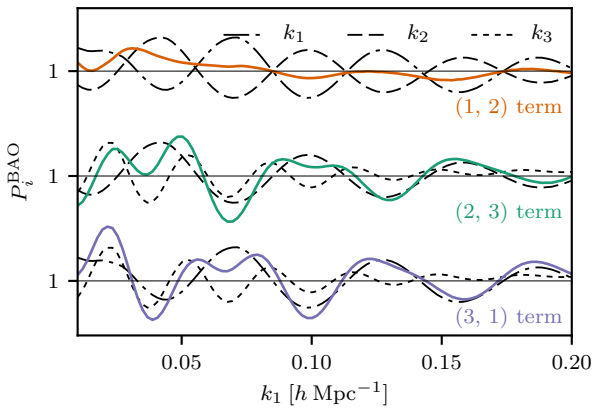

Figure 3.13: For a configuration with $\theta/\pi = 0.2$ and $\delta = 1$, the RMS amplitude in the R_{12} (uppermost set of curves above) term is driven by phase differences (i.e., interference, described in §3.6.1), while the pattern in the R_{23} (middle set) and R_{31} (lower set) terms is a result of wavelength differences (i.e., incoherence, described in §3.6.1). Black curves show the ratio of the linear to the no-wiggle power spectrum, $P_i^{\rm BAO} = P(k_i)/P_{\rm nw}(k_i)$, for each wavenumber as it varies with k_1; the product of each pair of ratios is shown in color ($P_1^{\rm BAO}P_2^{\rm BAO}$ in orange, $P_2^{\rm BAO}P_3^{\rm BAO}$ in teal, and $P_3^{\rm BAO}P_1^{\rm BAO}$ in lavender). For example, the oscillations in $P_1^{\rm BAO}$ and $P_2^{\rm BAO}$ are out of phase, so the power spectra interfere destructively in P_{12} (orange, discussed in §3.6.1). In contrast, P_{23} and P_{31} include $P_3^{\rm BAO}$. At low θ, k_3 can vary over more than twice the range of k_1, so the oscillations in $P_3^{\rm BAO}$ are compressed relative to the others (e.g., compare the short-dashed curve to the dot-dashed curve in the lower set of curves). As can be seen in the lavender P_{31} term, the two interfering oscillations have very different wavelengths, so their product is neither "constructive" nor "destructive." The behavior in P_{23} (teal) is similar; see §3.6.1 for further discussion. Figure is reproduced from Chapter 2.

in $P_3^{\rm BAO}$, the small-wavenumber contributions cancel, and BAO are minimized.

The $\theta \to 0$ region of the R_{31} RMS map behaves similarly to the same region in the R_{23} RMS map—the wavelength of $P_3^{\rm BAO}$ is again much shorter than the wavelength of $P_1^{\rm BAO}$. The phase of $P_1^{\rm BAO}$ is fixed, so the faint banding pattern is diminished in R_{31}. Silk damping still decreases the amplitude of oscillations in $P_3^{\rm BAO}$ as δ increases; at large δ, $P_3^{\rm BAO}$ becomes smooth and approaches unity. The RMS amplitude in R_{31} therefore approaches that of $P_1^{\rm BAO}$ only. Near $\delta = 2$, the small-wavenumber region is a minimum of $P_3^{\rm BAO}$. As in R_{12}, contribution of the low-wavenumber region to the amplitude is therefore minimized, resulting in a faint minimum in the RMS map.

Feathering

In Region C of the R_{23} and R_{31} RMS maps in Figure 3.11, small maxima and minima alternate as δ increases. We refer to this behavior as "feathering," a pattern of bright feathers alternating with regions of lower $\mathcal{A}$. As in Region B (§3.6.1), the behavior of $\mathcal{A}$ in this region results from the difference in wavelength between $P_3^{\rm BAO}$ and the other two power spectra. In Region C, θ approaches π, so dk_3/dk_1 (equation 3.54) is small and k_3 changes little with k_1. P_{31} is then $P_1^{\rm BAO}$ modulated by a stretched-out and slowly varying $P_3^{\rm BAO}$. Across the full range of k_1, $P_3^{\rm BAO}$ traverses half a wavelength. If this half wavelength starts from an extremum of $P_3^{\rm BAO}$ where k_1 is small, and ends at the other extremum of $P_3^{\rm BAO}$ where k_1 is large (see Figure 3.14), $\mathcal{A}$ is maximized. In contrast, $\mathcal{A}$ is minimized between the bright feathers, where instead $P_3^{\rm BAO} \sim 1$ for small k_1, and again $P_3^{\rm BAO} \sim 1$ for the highest k_1 in our range. In this case, the range of $P_3^{\rm BAO}$ is halved relative to the maximum case, and the amplitude contribution due to the k_3 modulation is minimized.

In region C of the R_{31} RMS map (right panel of Figure 3.11), bright feathers alternate with brighter feathers (for example, the maximum at $\delta = 2.5$, $\theta/\pi = 0.89$). This alternating pattern arises as $P_3^{\rm BAO}$ at small k_1 moves from a trough, to unity, to a peak. Because Silk damping reduces BAO amplitude at small scales, $P_1^{\rm BAO}$ is maximized for small k_1. If

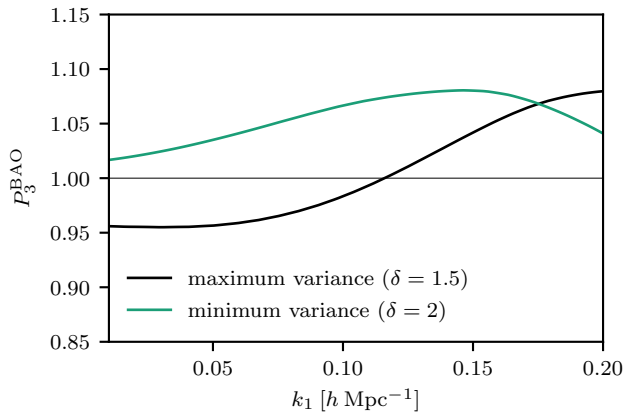

Figure 3.14: For configurations in the "feathering" region (Region C of Figure 3.11), BAO amplitude is driven by the long-wavelength oscillation in $P_3^{\rm BAO}$. The BAO amplitude $\mathcal{A}$ is maximized when $P_3^{\rm BAO}$ varies fully, from trough to peak (black), and minimized when $P_3^{\rm BAO}$ covers only half of that range (green). See §3.6.1 for further discussion.

this maximum coincides with a maximum in $P_3^{\rm BAO}$, its contribution to the RMS amplitude is larger than when it coincides with a minimum in $P_3^{\rm BAO}$. That is, $\mathcal{A}$ is greater when $P_3^{\rm BAO}$ travels from a peak to a trough than vice versa, because Silk damping reduces the contribution of peaks at high k_1 that coincide with the final peak in the latter case. Therefore, while all bright feathers occur where $P_3^{\rm BAO}$ starts from an extremum, they are brighter where that extremum is a maximum and dimmer where it is a minimum. As δ increases, $\mathcal{A}$ declines for the feathers, for the same reason as the R_{12} interference described in §3.6.1. At high δ, k_3 is larger, so Silk damping reduces the amplitude of oscillations.

Similar logic holds for R_{23} (middle panel of Figure 3.11). Bright feathers occur where $P_3^{\rm BAO}$ is either a peak or a trough at low k_1, and the opposite at high k_1; $\mathcal{A}$ is minimized where instead $P_3^{\rm BAO}$ is unity at both small and large wavenumber. However, unlike R_{31}, there is no pattern of alternating brighter and dimmer maxima. In R_{31}, $P_1^{\rm BAO}$ is held fixed while the starting point of the oscillation in $P_3^{\rm BAO}$ varies, so the initial peak in $P_1^{\rm BAO}$ can correspond to either a trough in $P_3^{\rm BAO}$ (as in Figure 3.14) or a peak. But in R_{23}, the starting points of both $P_2^{\rm BAO}$ and $P_3^{\rm BAO}$ depend on δ—and differ by k_1. In this region the magnitude of k_3 is determined by the difference $k_2 - k_1$. At the initial value of k_1 in

our range, $k_1 = 0.01\ h/\text{Mpc}$, the magnitude of k_3 is roughly $k_2 - 0.01\ h/\text{Mpc}$. k_2 and k_3 therefore differ by less than one-fourth of the BAO fundamental wavelength λ_f. Whether P_3^{BAO} is maximized or minimized at $k_1 = 0.01\ h/\text{Mpc}$, P_2^{BAO} at $k_1 = 0.01\ h/\text{Mpc}$ must fall in the same quarter wavelength, so its value must be close to that of P_3^{BAO} but closer to unity. Since the difference between k_2 and k_3 is fixed at small k_1 (unlike the difference between k_1 and k_3 at small k_1), an initial peak or trough in P_3^{BAO} can only correspond to a more limited range of values of P_2^{BAO}, removing the mechanism by which the brightest feathers appear in the RMS map for R_{31}.

Single Power Spectrum

Configurations in Region D (labeled in Figure 3.11) are squeezed, and only one power spectrum term contributes BAO. With only one oscillation, there can be no interference to amplify BAO, so $\mathcal{A}$ is fairly uniform in Region D of the R_{23} and R_{31} RMS maps.

When $\theta/\pi = 1$, equation (3.54) gives $dk_3/dk_1 = 0$; that is, k_3 is independent of k_1 for any choice of δ. R_{23} is then simply the oscillation from P_2^{BAO} alone, with no interference. Similarly, R_{31} reduces to P_1^{BAO}. In both terms, the bispectrum BAO come solely from the oscillation of one P^{BAO}. This oscillation is multiplied by P_3^{BAO}, which does introduce a slight dependence on the parameter δ. While P_3^{BAO} is constant as a function of k_1 for any value of δ, the value of that constant does depend on δ: as in equation (3.21), $k_3 = \delta\lambda_f/2$ for all k_1. The argument of P_3^{BAO} changes with δ, so the level of P_3^{BAO} oscillates up and down as δ increases. In R_{31}, $\mathcal{A}$ along the $\theta/\pi = 1$ line depends only on the level of P_3^{BAO}. When $P_3^{\text{BAO}} > 1$, the entire oscillation in P_1^{BAO} is stretched vertically by a factor greater than unity, slightly increasing $\mathcal{A}$. The opposite is true when $P_3^{\text{BAO}} < 1$, which compresses the amplitude of the P_1^{BAO} oscillation and decreases the RMS amplitude. This effect diminishes at higher δ, as P_3^{BAO} converges to unity.

In R_{23}, again the oscillation of P_3^{BAO} with changing δ causes the $\theta/\pi = 1$ RMS amplitude to depend on δ. Additionally, as δ increases, P_2^{BAO} is increasingly Silk-damped, smoothly

decreasing the RMS amplitude in $P_2^{\rm BAO}$. The faint banding in both R_{31} and R_{23} is visible along the rightmost edge of the middle and right panels of Figure 3.11.

3.6.2 Double Dominance

In the purple ($k_2 > k_1$ but $k_3 \approx k_2$) and green ($\theta/\pi \sim 1$ and δ small) regions of the dominance map (Figure 3.4), two terms are of comparable magnitude. In these regions we calculate the RMS amplitude (3.6), shown in the second panel of Figure 3.10, of the ratio of the sum of the two dominant terms to its no-wiggle analog (3.17).

B_{12} and B_{31} Dominant

The purple region of the dominance map (Figure 3.4), where B_{12} and B_{31} are both dominant, is a region of transition between B_{12} dominance at smaller θ and B_{31} dominance at higher θ. The curve of $k_2 = k_3$ passes through the center, as shown in Figure 3.15. Along this curve, B_{12} and B_{31} are very similar (but not identical, since the fact that k_2 is equal to k_3 for our representative $k_1 = 0.1$ h/Mpc does not imply that $k_2 = k_3$ for all k_1 in a configuration). At θ lower than the cutoff defined by the curve of $k_2 = k_3$ in Figure 3.15, B_{12} begins to dominate. While B_{31} is still large, k_3 is close to k_2, so the oscillations and interference behavior of the two terms are very similar. The RMS amplitude $\mathcal{A}$ is maximum on the lines where $k_2 = k_1 + n\lambda_f$. The reverse holds at θ higher than the $k_2 = k_3$ curve, where B_{31} grows to become dominant. Again, the oscillatory behavior of the two terms is similar, with B_{31} becoming dominant as θ continues to grow.

B_{23} and B_{31} Dominant

In the green region of the dominance map (Figure 3.4), B_{23} and B_{31} are dominant and $\mathcal{A}$ is maximized, as shown in the RMS map (Figure 3.2). These are the squeezed configurations: k_2 points back along k_1, and is slightly longer by $\delta\lambda_f/2$, so the magnitude of k_3 is constant

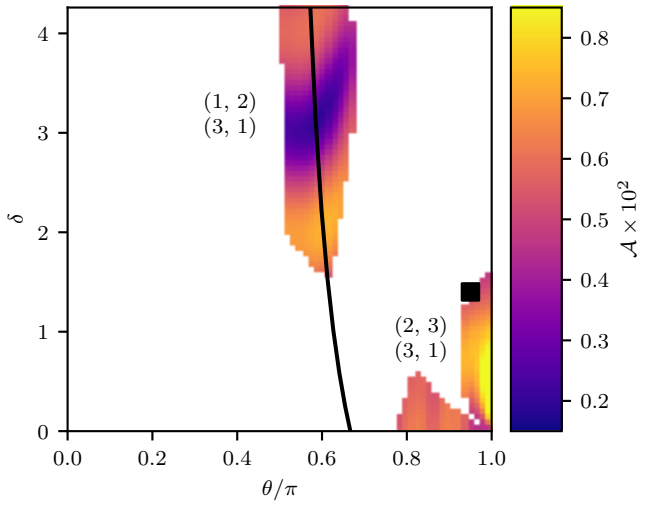

Figure 3.15: Regions where two terms dominate the bispectrum (same as second panel of Figure 3.10) are discussed in §3.6.2. The curve $k_2 = k_3$ for $k_1 = 0.1$ h/Mpc is shown in black. At θ to the left of this curve, the RMS map behaves like that for R_{12}, while for higher θ, it is more similar to that for R_{31}. In both R_{12} and R_{31} in this region, $\mathcal{A}$ is determined by the interference mechanism of §3.6.1. As δ increases above the black square symbol, only B_{31} dominates the cyclic sum.

for any configuration—that is, when $\theta/\pi = 1$ and δ is fixed, $k_3 = \delta\lambda_f/2$ for all k_1 as in equation (3.21). The RMS amplitude arises from the sum of B_{31}, which is large and positive, and B_{23}, which is large and negative (as discussed in §3.4.3). Because k_3 is constant for all k_1 in a configuration, $P_3^{\rm BAO}$ is a constant, so $R_{31} = P_3^{\rm BAO} P_1^{\rm BAO} \propto P_1^{\rm BAO}$ and $R_{23} = P_2^{\rm BAO} P_3^{\rm BAO} \propto P_2^{\rm BAO}$. In the ratio of the sum to its no-wiggle analog, R_{23+31} (equation 3.17), oscillations arise from the difference between the B_{23} and B_{31} contributions: a sine added to a negative sine, only slightly out of phase. As δ is positive, the negative B_{23} is always slightly smaller in magnitude than B_{31}, so R_{23+31} remains positive.

As δ increases, B_{23} shrinks and B_{31} becomes dominant near the square symbol in Figures 3.4 and 3.15, as discussed in §3.4.3. The RMS map transitions into the B_{31}-dominant region described in §3.6.1 above where only one power spectrum contributes BAO to the bispectrum.

3.6.3 No Term Negligible

In the final region around equilateral triangles (black region, Figure 3.4), all three sides are comparable and no term can be disregarded. We must simply reproduce the calculation of the full RMS map in this region, as shown in the middle right panel of Figure 3.10. The RMS amplitude $\mathcal{A}$ is maximized at the equilateral triangle ($\delta = 0$, $\theta/\pi = 2/3$), where all three sides of a triangle are equal—and therefore the $F^{(2)}$ kernels are all equal and the power spectra are all in phase.

3.7 Discussion

3.7.1 Implications for the Reduced Bispectrum

For many triangle configurations, we find that the full bispectrum RMS map is described well by the behavior of only one or two terms in the cyclic sum. The large dynamic range of the $F^{(2)}$ kernel can separate terms by an order of magnitude, allowing us to disregard smaller terms when computing the RMS amplitude. If this held for the reduced bispectrum (defined in e.g. Ref. [26])

$$Q(k_1, k_2, k_3) = \frac{B(k_1, k_2, k_3)}{P(k_1)P(k_2) + P(k_2)P(k_3) + P(k_3)P(k_1)}, \tag{3.55}$$

we could simplify its behavior as well. Our work does show that there are regions where the numerator of Q can indeed be simplified.

However, a more useful approximation would be if one could approximate the denominator by just one term rather than the full cyclic sum of power spectra, or even a pair of products. Unfortunately, though, it is primarily the $F^{(2)}$ kernel that drives the dominance of one term relative to the others (see §3.4.2), and it does not enter the denominator of Q. As the leftmost panel of Figure 3.6 shows, a large swatch of the δ-θ plane is black (no term negligble) for the relevant products of power spectra P_{ij}. While there are several regions

where two terms dominate the others (purple, blue, green), these do not seem to offer a significant simplification as they still produce a complicated denominator in equation (3.55).

It is only in the red (B_{12} dominant) and blue (B_{31} dominant) regions that the denominator greatly simplifies, to respectively P_{12} or P_{31}. Since the dominant term in $F^{(2)}$ is always the same as that in P_{ij} (see §3.4.4 and Figure 3.8), in these regions $F_{12}^{(2)}$ and $F_{31}^{(2)}$, respectively, will be much larger than the other two $F_{ij}^{(2)}$. The bispectrum is therefore well-approximated by respectively $F_{12}^{(2)}P_{12}$ and $F_{31}^{(2)}P_{31}$. Thus, in these limited regions, Q reduces to $F_{12}^{(2)}$ and $F_{31}^{(2)}$. In short, working in the δ-θ basis does highlight convenient triangles where one can directly measure the growth kernel $F^{(2)}$ alone and easily divide out the linear theory density field statistics. The contribution of gravitational growth can thereby be isolated from that of the linear theory density field. This isolation might be especially useful in using the 3PCF as a probe of modified gravity (e.g., Vernizzi et al. 2019, in prep.).

Especially insofar as high-wavenumber details of the power spectrum sourced by baryon physics remain challenging to model, canceling out the power spectrum from measurements of $F^{(2)}$ may be desirable. Of course this must be weighed against the reduction in number of usable configurations, as this cancellation happens only on limited regions of the δ-θ plane. Further, at the wavenumbers where baryons become relevant, a tree-level, linearly biased model of the bispectrum is likely already beginning to falter; the numerator is measuring higher-order perturbation theory kernels and higher-order biasing even in these "simpler," single-dominance regions.

3.7.2 Connection to Real Space

We now briefly discuss the connection of the present chapter to the 3PCF in configuration space (i.e., real space without redshift-space distortions). Ref. [137] further discuss differences and similiarites between bispectrum and 3PCF more generally, though with a focus on bias parameters, most relevant for smaller scales than the BAO scales investigated here. The wiggles in the bispectrum ultimately correspond to sharp features in configuration space, in

particular the BAO creases where one triangle side is the BAO scale or twice the BAO scale. These are visible in Figure 7 of Ref. [290], particularly in the linear bias ($\ell = 1$) panel but also more faintly in $\ell = 0$ and $\ell = 2$. The intuition is much the same as with the 2PCF and power spectrum, where a bump in configuration space leads to a harmonic series of oscillations in Fourier space. One important difference here is that the $F^{(2)}$ kernel, which weights the products of power spectrum by $k^{\pm 1}$, acts like a derivative in configuration space. The BAO feature in the 3PCF is thus essentially the derivative of a BAO bump: positive as the BAO bump rises, then zero at the BAO scale, and negative at larger separations as the BAO bump falls.

3.7.3 Simplification of Multipole Basis

We now point out an interesting additional implication of our work. The multipole basis (expanding the angular dependence of the 3PCF or bispectrum in Legendre polynomials), proposed in Ref. [303] (see also Ref. [237]), has recently been exploited in a series of works [91, 288, 289] to accelerate measurement of the 3PCF. However, in practice that approach truncates the multipole expansion of the 3PCF as one measures it. The works cited above chose a maximum multipole of $\ell_{\rm max} = 10$. In principle, however, even at tree level the 3PCF has support out to infinite ℓ, as the expansion is done with respect to $\hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2$ and k_3 and $1/k_3$ have an infinite multipole series in this variable.

In practice the 3PCF seems well-converged when summed into a function of opening angle using different numbers of multipoles (see Figure 8 of Ref. [288]). Our work shows that for certain configurations the multipole support is in principle finite. In the regions dominated by B_{12}, the bispectrum multipole expansion has compact support, requiring (at least at tree level) only $\ell = 0, 1$, and 2. The same will hold for the 3PCF; Ref. [303] shows that a given ℓ in Fourier space maps to only the same ℓ in configuration space. (This immediately follows from the plane wave expansion into spherical harmonics and spherical Bessel functions, use of the spherical harmonic addition theorem, and orthogonality of the spherical harmonics.)

There are two implications here: first, the adequacy of tree level perturbation theory can be easily tested using a very small set of multipoles in the red region of B_{12} dominance. Second, within this restricted region, the computational work and covariance matrix dimension can be greatly reduced by measuring the 3PCF in the multipole basis only to $\ell_{\rm max} = 2$. Of course, the price is the reduced number of configurations (and signal) available. While this level of compression may not be necessary for isotropic statistics, RSD introduce a much richer angular structure at a fixed $\ell_{\rm max}$ (see Refs. [291, 299]), so a reduction in $\ell_{\rm max}$ may be of particular value.

3.8 Conclusion

Our bispectrum basis (§3.2), designed to identify triangle configurations that amplify the BAO signal, also provides insight on the structure of BAO in the bispectrum. Our analysis in §3.4 shows that for certain triangle shapes, the bispectrum is dominated by only one or two terms of the cyclic sum (3.1). The dominance structure is driven primarily by the $F^{(2)}$ kernel of Eulerian standard perturbation theory (§3.4.2), which is highly sensitive to triangle shape. In §3.5 we show analytically that, because BAO are a small feature relative to the broadband bispectrum, the RMS BAO amplitude in the full bispectrum reduces to the RMS BAO amplitude in the dominant term or terms. The error in this approximation is suppressed by one order relative to the BAO amplitude itself. In §3.6, we build up the complete RMS map of the dependence of BAO amplitude on triangle parameters from simpler maps. These maps show the RMS BAO amplitude in each of the three terms contributing to the cyclic sum. In regions where the corresponding terms dominate the cyclic sum, the full RMS BAO amplitude is well approximated by the single-term-dominant or double-term-dominant maps. We reproduce the full bispectrum RMS map by stitching together these simpler maps in the regions where they provide the dominant contribution to bispectrum BAO, then fully discuss the mechanisms that drive BAO amplitude in each single-term-dominant (§3.6.1) and double-term-dominant (§3.6.2) RMS map.

The BAO amplitude in each single term is determined by one of four mechanisms: interference (§3.6.1), incoherence (§3.6.1), feathering (§3.6.1), or single power spectrum (§3.6.1). The first mechanism, interference, results from phase differences between two power spectra, and dramatically amplifies the BAO signal. The other three mechanisms occur where the wavelengths of the two BAO features are widely different, so the interaction between the two power spectra cannot consistently amplify BAO.

Finally, in §3.7 we outline implications of our work for the reduced bispectrum, its connection to the 3PCF, and the potential to simplify the multipole expansion of the bispectrum and 3PCF for certain triangle shapes.

In Chapter 2, we used the interferometric basis detailed here to obtain substantial improvement in BAO constraints over the power spectrum alone, using a relatively small number of bispectrum measurements that carry the most BAO information. Ideally, bispectrum measurements on all possible triangles would be used to constrain the BAO scale. However, the number of mock catalogs needed to accurately estimate and invert the covariance matrix scales with the number of triangles [241]. The number of triangles that can be measured is therefore limited by the number of mock catalogs available, and bispectrum BAO constraints like those of Ref. [239] are limited by the error in the covariance matrix. Since current resources limit the number of triangles that can be used to constrain BAO in the bispectrum, the best constraints will be obtained from the triangles that carry the most BAO information and are most independent from each other. One way to identify these triangles is by measuring the full covariance matrix, but such an approach faces the same initial problem of limited mock catalogs.

We therefore face a circular problem: because it is computationally prohibitive to use fully N-body mocks to constrain the covariance matrix of all bispectrum triangles, we wish to reduce the size of the covariance matrix by selecting a subset of optimal triangles for BAO constraints. But without the full covariance matrix, how can those triangles be identified? Our basis offers a compression to only those triangles that are most sensitive to BAO,

enabling a 15% improvement over power spectrum BAO constraints using relatively few bispectrum measurements, as shown in Chapter 2 [45].

Of course, the optimal set of triangles for BAO measurement depends not only on the amplitude of the BAO signal in each configuration, but also on its signal to noise ratio and its covariance with previously measured configurations. In future work, we will further develop an algorithm for selecting triangle configurations, assuming the number of mock catalogs available limit the number of bispectrum measurements that can be used.

In future work, we will use also use BAO-sensitive triangles to better understand the covariance structure of BAO in the bispectrum and power spectrum. Reconstruction [84, 226, 235, 236] is expected to affect the covariance between the power spectrum and bispectrum ([270]; see also Ref. [294], §8.2), but as reconstruction is a numerical procedure, its effect on covariance is difficult to model analytically. Like the bispectrum measurements discussed in the previous paragraph, a full numerical study of the covariance between the post-reconstruction power spectrum and the pre-reconstruction bispectrum is limited by the number of fully N-body mocks available. Fewer mocks are needed if analysis is restricted to the set of triangles most sensitive to BAO, reducing the dimension of the covariance matrix. We will study the effects of reconstruction on these triangle configurations. This effort will allow us to combine bispectrum measurements with the post-reconstruction power spectrum. Depending on the level of independence, the combination of bispectrum measurements and reconstruction may offer further improvement in BAO constraints over that offered by reconstruction alone.

Our approach offers many further applications to the study of BAO in the bispectrum, which we plan to address in future work. For example, the phase of BAO in the power spectrum is sensitive to $N_{\rm eff}$, the effective number of relativistic neutrino species [17, 19, 20, 90]. Our basis is very sensitive to phase effects, so it may be useful to constrain $N_{\rm eff}$ using the bispectrum (Child et al. 2019, in prep.). Other sources of a phase shift in power spectrum BAO such as relative velocities between baryons and dark matter [31, 61, 268, 311, 340],

constrained in the power spectrum by Refs. [341] and [28] and in the 3PCF by Ref. [294], may also be constrained using our interferometric basis. Last, our approach may enable study of massive spinning particles, which, if present during inflation, introduce oscillatory cosine terms in the bispectrum [210]. These terms depend on the wavenumbers, so they can interfere with each other when cyclically summed.

CHAPTER 4

HALO PROFILES AND THE CONCENTRATION-MASS RELATION FOR A ΛCDM UNIVERSE

Profiles of dark matter-dominated halos at the group and cluster scales play an important role in modern cosmology. Using results from two very large cosmological N-body simulations, which increase the available volume at their mass resolution by roughly two orders of magnitude, we robustly determine the halo concentration–mass (c–M) relation over a wide range of masses, employing multiple methods of concentration measurement. We characterize individual halo profiles, as well as stacked profiles, relevant for galaxy–galaxy lensing and next-generation cluster surveys; the redshift range covered is $0 \leq z \leq 4$, with a minimum halo mass of $M_{200c} \sim 2 \times 10^{11} M_{\odot}$. Despite the complexity of a proper description of a halo (environmental effects, merger history, nonsphericity, relaxation state), when the mass is scaled by the nonlinear mass scale $M_{\star}(z)$, we find that a simple non-power-law form for the c–$M/M_{\star}$ relation provides an excellent description of our simulation results across eight decades in $M/M_{\star}$ and for $0 \leq z \leq 4$. Over the mass range covered, the c–M relation has two asymptotic forms: an approximate power law below a mass threshold $M/M_{\star} \sim 500 - 1000$, transitioning to a constant value, $c_0 \sim 3$ at higher masses. The relaxed halo fraction decreases with mass, transitioning to a constant value of ~ 0.5 above the same mass threshold. We compare Navarro–Frenk–White (NFW) and Einasto fits to stacked profiles in narrow mass bins at different redshifts; as expected, the Einasto profile provides a better description of the simulation results. At cluster scales at low redshift, however, both NFW and Einasto profiles are in very good agreement with the simulation results, consistent with recent weak lensing observations.

This chapter appears as "Halo Profiles and the Concentration-Mass Relation for a ΛCDM Universe" [43] in the Astrophysical Journal. The measurements presented have also been used to study the contribution of smaller halos to strong lensing ("The importance of sec-

ondary halos for strong lensing in massive galaxy clusters across redshift," [176]) and to demonstrate the properties of the Outer Rim simulation (Heitmann et al. 2019, in prep.).

4.1 Introduction

The endpoints of structure formation in cold dark matter cosmologies are dark matter-dominated clumps called halos. In these cosmologies, initial density perturbations are amplified by the gravitational Jeans instability and small localized nonlinear structures form at high redshift. As the universe evolves, halos grow via mass accretion and halo mergers; galaxies form within halos. Halo abundance, evolution history, and properties such as mass, velocity, sub-structure, and phase space structure, as well as the halo gas and galaxy content, all play important roles in modern cosmology, as well as in the modeling of galaxy formation, whether by empirical or semi-analytic means [18, 21–23, 25, 50, 55, 116, 127, 151, 155, 156, 212, 238, 277, 296, 317, 331, 333, 348]. A number of observational probes based on strong and weak gravitational lensing, X-ray observations, and galaxy clustering are sensitive to the nature of halo density profiles [4, 29, 54, 154, 194, 195, 204, 207, 222, 224, 229–232, 313–316]. This is particularly true for cluster cosmology and galaxy–galaxy lensing, which focus at the upper end of the halo mass range.

In the remainder of the introduction, we briefly discuss halo profiles and concentrations and their importance for cosmology (Section 4.1.1), the current state of observed halo profiles (Section 4.1.2), and prior work on the profiles and concentrations of simulated halos (Section 4.1.3). We then outline the aims and primary results of this work (Section 4.1.4).

4.1.1 Halo Profiles

Although the formation of halos is a complex, hierarchical nonlinear dynamical process, the radial density profile of individual halos is robustly fit by a surprisingly simple form, as first described by Refs. [217, 218], using results from cosmological N-body simulations. The

two-parameter NFW density profile is given by

$$\rho(r) = \frac{\delta_c \rho_{\rm c}}{(r/r_s)\,(1 + r/r_s)^2}, \tag{4.1}$$

where δ_c is a characteristic dimensionless density parameter. The critical density is $\rho_c(z) = 3H^2(z)/8\pi G$; $H(z)$ is the Hubble parameter, and the NFW scale radius, r_s, is defined by the radius where the logarithmic profile slope $n_{\rm eff} = d\ln\rho/d\ln(r/r_s) = -2$. For $r/r_s \ll 1$, $n_{\rm eff} \to -1$, whereas for $r/r_s \gg 1$, $n_{\rm eff} \to -3$. A dimensionless shape parameter, the halo concentration, $c_\Delta \equiv r_\Delta/r_s$, is commonly used as one of the NFW parameters. The halo radius r_Δ is a radial scale set by the spherical overdensity (SO) halo mass definition: $M_\Delta \equiv (4/3)\pi r_\Delta^3 \rho_c \Delta$, where Δ is a dimensionless overdensity parameter. We choose the critical density as the reference density; the mean density of the universe is another common choice. We also make the conventional choice of $\Delta = 200$ (for X-ray work with clusters, higher values of Δ are often used, such as $\Delta = 500$ or $\Delta = 1000$), and refer to the corresponding concentration as c_{200c}.

Describing individual halos in terms of the NFW description is obviously a severe idealization. Halos are not spherical and can have complex shapes. In particular, a more realistic description of individual halos is as prolate ellipsoids with a major axis length roughly twice as long as the minor axis [152]. Additionally, at a fixed halo mass and more or less independent of the how the mass is defined, halo shapes and profiles can display considerable variability, with some dependence on whether the halos are dynamically relaxed [191, 332]. Observations that focus on stacked halos, such as galaxy–galaxy lensing or stacked cluster weak lensing, involve averaging over many individual halos and thus reduce bias due to the characteristics of individual lenses (see, e.g., Ref. [286] as an example of the current state of the art).

Despite some of these caveats, there are well-defined and observationally testable predictions for halo masses and profiles as a function of cosmological parameters. For instance,

the halo mass function is an essential cosmological quantity, relevant to determining cluster abundance [138] and to modeling of the observed galaxy distribution, to mention two obvious examples. The halo profile shape can also be predicted accurately in modern cosmological simulations and is known to be correlated with the halo mass. One aspect of this correlation is the existence of a well-determined c–M relation, as was already noted by NFW. Cosmological constraints delivered by ongoing surveys such as the Dark Energy Survey (DES[1]) and next-generation surveys such as the Dark Energy Spectroscopic Instrument (DESI[2]), Euclid [256], the Large Synoptic Survey Telescope (LSST) [183], and the Wide-field Infrared Survey Telescope (WFIRST) [297] will rely on having accurate predictions for halo profiles and masses.

4.1.2 Observed Profiles: Individual and Stacked Halos

Individual halo profiles can be measured using X-ray and strong and weak lensing measurements, as already mentioned. Because of a number of factors (e.g., observational limitations, selection bias, individual variability, line-of-sight dependence, analysis issues), there are difficulties in comparing these observations directly to theoretical predictions. Earlier measurements tended to have higher concentrations and a significantly steeper c–M relation than that predicted by simulations (see, e.g., Refs. [37, 229, 232, 269] and the discussion in Ref. [54]); however, the state of the art has been significantly enhanced by more recent group and cluster-scale observations. In Figure 4.1 we present a set of recent observational results for the c–M relation from measurements of individual clusters. (For another compilation, see Ref. [29].) We also show results from the simulation carried out in this chapter, which are discussed in detail in Section 4.4.1 below; we find good agreement between the observations and the ΛCDM predictions, despite uncertainties in accounting for selection biases and other measurement errors.

1. https://www.darkenergysurvey.org

2. http://desi.lbl.gov

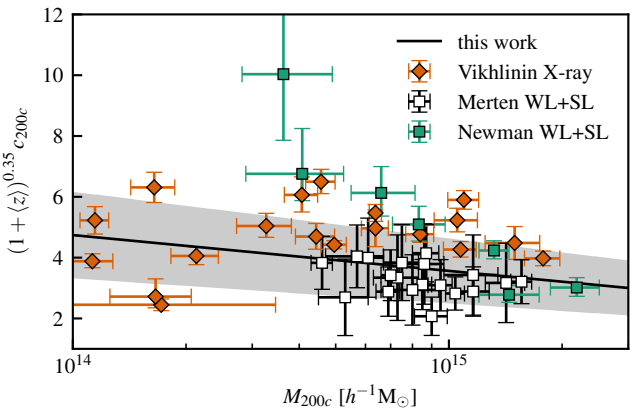

Figure 4.1: Individual-halos c–M relation with individual cluster observations using X-ray [29, 322] and weak and strong lensing [207, 222]; see Section 4.4.1 for details. The gray band represents the 1σ intrinsic scatter in the c–M relation, as found from the simulations.

Given sufficient statistics, stacking techniques can be used for both higher- (cluster weak lensing) and lower-mass halos (galaxy–galaxy lensing). The results from stacked observations average over intrinsic halo variability and lines of sight, but have different systematic issues compared to individual halo measurements. Moreover, stacked density profiles differ systematically from the NFW prescription, with potentially observable consequences. As discussed in Section 4.3.2, the Einasto profile [81] is a much better fit in this case. Next-generation surveys, and LSST in particular, will increase the number of known clusters by over an order of magnitude. Provided systematic errors can be sufficiently controlled, there is, therefore, sufficient motivation to consider the individual and stacked halo profiles separately. Figure 4.2 shows observational results for stacked observations using galaxy–galaxy lensing and cluster weak lensing. Here too, the results are in good agreement; significant improvements in the observational results are expected in the near future.

4.1.3 Concentrations of Simulated Halos

Even without including baryonic effects, which at cluster-scale masses could lead to changes at the $\sim 10\%$ level [79], the current status of theoretical predictions and comparison with

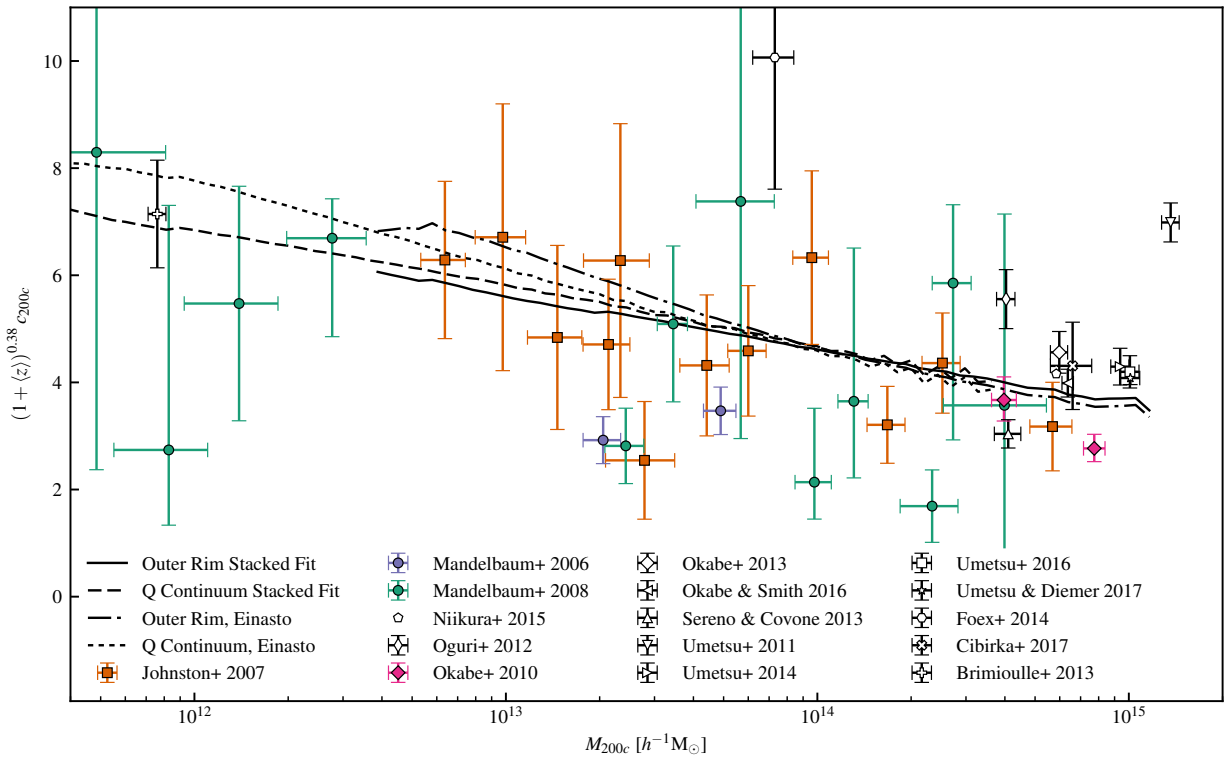

Figure 4.2: Observational results for the stacked-halo c–M relation; see Section 4.4.1 for details. The Einasto fit yields the same concentration as NFW $c_{\rm fit}$ at high masses, but Einasto concentration rises when only one side of the peak is available to fit. Each paper represented by a white point provides a single measurement. Refs. [89] and [313] discuss the effects of strong-lensing bias on their measurements of c_{200c}.

observations as shown in Section 4.4.1 is certainly satisfactory. However, a number of open issues still remain. Simulations have shown that the concentration depends on both mass and redshift, with massive halos less concentrated than lower-mass halos at the same redshift. The c–M relation has been measured in a large number of papers, in different ways [29, 38, 57, 73, 78, 85, 97, 161, 162, 187, 188, 190, 192, 246, 345]. There has been some disagreement in the results obtained, which appears to be largely due to the different ways in which the concentration has been *operationally* defined. Different fitting methods and binning choices can produce inconsistent c–M relations (see, e.g., discussions in Refs. [29, 80, 205]). One of the objectives of this chapter, therefore, is to present a robust set of results, arrived at by using methods with potentially different systematics, and by investigating several possible

sources of numerical error.

4.1.4 This Work

The primary aim of this chapter is to further study and robustly determine the c–M relation and halo profiles at group and cluster scales at low to medium redshifts, and to characterize the profile evolution over this redshift range for both relaxed and unrelaxed halos. The halo mass range considered here is mostly focused on masses significantly larger than the nonlinear (or "collapse") mass scale $M_\star(z)$, which gives the mass scale corresponding to peaks of the initial Gaussian random field collapsing at redshift z (see definition in Section 4.4): $M_\star \approx 10^{12.5}\,h^{-1}M_\odot$ at $z = 0$, $10^{11}\,h^{-1}M_\odot$ at $z = 1$, and $10^{9.5}\,h^{-1}M_\odot$ at $z = 2$. Following the discussion above, we split our c–M relation study into two parts: one for individual halos and the other for stacked halos, in order to be consistent with current and future observational strategies.

We determine the concentration by fitting the radial mass distribution, dM/dr, rather than the density profile, following the procedure outlined in [29]. In addition to this primary method, we introduce two alternative techniques for concentration estimation – one of which is independent of the assumed form of the density profile – that we employ to characterize the robustness with which the c–M relation can be determined. We find that the results are in good agreement over the mass ranges where all of our methods can be properly used. Each of the three methods is sensitive to a different range of the profile, and subject to different types of error – we use the agreement across methods to show that our concentration measurements are robust.

The two massive state-of-the-art ΛCDM simulations (Section 4.2) that form the basis of the results presented here share two essential characteristics: large volumes and very good mass resolution. At the mass resolutions considered, these simulations have roughly two orders of magnitude more volume than previous work. This allows us to robustly explore the c–M relation with excellent statistics over a relatively wide halo mass and redshift range,

and to use narrow mass bins in stacked analyses. Our results are consistent with the general idea that the concentration of halos should be set more or less by the mean density of the universe when the halos are assembled. (Unfortunately, there is no predictive theory for how the c–M relation should depend on redshift, because there is no real theory for the NFW profile either.) At $z = 0$, massive clusters, which are still forming today, would be expected to have a lower concentration than smaller-mass halos that have masses less than $M_\star$. Because $M_\star$ drops steeply with redshift, $\sim 10^8\ M_\odot$ at $z = 3$, one might expect the c–M relation to flatten over a significantly extended mass range as redshift increases and a large fraction of the halos in the upper mass range are still forming. Our results are very consistent with this expectation.

By combining results from multiple redshifts and scaling the halo mass by the nonlinear mass scale $M_\star$, we find that the c–M relation can be well-fit by a single expression; in agreement with Refs. [97, 188, 345], we find a concentration floor of $c_{200c} \sim 3$. The transition from a power-law behavior of the scaled c–M relation to an asymptotically flat regime occurs at a halo threshold mass, $M_T \simeq 500 - 1000 M_\star$, right at the upper end of cluster-scale halo masses at $z = 0$. We note that the unrelaxed halo fraction is roughly half at masses above M_T, and decreases with mass below M_T. Further discussion of these results can be found in Section 4.4.

As $M_\star$ is a function of cosmological parameters, we can compare our results to other simulations in the same family of ΛCDM cosmologies by assuming a certain level of universal behavior, as is known to hold for the mass function (see, e.g., Ref. [130]). Appendix 4.6.1 shows that the c–$M/M_\star$ relations of other simulations fall within the population variance $\bar{c} \pm \bar{c}/3$ of our results, and their results from redshifts between $z = 0$ and $z = 3$ broadly follow the shape of a power-law transitioning to constant c. We have checked that this cosmology-independent behavior does not hold sufficiently far away from our fiducial cosmology by comparison to the wCDM results presented in Ref. [29] and in Ref. [166].

The rest of this chapter is organized as follows. In Section 4.2, we describe the large

cosmological N-body simulations that are the source of the halo catalogs and halo profiles. Section 4.3 describes our methodology for measuring the halo concentrations using the radial profiles. Section 4.4 presents the resulting c–M relations and the new fitting form using scaled halo masses. Section 4.5 concludes with a final discussion of the results. The appendices contains the results of investigations of possible sources of numerical error and comparisons of some of our results to previous work.

4.2 Simulations

The results reported here use data from two very large, gravity-only N-body simulations run with the Hardware/Hybrid Accelerated Cosmology Code (HACC) framework [118]. These are the 'Q Continuum' [129] and 'Outer Rim' [118] simulations carried out on the CPU/GPU system Titan at Oak Ridge National Laboratory and the Blue Gene/Q (BG/Q) system Mira at Argonne National Laboratory, respectively. HACC uses a hybrid force calculation scheme, splitting the total force calculation into a long-range component and a short-range component. In both runs, the long-range forces are computed using the same high-order spectral particle mesh method, while the short-range forces are computed using different methods (albeit with the same hand-over scale to the short-range solver) in order to best exploit the available computational architecture. A direct particle-particle interaction technique is used for the CPU/GPU system and an RCB (Recursive Coordinate Bisection) tree method for the BG/Q system. Halo identification and characterization is carried out with HACC's parallel CosmoTools analysis framework, using a combination of in situ and offline analyses.

The Q Continuum and Outer Rim runs represent independent realizations of the same

shared WMAP-7 [163] cosmology:

$$\begin{aligned}
\omega_{\rm cdm} &= 0.1109 \stackrel{h=0.71}{\Rightarrow} \Omega_{\rm cdm} = 0.220, \\
\omega_{\rm b} &= 0.02258, \\
n_s &= 0.963, \\
h &= 0.71, \\
\sigma_8 &= 0.8, \\
w &= -1.0, \\
\Omega_\nu &= 0.0,
\end{aligned} \tag{4.2}$$

but with differing volumes and mass resolution. The box size for the Q Continuum run is $L_{\rm QC} = 1300$ Mpc $= 923\,h^{-1}$ Mpc, while that of Outer Rim is $L_{\rm OR} = 4225$ Mpc $= 3000\,h^{-1}$ Mpc. The number of particles in these simulations are $8192^3 = 0.55$ trillion (Q Continuum) and $10240^3 = 1.1$ trillion (Outer Rim); the associated mass resolutions are $m_p = 1.48 \times 10^8\ M_\odot = 1.05 \times 10^8\,h^{-1} M_\odot$ (Q Continuum) and $m_p = 2.6 \times 10^9\ M_\odot = 1.85 \times 10^9\ h^{-1} M_\odot$ (Outer Rim). The force resolutions are (comoving) $2\,h^{-1}$ kpc (Q Continuum) and $3\,h^{-1}$ kpc (Outer Rim). Both simulations are given a Zel'dovich approximation initial condition at $z = 200$ with transfer functions generated by the CAMB code [175]. The Outer Rim simulation has been used for several analyses of SDSS IV extended Baryon Oscillation Spectroscopic Survey data [104, 139, 342].

The large volumes and excellent mass resolution in these simulations lead to the following advantages in characterizing halo properties: (1) sufficiently large numbers of halos at high masses over the redshift ranges studied (at $z = 0$, ~ 20 million and ~ 10 million halos of at least 2000 particles in Q Continuum and Outer Rim, respectively); (2) excellent profile resolution for individual halos; and (3) the ability to study stacked halo profiles in narrow mass bins – hundreds of halos in mass bins of width $\pm 5\%$ at cluster scales, and hundreds of thousands at lower masses. Compared to our previous work in Ref. [29], the mass resolution

is improved by more than an order of magnitude. In addition, the overlapping volume and mass resolution coverage between the two boxes (which are run using different N-body algorithms) provides an automatic cross-check for certain types of systematic errors that can arise in cosmological simulations.

4.3 Concentration Measurement

In this section, we describe our methods for measuring halo concentrations. This requires first defining and measuring the halo mass, followed by a determination of the halo concentration. In the context of the c–M relation, the mass is usually defined in terms of an SO, Δ, as discussed in Section 4.1. A halo with mass M_Δ has a corresponding size, r_Δ, the radius within which the halo has an average overdensity of Δ with respect to ρ_c. Common choices of Δ include $\Delta = 200$ and $\Delta = \Delta_{\rm vir}$, where $\Delta_{\rm vir}$ follows from the spherical top-hat collapse model. It is also not uncommon to define the overdensity with respect to the mean density of the universe, rather than the critical density. As stated earlier, we will use the critical density as the reference, with $\Delta = 200$. It should be noted that the definition of the SO mass does not depend on the nature of the density profile.

A common alternative NFW parameterization, which we use here, describes the NFW profile in terms of an SO halo mass and the halo concentration, c_Δ. Written in terms of the SO radius r_Δ and the concentration, the NFW profile becomes

$$\rho(r) = \frac{\Delta\rho_c}{3A(c_\Delta)} \frac{1}{(r/r_\Delta)(1/c_\Delta + r/r_\Delta)^2}, \tag{4.3}$$

where $A(c_\Delta) \equiv \ln(1 + c_\Delta) - c_\Delta/(1 + c_\Delta)$.

Using the NFW profile as defined by Equation 4.3, it is clear that given r_Δ (or equivalently, M_Δ) and c_Δ (or equivalently, r_s), the NFW profile is uniquely determined. Given the NFW description, the SO mass and concentration together completely determine the spherically averaged halo profile.

Halo concentrations are computed using halo profiles built by HACC's parallel SO halo finder, which is part of the CosmoTools analysis framework. First, friends-of-friends (FOF) halos with dimensionless linking length $b = 0.168$ are found by a fast, parallel, tree-based algorithm, and their centers determined by finding the deepest potential minimum within the FOF halo. All particles (not just those in the original FOF halo) are counted in radial shells centered on the point of minimum potential, and the mass M_{200c} is calculated as the mass within a sphere whose average density is 200 times the critical density. No unbound particles are removed because we are interested in the density profile as measured and not in some idealized theoretical notion of what might constitute membership in a halo. Twenty shells are placed uniformly in log space between a minimum radius at the smoothing scale and a maximum radius greater than R_{200c}, and the differential mass profile dM/dr is calculated from the bin widths and particle counts. We note that the notion of an SO halo becomes problematic during major halo mergers; the halo center is also potentially not well-defined during such epochs.

The concentration measurement procedure uses three different methods, explained in detail below: profile fit (as in Ref. [29]), accumulated mass, and peak finding. The idea behind using these different methods is to explore the robustness with which the concentration can be determined in the presence of different types of systematic errors. We measure concentrations only for well-sampled halos, i.e., those above a conservative threshold of at least 2000 particles within r_{200c} (see Appendix 4.6.2), corresponding to a mass of $2.1 \times 10^{11}\ h^{-1}M_{\odot}$ for Q Continuum and $3.7 \times 10^{12}\ h^{-1}M_{\odot}$ for Outer Rim.

Following standard practice, for all three methods we keep the mass M_{200c} fixed as found by the SO algorithm described above and fit only for the concentration. In principle, it is possible to allow both quantities to float, but this leads to variability in the determined concentrations, even if the associated M_{200c} changes only by a small amount.

4.3.1 Individual Halos

Halos are dynamically evolving objects; a halo profile may not be well-described by the NFW profile if it is far from a dynamically relaxed state – for example, if it is the product of a recent merger. We identify these halos by a simple test [78, 221]: halos are labeled relaxed if the distance between the halo center and the center of mass of all particles in the SO halo is, at most, $0.07R_{200c}$. If the offset exceeds $0.07R_{200c}$, the halo is assumed to be unrelaxed. (In rare circumstances, characterized by accidental symmetry, unrelaxed halos can pass the relaxed halo test, but the reverse is not true.) At $z = 0$, $\sim 80\%$ of all halos of at least 2000 particles are relaxed, with a higher fraction of high-mass halos unrelaxed. As redshift increases, the relaxed fraction drops to 45%, and by $z = 2$ is independent of mass across the $10^{11}\ h^{-1}M_\odot < M_{200c} < 10^{15}\ h^{-1}M_\odot$ range (see Figure 4.3).

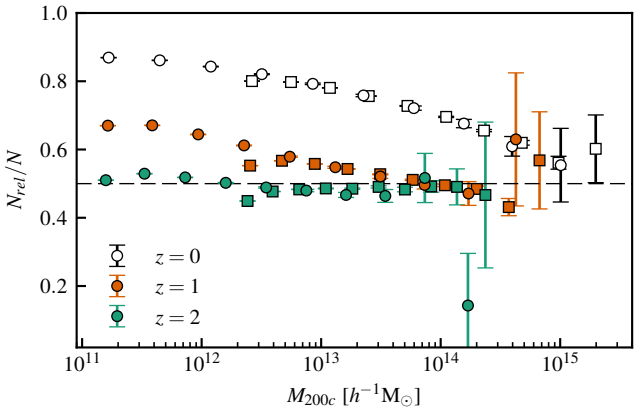

Figure 4.3: Relaxed fraction for Q Continuum (circles) and Outer Rim (squares) halos. A halo is considered relaxed if the distance between its most bound particle and SO center of mass is, at most, $0.07R_{200}$.

Figure 4.4 shows examples of relaxed and unrelaxed halo profiles at high and low redshifts. Note that much of the mass of the unrelaxed halo is far from the center, so the identified scale radius is large and concentration is small. In general, our relaxation criterion implies that unrelaxed halos have significant mass far from the potential minimum, so we expect and find unrelaxed halos to be less concentrated than relaxed halos of the same mass (Figure 4.5). The unrelaxed halo c–M relation has lower amplitude and slope than the c–M relation

of all halos, but as for relaxed halos, concentration decreases with mass at low redshift and is constant at high z. Notably, this holds true at high masses and redshifts – the relaxed halo c–M relation does not fall below the c–M relation for unrelaxed halos.

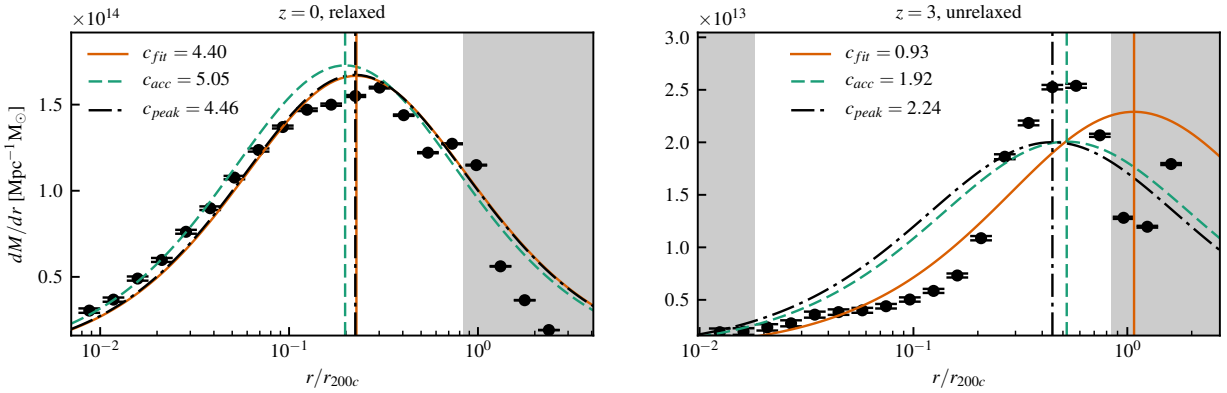

Figure 4.4: Example profiles of individual high-mass Q Continuum halos well- and poorly fit to an NFW profile: a relaxed halo at $z = 0$ (left, $M_{200c} = 10^{14}\,h^{-1}M_\odot$) and an unrelaxed halo at $z = 3$ (right, $M_{200c} = 9 \times 10^{12}\,h^{-1}M_\odot$). Note the small Poisson error. At $z = 0$ the outer radius of the innermost bin encloses at least 100 particles, so all points within r_{200c} are included in the profile fit. At $z = 3$, the first two bins do not meet this criterion and are dropped. More points must be dropped for lower-mass and less-concentrated halos. Shaded regions are not included in the fit, nor expected to follow an NFW profile.

Individual halos are fit to the NFW differential mass profile,

$$\frac{dM}{dr} = 4\pi r^2 \rho(r) = \frac{M_\Delta}{A(c_\Delta)\, r_\Delta} \frac{r/r_\Delta}{(1/c_\Delta + r/r_\Delta)}, \tag{4.4}$$

which rises as r at small r, peaks at $r = r_s$, and falls off as $1/r$ using the following three methods:

Profile Fit, $c_{\rm fit}$. The halo mass M_Δ and radius r_Δ are fixed by M_{200c} as found by the SO algorithm. The profile fit uses the Levenberg–Marquardt algorithm, which, weighted by the Poisson error in the number of particles in each shell, minimizes

$$\sum_i \frac{\left[dn_i/dr_i - (dn/dr)_i^{\rm NFW}\right]^2}{dn_i/dr_i^2}, \tag{4.5}$$

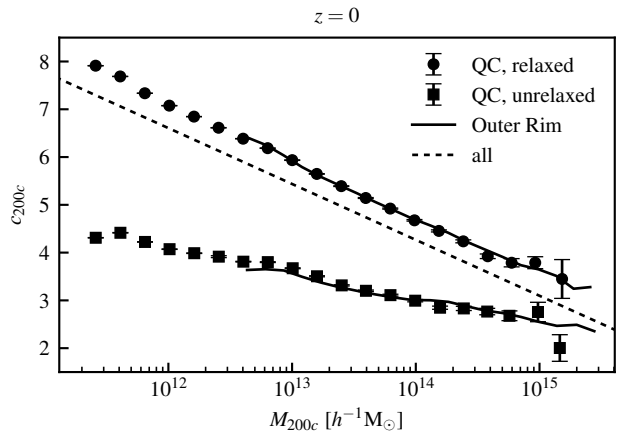

Figure 4.5: Mean fit concentrations at $z = 0$ for all relaxed and unrelaxed halos from the Outer Rim and Q Continuum (QC) runs using the profile fit method. At this redshift, 80% of halos of at least 2000 particles are relaxed. The dotted line shows the power-law fit to all halos; the relaxed halos have slightly higher concentration (upper data set), while the c–M relation for unrelaxed halos is lower and flatter (lower data set). Note the excellent agreement between the two simulation results.

where dr_i is the radial width of a shell that contains dn_i particles and $(dn/dr)_i^{\rm NFW}$ is evaluated at the midpoint of the bin. Only shells whose outer radius falls within r_{200} and encloses at least 100 particles are fit. Shells beyond r_{200}, with their high particle counts and low Poisson error, would have disproportionate influence on the fit, and the NFW form does not necessarily hold at the farther edges of a halo. The requirement of at least 100 particles in a shell also excludes the inner regions (roughly a tenth of the virial radius at cluster mass scales) that may suffer from numerical errors and not be modeled well by gravity-only simulations (due to missing baryonic/feedback effects).

Accumulated Mass, $c_{\rm acc}$. This method uses the fact that the mass enclosed by the NFW scale radius is

$$M(r_s) = \frac{M_\Delta}{A(c_\Delta)} \left(\ln 2 - \frac{1}{2} \right) . \tag{4.6}$$

The concentration is found iteratively by fixing c_Δ, interpolating the enclosed mass profile to solve Equation (4.6) for r_s, and updating $c_\Delta = r_\Delta / r_s$.

Peak Finding, $c_{\rm peak}$. The differential mass profile, Equation (4.4), peaks at $r = r_s$, so the

scale radius can be measured by simply locating the peak. To do this, profiles are smoothed using a three-point Hanning filter

$$f(r_i) = \frac{1}{4}\left[f(r_{i-1}) + 2f(r_i) + f(r_{i+1})\right], \tag{4.7}$$

and the scale radius is set to the location of maximum smoothed dM/dr, excluding the first and last radial bins. Note that this method makes no assumption about the specific form (NFW, Einasto, etc.) of the halo profile.

Our goal in using three different methods is to verify the robustness of our concentration measurements. These three methods are sensitive to different features of the profile. Both the profile fit and accumulated mass methods assume the NFW form, while peak finding does not. The accumulated mass method counts particles only to the scale radius, so it is most sensitive to the inner profile, $r < r_s$. The profile fit is more influenced by the outer profile, as the outer shells, with their higher particle counts, have lower Poisson error and are weighted more heavily in the fit. If halos were always well-described by the NFW form, all three methods would find the same concentration; when their results differ, the profile is not a perfect NFW profile. We use $c_{\rm fit}$ as the primary concentration measurement and the other two methods to check our results and to better understand changes in halo profiles with mass and redshift.

The three methods agree best on well-resolved halos at low redshift, as shown in Figure 4.6. At high redshift, the methods differ by as much as 10%: the fit and peak concentrations are essentially flat as a function of mass, while the accumulated mass c–M relation slopes slightly upward, still well within the range of the intrinsic concentration scatter.

At all mass ranges and redshifts with sufficient halos in a bin, the distribution of accumulated mass concentrations is approximately normal, as shown in Figure 4.7. The accumulated mass scale radius can only fall between the innermost bin in the profile and r_{200}, limiting the range of $c_{\rm acc}$. However, fit concentrations can be arbitrarily high, so the distribution of $c_{\rm fit}$ is

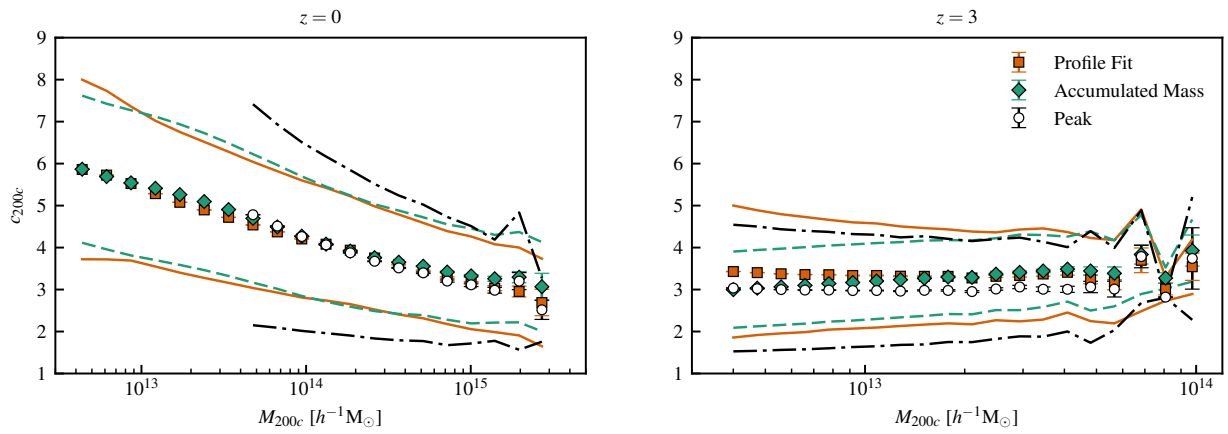

Figure 4.6: Profile fit, accumulated mass, and peak finding methods of concentration measurement for all Outer Rim halos at $z = 0$ (left panel) and $z = 3$ (right). The outer curves represent the 1σ intrinsic variation in the concentration at a fixed mass. Note the very small statistical error bars – the lowest and highest mass bins shown contain 10^6 and 11 halos, respectively, at $z = 0$, and 10^5 and 4 halos at $z = 3$. Here, $c_{\rm peak}$ is shown only for high-mass, low-concentration halos; see further discussion in Appendix 4.6.2. For high-mass halos at $z = 0$, there is little difference in mean concentration between methods; at $z = 3$, $c_{\rm peak}$ differs from $c_{\rm fit}$ by about 10%.

positively skewed, particularly at masses and redshifts where profiles are less well-described by the NFW form. See also Appendix 4.6.1.

4.3.2 Stacked Halos

Individual halo profiles can be noisy, but with 10 million halos of at least 2000 particles at $z = 0$ in Q Continuum and 20 million in Outer Rim, we can stack thousands of halos in a narrow mass bin (for example, $M \pm 1\%$, hereafter "a 1% stack") to obtain smooth profiles. Relaxed halos are stacked by interpolating and summing the individual enclosed mass profiles. The differential mass profile is calculated from the mean enclosed mass profile. Note that we fit the 3D halo profile; future work will compare the concentrations that would be found for the same stacks using observational methods like weak lensing, which must fit the 2D projected mass profile.

Assuming the particles in our simulated profiles are drawn from a true NFW distribution,

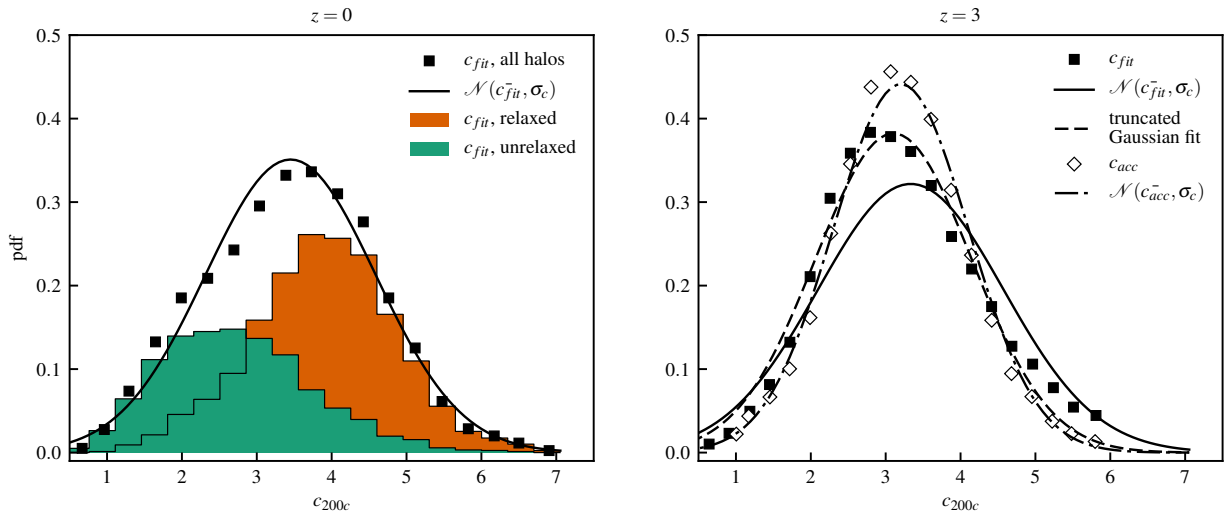

Figure 4.7: Distributions of Outer Rim fit and accumulated mass concentrations. *Left panel:* fit concentrations are normally distributed in a high mass, low-redshift bin with $M_{200c} = 5.07 \times 10^{14} h^{-1} M_\odot$ at $z = 0$; the excess at low concentrations is due to unrelaxed halos (in this high-mass bin, 40% of halos are unrelaxed). *Right panel:* high-redshift ($z = 3$) bin centered at $M_{200c} = 1.08 \times 10^{13}\, h^{-1} M_\odot$, all halos (relaxed and unrelaxed). Solid and dot–dash lines show normal distributions with the sample mean and standard deviation, while dashed line is fit to the truncated $c_{\rm fit}$ distribution shown. In this bin, the distribution of fit concentrations is positively skewed, while the distribution of accumulated mass concentrations is not.

the stacked profile is a sum of NFW profiles with different concentrations, which, in principle, is not describable as an NFW profile. To visualize how the stacked profiles compare to an NFW profile, we calculate the effective power-law index of $\rho(r)$, that is, the slope of $\ln\rho\,(\ln r)$. For an NFW profile,

$$n_{\rm eff}^{\rm NFW} = \frac{d\ln\rho^{\rm NFW}}{d\ln\,(r/r_s)} = -\frac{1+3r/r_s}{1+r/r_s}; \tag{4.8}$$

the density is proportional to r^{-1} at small r and r^{-3} at large r, crossing $n_{\rm eff}^{\rm NFW}(r_s) = -2$ at the scale radius. We calculate $n_{\rm eff}(r)$ for stacked halos using a low-noise Lanczos differentiator with $N = 5$,

$$f'(x_i) \approx \frac{1}{10h}(-2y_{-2} - y_{-1} + y_1 + 2y_2); \tag{4.9}$$

for the first two and last two bins,

$$f'(x_0) \approx \frac{1}{6h}(-11y_0 + 18y_1 - 9y_2 + 2y_3), \tag{4.10}$$

$$f'(x_1) \approx \frac{1}{6h}(-2y_0 - 3y_1 + 6y_2 - y_3). \tag{4.11}$$

Figure 4.8 shows examples of stacked profiles and their $n_{\rm eff}$ at $z = 0$ and $z = 3$. At high redshifts, the outer and inner profiles diverge from an NFW profile: by $r \sim r_{200c}$, density falls off more steeply than the NFW profile, while at small r, the density profile becomes shallower than the NFW profile (due to limitations imposed by force and mass resolution; however, results from our two simulations are still in good agreement, indicating little effect on the concentration itself). When we nevertheless calculate the NFW concentrations of the stacks, the three methods can find different concentrations, despite their agreement on the average concentration of individual halos. Additionally, the stacked fit concentrations differ from mean individual fit concentrations, as shown in Figure 4.9, but by no more than 5%,. The stacked concentrations are lower than individual means at low mass, but higher at high mass. These discrepancies reflect the fact that these stacked profiles deviate from the NFW form.

Previous studies have shown that profiles with a third parameter can better fit stacked profiles [97, 206, 219, 247]. After investigating several three-parameter profiles, we find the Einasto profile,

$$\ln\left(\frac{\rho_E(r)}{\rho_{-2}}\right) = -\frac{2}{\alpha}\left[\left(\frac{r}{r_{-2}}\right)^{\alpha} - 1\right], \tag{4.12}$$

to be the best. We fit stacked halos to the Einasto form, allowing the concentration and the shape parameter α to vary freely. The mass M_Δ is the mean mass of halos in the stack, as determined by the SO algorithm. We find that the Einasto shape parameter increases with mass and redshift, as shown in Figure 4.10. These results are in reasonable agreement with the model of Ref. [97] fit to results at $0 \leq z \leq 3$. The Einasto concentrations agree between the Outer Rim and Q Continuum runs, but the shape parameters differ at high redshift.

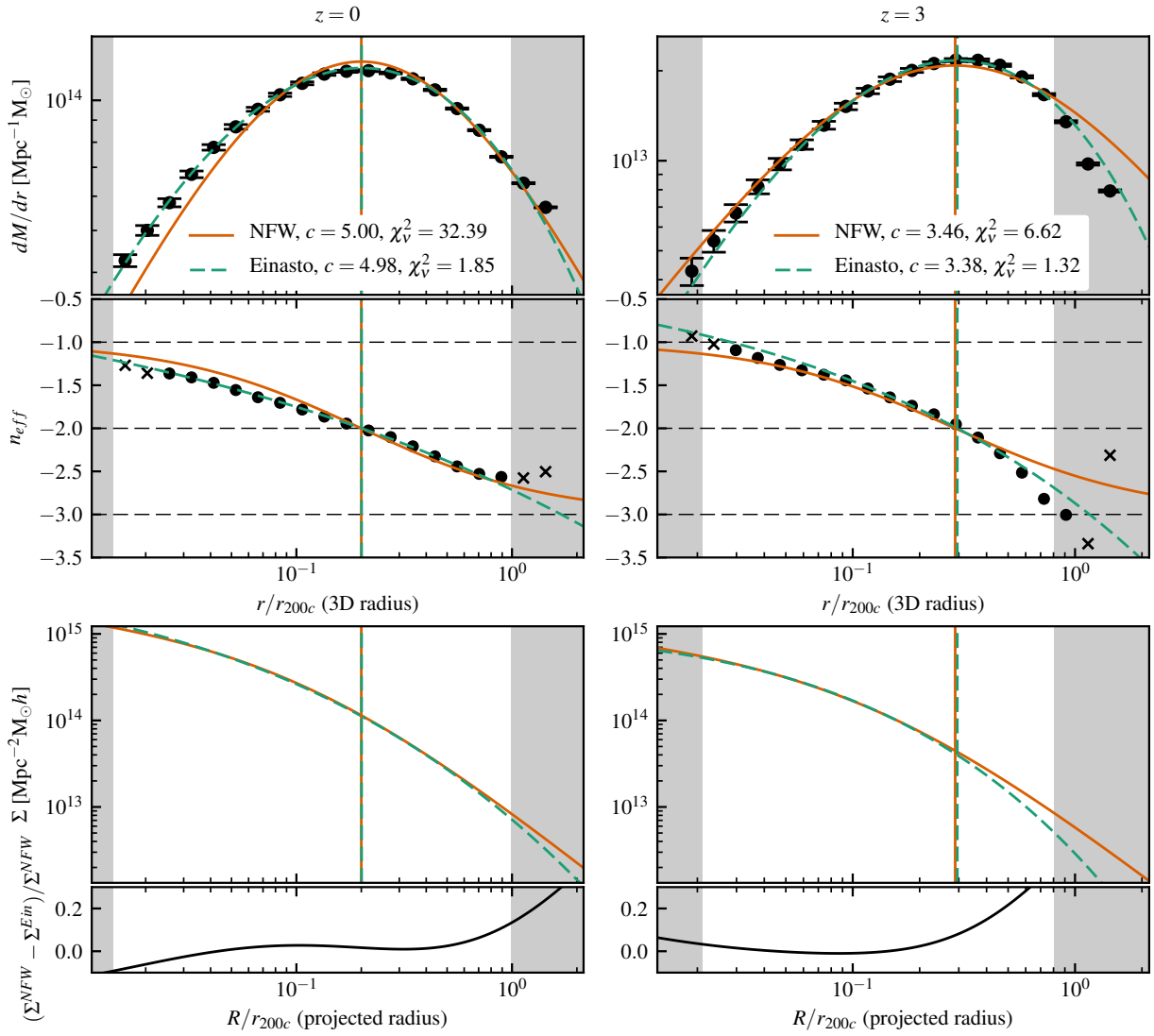

Figure 4.8: Stacks of relaxed Q Continuum halos at $z = 0$ (left panels, 116 halos, $M_{200c} = 6 \times 10^{13}\, h^{-1} M_\odot \pm 0.2\%$) and $z = 3$ (right, 81 halos, $M_{200c} = 9 \times 10^{12}\, h^{-1} M_\odot \pm 1\%$). Shaded regions are not included in the fit ($r > r_{200}$ or fewer than 100 particles enclosed); reduced χ^2_ν values are calculated only on the points used to fit. *Top row*: NFW and Einasto profiles are fit to the dM/dr profile as described in Section 4.3; vertical lines show the corresponding scale radii. At $z = 0$, the NFW fit concentration is $c_{\rm NFW} = 5.00$; Einasto fit concentration is $c_{\rm Ein} = 4.98$. At $z = 3$, $c_{\rm NFW} = 3.46$ and $c_{\rm Ein} = 3.38$. The Einasto fit captures the peak better than the NFW fit does at both redshifts; at $z = 3$, it also improves on the high-r behavior of the NFW profile. *Second row:* effective power-law index of the density profile. Slopes for the first and last two radial bins ($\times$ symbols) are less reliable than those with four neighboring points to include in the calculation. At high redshifts, our stacked profile is steeper than an NFW profile at high r. *Bottom rows:* surface mass density corresponding to the fit NFW and Einasto profiles and their relative difference; note that differences in the projected profiles are small, especially in the high-mass $z = 0$ case.

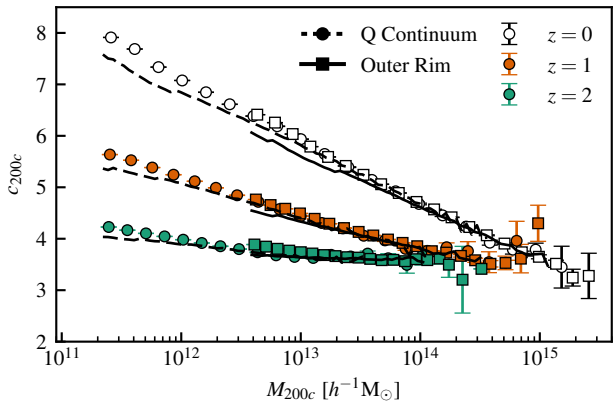

Figure 4.9: Concentrations found using the profile fit method, $c_{\rm fit}$, for relaxed individual and stacked halos from both simulations (circles are for Q Continuum, squares for Outer Rim). Points are means of individual fits; lines show concentrations found by fitting 5% stacks at the same redshifts to an NFW profile.

See Appendix 4.6.2 for further discussion of this discrepancy, and Appendix 4.6.1 for further comparison to other works.

Our results show that, at low mass, the stacked peak is to the left of the peak of the NFW profile fit to the stack, while at high masses at high redshift the opposite is true. At $z = 0$, Einasto concentrations are greater than NFW concentrations at low mass; the Einasto and NFW c–M relations cross around $M_{200c} = 10^{14}\,h^{-1}M_\odot$ (see Figure 4.8 for an example stack whose peak location is well captured by the NFW fit), and at the highest masses, the Einasto concentration begins to fall below the NFW concentration. Luckily, at cluster scales at $z = 0$, the NFW fit peak is transitioning from the left of the stacked peak to the right, and appears to be temporarily in agreement – so NFW and Einasto profiles both fit the stacks very well, as found for the CLASH dataset by Ref. [314]. These authors stack profiles in a wide mass bin; we do not find substantial differences in the quality of NFW vs. Einasto fits for wide mass bins, of up to 70%, compared to narrower 1% or 5% bins.

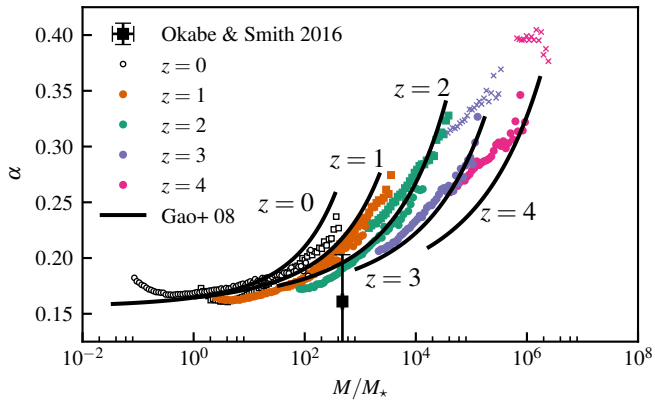

Figure 4.10: Relaxed halo Einasto $\alpha - M/M_\star$ relation for Q Continuum (circles) and Outer Rim (squares). Einasto profiles are fit to halos stacked in 5% mass bins, fixing mass to the SOD mass. High mass resolution is critical to measure the shape parameter α; at high z, Outer Rim halos are insufficiently resolved ($\times$ symbols; see Appendix 4.6.2 for discussion of high-redshift discrepancies between the two simulations). Black curves show the fit of Ref. [97] for $z = 0, 1, 2, 3, 4$. The single observational point is from a LoCuSS weak lensing measurement at $z = 0.23$ [230].

4.4 The Concentration–Mass Relation

The concentration measurements described above are now used to investigate the c–M relation. As shown in Figures 4.6 and 4.9, the c–M relation flattens with increasing redshift. At higher redshifts, halos of all masses have concentration $c \sim 3 - 4$ and the c–M relation falls no further; Ref. [345] found a similar floor. We do not present results for redshifts higher than $z = 4$ (see discussion in the appendix), but concentrations measured up to redshifts as high as $z = 10$ are consistent with this floor.

In order to represent the data in a z-independent form, we scale the mass by the redshift-dependent nonlinear (or "collapse") mass scale $M_\star = 4\pi/3\, \rho_c(z)\, \omega_m(z)\, R_\star^3$, which solves

$$\sigma(R_\star,\, z) = \delta_c, \tag{4.13}$$

where $\delta_c = 1.686$ is the linear critical density for collapse and $\sigma(M, z)$, the amplitude of mass fluctuations, is the power spectrum smoothed with a top-hat filter. The power spectrum

$P(k,\ z) = d^2(a)\ P(k,\ z = 0)$ is calculated from the growth factor

$$d(a) = \frac{D^+(a)}{D^+(a = 1)}, \tag{4.14}$$

$$D^+(a) = \frac{5\Omega_m}{2}\frac{H(a)}{H_0}\int_0^a \frac{da'}{[a'H(a')/H_0]^3} \tag{4.15}$$

and $P(k, z = 0)$ is as given by CAMB [175]. The mean square perturbation is

$$\sigma^2\,(R, z) = \frac{1}{2\pi^2}\int k^2\,dk\,W^2\,(k, R)\ \ P\,(k, z)\,, \tag{4.16}$$

where we choose the window function $W^2(k, R)$ to be the Fourier transform of a spherical top-hat filter of radius R,

$$W(k, R) = \frac{3}{(kR)^3}[\sin(kR) - kR\cos(kR)]. \tag{4.17}$$

The nonlinear mass $M_\star$ depends weakly on cosmology and falls steeply with redshift: for our cosmology, $\log(M_\star/h^{-1}M_\odot) = 12.5$ at $z = 0$, 11 at $z = 1$, and 9.5 at $z = 2$. Already at $z = 3$, $\log(M_\star/h^{-1}M_\odot) = 8$, and our least-massive halos exceed the nonlinear mass scale by three orders of magnitude at this redshift. Combining the results of the two simulations at redshifts up to $z = 4$, we probe eight decades in $M/M_\star$, as shown in Figure 4.11.

When represented in terms of $M/M_\star$, the individual c–M relations at different redshifts fall relatively tightly onto a single relation: the concentration behaves as a power law with $M/M_\star$ as long as $M/M_\star \lesssim 10^3$; at masses farther above the nonlinear mass scale, the concentration asymptotes to a constant value ~ 3. This behavior is shown in Figure 4.11 for both individual and stacked halos, and is consistent with analytic models which find concentration to be independent of mass above a threshold mass. Ref. [60] find concentration to approach a constant $c \sim 4$ for high-mass halos, those for which $\sigma(M) \ll \delta_c$ (implying $M \gg M_\star$), where Gaussian statistics are a good description of the slopes of the corresponding

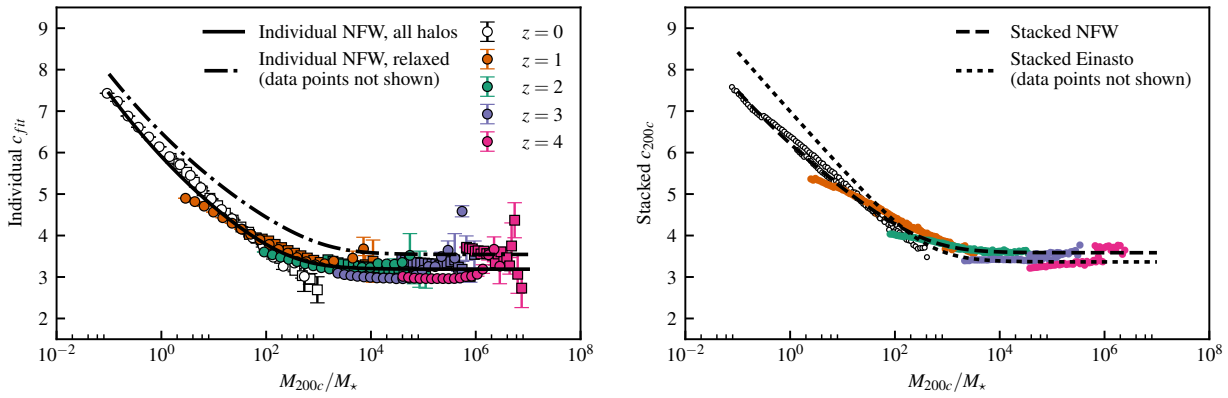

Figure 4.11: Individual mean $c_{\rm fit}$ for all halos (left) and stacked $c_{\rm fit}^{\rm NFW}$ for relaxed halos in 5% stacks (right) as a function of M/M_* across eight orders of magnitude for Q Continuum (circles) and Outer Rim (squares). Red solid (left) and dashed (right) curves are fit to all points shown, as well as $z = 0.502$ and $z = 1.494$. Dotted–dashed red curve is fit to individual mean concentrations of relaxed halos only (left panel, data points suppressed to avoid clutter); dotted is fit to Einasto concentrations of 5% stacks (right panel, data points suppressed).

peaks in the initial Gaussian random field. Ref. [233], similarly, predict a constant c–M relation, $c \sim 2.5$, in the high-mass regime where an assumption of spherical collapse is valid. See further discussion in Appendix 4.6.1.

We fit to a simple functional form that captures this behavior:

$$c_{200c} = A \left[\left(\frac{M_{200c}/M_\star}{b} \right)^m \left(1 + \frac{M_{200c}/M_\star}{b} \right)^{-m} - 1 \right] + c_0, \tag{4.18}$$

transitioning at $M = M_T \equiv bM_\star$ from a power law to a constant $c = c_0$. This plateau is found at a concentration between 3 and 4, depending on the halos included (all halos vs. relaxed halos only, for example) and the type of fit. Table 4.1 gives the fit parameters. We note that the fit should not be naïvely extrapolated to masses smaller than those considered here.

This form is not fully universal in the sense of being approximately independent of cosmology, as can be shown by checking against the wCDM results of Refs. [29] and [166]. We are currently investigating the detailed cosmology dependence of the c–M relation using the

Table 4.1: c–$M/M_\star$ fit parameters, $0 \leq z \leq 4$

Type of Fit	m	A	M_T/M_*	c_0
Individual, all	-0.10	3.44	430.49	3.19
Individual, relaxed	-0.09	2.88	1644.53	3.54
Stack, NFW	-0.07	4.61	638.65	3.59
Stack, Einasto	-0.01	63.2	431.48	3.36

Notes. Use Equation 4.18 with variance $\sigma_c = c_{200c}/3$.

Mira–Titan Universe suite of simulations [128, 171], which cover ∼100 cosmological models allowing for a (w_0, w_a) parameterization of dark energy as well as the effect of massive neutrinos. These results will be presented elsewhere.

The peak-height parameter ν is frequently used to find a redshift-independent c–ν relation, with similar motivations (see, e.g. Refs. [29, 73, 80, 188]). We find our results to be more universal with redshift as a function of $M/M_\star$ than as a function of ν (see Appendix 4.6.1 for further details). Moreover, we find z-independence to hold for other overdensities, although it does deteriorate for smaller choices of Δ (including $\Delta_{\rm vir}$).

Aside from the concentration, it is also important to investigate the fraction of relaxed halos as a function of $M/M_\star$. Because halos with $M \gg M_\star$ are likely to be in the "halo formation" phase, the unrelaxed fraction should increase with mass. However, one would anticipate that the existence of a finite asymptotic value of the concentration, c_0, suggests that the relaxed fraction also reaches a limiting value. We find that, like the concentration, the relaxed fraction is also independent of mass for $M_{200c} > 10^{11} M_\odot$ at redshifts higher than $z \sim 2$, or alternatively, for $M/M_\star \gtrsim 10^3$, as shown in Figure 4.12. At low redshift, up to 80% of low-mass halos are relaxed, falling to 50% at high masses, which is still a substantial fraction; at high redshifts, 50% of halos (that pass our minimum mass threshold) are relaxed at all masses, even for $M \gg M_\star$.

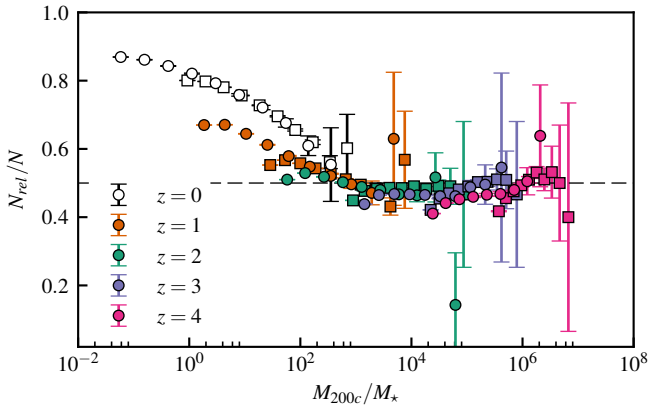

Figure 4.12: The fraction of Q Continuum (circles) and Outer Rim (squares) halos considered relaxed (distance between most bound particle and FOF center of mass at most $0.07R_{200}$) falls with mass until $M \sim M_T$, but is approximately constant at 0.5 for $M > M_T$.

4.4.1 Comparison with Observations

We compare our c–M relation at $z = 0$ to individual and stacked halo concentrations measured from individual X-ray [322], individual weak and strong lensing [207, 222], and stacked lensing [36, 48, 89, 154, 194, 195, 224, 229–232, 279, 313, 315, 316] observations. Concentrations and masses are reported in a variety of mass definitions (M_{500c}, M_{200c}, $M_{\rm vir}$); when M_{200c} is not provided, we assume an NFW profile and convert to M_{200c}, c_{200c} according to Ref. [142].

All observations considered here were carried out at redshifts between $z = 0$ and $z = 1$; in the case of stacked analyses, we assign a mean redshift to the sample. We fit the c–M relation at $z = 0, 1$ to

$$c_{200c} = A\,(1+z)^d\,M^m; \tag{4.19}$$

the fit parameters are shown in Table 4.2. Note that this simple fit is only valid for redshifts $0 \leq z \leq 1$, where a power-law dependence on mass and redshift is a reasonable description of our results. Individual (Figure 4.1) and stacked (Figure 4.2) measurements are scaled to $z = 0$ according to the fit redshift dependence and compared to the corresponding mean $c_{\rm fit}$ of individual relaxed halos or stacked $c_{\rm fit}$. Note that our Einasto and NFW fit concentrations

Table 4.2: Power-law c–M fit parameters, $0 \leq z \leq 1$

Type of fit	A	d	m
individual, all	75.4	-0.422	-0.089
individual, relaxed	68.4	-0.347	-0.083
stack, NFW	57.6	-0.376	-0.078
stack, Einasto	122	-0.446	-0.101

Notes. Use Equation 4.19 with variance $\sigma_c = c_{200c}/3$.

are identical around $M_{200c} \sim 10^{14} h^{-1} M_\odot$, and are only beginning to diverge at cluster scales, as mentioned in Section 4.3.2 above. This is the mass range where, as shown in Figure 4.8, Einasto and NFW profiles both provide very good fits and are difficult to distinguish. The CLASH measurement of Ref. [314], where both Einasto and NFW forms are able to fit the projected mass profile, also falls in this range.

The current agreements with observations are very good for both individual and stacked halo measurements, although there are a few outliers. Next-generation CMB, optical, and X-ray surveys will greatly increase the number of well-measured group and cluster-scale halos. This will lead to much better control on mass measurement for stacked observations (currently at 4%, as claimed by LoCuSS), as well as on the profile measurements [286].

4.5 Discussion

The high-volume, high-resolution simulations Q Continuum and Outer Rim simultaneously provide superior statistics (20 million and 10 million halos of at least 2000 particles at $z = 0$, respectively) and halo resolution, allowing us to use multiple methods of concentration measurement to study the c–M relation across eight decades in $M/M_\star$, the distribution of concentration within mass bins, and smooth stacked profiles of at least hundreds of halos in narrow (5% and smaller) mass bins.

Our three methods of concentration measurement (fit, accumulated mass, and peak finding) agree best on massive halos at low redshift; at higher z, mean measurements differ by 10–20%. This is consistent with the general expectation that concentration measure-

ments can have systematic differences depending on how the measurements are carried out (Refs. [29, 80, 205]; see Appendix 4.6.1 for further discussion). These caveats aside, our results are in excellent agreement with observations as well as with most recent simulations.

The c–M relation is not a precise, narrow correlation between the halo concentration and halo mass. There is a substantial amount of intrinsic variability, which appears, remarkably, to be cosmology-independent; it is specified by $\sigma/c \sim 1/3$, where σ is the standard deviation around the mean concentration at a given halo mass [29, 75]. Our results are consistent with a Gaussian distribution of concentration within mass bins [29, 191, 255]. Small deviations at lower concentrations can be explained by an unrelaxed halo population, and at higher concentrations due to the existence of a small high-concentration tail.

Because of the excellent statistics made possible by our simulations, we can study the stacked profiles of halos in narrow mass bins. We find that the Einasto profile is an excellent fit to the data, and that concentrations measured from our stacked profiles agree with the means of the concentrations of those halos measured individually.

Our well-characterized results for the c–M relation across a wide range of redshifts motivate searching for a simple description of the data in a redshift-independent form, much as in the case of the halo mass function. Although there is no basic theory for the c–M relation and its evolution with redshift, there exist a number of models and fits to numerical data. These include some simple analytical ideas, power-law models, and significantly more complex fits (see Appendix 4.6.1). We find that scaling the halo mass by the nonlinear mass scale $M_\star$ allows us to describe results from different redshifts by a single relation with a very simple form. This c–M relation has a power-law form at lower masses (over the mass range investigated) and transitions to a constant value at masses above a threshold mass $M_T \sim 500 - 1000 M_\star$. At redshifts higher than $z = 3$, all halos above the threshold of 2000 simulation particles exceed M_T and have concentration $c \sim 3$. At low redshifts (such as $z = 0$, where $M_T \sim 5 - 10 \times 10^{14.5}\ M_\odot$), only the most massive (cluster-scale) halos approach M_T. These halos, too, have concentrations $c \sim 3$.

Our results are in excellent agreement with current observational datasets. Future measurements will have significantly enhanced statistics for both individual and stacked halo profiles. This will bring the observational errors closer to the current error estimates from the simulations and allow for any differences to become more apparent. Increasing our understanding of the halo profiles will aid in controlling systematic errors in a number of cosmological analyses.

4.6 Appendix

In this appendix, we compare our results with those from other simulations commenting on agreements and differences. We also investigate various simulation errors that could affect halo concentrations, demonstrating that our results are robust to these potential error sources. We note that the Outer Rim and Q Continuum simulations have mass resolutions that are different by a factor of 20 and were run using different short-range force algorithms (P^3M versus tree). Achieving very consistent results, given these differences, is itself a good test of the overall methodology.

4.6.1 Other Simulations

Our results are in reasonably good agreement with those of recent simulations; differences often arise due to choices made in the operational definition of the concentration.

Profile Shape

To illustrate the closeness of obtained profiles across simulations, we compare our results with Refs. [162] and [246] who calculate the concentrations of halos from the MultiDark and Bolshoi project suite of simulations. These halo profiles are publicly available in the CosmoSim database[3]. The cosmologies of the Bolshoi and MultiDark simulations differ from

3. https://www.cosmosim.org/cms/simulations

ours, but the mass resolutions of Bolshoi ($1.35 \times 10^8\ h^{-1}M_\odot$) and Q Continuum ($1.05 \times 10^8\ h^{-1}M_\odot$) are similar. A stacked profile of high-redshift Bolshoi halos overlaps with a Q Continuum profile at the same mass and redshift (Figure 4.13); the shape of the halo profiles is essentially the same.

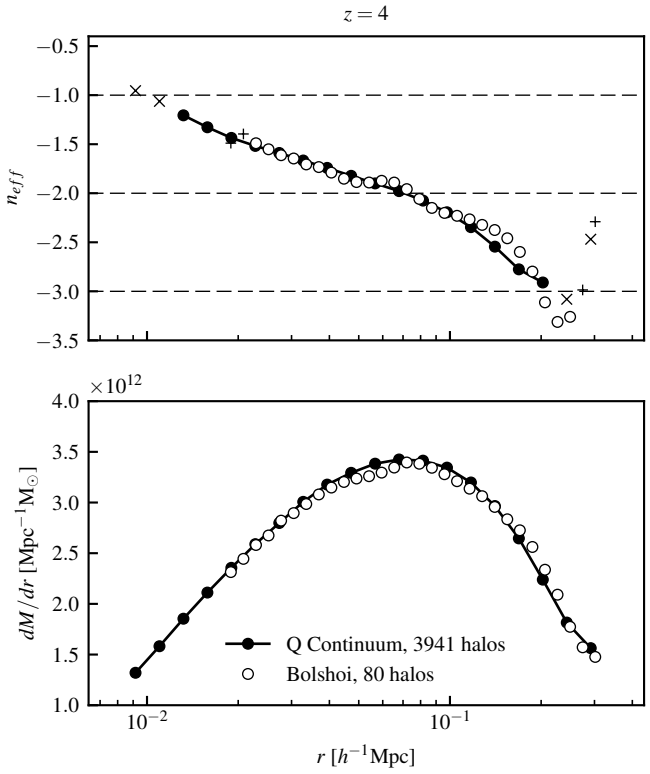

Figure 4.13: Stacked Q Continuum and Bolshoi profiles in a narrow mass bin at $z = 4$, $M_{200c} = 6 \times 10^{11}\ h^{-1}M_\odot \pm 1\%$, relaxed halos only. Slopes for the first and last two radial bins (shown as + symbols for Bolshoi, × for Q Continuum) are less trustworthy than those calculated from the full five points.

The c–M Relation in Other Simulations

Many different methods have been used to measure concentrations from a large set of simulations. The results fall in one of three categories: concentration is power-law at all masses and redshifts; concentration flattens at high masses and redshifts; or concentration increases at high mass and redshift. Our results are most consistent with the second category; we do not find evidence that concentration increases with mass at high masses and redshifts, but

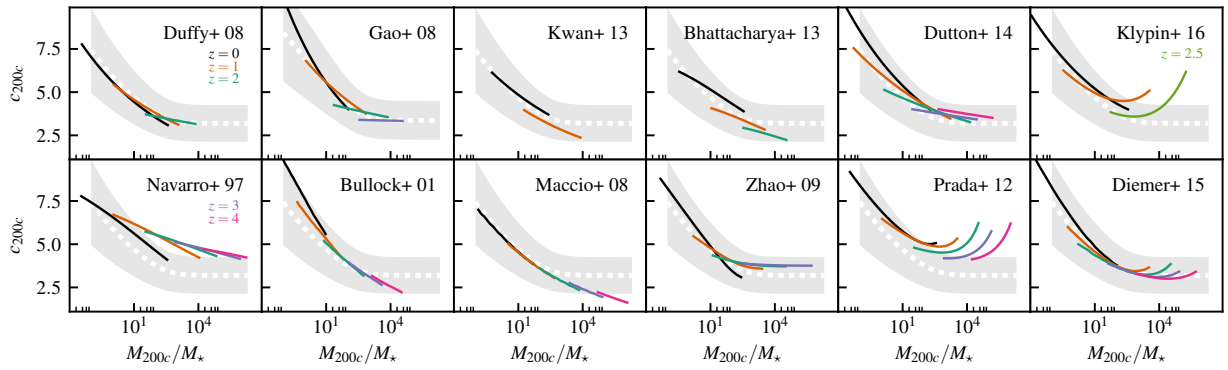

Figure 4.14: Other fits and models generally fall within one standard deviation (shaded region, $\sigma_c = c/3$) from our fit to concentrations at $0 \leq z \leq 4$ (white, dashed); all but Gao et al. (2008) [97] are results for all (relaxed and unrelaxed) individual halos. Note that each panel shows results from multiple redshifts, to illustrate redshift-independent behavior when mass is scaled by $M_\star$. *Top row:* fits. Duffy et al. (2008) [78] (WMAP-5 cosmology), Gao et al. (2008) [97] ($\sigma_8 = 0.9$, Einasto concentrations of relaxed halos only), and Dutton & Macciò (2014) [80] (Planck cosmology) results are power-law fits to cosmologies that differ from ours; $M_\star$ is calculated accordingly, but note that $M_\star$ scaling does not provide full cosmology independence. The other three fits are shown for our WMAP-7 cosmology. *Bottom row:* models. The predictions of Navarro, Frenk, & White (1997) [218] (with the free parameter $f = 0.01$), Zhao et al. (2009) [346], Prada et al. (2012) [246], and Diemer & Kravtsov (2015) [73] are calculated for our cosmology. The cosmologies of Bullock et al. (2001) [38] ($\sigma_8 = 1$; $F = 0.01$, $K = 4.0$) and Macciò et al. (2008) [192] (WMAP-5; $F = 0.01$; $K = 3.6$) differ from ours; $M_\star$ is calculated accordingly. See text for further description of the models.

instead find that it approaches a constant $c \sim 3$. Selected fits and analytical models of all three types are compared to our results in Figure 4.14. Masses are scaled by $M_\star$ and multiple redshifts are shown to check for redshift-independent behavior in others' results; note, however, that the c–$M/M_\star$ relation is not fully cosmology-independent – concentrations are higher in high-σ_8 cosmologies.

The first papers to note a deviation from power-law behavior at high mass and redshift described a flattening of the c–M relation. Ref. [345] describes a concentration floor at $c \sim 3$ and a mass-independent c–M relation at high redshift; their $z = 4$ plot may show a slight upward slope. Ref [221] provides results only for $z = 0$, but comments on a tendency toward constant concentration at high mass. Ref. [97] notes a flattening at high mass and redshift. The model of Ref. [346], which predicts concentration from the time when a halo reaches a

fraction of its final mass, predicts a flat regime at high masses and redshifts; again, the results at $z = 4$ may slope slightly upward, but are fit to a model that predicts mass-independent concentration. Two results at high redshift also find very shallow c–M relations: Ref. [132] provides results for relaxed halos only, but find that the c–M relation flattens at redshifts up to $z = 9$. Ref. [5] finds a shallow c–M relation for relaxed halos at $z > 5$; the c–M relation for the full halo sample is shown to be consistent with the upturn of Ref. [73], but also appears to be consistent with a flat relation.

Using a new method to measure concentration, based on maximum circular velocity $V_{\rm circ}$, Ref. [162] notes an upturn in the c–M relation. The upturn is stronger for relaxed halos (in addition to our criterion that the distance between halo centers fall within $0.07R_{200}$, their halos are only relaxed if spin parameter $\lambda < 0.1$). The reasons for the upturn are unknown, but the statistics of the highest density peaks are suggested as a possible explanation. Using a modified version of the V_{max} method and including only bound particles in halo profiles, Ref. [246] also finds an upturn. The upturn is also seen in relaxed halos selected using the spin parameter and center-offset criteria as well as an additional requirement that the virial ratio $2K/|U| < 0.5$. Elaborating on the statistics of high-density peaks, Ref. [246] shows that infall velocity is greatest for the most massive halos. Ref. [161] returns to this point, suggesting that greater infall velocity will produce more concentrated halos, and notes a relationship between true concentration and the Einasto shape parameter α. Ref. [214] uses a profile fitting method rather than V_{max}; power-law fits are used to describe their results, but a flattening of the $z = 2$ c–M relation at high mass is noted to be consistent with Ref. [162]. Ref. [57] fits density profiles, and find an upturn or flattening when all halos are included in the c–M relation, but not when the sample is restricted to relaxed halos. Ref. [73] also uses a density profile fitting method, finds an upturn, and proposes a model that predicts a positive slope for the high-ν c–ν relation.

The methods of the above papers differ in several respects. Other papers discuss the effects of these differences in methodology on the c–M relation. Ref. [190] finds an upturn

using the V_{max} method, but the upturn is not seen when the halo sample is restricted to relaxed halos selected using two criteria (virial ratio and fraction of mass in substructure) in addition to our restriction on the distance between halo centers (the c–M relation for our unrelaxed halo sample does not differ from that of relaxed halos in shape, only in amplitude). The additional criteria are relevant, as halos are seen to pass through phases of high concentration immediately following major mergers: when the infalling progenitor passes near the center of the halo, the halo is highly concentrated and is classified as unrelaxed according to the virial ratio test. Refs. [188], [5], and [132] also see no upturn in the relaxed-halo c–M relation. Ref. [205] are able to bring the results of Ref. [246] into agreement with those of Ref. [78] by binning halos in M_{200} as opposed to V_{max}, and by measuring concentration through profile fitting as opposed to the V_{max} method. Ref. [80] do find that the c–M relation slopes upward at $z = 3$ when the V_{max} method is used, but is flat when concentrations are determined by a profile fit method. The c–M relations are fit to power laws, but do show either a slight upturn or plateau at high masses and redshifts. Ref. [259], however, do find an upturn in concentrations measured by profile fitting as well as using the V_{max} method; Ref. [161] find an upturn both when binning in V_{max} and in virial mass.

Models of the c–M Relation

Several different fits and models, with varying levels of physical motivation, have been used to describe the c–M relation. The simplest description is a power-law fit at each redshift to concentration as a function of mass (top row of Figure 4.14: Refs. [78, 80, 97, 166]). In the case of the upturn, a third parameter is needed to fit $c(M)$, as in Ref. [161]. The peak-height parameter ν has been used as the independent parameter to find a redshift-independent c–ν relation; the shape may simply be fit, as in e.g. Ref. [29], or described by a model (bottom row of Figure 4.14).

One set of models assumes that the density of the central regions of a halo is determined by the background density at the time the halo forms. The model of Ref. [218] predicts

the concentration of a halo of mass M from the collapse redshift $z_{\rm coll}$, the redshift when the total mass of progenitors of mass fM ($f < 1$) or greater first reaches $0.5M$, calculated using the extended Press–Schechter formalism. Concentration is found by assuming that the mean density at $z_{\rm coll}$ determines the density of the final halo within R_s. The model is found to describe simulated c–M relations for small values of $f \sim 0.01$. Ref. [38] revises the model of Ref. [218]; the collapse redshift is found as the time when the collapse mass, $M_\star$, first reaches a fraction of the final halo mass, FM. In addition to the free parameter, F, a second parameter, K, relates a density of the halo to the background density at $z_{\rm coll}$. Ref. [192] modify the Ref. [38] model, allowing the halo overdensity with respect to the mean background density to change with redshift.

Using mass accretion histories, Ref. [346] finds that concentration of a halo of mass M can be calculated from the age of the universe, $t_{0.04}$, when its main progenitor first reached a mass of $0.04M$. This time is calculated from a model of mass accretion history. The fit to concentration as a function of t and $t_{0.04}$ predicts a constant concentration of $c = 4$ when $t/t_{0.04} < 3.75$. Combined with the mass accretion histories, this implies a flat c–M relation at high redshift, where halos rapidly accrete mass; by $z \sim 3$, a halo of mass $10^{12}h^{-1}M_\odot$ is approaching the stage of its evolution where concentration begins to grow. This model is also used by Ref. [320], with modified parameters, to describe $z = 0$ concentrations found using the maximum circular velocity method of Ref. [162]. Refs. [328] and [345] also provide concentrations based on mass accretion history fits. Ref. [57] combines mass accretion histories, extended Press–Schechter theory, and a fit relationship between concentration and formation time, defined as the time the mass of a halo's main progenitor reaches the mass within the scale radius of the descendant halo. This model predicts concentration that falls with mass for relaxed halos at all masses and redshifts.

Two other models use an additional cosmology-dependent variable to fit the c–ν relation. Ref. [246] describes concentration as a function of σ and a time variable x, motivated by its appearance in the growth factor D. Ref. [73] develops a model in which concentration is a

function of both ν and the local slope of the power spectrum, n.

Two papers that predict halo concentrations from the initial Gaussian random field are consistent with our results. Ref. [60] predict halo profiles from the shapes of initial peaks in the Gaussian random field, then fit those profiles to an NFW form to find the concentration. The shape of the c–M relation is determined by the outer slopes of these peaks. For high-mass halos, where slope is shallow and consistent with a prediction from Gaussian statistics, concentration $c \sim 4$ is independent of halo mass. The outer slopes of simulated peaks corresponding to low-mass halos differ from the Gaussian statistics predictions and produce halo profiles with higher concentrations. The transition between these two regimes is redshift-dependent, but falls between $M_\star$ and $1000M_\star$ for $0 \leq z \leq 4$. Ref. [233] calculates the initial energy of a region before collapse, then assume energy is conserved and use the Jeans equation to find concentration from the final energy. In the case of spherical collapse (valid for the largest halos), the resulting concentration $c \sim 2.5$ is mass-independent. For smaller halos, ellipsoidal collapse predicts concentrations that fall with mass.

Gaussian Distribution of Concentrations

As discussed in Section 4.3.1 above, we find that concentrations within a mass bin are normally distributed. Refs. [255] and [29] also discuss the Gaussian distribution of their concentration measurements. Figure 4.15 compares our results in corresponding mass bins. There is very good agreement, in the lowest mass bin, with the results of Ref. [29]; the corresponding simulations have very similar mass resolutions. The two higher mass bins of Ref. [29] use a simulation with significantly worse mass resolution, which may explain the worse agreement with our results. The agreement with Ref. [255] is very interesting, considering that the underlying cosmology differs considerably from the one used here. We note that the mass bin here is by no means narrow; our c–M relation falls by 25% from the lowest mass in this mass bin to the highest.

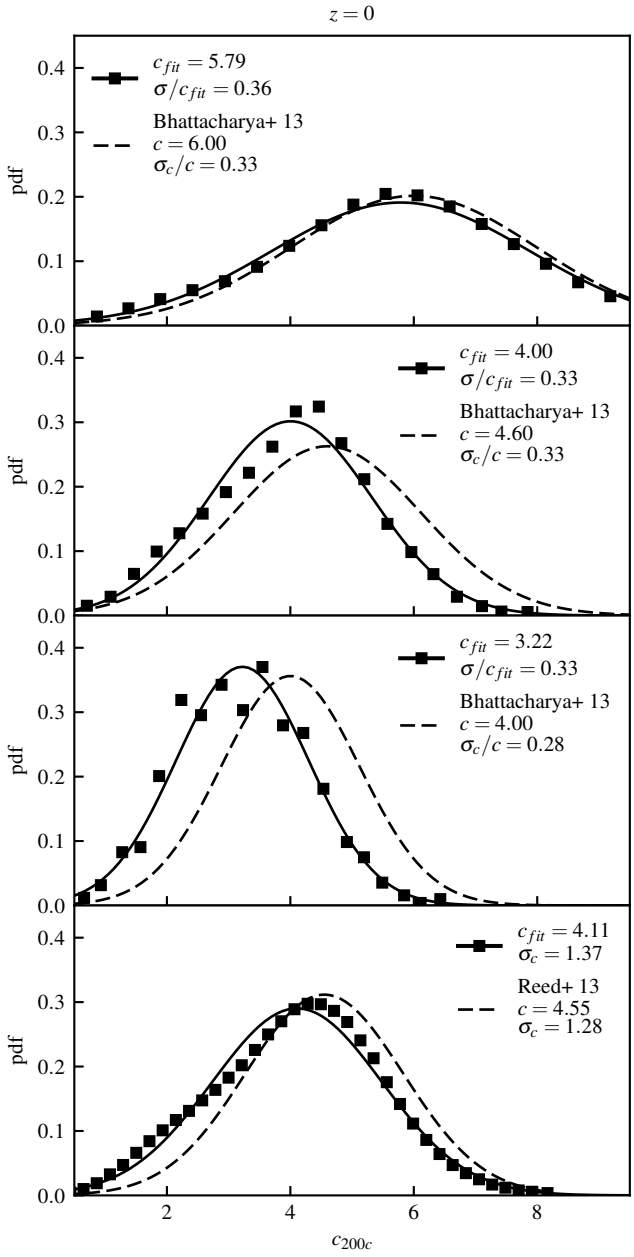

Figure 4.15: *Top three panels:* distributions of Outer Rim fit and accumulated mass concentrations for three mass bins at $z = 0$: $M_{200c} = 5 \times 10^{12}\ h^{-1}M_\odot \pm 1\%$ (top), $M_{200c} = 1.5 \times 10^{14}\, h^{-1}M_\odot \pm 1\%$ (middle), and $M_{200c} = 8 \times 10^{14}\, h^{-1}M_\odot \pm 5\%$ (bottom). All halos, relaxed and unrelaxed, are included. The results in these three panels are analogous to those presented in Figure A13 of Ref. [29] (black dashed curve). *Bottom panel:* distribution of Outer Rim fit concentrations for all 641 779 halos of mass $M_{200c} \geq 6.794 \times 10^{14}\, h^{-1}M_\odot$ at $z = 0$. The comparison is with Ref. [255], 3501 halos from the Millennium Simulation (black dashed curve; our $6.794 \times 10^{14}\ h^{-1}M_\odot$ with $c_{200c} = 4.367$ corresponds to $M_{\rm vir} = M_{95.4c} = 8.600 \times 10^{14}\ h^{-1}M_\odot$)). For further discussion, see text.

Peak Height Parameter ν

Efforts to find a redshift-independent expression for profile parameters often (e.g. Refs. [80, 97, 161, 246]) convert mass to the "peak height" parameter ν. Like $M_\star$, ν is defined from the mean square perturbation σ (Equation 4.16):

$$\nu(M, z) = \delta_c/\sigma(M, z) \tag{4.20}$$

We again use $\delta_c = 1.686$. Note that $\nu(M_\star, z) = 1$, but ν is much less sensitive to mass and redshift than $M/M_\star$. In Figure 4.16, we see disagreement between concentrations at different redshifts at low ν. Figure 4.10 shows that scaling mass by $M_\star$ does not suggest a redshift-independent scaling for the Einasto shape parameter, but we do not find significant improvement when α is expressed as a function of ν. The $\alpha - \nu$ relation is too flat at high z to fall on the same line as the lower-z relations, as seen in Figure 4.16. However, our results are in reasonable agreement with those of Ref. [97] for $0 \leq z \leq 3$, $\sigma_8 = 0.9$ (and slightly less so with the Planck cosmology results of Ref. [161]). A discrepancy with the model of Ref. [97] arises at $z = 4$, but the model was developed using results only at $z \leq 3$.

4.6.2 Sources of Error

Characteristics of the simulation, such as timestep or initial redshift, or of the halo finder, such as the choice of halo center, may affect the shape or amplitude of the c–M relation. We check some of the error scenarios in small simulations, and others using the different properties of the Q Continuum and Outer Rim simulations. We find that the c–M relation as obtained here is robust for $z \leq 4$. However, there is more distortion at higher redshifts, where, as expected, the agreement between our two simulations with different mass resolutions deteriorates. Below, we present our tests and comment on the various results obtained.

Unfortunately, none of these tests are able to reproduce the differently shaped c–M relations found by other works in Appendix 4.6.1 above. No two works use the same procedures,

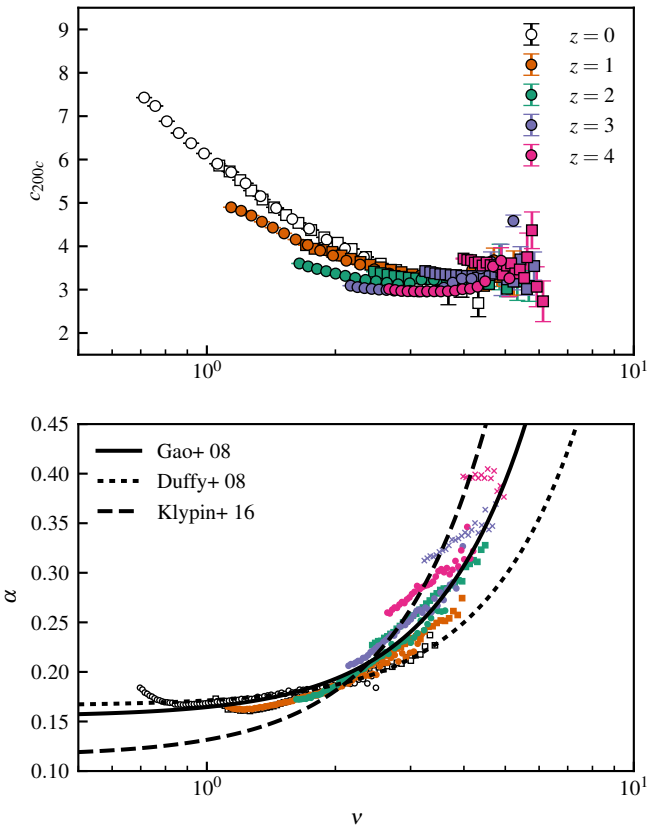

Figure 4.16: Individual $c_{\rm fit}$ (top panel) and stacked Einasto shape parameter α (bottom panel) as a function of ν for Q Continuum (circles) and Outer Rim (squares). As in Figure 4.10, at high z, Outer Rim halos are insufficiently resolved ($\times$ symbols). Curves show the fits given by Gao et al. (2008) [97] (used also by Dutton & Macciò 2014 [80]) and Klypin et al. (2016) [161].

so the many remaining methodological differences we are unable to test must contribute to the discrepancies between the resulting c–M relations. These include halo identification algorithms, the exclusion of unbound particles from halo profiles, and methods of calculating concentrations, among others.

Initial Redshift

The large simulations are initialized at $z_{\rm in} = 200$. (In general, the simulations follow the guidelines laid down in Ref. [131].) We compare a small simulation (same cosmology with box size $L = 1000\,h^{-1}$Mpc and $m_p = 8.56 \times 10^9\,h^{-1}M_\odot$) with this initial redshift to one with a deliberately extreme value of $z_{\rm in} = 30$. While differences in the profiles are visible at $z = 2$, the c–M relations differ by only a few percent. The two small simulations have the same seed, so halos form at the same locations; two halos with the same center are the same halo and can be directly compared. At $z = 2$, 35,551 halos of at least 2000 particles are found in the simulation with $z_{\rm in} = 200$; 28,874 are found in the simulation with $z_{\rm in} = 30$. More halos form in the simulation with higher initial redshift, so we select relaxed halos from $z_{\rm in} = 30$ and pair each with the $z_{\rm in} = 200$ halo whose center is closest, then stack in narrow mass bins. As seen in the stacked profile of Figure 4.17, the $z_{\rm in} = 30$ halos lose mass at all radii (more than 90% of the $z_{\rm in} = 200$ halos have higher mass than their $z_{\rm in} = 30$ pairs), while the slope of the profile is unchanged. In particular, note that the radius at which $n_{\rm eff} = -2$ is identical; the small change in mass, thus r_{200}, does produce a slightly higher concentration when $z_{\rm in} = 200$. For the example in Figure 4.17, $c_{\rm fit}$ increases from 3.44 when $z_{\rm in} = 30$ to 3.49 when $z_{\rm in} = 200$.

Timesteps

In this test, we run smaller simulations (same cosmology with box size $L = 115.375\,h^{-1}\,$Mpc and same mass resolution as Q Continuum, $m_p = 1.05 \times 10^8\,h^{-1}M_\odot$) with half, twice, and the same number of timesteps as for the larger simulations. Again, all three test simulations

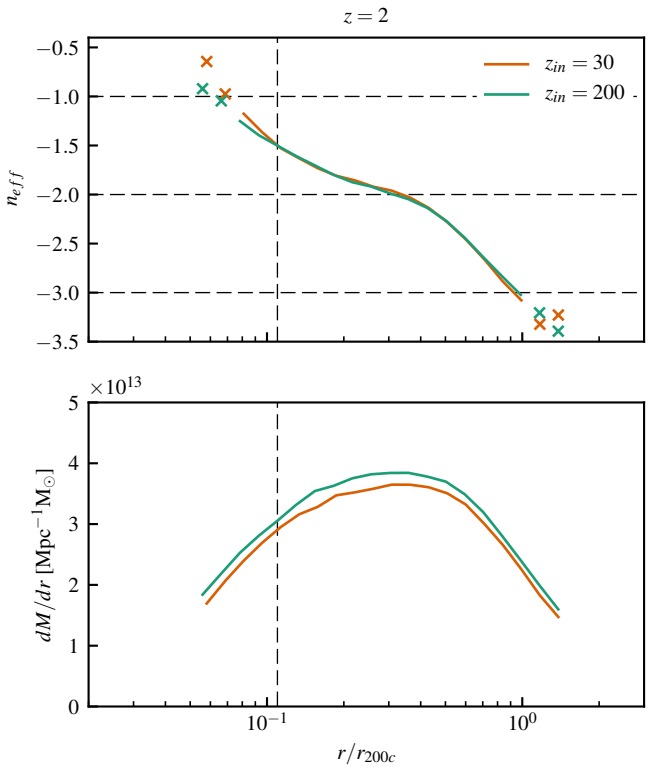

Figure 4.17: Effect of initial redshift on stacked halo profiles using two simulations with the same realization, but with initial redshifts $z_{\rm in} = 200$ and $z_{\rm in} = 30$. The figure shows a 1% stacked profile at $z = 2$, with 115 relaxed halos. Mass is lost across the profile (bottom panel) with little effect on $n_{\rm eff}$ (top); the vertical line shows the innermost radial bin that would be used in fitting this profile. Slopes for the first and last two radial bins ($\times$ symbols) are less reliable than those with four neighboring points to include in the calculation. Mean mass of the stacked $z_{\rm in} = 30$ halos is $2.000 \times 10^{13}\ h^{-1} M_\odot$; mean mass of their $z_{\rm in} = 200$ pairs is $2.169 \times 10^{13}\ h^{-1} M_\odot$.

use the same seed, so the profiles of paired halos can be compared. At $z = 4$, 2782 halos of at least 2000 particles are found in the simulation with the standard number of timesteps; 2625 are found in the simulation with halved timesteps, and 2705 when timesteps are doubled. The fewest halos form with halved timesteps, so relaxed halos are selected from that simulation and paired with the closest halos in the others. As shown in Figure 4.18, the profile at large radius is unchanged, but the inner profile requires more timesteps to converge. Note that all three profiles cross $n_{\rm eff} = -2$ and peak at the same point, indicating no change in scale radius. The shape of the c–M relation is essentially unaffected; at $z = 3$, doubling the number of timesteps shifts the c–M relation uniformly up by less than 5% for both fit and

accumulated mass methods, while halving timesteps shifts it down by about 8%.

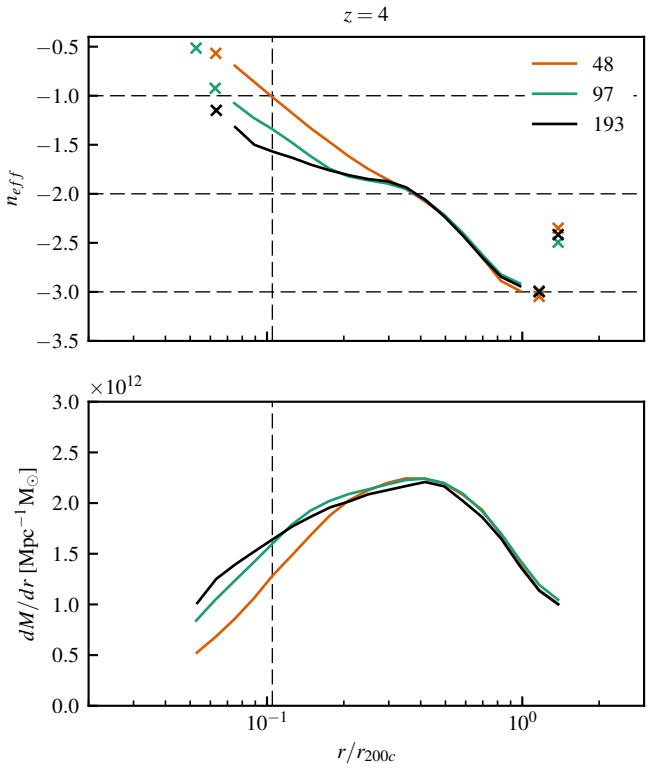

Figure 4.18: Effect of number of timesteps on stacked halo profiles: a 5% stacked profile at $z = 4$, 116 relaxed halos. The inner profile converges with more timesteps; the vertical line shows the innermost radial bin that would be used in fitting this profile. Slopes for the first and last two radial bins ($\times$ symbols) are less reliable than those with four neighboring points to include in the calculation. The mean mass of the stacked halos with 48 timesteps is $2.983 \times 10^{11}\, h^{-1} M_\odot$; mean mass of their pairs is $3.069 \times 10^{11}\, h^{-1} M_\odot$ with 97 timesteps, and $3.011 \times 10^{11}\, h^{-1} M_\odot$ with 193 timesteps.

Halo Centering

We identify the center of a halo as the most bound particle (MBP), locating the local minimum of the potential. Other methods select the most connected particle (MCP, the particle with the greatest number of FOF neighbors), or use a histogram method to find the maximum density. We compare these three methods for a test simulation at $z = 0$. For relaxed halos, where the MBP and MCP centers can differ by $60\, h^{-1}$ kpc, the concentrations vary by less than 5%. The distance between centers is greater for unrelaxed halos – an

average of 160 h^{-1} kpc. At low masses, their fit concentrations differ as well; MCP and histogram fit concentrations are up to 20% greater than MBP for halos of 2000 particles, $M_{200c} = 2 \times 10^{13}\, h^{-1} M_\odot$, while the difference returns to less than 5% at high masses.

Inner Profile

As seen in Figures 4.8 and 4.18, differences from the 'true' profile are more pronounced in the inner profile than at large radius. Error in the inner profile may have little effect on measurements of NFW concentration, but the Einasto profiles are more sensitive to profile shape; a small change in the inner profile produces a larger change in the shape parameter α than in the concentration, as seen in Figures 4.10 and 4.19. Inner-profile discrepancies between Outer Rim and Q Continuum arise only at high redshift; where measurements of the shape parameter are inconsistent, the higher-resolution Q Continuum profiles are closer to the 'true' shape. The convergence in the inner profile is consistent with the considerations presented in Ref. [245].

Peak-finding Method

We present c_{peak} only for lower-concentration halos; this is because of a systematic error that can arise due to insufficient mass resolution. Figure 4.20 shows the artificially high c_{peak} measurements found at lower masses at $z = 0$: when concentration is high, r_s is found at smaller radii, where error in particle counts per radial bin may be substantial. For a low-mass Outer Rim halo at $z = 0$ with mass $4 \times 10^{12}\, h^{-1} M_\odot$ and concentration $c \sim 6$, r_s is found around $0.07\, h^{-1}$Mpc. A single radial bin at this radius will contain about 100 particles; even with three-point smoothing, the peak is noisy. The other two methods are less susceptible to this error: accumulated mass sums over neighboring bins, and profile fit takes the Poisson error in each bin into account. At high masses and low concentrations, however, the relevant radial bins contain 1–2 orders of magnitude more particles; the peak is better defined and the peak-finding method finds concentrations more similar to the results

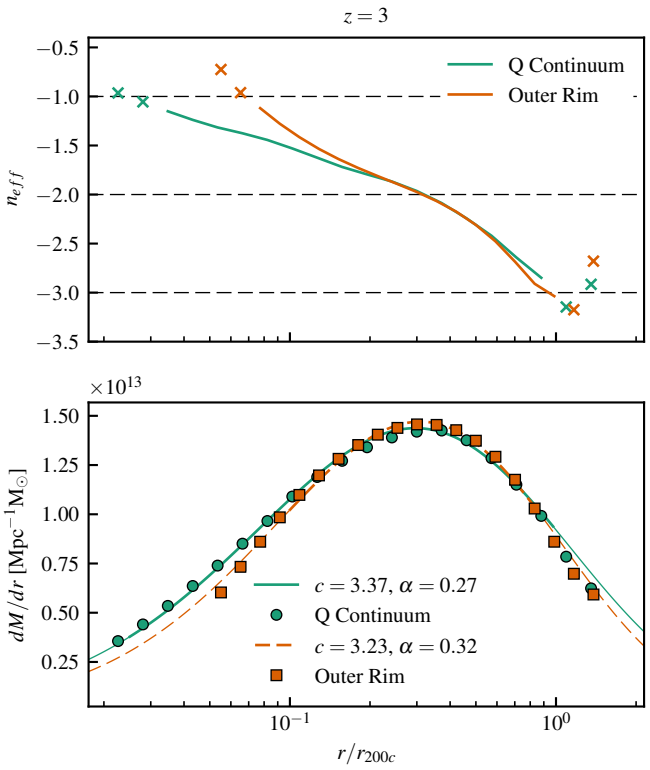

Figure 4.19: Einasto fit to 9556 stacked Outer Rim and 254 Q Continuum profiles at $z = 3$, $M_{200c} = 5 \times 10^{12}\, h^{-1} M_\odot \pm 1\%$. Slopes for the first and last two radial bins ($\times$ symbols, top panel) are less reliable than those with four neighboring points to include in the calculation. Einasto curves are thin in the regions not included in the fit ($r > r_{200}$ or fewer than 100 particles enclosed). In this example, the change in the inner profile causes a 20% discrepancy in fit shape parameter α (see the gap between Outer Rim and Q Continuum high-z shape parameters in Fig. 4.10), but a less than 5% difference in fit concentration.

of the other methods.

Minimum Particle Count

We present concentrations only for halos of at least 2000 particles, as in Ref. [29]. Ref. [221] discusses conservative minimum particle counts of $10^3 - 10^4$ to ensure high-quality fits and agreement between different mass resolutions. Ref. [243], finding statistical bias in concentration measurements of halos resolved with hundreds of particles, advocates a minimum particle count of approximately 4000 particles. Because the mass resolutions of Outer Rim and Q Continuum differ by a factor of 20, the two simulations can be used to check for

convergence at fixed mass. (The concentrations of 100-particle Outer Rim halos, for example, can be compared to 2000-particle Q Continuum halos of the same mass.) As seen in Figure 4.20, 2000 particles is sufficient for fit and accumulated mass methods to agree, but, as discussed above, the peak method can be used safely only on larger, less concentrated halos.

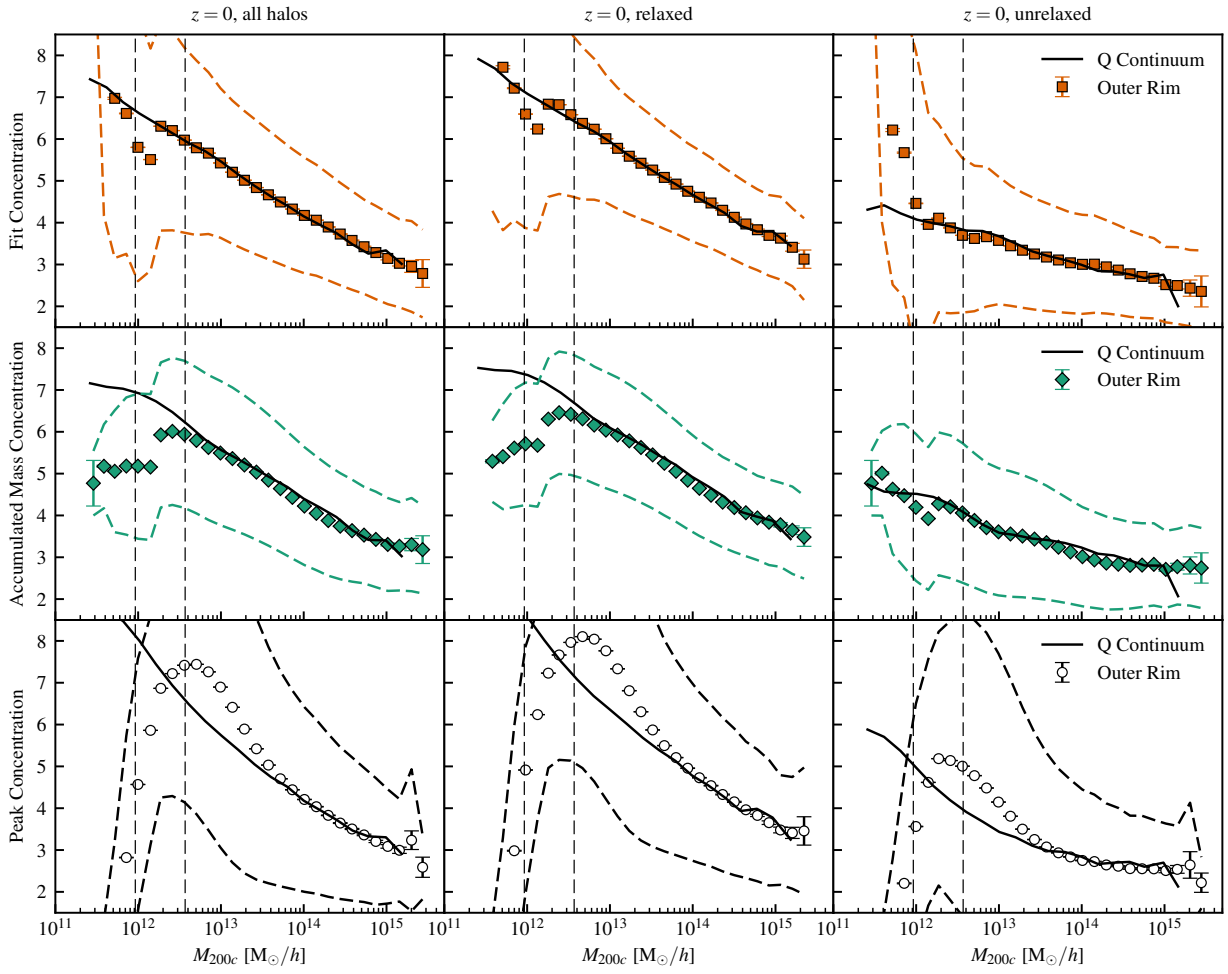

Figure 4.20: At $z = 0$, all three methods agree across simulations when particle count is high, but not on halos containing fewer particles. A Q Continuum halo of mass $M_{200c} = 2.1 \times 10^{11}\ h^{-1}M_\odot$ contains 2000 particles, while an Outer Rim halo of the same mass contains only 114. Dashed vertical lines show mass corresponding to 500 and 2000 Outer Rim particles; the minimum particle count of Q Continuum halos shown is 2000. Mean Outer Rim fit and accumulated mass concentrations (points) diverge from Q Continuum concentrations (lines) below 2000 particles; the peak method is only reliable on halos with an order of magnitude more particles. Dashed curves show 1σ intrinsic variance of Outer Rim halo concentrations.

CHAPTER 5

FORMATION TIME DEFINITIONS AND HALO PSEUDO-EVOLUTION

The connection between halo formation history and structure is important for producing synthetic sky catalogs and understanding structure formation in the Universe. We analyze halos of mass $3 \times 10^{12} h^{-1}\mathrm{M}_{\odot}$ to $9 \times 10^{14} h^{-1}\mathrm{M}_{\odot}$ at $z = 0$ from a cosmological N-body simulation to identify definitions of halo formation time that are best correlated with final halo concentration and thus which features of halo evolution have the greatest effect on the final halo profile. We then show that an assumption of pseudo-evolution predicts concentrations best for small, old, slowly growing halos, while for more massive, rapidly growing halos, pseudo-evolution does not provide a useful prediction of concentration.

5.1 Introduction

According to the hierarchical model of structure formation in a ΛCDM universe, post-inflation fluctuations in the matter density field collapse into a cosmic web of halos. Halos grow through a highly nonlinear process of mass accretion and mergers with other halos. Despite the complexity of the halo growth process, the spherically-averaged radial density profiles of both observed and simulated halos at all redshifts are well fit by the Navarro-Frenk-White (NFW) profile [217, 218]

$$\rho(r) = \frac{\Delta\rho_c}{3A(c_\Delta)} \frac{1}{(r/r_\Delta)(1/c_\Delta + r/r_\Delta)^2} \tag{5.1}$$

with

$$A(c_\Delta) \equiv \ln(1 + c_\Delta) - c_\Delta/(1 + c_\Delta). \tag{5.2}$$

The halo radius r_Δ is the radius within which the average density is Δ times the reference density, which we choose to be the critical (as opposed to mean) density

$$\rho_c(z) = 3H^2(z)/8\pi G, \tag{5.3}$$

where $H(z)$ is the Hubble parameter. Halo radius r_Δ is related to halo mass M_Δ by the spherical overdensity (SO) definition of halo mass:

$$M_\Delta \equiv (4/3)\pi r_\Delta^3 \rho_c \Delta. \tag{5.4}$$

The remaining parameter in Equation (5.1) is the concentration $c_\Delta \equiv r_\Delta/r_s$, which depends weakly on halo mass and redshift through a concentration-mass (c–M) relation. The NFW scale radius r_s is the radius at which the logarithmic slope of the density profile $n_{\rm eff} = d\ln\rho/d\ln(r/r_s)$ takes a value of $n_{\rm eff} = -2$ as it transitions from its value of $n_{\rm eff} = -1$ at small radii to $n_{\rm eff} = -3$ at large radii. Throughout this work we choose $\Delta = 200$ and label halo masses, radii, and concentrations accordingly as M_{200c}, r_{200c}, and c_{200c}.

It is not clear why all halos should be well described by such a simple universal form, and details of halo structure and formation (for example, the relationship between halo mass and concentration, the influence of the environment on halo formation, and the role of mergers in halo structure) are not yet fully understood. In the remainder of the introduction, we briefly outline the current general picture of halo evolution (§5.1.1), its development through models and simulations (§5.1.2), and our work in this chapter (§5.1.3).

5.1.1 Halo Evolution

The shape of a halo depends on its assembly history. Halos evolve through a combination of spherical infall and a series of mergers with violent relaxation to equilibrium [298, 310]. The density of a halo's central regions is expected to correlate with the (mean or critical) background density of the Universe at the time it formed, and low-mass halos coalesce earlier

than more massive halos, so small halos have higher concentrations. During a phase of fast accretion early in their history [178], halos evolve rapidly with fairly constant concentration, as their central density remains set by the background density; major mergers, violent events where two halos of similar mass collide and coalesce into a much larger halo, disturb the inner regions of the halo and are expected to reset the central structure. These rapidly-forming halos have not yet fully disrupted their substructure, and are less relaxed than small, slowly-growing halos. Later, concentration grows as the halo slowly accretes additional mass; small halos may be accreted, but are disrupted by the primary potential well and do not disturb the central profile. The transition between these two phases can be difficult to identify for individual halos [221].

5.1.2 Analytic and Computational Models of Halo Evolution

Regardless of mass, redshift, and mass accretion history [53, 143, 144, 209], halo profiles are all described by a single universal form. The effective slope of the inner profile has been found to be best described by the NFW profile, although other profiles were once considered possible [92, 93, 111, 209, 257]. Numerous semianalytic and analytic works offer partial explanations and predictions of halo shape, profile, evolution, and concentration (e.g. Refs. [8–10, 13, 14, 51, 60, 68, 108, 133, 136, 165, 179, 181, 182, 184, 186, 196, 218, 227, 233, 248, 250, 265, 266, 328, 334]), while computational works consistently show the suitability of the NFW profile.

Simulations of smaller numbers of halos or specific events like mergers can reveal more details about the development of the halo profile than can be modeled with cosmological volumes. For example, mergers play a major role in the development of halo profiles, but these violent events are difficult to model analytically. Merging halos, disrupted by dynamical friction and relaxing into the inner regions of the larger halo [64, 110, 281, 304], establish cuspy halo profiles that are better described by NFW than by profiles with cores [33, 66, 228, 302]. Mergers do redistribute particles between radial bins, without substantially changing

the inner profile [96, 318]. While major mergers typically account for only 15-30% of a halo's mass growth [103, 117], they are the only way to introduce new material into the inner profile [325]; particles accreted in minor mergers are more likely to be found in the outer regions of the halo. The effect of a merger on concentration depends not only on the mass ratio of the progenitors, but also on the angular momentum of the merger [208]. In addition to concentration, mergers affect the shape and spin of the final halo [77].

Zoom-in simulations model the evolution of small numbers of halos at high mass resolution. Simulations of small numbers of halos have shown that differences in mass accretion histories can lead to differences in the resulting halo profiles [160]. For the largest clusters, halo profiles evolve little after $z \sim 1$, regardless of mergers, and are consistent with NFW profiles with concentration $c_{200c} \sim 3.5$ [172, 213]. Constrained simulations, which similarly evolve a small number of halos with initial conditions tuned to produce mass accretion histories with varying levels of merger activity, support the two-phase picture of halo formation: when halos are in the rapid accretion phase and far from virial equilibrium, both the scale radius and virial radius can grow rapidly. In phases of slow accretion, halos are well described by an NFW profile, and only the virial radius grows with the scale factor a [135, 260, 261, 263].

Halo concentrations depend on cosmology [75] and initial conditions [12]; the largest halos evolve from the rarest, highest peaks in the initial density field [71]. Environment effects also play a role in determining concentrations and formation times; some of the scatter in the c–M relation can be attributed to environment effects [24]. Older halos are usually surrounded by other old halos [35]. Interactions of small halos with larger neighbors can strip mass, making the halos seem older than had they never been tidally stripped [70, 122], and concentration may be influenced by distance to massive neighbors [264]. Similarly, old, small halos are likely to have larger neighbors [324]. Halos in overdense cluster environments, where mergers are more common [87], are typically more concentrated [11, 14, 193] and form earlier [121, 124, 284] than halos in voids. Halo clustering has been shown to depend on

concentration and formation time among other halo properties (e.g. Refs. [6, 34, 68, 86, 153, 168, 329, 330]), an effect known as assembly bias, but the exact nature and source of this dependence are not yet established.

5.1.3 This Work

While an understanding of halo assembly is important for developing mock catalogs for survey analysis and for understanding assembly bias, the details of halo formation are clearly not yet fully understood. This chapter is a study of the relationship between halo internal structure, parameterized by concentration, and formation time. In Chapter 4 [43], we presented robust concentration measurements and average c–M relations at a range of redshifts. We found a transition in the shape of the c–M relation around three orders of magnitude above the nonlinear mass scale $M_\star$; previous works have also posited that this mass scale introduces a difference in the rate of mass accretion of halos (e.g. Ref. [3]).

In this chapter, we use these concentration measurements to study the relationship between halo formation time and final halo concentration. In Section 5.2 we outline the simulations used throughout. We summarize definitions of formation time used in other works in Section 5.3, and test their correlation with concentration in Section 5.4. In Section 5.5 we present the evolution of mass and concentration for individual halos and stacks of similar halos and show that pseudo-evolution predicts concentration well for some halos. We conclude in Section 5.6.

5.2 Simulations and Merger Trees

This chapter analyzes a large gravity-only N-body simulation, 'Alpha Quadrant,' run with the Hardware/Hybrid Accelerated Cosmology Code (HACC) framework [118]. The cosmo-

logical parameters are consistent with WMAP-7 [163]:

$$\begin{aligned}
\omega_{\rm cdm} &= 0.1109 \stackrel{h=0.71}{\Rightarrow} \Omega_{\rm cdm} = 0.220, \\
\omega_{\rm b} &= 0.02258, \\
n_s &= 0.963, \\
h &= 0.71, \\
\sigma_8 &= 0.8, \\
w &= -1.0, \\
\Omega_\nu &= 0.0. \qquad (5.5)
\end{aligned}$$

Alpha Quadrant evolves $1024^3 = 1$ billion particles in a box with side length $L_{AQ} = 256\ h^{-1}$Mpc, for a mass resolution of $m_p = 1.62 \times 10^9\ \mathrm{M}_\odot = 1.15 \times 10^9\ h^{-1}\mathrm{M}_\odot$. The initial conditions are a Zel'dovich approximation at $z = 200$ with transfer functions generated by CAMB [175]. Snapshots are output at 99 time steps between $z = 10$ and $z = 0$. We calculate concentrations as $c_{\rm fit}$ of Ref. [43] for halos with $M_{200c} \geq 2.3 \times 10^{12} h^{-1}\mathrm{M}_\odot$ at $z = 0$, corresponding to at least 2000 particles in the SOD profile.

Halos are identified using a friends-of-friends (FOF) finder with linking length $b = 0.168$ in HACC's parallel CosmoTools analysis. Spherical overdensity (SOD) profiles are built around the location of the most bound particle in the FOF group. Halo formation histories are tracked using merger trees built with the algorithm described in Ref. [252]. These merger trees identify halo progenitors as the halos in the previous time snapshot that contribute most of their particles to the descendant halo; examples are shown in Figures 5.1 and 5.2. We analyze full merger trees as well as 'backbones' formed of the $z = 0$ halo, its main progenitor (i.e., the progenitor at the previous step with the largest FOF mass), that halo's main progenitor, and so on to the first redshift at which the backbone halo exceeds 500 simulation particles.

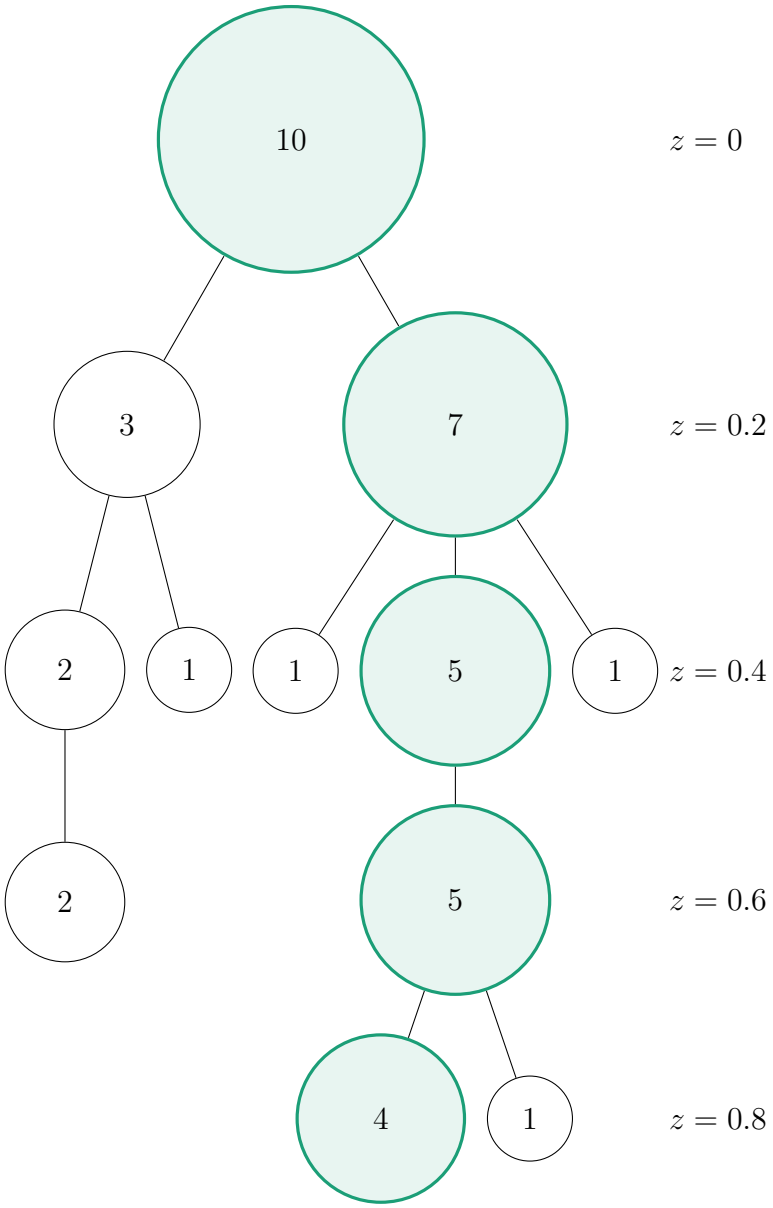

Figure 5.1: A simple example of a merger tree. Node area is proportional to halo mass, and the mass of each halo is given in the node (in units of, for example, 10^{13} h/Mpc). Backbone halos—the most massive progenitor of each most massive progenitor—are highlighted in green. In this example the most massive halo in the merger tree at each step is the backbone halo, but this need not always be the case.

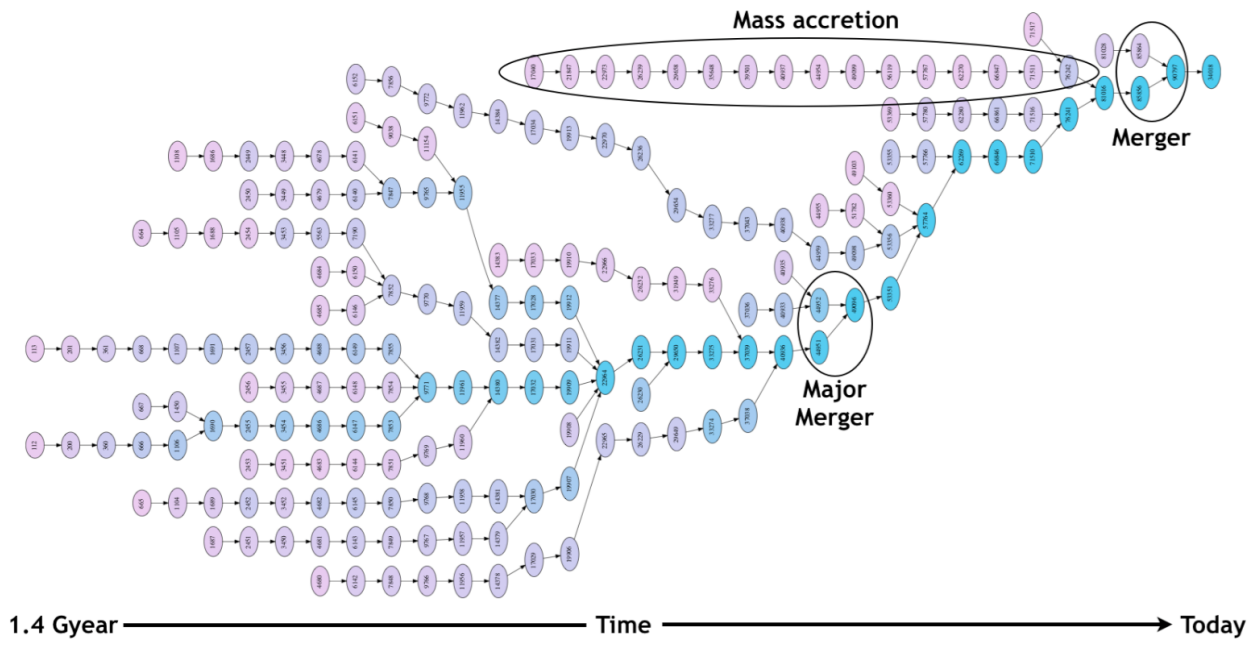

Figure 5.2: Merger trees in simulations are much more complex than the simple example of Figure 5.1. Nodes are color coded by halo mass, increasing from pink to blue. The final (rightmost) halo at $z = 0$ grew through a series of mergers of smaller halos, each of which formed through a combination of mergers and slow accretion of small amounts of matter. Figure courtesy of Eve Kovacs and Esteban M. Rangel.

5.3 Definitions of Formation Time

Many different definitions have been used to calculate halo formation times. In this section, we outline the various options and give examples of their applications. Several works have compared multiple definitions of formation time and are listed in more than one section (e.g. Refs. [3, 49, 177, 178, 221]). Each definition is sensitive to a different stage in the mass accretion history; the definitions that are most predictive of concentrations reflect the features of halo assembly that are most influential in determining internal structure.

5.3.1 Power Spectrum Collapse Mass

Some definitions of formation time do not make use of halo assembly histories, but instead calculate the time when the halo should have collapsed. These definitions do not allow for any scatter in the concentration of a halo; concentration is set solely by the mass of a halo

through the time when its mass scale was collapsing. For this reason we do not include them in our tests of formation time definitions. We do, however, use a similar definition personalized to halo formation histories. In Chapter 4 we showed that the c–M relation is constant with mass above a redshift-dependent threshold mass $\alpha M_\star(z)$ with $\alpha \sim 1000$. The nonlinear mass scale $M_\star(z)$ (defined in §4.4) describes the halo mass at which the corresponding peaks in the initial Gaussian random field are collapsing at redshift z, and is strongly dependent on redshift: $M_\star \approx 10^{12.5}\,h^{-1}M_\odot$ at $z = 0$, $10^{11}\,h^{-1}M_\odot$ at $z = 1$, and $10^{9.5}\,h^{-1}M_\odot$ at $z = 2$. For halos smaller than $\alpha M_\star(z)$, concentration falls with mass, so $\alpha M_\star$ represents a transition between two concentration regimes.

In Ref. [38], formation time is found as the redshift at which the nonlinear mass scale $M_\star$ reaches a fraction F of the final halo mass. A small $F = 0.01$ is found to be best for the ΛCDM model tested. Refs. [53, 75, 187] also test this definition. Refs. [75, 85] calculate the redshift at which the power spectrum amplitude at mass scale M reaches a constant $C = 1/28$.

5.3.2 Fraction of Final Mass Assembled into Progenitors

The first papers to introduce the NFW profile [217, 218] define the formation time of a halo based on the full collapsed mass history of the halo. A halo formed when a fraction F of the final mass of the halo was found in progenitors each exceeding a small fraction f of the final mass. The two fractions F and f are free parameters; many works have chosen $F = 0.5$, indicating half the mass of the final halo was contained in progenitors. Refs. [217, 218] used $F = 0.5$ with $f = 0.5$, 0.1, and 0.01. The largest value of f indicates of course that a single large progenitor crossed half the final mass of the halo; the agreement with simulations was best for smaller values of f. Ref. [148] uses $F = 0.5$ and $f = 0.1$ to measure halo age for comparison with other properties, including concentration. Refs. [75, 147, 284] use $F = 0.5$ and $f = 0.01$; Ref. [221] uses $F = 0.5$ and $f = 0.1$; and Ref. [177], $F = 0.5$ and $f = 0.02$.

Some papers generalize to different values of F. Ref. [97] used a smaller $F = 0.1$ with

$f = 0.01$; this definition was also tested in Ref. [187].

Rather than a numerical value of F, Ref. [187] connects it to the concentration of the final halo: a halo forms when the masses of all progenitors (each exceeding a fraction f of the final virial mass) sum to M_{-2}, the mass contained within the final halo's scale radius. This definition translates to $F = M_{-2}(z = 0)/M_{vir}(z = 0)$. For the minimum mass of progenitors, $f = 0.005$, 0.01, and 0.02 are tested; $f = 0.02$ performs best. This definition is also used by Ref. [24].

Motivated by the physics of star formation, Ref. [220] requires progenitors to pass a redshift-dependent threshold mass, rather than a fraction of the final halo mass. Ref. [177] uses a similar definition, but fixes the threshold mass as $M_c = 10^{11.5}\, h^{-1}\, M_\odot$; their $z_{1/2,t_2}$ is defined as the redshift when the mass of these progenitors exceeds $F = 0.5$ of the final halo mass. In another definition motivated by star formation, $z_{M/L}$, Ref. [177] allows the fraction F to depend on the final halo mass as a power law; progenitors must still exceed the threshold mass M_c. We do not test these definitions, as their motivations lie outside the dark-matter-only simulations used here, and they are not commonly used for concentration studies.

5.3.3 Fraction of Final Mass in Main Progenitor

By far the most common definition of formation time is the time when the main progenitor (or a progenitor) first reaches half the final (or, less often, maximum) halo mass. This definition appears in Refs. [13, 14, 24, 32, 49, 52, 58, 68, 87, 95, 98–101, 108, 110, 120–122, 124, 132, 132, 146, 153, 159, 167, 169, 170, 173, 177–179, 186, 187, 190, 197, 220, 227, 244, 249, 261, 267, 283, 284, 306, 308, 309, 319, 321, 323, 324, 326, 330, 335–337, 343, 349]. Ref. [242], instead, considers the last time when half the final mass was reached, and Ref. [264] uses the peak mass of the halo throughout its history in place of its final mass.

Other fractions are also possible. Ref. [346] chooses $f_M = 0.04$, and Ref. [320] uses their model; Ref. [109] tests $f_M = 0.01$, 0.1, 0.5, and 0.9; Ref. [35] uses $f_M = 0.25$, 0.5,

and 0.75; Ref. [221] tests $f_M = 0.1$, 0.25, and 0.5; Ref. [65], 0.5 and 0.85; Ref. [307], 0.5 and 0.9; Ref. [123], 0.05, 0.1, 0.3, 0.5, and 0.8. Ref. [34] uses $f_M = 0.25$ in addition to $f_M = 0.5$. Ref. [47] tests a range of f in addition to their definition of age calculated from mass accretion histories and Gaussian statistics.

Similar to Ref. [187] in Section 5.3.2 above, Ref. [189] defines formation time as the redshift when the virial mass of the halo first reaches the mass within its $z = 0$ scale radius; that is,

$$M(r_{200}, z = z_{\rm form}) = M(r_s, z = 0), \tag{5.6}$$

so $f_M = M(r_s, z = 0)/M(r_{vir}, z = 0)$. This definition is also tested in Refs. [24, 56, 57, 188] and in Ref. [187], which finds concentrations to be directly proportional to formation times defined as in equation (5.6).

The main progenitor can be defined in different ways. At each timestep, we select the most massive of all the progenitors from the previous time step. Alternative definitions are the most massive progenitor of the final halo at any time step, regardless of whether its direct descendent is on the main branch of the merger tree (e.g. Refs. [109, 177, 254]); the halo that contributes the most particles to its descendant (e.g. Ref. [261, 283, 347]) or maximizes another merit function of the number of particles found in both haloes (e.g. Ref. [172]); the location of particles that at $z = 0$ are found in the density peak of the final halo (e.g. Refs [253]); or the branch that contributes the most mass to the final halo (e.g. Ref. [35]).

5.3.4 Threshold Mass

A simple measure of formation time considers a halo to form when its mass first exceeds a fixed threshold. This definition of formation time is used in special cases where other definitions are not feasible, like subhalo formation (Ref. [76], threshold mass of $3 \times 10^8\, h^{-1}{\rm M}_\odot$) or microhalos (Ref. [145], threshold mass of $10^{-6}\, h^{-1}{\rm M}_\odot$). Ref. [49] employs a threshold mass of $10^{14}\, h^{-1}{\rm M}_\odot$ for cluster formation; Ref. [126] uses $10^{12}\, h^{-1}{\rm M}_\odot$. Motivated by star

formation, Ref. [177] uses a threshold mass of $M_c = 10^{11.5}\,h^{-1}\,M_\odot$ to define their $z_{\rm core,mb}$ (main branch mass is compared to M_c) and $z_{\rm core,mp}$ (mass of the most massive progenitor is compared to M_c). We test only the larger threshold masses, as we do not resolve microhalos and these masses represent much less than one percent of the final mass of the halos we consider.

5.3.5 Accretion History Fits

Early in its history, a halo accretes mass rapidly before transitioning to a slow accretion regime. This transition is visible as a bend in a halo's mass accretion history. Unlike the other definitions of formation time, fits to mass accretion histories use information from more than a single point in the halo's history.

Ref. [328] fits accretion histories to the form

$$M(a) = M_0 \exp\left[-a_c S\left(\frac{a_0}{a} - 1\right)\right], \tag{5.7}$$

where a_c is the formation scale factor, a_0 is the scale factor when the halo is observed with mass M_0 at $z = 0$, and S is an arbitrary constant chosen as $S = 2$. Refs. [126, 172, 184, 263, 280, 329, 330] use the same definition of formation time; Ref. [3] also allows M_0 to be a free parameter. Ref. [306] uses equation (5.7), as well as a generalized form

$$M(a) = M_0 \left(\frac{a}{a_0}\right)^p \exp\left[\alpha\left(1 - a_0/a\right)\right], \tag{5.8}$$

with $\alpha = a_c S/a_0$, used also by Ref. [203].

5.3.6 Circular Velocity

Refs. [70, 254] calculate rotation curves according to

$$V_c(r) = (GM(<r)/r)^{1/2}. \tag{5.9}$$

A halo forms when $V_{\rm peak}$, the peak value of V_c for the largest progenitor, reaches 85% of its maximum value. The definition of Ref. [200] is similar, but generalized to other fractions of the maximum $V_{\rm peak}$: 0.9, 0.8, and 0.6. Ref. [178] finds the redshift where the halo's circular velocity at its virial radius, V_h, first reaches the $z = 0$ value, as well as the time where V_h reaches its maximum value over the entire history of the halo and transitions from fast accretion to slow accretion (the latter is also used by Ref. [177]).

Ref. [347] uses the turning point between fast and slow accretion phases, identified as the redshift where

$$\log(V_h(z)) - \gamma \log(H(z)) \tag{5.10}$$

reaches its maximum with $\gamma \sim -1/4$. At this redshift, a halo forms with concentration 4. With the same definition of formation time and more halos, Ref. [345] finds an initial concentration of 3.5 to better predict concentrations.

5.3.7 Last Major Merger

A major merger is a violent event that can disrupt the inner regions of a halo. Refs. [250, 265] consider a major merger to have occurred if the mass of a halo increases by a factor $(1+\Delta_m)$ and find concentrations are well predicted with $\Delta_m = 0.6$. After the merger, the scale radius remains fixed; as the halo continues to slowly accrete mass, the virial radius expands, increasing the concentration. The characteristic density of the halo δ_c is a constant multiple of the mean density at the time of formation. Ref. [196] allows the effective threshold for mergers Δ_m to vary between $0.3 \leq \Delta_m \leq 0.7$, with a best value of $\Delta_m = 0.5$. Refs. [177, 178, 197, 242, 264] define a major merger as the mass ratio of progenitors exceeding 1/3. Ref. [49] identifies the last large increase in mass, such as an increase of 20% in a time interval $\delta\tau = 100h^{-1}\,M_\odot$; Ref. [330] also uses an increase of 20%, but in a time interval corresponding to the light crossing time. Ref. [261] finds that halos with the most recent major mergers also have recent half-mass formation times. Ref. [146] finds weak correlation

between concentration and the time of the last major merger. Ref. [336] considers a merger to be major if the ratio of the merging mass to the final mass exceeds 0.3.

5.4 Comparison of Formation Time Definitions

From the works listed above, we collect the following list of formation time definitions:

- $z_{\rm NFW}(F, f)$: A fraction F of the halo's final FOF mass has assembled in progenitors each exceeding a fraction f of the final mass (Section 5.3.2). We test all combinations of $F = 0.1,\ 0.5,\ M_{-2}(z = 0)/M_{200c}(z = 0)$ and $f = 0.005,\ 0.01,\ 0.02,\ 0.1,\ 0.5$ such that $F \geq f$. In our notation, $z_{\rm NFW}(f = 0.1,\ F = 0.5)$ denotes the redshift corresponding to the first step when progenitors of mass at least $0.1M_{200c}(z = 0)$ sum to $0.5M_{200c}(z = 0)$. We note that when $F = f$, $z_{\rm NFW}$ is almost the FOF analog of the next definition, $z_{\rm M}(F_{\rm M})$, but considers the most massive progenitor and not the main progenitor.

- $z_{\rm M}(F_{\rm M})$: The halo's main progenitor reaches a fraction $F_{\rm M}$ of the halo's final mass M_{200c} (Section 5.3.3). We test $F_{\rm M}$ in steps of 0.01 from 0.01 to 0.99 and $F_M = M_{-2}(z = 0)/M_{200c}(z = 0)$.

- $z_{\rm T}(M_T)$: Threshold mass: when the mass of a halo's main progenitor first attains a fixed mass M_T (Section 5.3.4). We test $M_T/(h^{-1}M_\odot) = 10^{11.5}$, 10^{12}, and 10^{14}.

- $z_{\rm W02}$ and $z_{\rm T04}$: Both definitions find the turning point from fast to slow accretion by fitting the mass accretion history to a functional form (Section 5.3.5; equation (5.7) for $z_{\rm W02}$ and equation (5.8) for $z_{\rm T04}$, both with $S = 2$). The formation time is calculated from the fit value of the formation scale factor a_c.

- $z_{vpeak}(f_V)$: For a halo along the backbone, the peak value $V_{\rm peak}$ of the circular velocity V_c (defined in Section 5.3.6 above) first reaches a fraction f_V of its maximum value. We test values of $f_V = 0.6$, 0.8, 0.85, 0.9, and 1.

- z_{vmax}, z_{vvir}, and z_{Z03}: Using instead V_h, the circular velocity at r_{200c} of the main progenitor, we calculate the time z_{vmax} when V_h reaches its maximum value over the halo history; z_{vvir} when V_h first reaches its value at $z = 0$; and z_{Z03} according to equation (5.10).

- $z_{\rm merge}(\Delta)$: The most recent time when a halo's mass increased by a factor of at least $(1 + \Delta)$ in one timestep; that is, the final halo of mass at least $(1 + \Delta)M$ has a main progenitor with mass M (Section 5.3.7) at the previous timestep. We test $\Delta = 0.2$, 0.3, 0.4, 0.5, 0.6, and 0.7.

- $z_{\alpha M_\star}(\alpha)$: The redshift $z_{\alpha M_\star}$ at which the halo's main progenitor last fell below the redshift-dependent mass threshold $\alpha M_\star(z)$. We test $\alpha = 500$, 1000, 1500.

We use concentrations $c_{\rm fit}$ of [43]. We consider both relaxed and unrelaxed halos resolved with at least 2000 particles at $z = 0$, that is, of SOD mass at least $2.3 \times 10^{12} h^{-1} {\rm M}_\odot$. In this section, we calculate the correlation between concentration and each definition of formation time (§5.4.1), comment on patterns in correlation coefficients and the definitions that are best correlated with concentration (§5.4.2), and discuss the behavior of halo age as a function of mass and concentration for different definitions (§5.4.3).

5.4.1 Correlations Between Formation Time and Concentration

To quantify the relationship between final halo concentration and each definition of formation time, we measure the Spearman's rank-order correlation coefficient ρ between concentration and formation time, given in Table 5.1. Spearman's rank-order correlation coefficient, used by e.g. Refs. [122, 148, 197, 335] to study the relationships between various sets of halo parameters, nonparametrically measures the strength of a monotonic relationship between two data sets. It does not assume a linear relationship, but tests whether the earliest-forming halos are most concentrated or vice versa. Formation times and concentrations are ranked, and ρ is computed as the Pearson correlation coefficient of the ranked concentrations c_R and

ranked formation times z_R. That is,

$$\rho = \frac{\mathrm{cov}(z_R, c_R)}{\sigma(z_R)\sigma(c_R)}, \tag{5.11}$$

where $\mathrm{cov}(z_R, c_R)$ is the covariance of the rank variables z_R and c_R and $\sigma(z_R)$ and $\sigma(c_R)$ are their standard deviations.

We compute ρ in twenty mass bins. The correlation does vary with mass, but not strongly, and not in a way that affects the order of best definitions. That is, the formation time that is most correlated with concentration for halos of mass $10^{12}h^{-1}\mathrm{M}_\odot$ is also the most correlated for halos of mass $10^{14}h^{-1}\mathrm{M}_\odot$. We compute a single value of ρ for each definition by calculating the average value of ρ across all mass bins, weighted by the number of halos in the bin.

Some definitions of formation time cannot be measured for all halos, as shown in Table 5.1. Formation times based on circular velocity and accretion history fits are measured for all halos, but for other definitions, the fraction of halos for which the formation time is found depends on the threshold chosen. Additionally, the fraction of halos for which a formation time is defined varies with mass. For example, for low-mass halos resolved with only 2000 particles, a progenitor whose mass is 1% of the final halo mass contains only 20 particles and is too small to be included in the halo catalogs. Some threshold masses are similarly too small to be calculated for all halos, while others are too large (for example, few halos have passed the 10^{14} mass threshold even at $z = 0$). Only about one-fifth of halos have experienced mergers with mass ratios exceeding 0.6, but more than 90% of halos have merged with a mass ratio of at least 0.2.

5.4.2 Comparison of Definitions: Correlation and Completeness

The best definitions are those that are well-defined for the majority of halos and highly correlated with formation time. Luckily, the definitions that are best correlated with con-

Table 5.1: Correlation of formation time with concentration

Definition	percent of halos	correlation coefficient
$z_{\rm NFW}$		
$0.1, 0.005$	100	0.60
$0.1, 0.01$	100	0.60
$0.1, 0.02$	100	0.60
$0.1, 0.1$	100	0.58
$0.5, 0.005$	100	0.52
$0.5, 0.01$	100	0.52
$0.5, 0.02$	100	0.52
$0.5, 0.1$	100	0.56
$0.5, 0.5$	100	0.60
$M_{-2}/M_{\rm fof}, 0.005$	99.9	0.89
$M_{-2}/M_{\rm fof}, 0.01$	99.9	0.88
$M_{-2}/M_{\rm fof}, 0.02$	99.9	0.86
$M_{-2}/M_{\rm fof}, 0.1$	93	0.82
$M_{-2}/M_{\rm fof}, 0.5$	1.2	0.26
z_M		
0.01	3	0.37
0.04	16	0.44
0.1	44	0.52
0.25	96	0.60
0.5	100	0.65
0.75	100	0.56
0.8	100	0.52
0.9	100	0.37
M_{-2}/M_{200c}	100	0.56
z_T		
$10^{11.5}$	4	0.57
10^{12}	100	0.56
10^{14}	1	0.37
z_{W02}	100	0.71
z_{T04}	100	0.26
z_{vpeak}		
0.6	100	0.51
0.8	100	0.33
0.85	100	0.25
0.9	100	0.15
1	100	-0.30
z_{vmax}	100	-0.23
z_{vvir}	100	-0.23
z_{Z03}	100	-0.21
z_{merge}		
0.2	92	0.52
0.3	73	0.46
0.4	52	0.44
0.5	34	0.44
0.6	22	0.44
0.7	14	0.45
$z_{\alpha M_\star}$		
500	65	-0.53
1000	55	-0.50
1500	49	-0.47

Notes. The fraction of halos for which each definition of formation time can be computed, and the Spearman's rank-order correlation coefficient between concentration and each formation time. Formation times $z_M(F_M)$ for several representative values of F_M are included in the table; the results for all measured F_M are shown in Figure 5.3.

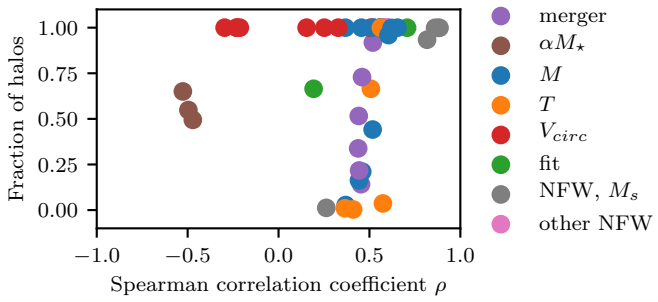

Figure 5.3: The best definitions of formation time are highly correlated with concentration and can be computed for a large fraction of halos. Many definitions cluster around $\rho = 0.5$, indicating only a moderate correlation with concentration. Definitions based on the mass of all progenitors $z_{\rm NFW}$ perform the best: they are defined for nearly all halos and very strongly correlated with concentration ($\rho \gtrsim 0.85$).

centration are well-defined for the vast majority of halos. Figure 5.3 shows the correlation coefficient ρ and fraction of halos on which it is calculated for each definition of formation time. The upper-right corner represents optimal definitions of formation time: those that can be measured on all halos and correlate strongly with concentration. The strongest correlations are for $z_{\rm NFW}$ with certain choices of F, the accretion history fit z_{W02} of Ref. [328], and half-mass formation time $z_M(0.5)$. Each of these three definitions considers a different subset of the information contained in a mass accretion history. We now discuss patterns in correlations between concentrations and several formation times.

$z_{\rm NFW}$— Of all definitions of formation time tested, $z_{\rm NFW(M_s/M_{fof},0.005)}$ is most correlated with final halo concentration, with a Spearman correlation coefficient of $\rho = 0.89$ (see Table 5.1). Ref. [221] similarly found better predictions of concentration using all progenitors rather than only the backbone. We find that collapsed mass definitions of formation time can improve on backbone definitions, but not for all choices of the mass fraction F. Correlation is much more dependent on F than on the minimum mass fraction f for a progenitor to be included in the sum. The value of F for which formation times are most correlated with concentration is $F = M_s/M_{\rm fof}$—that is, the time when the sum of progenitor masses first reached the mass that at $z = 0$ is contained within the scale radius. With this value of F, the correlation coefficient is highest for the smallest values of f.

By considering the mass of all halo progenitors, $z_{\rm NFW}$ counts the mass in the full Lagrangian volume that eventually becomes the $z = 0$ halo. The smallest values of f perform best because the sum over progenitors captures a larger fraction of the Lagrangian space. Additionally, smaller values of f allow halos to form earlier, enabling a larger dynamic range of formation times. As f increases, a smaller fraction of the Lagrangian volume is considered. For the largest values, f can exceed $M_s/M_{\rm fof}$ for many high-concentration halos. For these halos, no formation time can be found; for example when $F = M_s/M_{\rm fof}, f = 0.5$, only about 1% of halos have a formation time.

The formation times best correlated with concentration, those for which $F = M_s/M_{\rm fof}$ and f is small, therefore appear to perform so well for three reasons: they capture a larger fraction of the Lagrangian volume; they allow a larger range of formation times and more differentiation between halos; and they use the final halo concentration in determining formation time.

z_{W02}— The accretion history fit of Ref. [328] does not consider all progenitors, but does use the full (smoothed) accretion history. The other definition which calculates the turning point between fast and slow accretion by fitting the full accretion history, z_{T04} of Ref. [306], is, in contrast, much more weakly correlated with concentration: for z_{T04}, $\rho = 0.26$, one of the lowest correlation coefficients among formation times tested here. While the fitting form of Ref. [306] does provide better fits to some mass accretion histories, the tails of the distribution of z_{T04} are very fat, with formation times varying between 10^3 and -10^4, and the peak of the distribution is approximately $z = 0$. This fitting form predicts many negative formation times, which are reasonable because they describe halos that are still in the rapid phase of mass accretion and have not yet transitioned to a phase of slow mass accretion. However, large positive or negative formation times require significant extrapolation of the measured mass accretion histories, and formation times near $z = 0$ differentiate little between different halos. When compared to the results for z_{W02}, the additional free parameter in the fitting form for z_{T04} appears to allow overfitting of formation histories for the purposes

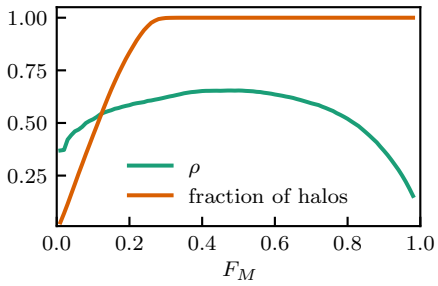

Figure 5.4: For formation time definitions based on the time z_M when the halo's main progenitor reaches a fraction F_M of the final halo mass, $F_M = 0.5$ is better correlated with final concentration than any other value of F_M. The Spearman correlation coefficient rises with F_M at small F_M, as very early progenitors carry little information about the final halo. More recent progenitors are better correlated with the halo's final concentration, but for fractions that are too large, z_M is comparable for all halos and does not distinguish between halos. For small values of $F_M \lesssim 0.3$, formation time is not found for all halos (orange curve) as the first identified progenitor already exceeded F_M times the final halo mass.

of concentration predictions, while providing good fits of the formation histories.

z_M— The half-mass formation time $z_M(0.5)$ measures only a single point in the mass accretion history, but nevertheless is highly correlated with final concentration. The correlation between $z_M(F_M)$ and concentration depends on F_M. The most commonly used fraction, $F_M = 0.5$, performs best of all fractions we test; its Spearman correlation coefficient is 0.65 for $z_M(0.5)$, similar to the value of 0.63 found by Ref. [197]. Figure 5.4 shows that for other choices of F_M around 0.5, the correlation coefficient is not highly sensitive to F_M, but at much larger and smaller values of F_M this definition of formation time performs much worse. The correlation coefficient increases monotonically with F_M until reaching a maximum at $F_M = 0.5$, then declines with F_M. The characteristic time in the evolution of a halo's main progenitor appears to be near the middle of its mass accretion—neither too early, when the mass is very small compared to the final mass, nor too late, when most of the halo's final mass has already been assembled. Of course, halos that only very recently attained half of their final mass are rapidly growing, and likely still in the fast accretion phase.

Any one measure of assembly history will not account for all the scatter in final halo concentration (e.g. Ref. [335], who find halo age and the late-time rate of growth to together

explain the majority, 70%, of concentration scatter). Future work will study how much variation in concentration can be explained by combinations of multiple measures of accretion history.

5.4.3 Visualizing Halo Age, Concentration, and Mass

Figure 5.5 (continued) displays average formation times in bins of mass and concentration for several definitions of formation time. For most definitions, the Spearman rank correlation coefficient ρ between formation time and concentration is positive, indicating that the earliest-forming halos in a mass bin are also the most highly concentrated. For the most highly-correlated definitions ($z_{\rm NFW}$, Z_{W02}, and z_M), the earliest-forming halos are the low-mass halos with high concentration. Halos with similar concentration have similar formation times, regardless of mass. The threshold mass definition, in contrast, conflates mass and concentration; clearly the most massive halos were the first to cross any given mass threshold, so these halos are assigned the earliest formation times. Most halos first reached their maximum value of V_c very recently, so formation times $z_{vpeak}(1)$ are very small and not highly correlated with concentration.

5.5 Evolution of Concentrations

We track halo concentrations as they evolve along the backbones of halos. In this section, we define the backbone by the chain of the most massive progenitors of the most massive progenitor at each timestep. We measure concentrations $c_{\rm fit}$ for each halo in the backbone and calculate physical scale radii r_s according to $r_s(z) = r_{200c}(z)/c_{200c}(z)$. Two examples of M_{200c}, c_{200c}, and r_s evolution are shown in Figure 5.6. One of the accretion histories shows a halo that is cluster scale at $z = 0$ (mass $9 \times 10^{14} h^{-1}{\rm M}_\odot$), while the other is a small, old halo (mass $7 \times 10^{12} h^{-1}{\rm M}_\odot$ and half-mass formation time $z_M(0.5) = 2$). Figure 5.6 also shows half-mass formation times for both halos; the more massive halo did not attain half

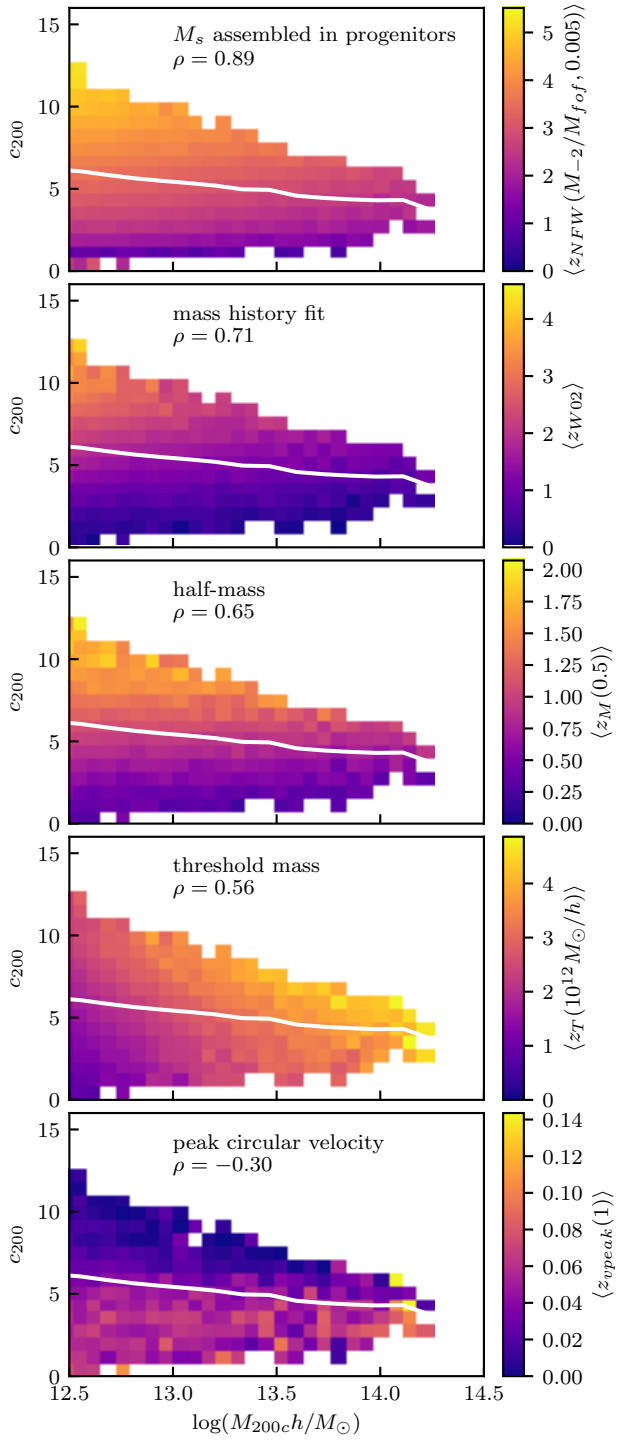
M_s assembled in progenitors
$\rho = 0.89$
mass history fit
$\rho = 0.71$
half-mass
$\rho = 0.65$
threshold mass
$\rho = 0.56$
peak circular velocity
$\rho = -0.30$
c_{200}
$\log(M_{200c}h/M_\odot)$
$\langle z_{NFW}(M_{-2}/M_{fof}, 0.005)\rangle$
$\langle z_{W02}\rangle$
$\langle z_M(0.5)\rangle$
$\langle z_T(10^{12}M_\odot/h)\rangle$
$\langle z_{vpeak}(1)\rangle$

Figure 5.5

Figure 5.5 (continued): By most definitions of formation time, the most concentrated halos in a mass bin are older than halos with lower concentration. Each panel shows, for a given formation time definition, the average formation time in bins of mass and concentration; white curves are the average c–M relation for all halos (same in all panels), and ρ is the Spearman rank correlation coefficient (§5.4). A formation time definition $z_{\rm NFW}(M_{-2}/M_{\rm fof}, 0.005)$ based on the time when the mass of progenitors first summed to M_s (top panel), the accretion history fit of Ref. [328] (second panel) and the half-mass formation time $z_M(0.5)$ (third panel) assign the highest formation times to the smallest, most concentrated halos, and are highly correlated with concentration. Under a threshold mass definition like that of the fourth panel ($M_T = 10^{12}h^{-1}{\rm M}_\odot$), high-mass halos are oldest, because a currently high-mass halo passed any given threshold mass earlier in its history. Some definitions assign the highest formation times to less concentrated halos, such as the time $z_{vpeak}(1)$ when the halo's peak circular velocity first reached its maximum value (bottom panel). This last definition shows little correlation with concentration.

its final mass until $z_M(0.5) = 0.46$, and grew by a factor of more than 100 from $z = 2$ to $z = 0$.

While the two halos shown in Figure 5.6 were of similar mass at $z \sim 2$, by $z = 0$, one is larger than the other by two orders of magnitude. As it rapidly accreted mass, its scale radius grew by more than an order of magnitude and its concentration remained low ($c \sim 3.5$). The halo that is much less massive at $z = 0$ accreted mass much more slowly and its scale radius remained small. In contrast, its concentration grew steadily from $z = 2$ to the present epoch. As seen in this example, halo mass evolves fairly smoothly along backbones, although major mergers cause M_{200} to jump rapidly and halos can occasionally lose mass. Concentration is a noisier measurement. In this section we show how an assumption of pseudo-evolution can predict the evolution of concentration for some halos.

5.5.1 Pseudo-evolution

The boundary of a halo defined by any overdensity criterion does not correspond to a physical choice (for $z \geq 0$; in the future, halos will be so far from their neighbors that halo boundaries will clear [39]). Because the critical density of the Universe evolves with time (equation 5.3), concentration must evolve even if the mass distribution and therefore r_s of a halo remain

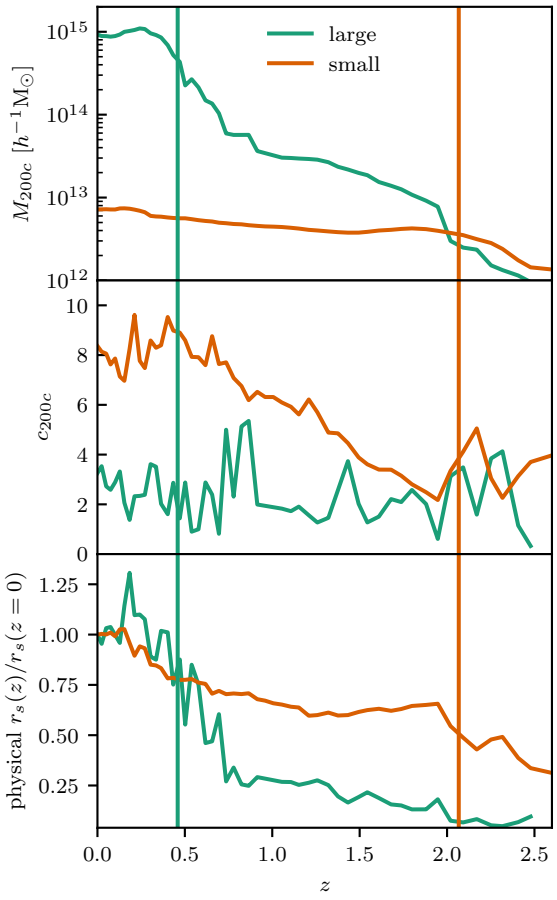

Figure 5.6: Two halos with similar mass at high redshift can evolve very differently. *Top*—Mass accretion histories for two halos whose mass at $z = 0$ differs by a factor of 100. Vertical lines indicate the half-mass formation times $z_M(0.5)$ corresponding to the mass accretion histories of the small halo (orange) and large halo (green). The small halo has doubled in mass since $z = 2$, while the large halo has grown by a factor of more than 100. The mass of the large halo has increased in several large jumps, indicating major mergers, while the mass accretion history of the smaller halo is relatively smooth. *Middle*—When a halo rapidly gains mass, its concentration remains low, around $c \sim 3$. When it instead slowly accretes mass, concentration grows, in this case to $c \sim 9$ by $z = 0$. *Bottom*—The physical scale radius r_s of the large halo grows substantially, while r_s changes relatively little across the history of the slowly-growing halo.

fixed. The halo radius r_{200c} is defined relative to the critical density of the Universe, so as the critical density evolves, so will the calculated radius of the halo. Therefore concentration will increase slowly and predictably with time even if the halo itself is not changing: it pseudo-evolves. More physical definitions of halo radius such as a true virial radius [52], splashback radius [211] or static mass [59] are less susceptible to pseudo-evolution.

Ref. [265] suggests that halo scale radius remains fixed; only the virial radius expands, increasing concentration. Several works find concentration to evolve after formation proportional to the scale factor, predicting the scale radius r_s to be constant with redshift (e.g. Refs. [57, 85, 258, 345, 347]). This description is most accurate for small-mass halos, where much of the evolution of the c–M relation can be attributed to pseudo-evolution [74].

Of course, major mergers interfere with pseudo-evolution by disturbing the density profile. Most low-mass halos (for example, $M_{200c} \lesssim 10^{13} h^{-1} \mathrm{M}_\odot$ at $z = 0$) have not undergone major mergers since before $z = 2$ [178], while high-mass halos underwent major mergers much more recently, suggesting that low-mass and early-forming halos will be best described by pseudo-evolution [336, 344], while the concentrations of more massive and rapidly growing halos are more sensitive to transient effects of mergers. For example, immediately after a merger, the halo will have a large amount of substructure in its outer regions, decreasing the measured concentration. The merging mass will travel toward the center of the halo; as it passes closest to the center, it can temporarily increase the measured concentration [88]. Ref. [190], for example, found that halos less than one crossing time after a merger are likely to have high concentrations.

This picture has been supported by high-resolution simulations of individual halos. For example, simulated mergers show that the scale radius grows proportional to the growth in the virial radius in major mergers, while concentration grows proportional to the scale factor in more quiescent periods without major mergers [158]. Ref. [344] finds that low-mass halos accrete little mass and are more subject to pseudo-evolution than are larger halos. Ref. [70] finds that after the final major merger, the inner regions of the halo are stationary while only

the outer regions accrete a small amount of additional mass. Ref. [312] simulates relatively small haloes ($10^{10} - 10^{12}$), and finds that at redshifts smaller than $z \sim 7$ mass profiles change only at large radii, while after $z \sim 3$ the mass profile is stable and only pseudo-evolution continues to change the virial radius and concentration. The profiles of larger cluster-scale haloes only stabilize after $z \sim 1$ [172, 213].

Tests of pseudo-evolution usually focus either on high-resolution individual halo profiles, or on the evolution of average concentrations (e.g. Ref. [72]). In this work we compare the concentrations predicted by pseudo-evolution to the average mass accretion histories of halos in narrow mass and formation time bins. To isolate the contribution to pseudo-evolution from the evolving background density, we first assume that the entire mass of the halo is a located at $r = 0$. With this assumption, r_{200c} expands solely through the change in ρ_c; as r_{200c} grows, M_{200c} remains fixed at its initial value, because the regions being added to the halo contain no mass. Holding the scale radius constant as the inner profile does not change, concentration evolves according to

$$c(z) = \frac{r_\Delta(z)}{r_\Delta(z_{\rm form})} c(z_{\rm form}) \tag{5.12}$$

where $c(z_{\rm form})$ is measured at $z_{\rm form}$, $r_\Delta(z_{\rm form})$ is calculated from the measured $M_\Delta(z_{\rm form})$, and

$$r_\Delta(z) = \frac{3}{4\pi\Delta} \frac{M_\Delta(z_{\rm form})}{\rho_c(z)}. \tag{5.13}$$

Of course the mass distribution also changes as mass collapses onto the halo. In the absence of violent events, halos accrete mass slowly. The new mass is only at the outer edges of the halo, and does not disturb the inner structure of the halo. With this assumption, concentration still evolves according to equation (5.12), but $r_\Delta(z)$ also evolves according to the growth in $M_\Delta(z)$:

$$r_\Delta(z) = \frac{3}{4\pi\Delta} \frac{M_\Delta(z)}{\rho_c(z)}, \tag{5.14}$$

where $M_\Delta(z)$ is measured for each backbone halo. We note that there are two mechanisms

for $M_\Delta(z)$ to grow: as r_Δ grows, it will enclose more of the halo profile, adding mass to M_Δ whether or not the halo profile is changing; but the halo may also continue to slowly accrete mass in its outer regions.

The assumption of passive accretion is valid only for a subset of halos, those that do not undergo major mergers—for example, halos that are old and yet still small. For these halos, Eq. 5.12 is a good description of concentration evolution. For the individual halos shown in the left column of Fig. 5.7, concentrations predicted from measured mass accretion histories and Eq. 5.12 are comparable to measured concentrations. For halos that rapidly gained mass such as those in the right column of Fig. 5.7, however, half of the final mass is not attained until $z \lesssim 1$, and the halos continue to rapidly grow. Across the entire history, measured concentrations stay around $c \sim 3.5$ while the rapid mass growth would predict much higher concentrations if the mass were accreted passively. Of course, these halos did not gain most of their mass by passive accretion, but instead through major mergers.

To reduce the noise in concentration measurements for individual halos, we stack halo backbones. Halos are binned in $z = 0$ mass and the time at which they first passed a threshold mass of $M_{200c} = 2.3 \times 10^{12} h^{-1} \mathrm{M}_\odot$, corresponding to 2000 particles in the SOD halo. At each timestep, the mass and concentration of the stack are set to their average values in the bin. Halos that are small at $z = 0$, like those shown in the left column of Fig. 5.8, are well described by passive accretion—so long as they grew only slowly. The youngest small halos, like the more massive halos of the right column of Fig. 5.8, grow rapidly and their concentrations remain low.

Figure 5.9 shows that the concentration at formation, $c(z_M(0.5))$, averages approximately 3.5 across all redshifts. The increase in formation concentration for $z_M(0.5) \lesssim 1$ may be related to the transition to a dark-energy-dominated regime. In future work we will test average formation concentration as a function of formation time using the Mira-Titan suite of simulations, in which the dark energy equation of state is modeled as $w(z) = w_0 + w_a[1-a(t)]$.

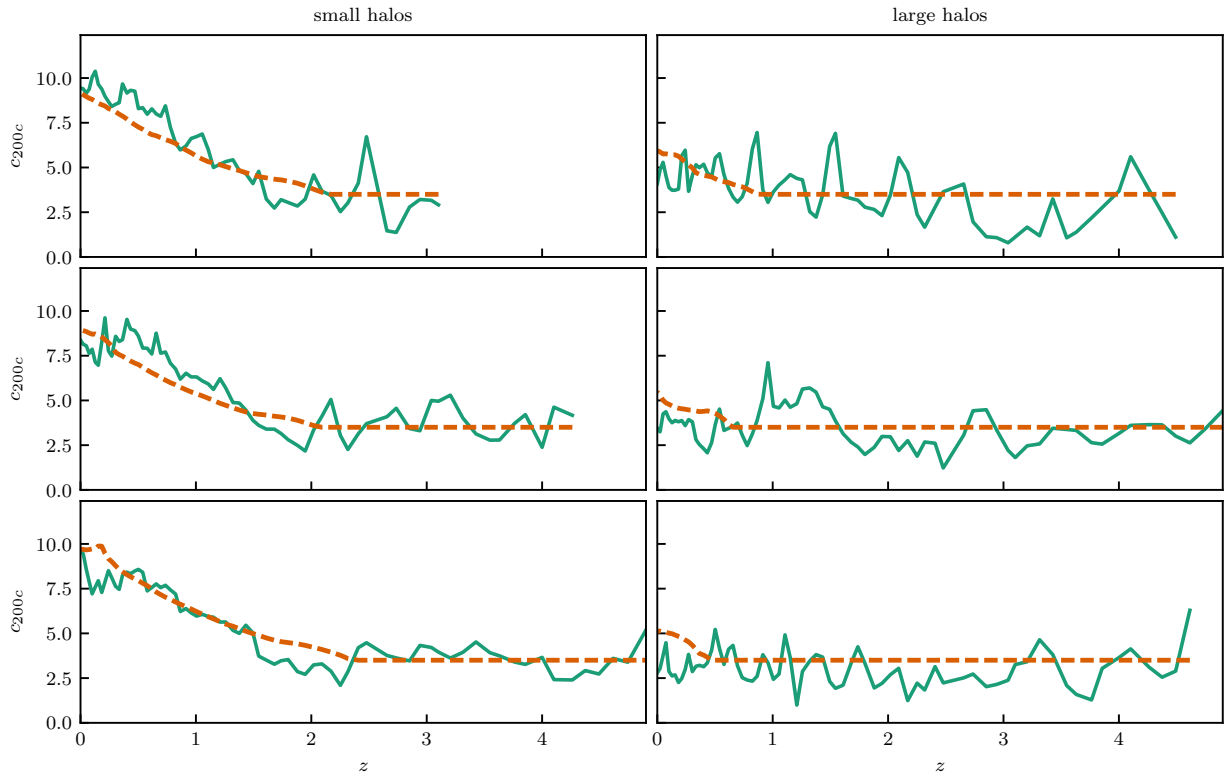

Figure 5.7: For a selection of halos with small final mass (left column), pseudo-evolution (dashed orange curves) predicts the evolution of concentration (solid green curves) from the half-mass formation time $z_M(0.5)$. For larger halos (right column), the half-mass formation time is much more recent. These halos accrete mass rapidly, and in the short time since $z_M(0.5)$, pseudo-evolution predicts rising concentration while measured concentrations remain low.

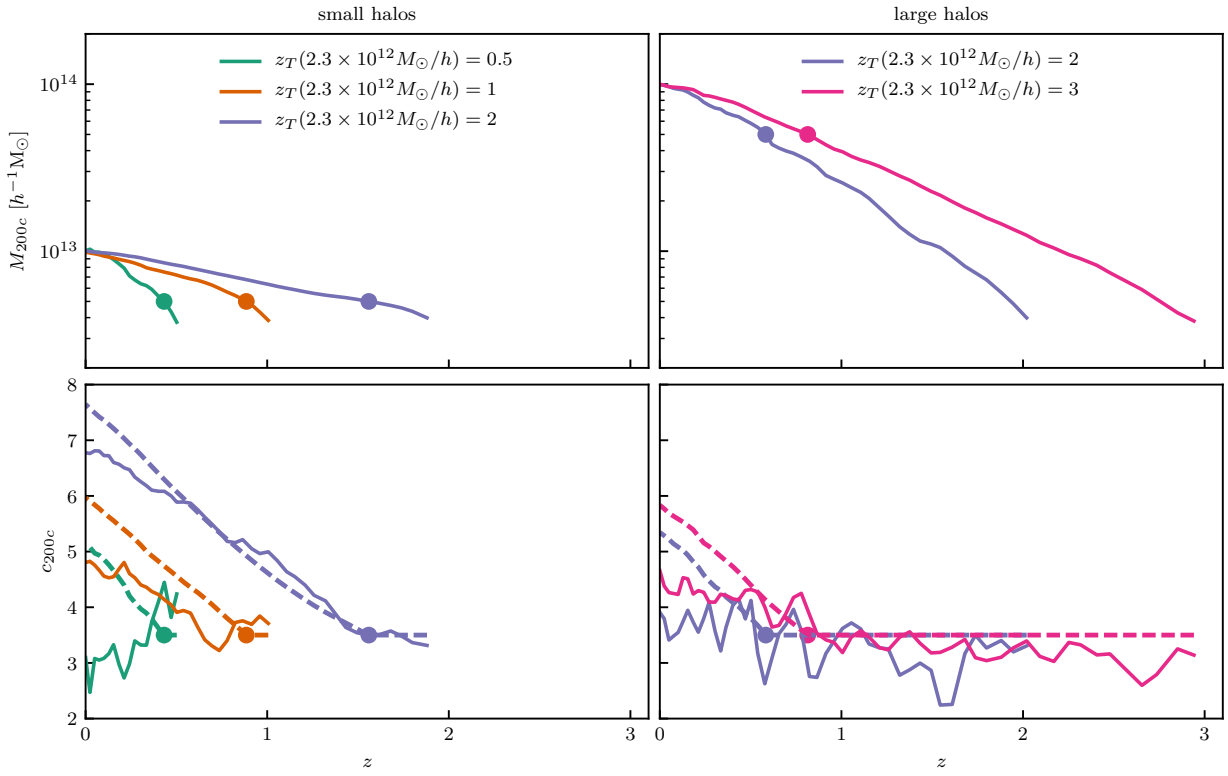

Figure 5.8: Stacked mass accretion (top row) and concentration (bottom row, solid curves) histories for halos binned by final mass and the time when they first attained a mass of $2.3 \times 10^{12} h^{-1}\mathrm{M}_\odot$. The concentrations of small, slow-growing halos (e.g., purple curves in the left column) match the prediction of pseudo-evolution (dashed curves) reasonably well. In contrast, for larger halos and halos that accreted their mass over a shorter period of time, concentrations remain constant around $c \sim 3-4$ while the concentrations predicted by pseudo-evolution grow rapidly. Pseudo-evolution concentrations are calculated based on the half-mass formation time $z_M(0.5)$, shown as points; before $z_M(0.5)$, $c = 3.5$, while after $z_M(0.5)$ concentration obeys equations (5.12) and (5.14).

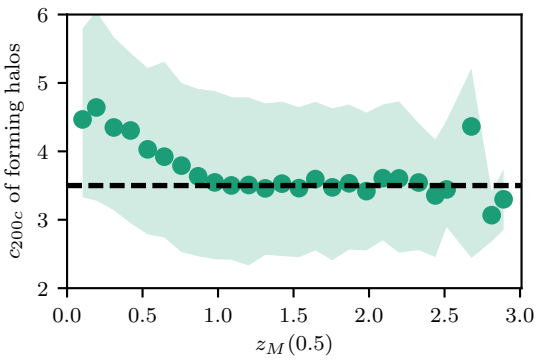

Figure 5.9: At formation redshifts greater than $z_M(0.5) \sim 1$, the median concentration of progenitors accreting half of the final halo mass (points) is roughly 3.5 (dashed line). Shaded region shows the quartiles of the concentration distribution. In contrast, the youngest halos formed with slightly higher concentrations.

5.6 Conclusion

We have tested all definitions of halo formation time commonly used in the literature. We find that the most commonly used definition—$z_M(0.5)$, which considers a halo to form when its main progenitor first reaches half the final halo mass—is well correlated with final concentration, but the correlation coefficient is highest for definitions of formation time that account for the mass of all progenitors. Formation times based on the time when the sum of all progenitor masses reach the mass contained within the $z = 0$ scale radius are best correlated with concentration and assign the highest formation times to the smallest, most highly concentrated halos.

We find that an assumption of pseudo-evolution provides an upper bound for concentration evolution; for halos that slowly accrete mass, the inner profile changes little. As the halo evolves, the background density ρ_c used to define masses and concentrations changes, and mass is accreted in the outer regions of the halo, but the radius r_s at which the effective slope of the profile $n_{\rm eff} = -2$ does not change except in violent events like major mergers. In contrast, halos that undergo frequent major mergers and only recently reached half of their final mass do not have much time to pseudo-evolve. Their concentration changes with time, but never departs far from the floor of $c \sim 3.5$.

In future work, we will analyze the Mira-Titan Universe suite of simulations to test the effects of dynamical dark energy on concentrations and their correlation with formation time and combine multiple definitions of formation time. Formation times are also related to subhalo abundance and the fraction of mass in substructure (e.g. Refs. [110, 146]), sphericity and velocity dispersion (e.g. Ref. [249]), and dynamical state (e.g. Ref. [244]); we can compare the relative importance of different stages of mass accretion histories to different halo properties.

CHAPTER 6

CONCLUSION

We have considered two topics in the formation of large scale structure: in the mildly nonlinear regime, we used bispectrum measurements to improve constraints on the baryon acoustic oscillation (BAO) length scale, while in the deeply nonlinear regime, we presented improved measurements of average halo profiles and the evolution of internal halo structure. In combination with future survey data, these results can be used to better constrain cosmological parameters. In the remainder of the Conclusion, we outline our primary results and directions for future work in both topics.

6.1 Primary Results

6.1.1 Bispectrum BAO

In Chapter 2 we presented a new technique to select bispectrum configurations that are most sensitive to BAO. In these "constructive" configurations, two sides of the triangle differ by a multiple of the fundamental BAO wavelength. The oscillations in a pair of power spectra entering the bispectrum then interfere constructively. Our parameterization enables us to visualize BAO in the bispectrum as a function of a single wavenumber. We first calculate the root-mean-square (RMS) amplitude of BAO in the perturbation theory bispectrum for a selection of triangle configurations and plot the result (Figure 2.1). We then use this RMS map to select a set of ten configurations (seven constructive and three destructive) which we measure in simulations in order to estimate their ability to constrain the BAO distance scale. Fisher analysis shows that the full set of ten configurations can improve BAO constraints by about 15% over the unreconstructed power spectrum alone; in comparison, reconstruction improves BAO constraints by about 30%. As we use relatively few bispectrum measurements, our covariance matrices are well estimated and are not a limiting factor in our constraints. As expected, we find that constructive configurations offer more constraining

power: three constructive configurations improve BAO constraints by about 8%, while three destructive configurations offer only a less than 3% improvement. This is a primary message of Chapter 2: the improvement in BAO constraints offered by bispectrum measurements depends not only on the number of measurements, but also on the choice of measurements. In other words, not all triangles are equally useful.

In Chapter 3 we presented a full physical explanation of the RMS map, as full understanding of the RMS map will be critical for further applications of the interferometric basis. We show that for many triangle configurations, the full bispectrum is dominated by only a single term in the cyclic sum, and for many more configurations, a pair of terms dominates the bispectrum while the third term is negligible in comparison. This dominance structure is largely driven by the $F^{(2)}$ kernel of perturbation theory, rather than by the power spectrum that also enters the bispectrum. We show analytically that in regions where one or more terms in the bispectrum are negligible, the RMS amplitude of the BAO feature is well-approximated by the RMS amplitude of the dominant term or terms alone. We can therefore break the full RMS map into regions where it can be understood through the RMS map corresponding to the dominant term or terms. When these maps are stitched together in the regions where the corresponding terms dominate, we reproduce the full RMS map. In each individual term, features in the RMS map are driven by one of four mechanisms. One mechanism, interference, provides the strongest amplification of BAO: when two oscillations are shifted relative to each other by an integer multiple of the BAO fundamental wavelength, they interfere constructively to enhance the BAO signal. In many triangle configurations, however, the wavelengths of the two oscillations are too different to produce interference. Instead, other mechanisms (which we term "incoherence," "feathering," and "single power spectrum") drive the patterns in the RMS map. In these configurations, the interactions between the two power spectra cannot consistently amplify or suppress BAO. Overall, Chapter 3 develops our understanding of the physics behind the RMS map and will enable further use of our interferometric basis to address the topics outlined below in §6.2.1.

6.1.2 Halo Profiles and Assembly

In Chapter 4 we measured the concentration-mass (c–M) relation for halos in simulations of a ΛCDM universe with extremely powerful statistics and high mass resolution. We employ three different methods of concentration measurement—one classic fit to the halo profile, and two new methods sensitive to different features of the profile—to obtain robust results. Additionally, we perform numerous tests of our simulations to identify possible sources of error. We find that at low redshifts ($z \lesssim 3$), the three methods are in good agreement, indicating that the NFW profile is a good description of halo profiles. At all masses the average concentration of relaxed halos is higher than that of unrelaxed halos, reflecting the significant substructure found at relatively large distances from the centers of unrelaxed halos. We also measure concentrations of stacks of relaxed halos, showing that they are better described by the three-parameter Einasto profile than by the less-flexible NFW profile and enabling comparison to multiple measurements of observed halo concentrations. Our results are generally in agreement with existing measurements as well as other simulation results, and we anticipate that future surveys will reduce the measurement uncertainties and offer more stringent tests of cosmological parameters. At all masses and redshifts, the distribution of concentrations within a mass bin is well described by a normal distribution with standard deviation equal to one-third of the mean concentration. When masses are scaled by the redshift-dependent nonlinear mass scale $M_\star$, the c–M relation is independent of redshift across eight decades in $M/M_\star$. Concentrations fall with mass at low $M/M_\star$, while at high $M/M_\star$, concentration is constant around $c \sim 3.5$. We provide a fitting formula for concentration as a function of $M/M_\star$ and redshift, calibrated using our robust concentration measurements, which can be applied in preparation for and analysis of future surveys.

In Chapter 5 we combined the measurements from Chapter 4 with the formation histories of halos to study the relationship between halo assembly history and structure. We collect a variety of formation time definitions commonly used in the literature and calculate their correlations with the concentration of the halo at $z = 0$. We find that while all the definitions

of formation time do correlate with concentration, the best definitions are the half-mass formation time (when a halo first attained half of its final mass, the most commonly used definition of formation time), a fit to the full assembly history, or ones based on the total mass of all progenitors. Other mass fractions do not perform as well as the half-mass formation time, likely because much smaller fractions are attained very early in the halo's assembly history, while much larger fractions were reached only recently after the majority of the halo profile is already in place. Next, we analyze stacks of mass accretion histories for sets of halos with similar formation time and final mass. We show that for slowly growing halos, an assumption of pseudoevolution (that is, halos accrete mass only in their outer regions without changes to the inner profile) is able to predict concentration evolution for slowly-growing halos. In contrast, young, rapidly-growing halos have undergone major mergers that disrupted the inner profile. For these mass accretion histories, pseudoevolution predicts extremely high concentrations corresponding to the rapid growth in mass. In fact, however, concentration remains low, around the floor of $c \sim 3.5$.

6.2 Future Work

6.2.1 Bispectrum BAO

Our approach to measuring bispectrum triangles can be applied in four primary areas: understanding the covariance structure of the bispectrum, understanding reconstruction, measuring the BAO scale in survey data from BOSS or DESI, and constraining sources of a phase shift in the power spectrum or bispectrum. Finally, with further understanding of bispectrum covariance and reconstruction, our method for triangle selection may be improved.

Bispectrum Covariance

Our method opens a new avenue for numerically obtaining the cross covariance between the power spectrum and bispectrum, which is less easily treated with an analytic template than

is the auto covariance. At the level of a simply Gaussian random field analytic model, both the bispectrum and its cross-covariance with the power spectrum identically vanish as they are odd correlators, but this is clearly not the case in reality.

Measuring the full covariance is challenging for the same reason that we wish to restrict to a smaller number of bispectrum measurements: the number of fully N-body mocks available limits the number of bispectrum triangles that can be measured. A restricted set of triangles will enable further pursuit of this question, and there may be further opportunities to choose tailored triangles for BAO.

Reconstruction

Reconstruction [84, 226, 235, 236] has been shown to bring information from higher-order statistics into the 2PCF and power spectrum [270], which would alter the covariance between the power spectrum and bispectrum. But reconstruction is a numerical procedure, so its effect on both the bispectrum and the covariance of the bispectrum with the power spectrum is difficult to model analytically.

According to theory and our simulation results, reconstruction dramatically alters the broadband shape of the bispectrum, decreasing its amplitude and causing it to change sign at some scales for certain choices of smoothing scale. These effects are predicted by perturbation theory [134], but while we measure the post-reconstruction bispectrum in simulations, the currently available theoretical models are not in sufficiently good agreement with the broadband bispectrum to isolate a BAO feature. That is, the perturbation theory model of the post-reconstruction bispectrum differs from measurements too much to offer a post-reconstruction analog of the no-wiggle bispectrum. In order to identify a BAO feature in the post-reconstruction bispectrum, further theoretical work must be done either to improve the perturbation theory predictions or to develop a different no-wiggle template.

Like the bispectrum measurements discussed in the previous section, a full numerical study of the covariance between the post-reconstruction power spectrum and the pre-

reconstruction bispectrum is limited by the number of fully N-body mocks available. Fewer mocks are needed if analysis is restricted to the set of triangles most sensitive to BAO, reducing the dimension of the covariance matrix. We can study the effects of reconstruction on these triangle configurations, both to better understand which triangle configurations most contribute to reconstruction and to combine bispectrum measurements with the post-reconstruction power spectrum. Depending on the level of independence, the combination of bispectrum measurements and reconstruction may offer further improvement in BAO constraints over that offered by reconstruction alone.

Application to Data

Our Fisher forecast predicts that a small number of bispectrum measurements offer further constraints on the BAO length scale. The method should be applied to survey data such as that of BOSS or DESI to serve as a cross-check for earlier BAO constraints. Even if the bispectrum does not offer much information independent from the reconstructed power spectrum, it is an operationally independent method to obtain the same information. Unlike reconstruction, bispectrum constraints on the BAO scale involve no backward modeling. Though erroneous assumptions like incorrect bias, distance, and RSD parameters have recently been found to have little effect on the location of the reconstructed BAO peak [282], they can change the shape of the peak. As bispectrum measurements do not require these assumptions, they provide an additional valuable cross-check on reconstruction. Added to analyses like the final BOSS results of Ref. [2] (e.g. their Figure 13), where BAO measurements from many different probes are compared, a bispectrum measurement of the BAO distance scale will further confirm the consistency and robustness of BAO measurements across a number of operationally independent methods.

Phase Shifts

Our method identifies constructive and destructive configurations by enforcing a phase shift between the oscillating power spectra as a function of two triangle sides. Our method may provide opportunities to constrain several physical processes that introduce an additional phase shift to the power spectrum or bispectrum.

In one example, an isocurvature perturbation introduces an oscillation proportional to $\cos k\tilde{s}$ rather than $\sin k\tilde{s}$ [140]. We expect that in the presence of isocurvature, destructive configurations may once again show a BAO feature. A preliminary investigation has shown that within the bounds on isocurvature set by Planck [1], the presence of an isocurvature perturbation suppresses BAO for all triangle configurations; for no configuration is the BAO feature more pronounced in the mixed model than in the adiabatic case. While the oscillatory feature does not seem to reappear in destructive configurations in the presence of isocurvature, the bispectrum may yet provide constraints on isocurvature through, for example, the overall suppression of bispectrum BAO.

In another example, the phase of BAO in the power spectrum is sensitive to $N_{\rm eff}$, the effective number of relativistic neutrino species [17, 19, 20, 90]. Ref. [20] showed that this signature is not altered by nonlinear evolution in the time between the early Universe and observation, and Ref. [19] found in the power spectrum evidence for a phase shift due to the cosmic neutrino background. If the phase shift were simply a constant shift, our method would not be useful, as our method only probes the phase differences between two power spectra. But instead the phase shift is chromatic, increasing with wavenumber for $k \lesssim 0.2\ h/\mathrm{Mpc}$ (as shown in Figure 1 of Ref. [19]), so changes in the effective number of relativistic neutrino species do change the RMS map. Developing constraints on $N_{\rm eff}$ requires two steps: understanding the behavior of the RMS map under changes in $N_{\rm eff}$ in our current basis and investigating whether a redefinition of our basis (for example, a change in δ) would be more appropriate.

Relative velocities between baryons and dark matter [31, 61, 268, 311, 340] may also

introduce a phase shift in power spectrum BAO. While these have already been constrained in the power spectrum by Refs. [341] and [28] and in the 3PCF by Ref. [294], the interferometric basis may offer an additional or complementary constraint.

If massive spinning particles are present during inflation, they may introduce a phase directly in the bispectrum (e.g., Equation (2.17) of Ref. [210]). This template for the bispectrum includes a term proportional to $\cos(\ln k_i/k_j)$, a function of the ratio of two triangle sides. An analog of the interferometric basis may be developed for this application and used to constrain the presence of massive spinning particles.

Improving the Interferometric Method

With better understanding of the covariance structure of the bispectrum and the triangles, we can improve our method for triangle selection.

Any study of bispectrum BAO will be limited by the number of mock catalogs available. Given this limit, our method offers one way to efficiently select triangles that will offer the strongest BAO constraints; that is, our method identifies the triangles that are most sensitive to BAO. However, the contribution of each measurement to BAO constraints depends not only on the amplitude of the BAO signal, but also on the signal-to-noise ratio and the covariance of the new measurement both with prior measurements and with the power spectrum (pre- or post-reconstruction). Clearly one way to select the optimal triangle measurements is to perform all the measurements and identify the most important triangles, but short of this, we wish to develop a prescription for selecting triangles.

We expect a point of diminishing returns, where further measurements do little to improve the constraints. The last configuration in our set of ten added 20% to the constraints, which remains substantial, but several prior configurations added relatively little due to their high covariance with configurations that had already been included. Additionally, while the amplitude of the BAO signal is smaller in destructive configurations, these configurations do still carry information about the BAO scale. After some number of constructive configuration

measurements, destructive configurations may offer more improvement in constraints than yet another constructive configuration.

6.2.2 Halo Profiles and Assembly

While haloes are generally understood to assemble into an NFW profile through accretion and mergers, the details remain elusive. The simulations analyzed in Chapters 4 and 5 can be used to further develop our understanding of the role of assembly history in halo structure.

Our work has analyzed many commonly-used definitions of formation time, but many of them are calculated from only the backbone halo. The one definition that does consider the less-massive progenitors ($z_{\rm NFW}$) is better correlated with concentration than any backbone-only definition. Two further questions are immediately apparent: does the assembly history of smaller merging halos affect the final halo structure? And what is different about the structure of halos that merge into a larger halo and so are never analyzed as final halos? We can develop further measures of assembly history that consider the history of less-massive progenitors and study both the backbones and less-massive progenitors of those halos that do not survive to $z = 0$.

Our work has treated all formation time definitions independently, but clearly no single definition carries full information about the halo assembly history. Many definitions consider only a single time (e.g., the time when the halo first crossed a certain threshold in mass). Others, like the fit definitions, do include information from the full history, but the full mass accretion history must be smoothed to perform the fit. The definitions which include less-massive branches of the merger tree only consider the sum of progenitor masses. We can study combinations of different measures of accretion history to identify, for example, which are most orthogonal and which combinations are most correlated with final halo concentration.

Major merger events reset concentrations to the floor level of $c \sim 3 - 4$. Simulations [190] support the picture that concentrations decrease immediately after a merger (when the

recently-merged halo adds substructure to the outer regions), then peak when the new matter passes through the center of the profile, and finally relax to equilibrium. The DeltaQuadrant simulation, which has a time resolution almost a factor of eight better than that of Q Continuum and Outer Rim, will allow us to study the effects of mergers on concentration in further detail. For example, for rapidly-growing halos whose concentration remains low, is the halo profile ever allowed to fully virialize? If not, the constant low concentration may be a result of frequent mergers that never allow the substructure in the halo outskirts to fully relax.

Mergers are major events that clearly disturb halo structure and are easy to identify, but they are not the only interactions with surrounding halos that can affect halo evolution. For example, a halo may have lost mass when it passed near another halo ("flybys" [287]) or through another halo before being ejected [300]. We can identify halos that have undergone these processes and evaluate their effects on final mass and concentration.

At all masses and redshifts, halo concentration is normally distributed about the mean concentration with standard deviation one-third of the mean concentration—that is, $\sigma_c/c \sim 1/3$. This relation remains difficult to explain, but mass accretion histories may offer some insight. The Q Continuum and Outer Rim simulations provide sufficient statistics to break halos into narrow bins in both mass and final concentration. As a first investigation, we have shown that high-concentration halos are older than low-concentration halos in the same mass bin, but a more detailed comparison of mass accretion histories (in bins of mass, concentration, and age) may show which differences are most important.

Our analysis of merger trees has focused only on concentration as a measure of the interior structure of a halo. We have shown the Einasto profile to be a better description of stacked halo profiles, and the evolution of the Einasto shape parameter also carries information about the evolution of the halo (e.g. Ref. [185]). Correlations between formation times and Einasto parameters may reveal which stages of halo evolution are most influential in determining the shape of the final halo profile.

This work has assumed a ΛCDM cosmology with WMAP-7 parameters. The Mira-Titan suite of simulations model the evolution of structure in dark energy models with a non-constant equation of state where $w(a) = w_0 + w_a(1 - a)$. The halo catalogs for these simulations are complete with concentrations. Both the c–M relation and relationship between formation time and final concentration may depend on the time when dark energy first becomes dominant, so the analysis of this work should be repeated for these other simulations.

REFERENCES

[1] P. A. R. Ade et al. Planck 2015 results. XX. Constraints on inflation. A&A, 594:A20, 2016.

[2] S. Alam et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. MNRAS, 470(3):2617–2652, 2017.

[3] B. Allgood, R. A. Flores, J. R. Primack, A. V. Kravtsov, R. H. Wechsler, A. Faltenbacher, and J. S. Bullock. The shape of dark matter haloes: dependence on mass, redshift, radius and formation. MNRAS, 367:1781–1796, April 2006.

[4] S. Amodeo, S. Ettori, R. Capasso, and M. Sereno. The relation between mass and concentration in X-ray galaxy clusters at high redshift. A&A, 590:A126, May 2016.

[5] P. W. Angel, G. B. Poole, A. D. Ludlow, A. R. Duffy, P. M. Geil, S. J. Mutch, A. Mesinger, and J. S. B. Wyithe. Dark-ages reionization and galaxy formation simulation - II. Spin and concentration parameters for dark matter haloes during the epoch of reionization. MNRAS, 459:2106–2117, June 2016.

[6] R. E. Angulo, C. M. Baugh, and C. G. Lacey. The assembly bias of dark matter haloes to higher orders. MNRAS, 387:921–932, June 2008.

[7] R. E. Angulo, S. Foreman, M. Schmittfull, and L. Senatore. The One-Loop Matter Bispectrum in the Effective Field Theory of Large Scale Structures. JCAP, 1510(10):039, 2015.

[8] Y. Ascasibar, G. Yepes, S. Gottlöber, and V. Müller. On the physical origin of dark matter density profiles. MNRAS, 352:1109–1120, August 2004.

[9] E. Audit, R. Teyssier, and J. M. Alimi. Non-linear dynamics and mass function of cosmic structures. I. Analytical results. A&A, 325:439–449, September 1997.

[10] E. Audit, R. Teyssier, and J.-M. Alimi. Non-linear dynamics and mass function of cosmic structures. II. Numerical results. A&A, 333:779–789, May 1998.

[11] V. Avila-Reese, P. Colín, S. Gottlöber, C. Firmani, and C. Maulbetsch. The Dependence on Environment of Cold Dark Matter Halo Properties. ApJ, 634:51–69, November 2005.

[12] V. Avila-Reese, P. Colín, G. Piccinelli, and C. Firmani. The Effects of Non-Gaussian Initial Conditions on the Structure and Substructure of Cold Dark Matter Halos. ApJ, 598:36–48, November 2003.

[13] V. Avila-Reese, C. Firmani, and X. Hernández. On the Formation and Evolution of Disk Galaxies: Cosmological Initial Conditions and the Gravitational Collapse. ApJ, 505:37–49, September 1998.

[14] V. Avila-Reese, C. Firmani, A. Klypin, and A. V. Kravtsov. Density profiles of dark matter haloes: diversity and dependence on environment. MNRAS, 310:527–539, December 1999.

[15] T. Baldauf, L. Mercolli, M. Mirbabayi, and E. Pajer. The bispectrum in the Effective Field Theory of Large Scale Structure. JCAP, 5:007, May 2015.

[16] T. Baldauf, U. Seljak, V. Desjacques, and P. McDonald. Evidence for quadratic tidal tensor bias from the halo bispectrum. Phys. Rev. D, 86(8):083540, October 2012.

[17] S. Bashinsky and U. Seljak. Neutrino perturbations in CMB anisotropy and matter clustering. Phys. Rev. D, 69:083002, 2004.

[18] C. M. Baugh. A primer on hierarchical galaxy formation: the semi-analytical approach. *Reports on Progress in Physics*, 69:3101–3156, December 2006.

[19] D. Baumann, F. Beutler, R. Flauger, D. Green, M. Vargas-Magaña, A. Slosar, B. Wallisch, and C. Yèche. First Measurement of Neutrinos in the BAO Spectrum. *arXiv e-prints*, March 2018.

[20] D. Baumann, D. Green, and M. Zaldarriaga. Phases of New Physics in the BAO Spectrum. JCAP, 11:007, November 2017.

[21] A. J. Benson. Galaxy formation theory. Phys. Rep., 495:33–86, October 2010.

[22] A. J. Benson, R. G. Bower, C. S. Frenk, C. G. Lacey, C. M. Baugh, and S. Cole. What Shapes the Luminosity Function of Galaxies? ApJ, 599:38–49, December 2003.

[23] A. J. Benson, S. Cole, C. S. Frenk, C. M. Baugh, and C. G. Lacey. The nature of galaxy bias and clustering. MNRAS, 311:793–808, February 2000.

[24] A. J. Benson, A. Ludlow, and S. Cole. Halo concentrations from extended Press-Schechter merger histories. *arXiv e-prints*, page arXiv:1812.06026, December 2018.

[25] A. A. Berlind and D. H. Weinberg. The Halo Occupation Distribution: Toward an Empirical Determination of the Relation between Galaxies and Mass. ApJ, 575:587–616, August 2002.

[26] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro. Large-scale structure of the Universe and cosmological perturbation theory. Phys. Rep., 367:1–248, September 2002.

[27] D. Bertolini and M. P. Solon. Principal Shapes and Squeezed Limits in the Effective Field Theory of Large Scale Structure. JCAP, 1611(11):030, 2016.

[28] F. Beutler, U. Seljak, and Z. Vlah. Constraining the relative velocity effect using the Baryon Oscillation Spectroscopic Survey. MNRAS, 470:2723–2735, September 2017.

[29] S. Bhattacharya, S. Habib, K. Heitmann, and A. Vikhlinin. Dark Matter Halo Profiles of Massive Clusters: Theory versus Observations. ApJ, 766:32, March 2013.

[30] C. Blake and K. Glazebrook. Probing Dark Energy Using Baryonic Oscillations in the Galaxy Power Spectrum as a Cosmological Ruler. ApJ, 594:665–673, September 2003.

[31] J. A. Blazek, J. E. McEwen, and C. M. Hirata. Streaming Velocities and the Baryon Acoustic Oscillation Scale. Phys. Rev. Lett., 116(12):121303, March 2016.

[32] S. Bose, W. A. Hellwing, C. S. Frenk, A. Jenkins, M. R. Lovell, J. C. Helly, and B. Li. The Copernicus Complexio: statistical properties of warm dark matter haloes. MNRAS, 455:318–333, January 2016.

[33] M. Boylan-Kolchin and C.-P. Ma. Major mergers of galaxy haloes: cuspy or cored inner density profile? MNRAS, 349:1117–1129, April 2004.

[34] M. Boylan-Kolchin, V. Springel, S. D. M. White, A. Jenkins, and G. Lemson. Resolving cosmic structure formation with the Millennium-II Simulation. MNRAS, 398:1150–1164, September 2009.

[35] A. D. Bray et al. Modelling galactic conformity with the colourhalo age relation in the Illustris simulation. MNRAS, 455(1):185–198, 2016.

[36] F. Brimioulle, S. Seitz, M. Lerchster, R. Bender, and J. Snigula. Dark matter halo properties from galaxy-galaxy lensing. MNRAS, 432:1046–1102, June 2013.

[37] T. Broadhurst, K. Umetsu, E. Medezinski, M. Oguri, and Y. Rephaeli. Comparison of Cluster Lensing Profiles with ΛCDM Predictions. ApJ, 685:L9, September 2008.

[38] J. S. Bullock, T. S. Kolatt, Y. Sigad, R. S. Somerville, A. V. Kravtsov, A. A. Klypin, J. R. Primack, and A. Dekel. Profiles of dark haloes: evolution, scatter and environment. MNRAS, 321:559–575, March 2001.

[39] M. T. Busha, F. C. Adams, R. H. Wechsler, and A. E. Evrard. Future Evolution of Cosmic Structure in an Accelerating Universe. ApJ, 596:713–724, October 2003.

[40] J. J. M. Carrasco, M. P. Hertzberg, and L. Senatore. The Effective Field Theory of Cosmological Large Scale Structures. *JHEP*, 09:082, 2012.

[41] E. Castorina and M. White. Beyond the plane-parallel approximation for redshift surveys. MNRAS, 476:4403–4417, June 2018.

[42] C.-T. Chiang, A. M. Cieplak, F. Schmidt, and A. Slosar. Response approach to the squeezed-limit bispectrum: application to the correlation of quasar and Lyman-α forest power spectrum. JCAP, 6:022, June 2017.

[43] H. L. Child, S. Habib, K. Heitmann, N. Frontiere, H. Finkel, A. Pope, and V. Morozov. Halo Profiles and the Concentration-Mass Relation for a ΛCDM Universe. ApJ, 859:55, May 2018.

[44] H. L. Child, Z. Slepian, and M. Takada. A Physical Picture of Bispectrum Baryon Acoustic Oscillations in the Interferometric Basis. *ArXiv e-prints*, page arXiv:1811.12396, November 2018.

[45] H. L. Child, M. Takada, T. Nishimichi, T. Sunayama, Z. Slepian, S. Habib, and K. Heitmann. Bispectrum as baryon acoustic oscillation interferometer. Phys. Rev. D, 98:123521, Dec 2018.

[46] N. E. Chisari, M. L. A. Richardson, J. Devriendt, Y. Dubois, A. Schneider, A. M. C. Le Brun, R. S. Beckmann, S. Peirani, A. Slyz, and C. Pichon. The impact of baryons on the matter power spectrum from the Horizon-AGN cosmological hydrodynamical simulation. MNRAS, 480:3962–3977, November 2018.

[47] C. Y. R. Chue, N. Dalal, and M. White. Some assembly required: assembly bias in massive dark matter halos. JCAP, 1810(10):012, October 2018.

[48] N. Cibirka, E. S. Cypriano, F. Brimioulle, D. Gruen, T. Erben, L. van Waerbeke, L. Miller, A. Finoguenov, C. Kirkpatrick, J. P. Henry, E. Rykoff, E. Rozo, R. Dupke, J.-P. Kneib, H. Shan, and P. Spinelli. CODEX weak lensing: concentration of galaxy clusters at $z \sim 0.5$. MNRAS, 468:1092–1116, June 2017.

[49] J. D. Cohn and M. J. White. The Formation histories of galaxy clusters. *Astropart. Phys.*, 24:316–333, 2005.

[50] S. Cole, A. Aragon-Salamanca, C. S. Frenk, J. F. Navarro, and S. E. Zepf. A Recipe for Galaxy Formation. MNRAS, 271:781, December 1994.

[51] S. Cole, J. Helly, C. S. Frenk, and H. Parkinson. The statistical properties of Λ cold dark matter halo formation. MNRAS, 383:546–556, January 2008.

[52] S. Cole and C. G. Lacey. The Structure of dark matter halos in hierarchical clustering models. MNRAS, 281:716, 1996.

[53] P. Colín, A. Klypin, O. Valenzuela, and S. Gottlöber. Dwarf Dark Matter Halos. ApJ, 612:50–57, September 2004.

[54] J. M. Comerford and P. Natarajan. The observed concentration-mass relation for galaxy clusters. MNRAS, 379:190–200, July 2007.

[55] C. Conroy, R. H. Wechsler, and A. V. Kravtsov. Modeling Luminosity-dependent Galaxy Clustering through Cosmic Time. ApJ, 647:201–214, August 2006.

[56] C. A. Correa, J. S. B. Wyithe, J. Schaye, and A. R. Duffy. The accretion history of dark matter haloes - II. The connections with the mass power spectrum and the density profile. MNRAS, 450(2):1521–1537, 2015.

[57] C. A. Correa, J. S. B. Wyithe, J. Schaye, and A. R. Duffy. The accretion history of dark matter haloes - III. A physical model for the concentration-mass relation. MNRAS, 452:1217–1232, September 2015.

[58] D. J. Croton, L. Gao, and S. D. M. White. Halo assembly bias and its effects on galaxy clustering. MNRAS, 374:1303–1309, 2007.

[59] A. J. Cuesta, F. Prada, A. Klypin, and M. Moles. The virialized mass of dark matter haloes. MNRAS, 389:385–397, September 2008.

[60] N. Dalal, Y. Lithwick, and M. Kuhlen. The Origin of Dark Matter Halo Profiles. *ArXiv e-prints*, October 2010.

[61] N. Dalal, U.-L. Pen, and U. Seljak. Large-scale BAO signatures of the smallest galaxies. JCAP, 11:007, November 2010.

[62] K. S. Dawson et al. The Baryon Oscillation Spectroscopic Survey of SDSS-III. AJ, 145:10, January 2013.

[63] R. de Belsunce and L. Senatore. Tree-Level Bispectrum in the Effective Field Theory of Large-Scale Structure extended to Massive Neutrinos. *ArXiv e-prints*, April 2018.

[64] G. De Lucia, G. Kauffmann, V. Springel, S. D. M. White, B. Lanzoni, F. Stoehr, G. Tormen, and N. Yoshida. Substructures in cold dark matter haloes. MNRAS, 348:333–344, February 2004.

[65] A. J. Deason, C. Conroy, A. R. Wetzel, and J. L. Tinker. Stellar mass-gap as a probe of halo assembly history and concentration: youth hidden among old fossils. ApJ, 777:154, 2013.

[66] A. Dekel, J. Devor, and G. Hetzroni. Galactic halo cusp-core: tidal compression in mergers. MNRAS, 341:326–342, May 2003.

[67] DESI Collaboration. The DESI Experiment Part I: Science,Targeting, and Survey Design. *ArXiv e-prints*, page arXiv:1611.00036, October 2016.

[68] V. Desjacques. Environmental dependence in the ellipsoidal collapse model. MNRAS, 388:638, 2008.

[69] V. Desjacques, D. Jeong, and F. Schmidt. The Galaxy Power Spectrum and Bispectrum in Redshift Space. *ArXiv e-prints*, June 2018.

[70] J. Diemand, M. Kuhlen, and P. Madau. Formation and evolution of galaxy dark matter halos and their substructure. ApJ, 667:859–877, 2007.

[71] J. Diemand, P. Madau, and B. Moore. The distribution and kinematics of early high-σ peaks in present-day haloes: implications for rare objects and old stellar populations. MNRAS, 364:367–383, December 2005.

[72] B. Diemer and M. Joyce. An accurate physical model for halo concentrations. *arXiv e-prints*, page arXiv:1809.07326, September 2018.

[73] B. Diemer and A. V. Kravtsov. A Universal Model for Halo Concentrations. ApJ, 799:108, January 2015.

[74] B. Diemer, S. More, and A. V. Kravtsov. The Pseudo-evolution of Halo Mass. ApJ, 766:25, March 2013.

[75] K. Dolag, M. Bartelmann, F. Perrotta, C. Baccigalupi, L. Moscardini, M. Meneghetti, and G. Tormen. Numerical study of halo concentrations in dark-energy cosmologies. A&A, 416:853–864, March 2004.

[76] G. A. Dooley, B. F. Griffen, P. Zukin, A. P. Ji, M. Vogelsberger, L. E. Hernquist, and A. Frebel. The Effects of Varying Cosmological Parameters on Halo Substructure. ApJ, 786:50, 2014.

[77] N. E. Drakos, J. E. Taylor, A. Berrouet, A. S. G. Robotham, and C. Power. Major Mergers between Dark Matter Haloes – I: Predictions for Size, Shape and Spin. *arXiv e-prints*, page arXiv:1811.12839, November 2018.

[78] A. R. Duffy, J. Schaye, S. T. Kay, and C. Dalla Vecchia. Dark matter halo concentrations in the Wilkinson Microwave Anisotropy Probe year 5 cosmology. MNRAS, 390:L64–L68, October 2008.

[79] A. R. Duffy, J. Schaye, S. T. Kay, C. Dalla Vecchia, R. A. Battye, and C. M. Booth. Impact of baryon physics on dark matter structures: a detailed simulation study of halo density profiles. MNRAS, 405:2161–2178, July 2010.

[80] A. A. Dutton and A. V. Macciò. Cold dark matter haloes in the Planck era: evolution of structural parameters for Einasto and NFW profiles. MNRAS, 441:3359–3374, July 2014.

[81] J. Einasto. On the Construction of a Composite Model for the Galaxy and on the Determination of the System of Galactic Parameters. *Trudy Astrofizicheskogo Instituta Alma-Ata*, 5:87–100, 1965.

[82] D. J. Eisenstein and W. Hu. Baryonic Features in the Matter Transfer Function. ApJ, 496:605–+, March 1998.

[83] D. J. Eisenstein, W. Hu, and M. Tegmark. Cosmic Complementarity: H_0 and Ω_m from Combining Cosmic Microwave Background Experiments and Redshift Surveys. ApJ, 504:L57–L60, September 1998.

[84] D. J. Eisenstein, H.-J. Seo, E. Sirko, and D. N. Spergel. Improving Cosmological Distance Measurements by Reconstruction of the Baryon Acoustic Peak. ApJ, 664:675–679, August 2007.

[85] V. R. Eke, J. F. Navarro, and M. Steinmetz. The Power Spectrum Dependence of Dark Matter Halo Concentrations. ApJ, 554:114–125, June 2001.

[86] N. Espino-Briones, M. Plionis, and C. Ragone-Figueroa. Environmental Effects of Dark Matter Halos: The Clustering-Substructure Relation of Group-Size Halos. ApJ, 666:L5–L8, September 2007.

[87] O. Fakhouri and C.-P. Ma. Dark matter halo growth - II. Diffuse accretion and its environmental dependence. MNRAS, 401:2245–2256, February 2010.

[88] A. Faltenbacher, S. Gottloeber, and W. G. Mathews. Oscillatory relaxation of a merging galaxy cluster. *arXiv e-prints*, pages astro–ph/0609615, September 2006.

[89] G. Foëx, V. Motta, E. Jullo, M. Limousin, and T. Verdugo. SARCS strong-lensing galaxy groups. II. Mass-concentration relation and strong-lensing bias. A&A, 572:A19, December 2014.

[90] B. Follin, L. Knox, M. Millea, and Z. Pan. First Detection of the Acoustic Oscillation Phase Shift Expected from the Cosmic Neutrino Background. Phys. Rev. Lett., 115(9):091301, August 2015.

[91] B. Friesen, M. M. A. Patwary, B. Austin, N. Satish, Z. Slepian, N. Sundaram, D. Bard, D. J Eisenstein, J. Deslippe, P. Dubey, and Prabhat. Galactos: Computing the Anisotropic 3-Point Correlation Function for 2 Billion Galaxies. *ArXiv e-prints*, August 2017.

[92] T. Fukushige and J. Makino. Structure of Dark Matter Halos from Hierarchical Clustering. ApJ, 557:533–545, August 2001.

[93] T. Fukushige and J. Makino. Structure of Dark Matter Halos From Hierarchical Clustering: II. Universality and Self-Similarity in Cluster-Sized Halos. *arXiv e-prints*, pages astro–ph/0108014, August 2001.

[94] P. Gagrani and L. Samushia. Information Content of the Angular Multipoles of Redshift-Space Galaxy Bispectrum. MNRAS, 467:928–935, May 2017.

[95] L. Gao, C. S. Frenk, M. Boylan-Kolchin, A. Jenkins, V. Springel, and S. D. M. White. The statistics of the subhalo abundance of dark matter haloes. MNRAS, 410:2309–2314, February 2011.

[96] L. Gao, A. Loeb, P. J. E. Peebles, S. D. M. White, and A. Jenkins. Early Formation and Late Merging of the Giant Galaxies. ApJ, 614:17–25, October 2004.

[97] L. Gao, J. F. Navarro, S. Cole, C. S. Frenk, S. D. M. White, V. Springel, A. Jenkins, and A. F. Neto. The redshift dependence of the structure of massive Λ cold dark matter haloes. MNRAS, 387:536–544, June 2008.

[98] L. Gao, J. F. Navarro, C. S. Frenk, A. Jenkins, V. Springel, and S. D. M. White. The Phoenix Project: the dark side of rich Galaxy clusters. MNRAS, 425:2169–2186, September 2012.

[99] L. Gao, V. Springel, and S. D. M. White. The age dependence of halo clustering. MNRAS, 363:L66–L70, October 2005.

[100] L. Gao and S. D. M. White. Assembly bias in the clustering of dark matter haloes. MNRAS, 377:L5–L9, April 2007.

[101] J. P. Gardner. Dependence of Halo Properties on Interaction History, Environment, and Cosmology. ApJ, 557:616–625, August 2001.

[102] E. Gaztanaga, A. Cabre, F. Castander, M. Crocce, and P. Fosalba. Clustering of Luminous Red Galaxies III: Detection of the Baryon Acoustic Peak in the 3-point Correlation Function. MNRAS, 399:801, 2009.

[103] S. Genel, N. Bouché, T. Naab, A. Sternberg, and R. Genzel. The Growth of Dark Matter Halos: Evidence for Significant Smooth Accretion. ApJ, 719:229–239, August 2010.

[104] H. Gil-Marín, J. Guy, P. Zarrouk, E. Burtin, C.-H. Chuang, W. J. Percival, A. J. Ross, R. Ruggeri, R. Tojerio, G.-B. Zhao, Y. Wang, J. Bautista, J. Hou, A. G. Sánchez, I. Pâris, F. Baumgarten, J. R. Brownstein, K. S. Dawson, S. Eftekharzadeh, V. González-Pérez, S. Habib, K. Heitmann, A. D. Myers, G. Rossi, D. P. Schneider, H.-J. Seo, J. L. Tinker, and C. Zhao. The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample: structure growth rate measurement from the anisotropic quasar power spectrum in the redshift range $0.8 < z < 2.2$. MNRAS, 477:1604–1638, June 2018.

[105] H. Gil-Marín, J. Noreña, L. Verde, W. J. Percival, C. Wagner, M. Manera, and D. P. Schneider. The power spectrum and bispectrum of SDSS DR11 BOSS galaxies - I. Bias and gravity. MNRAS, 451:539–580, July 2015.

[106] H. Gil-Marín, W. J. Percival, L. Verde, J. R. Brownstein, C.-H. Chuang, F.-S. Kitaura, S. A. Rodríguez-Torres, and M. D. Olmstead. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: RSD measurement from the power spectrum and bispectrum of the DR12 BOSS galaxies. MNRAS, 465:1757–1788, February 2017.

[107] H. Gil-Marín, C. Wagner, F. Fragkoudi, R. Jimenez, and L. Verde. An improved fitting formula for the dark matter bispectrum. JCAP, 2:047, February 2012.

[108] C. Giocoli, J. Moreno, R. K. Sheth, and G. Tormen. An improved model for the formation times of dark matter haloes. MNRAS, 376:977–983, April 2007.

[109] C. Giocoli, G. Tormen, and R. K. Sheth. Formation times, mass growth histories and concentrations of dark matter haloes. MNRAS, 422:185–198, May 2012.

[110] C. Giocoli, G. Tormen, R. K. Sheth, and F. C. van den Bosch. The substructure hierarchy in dark matter haloes. MNRAS, 404:502–517, May 2010.

[111] A. W. Graham, D. Merritt, B. Moore, J. Diemand, and B. Terzić. Empirical Models for Dark Matter Halos. II. Inner Profile Slopes, Dynamical Profiles, and ρ/σ^3. AJ, 132:2701–2710, December 2006.

[112] E. J. Groth and P. J. E. Peebles. Statistical analysis of catalogs of extragalactic objects. VII - Two- and three-point correlation functions for the high-resolution Shane-Wirtanen catalog of galaxies. ApJ, 217:385–405, October 1977.

[113] D. Gualdi, H. Gil-Marín, R. L. Schuhmann, M. Manera, B. Joachimi, and O. Lahav. Enhancing BOSS bispectrum cosmological constraints with maximal compression. *ArXiv e-prints*, June 2018.

[114] H. Guo, C. Li, Y. P. Jing, and G. Börner. Stellar Mass and Color Dependence of the Three-point Correlation Function of Galaxies in the Local Universe. ApJ, 780:139, January 2014.

[115] H. Guo, Z. Zheng, Y. P. Jing, I. Zehavi, C. Li, D. H. Weinberg, R. A. Skibba, R. C. Nichol, G. Rossi, C. G. Sabiu, D. P. Schneider, and C. K. McBride. Modelling the redshift-space three-point correlation function in SDSS-III. MNRAS, 449:L95–L99, April 2015.

[116] Q. Guo, S. White, C. Li, and M. Boylan-Kolchin. How do galaxies populate dark matter haloes? MNRAS, 404:1111–1120, May 2010.

[117] Q. Guo and S. D. M. White. Galaxy growth in the concordance ΛCDM cosmology. MNRAS, 384:2–10, February 2008.

[118] S. Habib et al. HACC: Simulating Sky Surveys on State-of-the-Art Supercomputing Architectures. New A, 42:49–65, 2016.

[119] S. Habib, A. Pope, H. Finkel, N. Frontiere, K. Heitmann, D. Daniel, P. Fasel, V. Morozov, G. Zagaris, T. Peterka, V. Vishwanath, Z. Lukić, S. Sehrish, and W.-k. Liao. HACC: Simulating sky surveys on state-of-the-art supercomputing architectures. New A, 42:49–65, January 2016.

[120] O. Hahn, C. M. Carollo, C. Porciani, and A. Dekel. The Evolution of Dark Matter Halo Properties in Clusters, Filaments, Sheets and Voids. MNRAS, 381:41, 2007.

[121] O. Hahn, C. Porciani, C. M. Carollo, and A. Dekel. Properties of dark matter haloes in clusters, filaments, sheets and voids. MNRAS, 375:489–499, February 2007.

[122] O. Hahn, C. Porciani, A. Dekel, and C. M. Carollo. The Tidal Origin of the Environment Dependence of Halo Assembly. MNRAS, 398:1742, 2009.

[123] J. Han, Y. Li, Y. Jing, T. Nishimichi, W. Wang, and C. Jiang. The multidimensional dependence of halo bias in the eye of a machine: a tale of halo structure, assembly, and environment. MNRAS, 482:1900–1919, January 2019.

[124] G. Harker, S. Cole, J. Helly, C. Frenk, and A. Jenkins. A marked correlation function analysis of halo formation times in the millennium simulation. MNRAS, 367:1039–1049, 2006.

[125] J. Hartlap, P. Simon, and P. Schneider. Why your model parameter confidences might be too optimistic: Unbiased estimation of the inverse covariance matrix. A&A, 464:399–404, 2007.

[126] A. P. Hearin and D. F. Watson. The Dark Side of Galaxy Color. MNRAS, 435:1313–1324, 2013.

[127] A. P. Hearin, A. R. Zentner, F. C. van den Bosch, D. Campbell, and E. Tollerud. Introducing decorated HODs: modelling assembly bias in the galaxy-halo connection. MNRAS, 460:2552–2570, August 2016.

[128] K. Heitmann, D. Bingham, E. Lawrence, S. Bergner, S. Habib, D. Higdon, A. Pope, R. Biswas, H. Finkel, N. Frontiere, and S. Bhattacharya. The Mira-Titan Universe: Precision Predictions for Dark Energy Surveys. ApJ, 820:108, April 2016.

[129] K. Heitmann, N. Frontiere, C. Sewell, S. Habib, A. Pope, H. Finkel, S. Rizzi, J. Insley, and S. Bhattacharya. The Q Continuum Simulation: Harnessing the Power of GPU Accelerated Supercomputers. ApJS, 219:34, August 2015.

[130] K. Heitmann, Z. Lukić, S. Habib, and P. M. Ricker. Capturing Halos at High Redshifts. ApJ, 642:L85–L88, May 2006.

[131] K. Heitmann, M. White, C. Wagner, S. Habib, and D. Higdon. The Coyote Universe. I. Precision Determination of the Nonlinear Matter Power Spectrum. ApJ, 715:104–121, May 2010.

[132] W. A. Hellwing, C. S. Frenk, M. Cautun, S. Bose, J. Helly, A. Jenkins, T. Sawala, and M. Cytowski. The Copernicus Complexio: a high-resolution view of the small-scale Universe. MNRAS, 457:3492–3509, April 2016.

[133] R. N. Henriksen and L. M. Widrow. Relaxing and virializing a dark matter halo. MNRAS, 302:321–336, January 1999.

[134] C. Hikage, K. Koyama, and A. Heavens. Perturbation theory for BAO reconstructed fields: One-loop results in the real-space matter density field. Phys. Rev. D, 96(4):043513, August 2017.

[135] Y. Hoffman, E. Romano-Díaz, I. Shlosman, and C. Heller. Evolution of the Phase-Space Density in Dark Matter Halos. ApJ, 671:1108–1114, December 2007.

[136] Y. Hoffman and J. Shaham. Local density maxima - Progenitors of structure. ApJ, 297:16–22, October 1985.

[137] K. Hoffmann, E. Gaztanaga, R. Scoccimarro, and M. Crocce. Testing the consistency of three-point halo clustering in Fourier and configuration space. MNRAS, 476(1):814–829, 2018.

[138] G. Holder, Z. Haiman, and J. J. Mohr. Constraints on Ω_m, Ω_Λ, and σ_8 from Galaxy Cluster Redshift Distributions. ApJ, 560:L111–L114, October 2001.

[139] J. Hou, A. G. Sánchez, R. Scoccimarro, S. Salazar-Albornoz, E. Burtin, H. Gil-Marín, W. J. Percival, R. Ruggeri, P. Zarrouk, G.-B. Zhao, J. Bautista, J. Brinkmann, J. R. Brownstein, K. S. Dawson, N. Chandrachani Devi, A. D. Myers, S. Habib, K. Heitmann, R. Tojeiro, G. Rossi, D. P. Schneider, H.-J. Seo, and Y. Wang. The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample: anisotropic clustering analysis in configuration-space. MNRAS, July 2018.

[140] W. Hu and D. J. Eisenstein. Structure of structure formation theories. Phys. Rev. D, 59(8):083509, April 1999.

[141] W. Hu and Z. Haiman. Redshifting rings of power. Phys. Rev. D, 68(6):063004, September 2003.

[142] W. Hu and A. V. Kravtsov. Sample Variance Considerations for Cluster Surveys. ApJ, 584:702–715, February 2003.

[143] A. Huss, B. Jain, and M. Steinmetz. The Formation and Evolution of Clusters of Galaxies in Different Cosmogonies. *arXiv e-prints*, pages astro–ph/9703014, March 1997.

[144] A. Huss, B. Jain, and M. Steinmetz. How Universal Are the Density Profiles of Dark Halos? ApJ, 517:64–69, May 1999.

[145] T. Ishiyama. Hierarchical Formation of Dark Matter Halos and the Free Streaming Scale. ApJ, 788:27, 2014.

[146] T. Ishiyama, T. Fukushige, and J. Makino. Variation of the subhalo abundance in dark matter halos. ApJ, 696:2115–2125, 2009.

[147] T. Ishiyama, S. Rieder, J. Makino, S. Portegies Zwart, D. Groen, K. Nitadori, C. de Laat, S. McMillan, K. Hiraki, and S. Harfst. The Cosmogrid Simulation: Statistical Properties of Small Dark Matter Halos. ApJ, 767:146, April 2013.

[148] A. Jeeson-Daniel, C. Dalla Vecchia, M. R. Haas, and J. Schaye. The correlation structure of dark matter halo properties. MNRAS, 415:L69–L73, July 2011.

[149] D. Jeong and E. Komatsu. Primordial Non-Gaussianity, Scale-dependent Bias, and the Bispectrum of Galaxies. ApJ, 703:1230–1248, October 2009.

[150] Y. P. Jing and G. Börner. The Three-Point Correlation Function of Galaxies Determined from the Two-Degree Field Galaxy Redshift Survey. ApJ, 607:140–163, May 2004.

[151] Y. P. Jing, H. J. Mo, and G. Börner. Spatial Correlation Function and Pairwise Velocity Dispersion of Galaxies: Cold Dark Matter Models versus the Las Campanas Survey. ApJ, 494:1–12, February 1998.

[152] Y. P. Jing and Y. Suto. Triaxial Modeling of Halo Density Profiles with High-Resolution N-Body Simulations. ApJ, 574:538–553, August 2002.

[153] Y. P. Jing, Y. Suto, and H. J. Mo. The Dependence of Dark Halo Clustering on Formation Epoch and Concentration Parameter. ApJ, 657:664–668, March 2007.

[154] D. E. Johnston, E. S. Sheldon, R. H. Wechsler, E. Rozo, B. P. Koester, J. A. Frieman, T. A. McKay, A. E. Evrard, M. R. Becker, and J. Annis. Cross-correlation Weak Lensing of SDSS galaxy Clusters II: Cluster Density Profiles and the Mass–Richness Relation. *ArXiv e-prints*, September 2007.

[155] G. Kauffmann, A. Nusser, and M. Steinmetz. Galaxy formation and large-scale bias. MNRAS, 286:795–811, April 1997.

[156] G. Kauffmann, S. D. M. White, and B. Guiderdoni. The Formation and Evolution of Galaxies Within Merging Dark Matter Haloes. MNRAS, 264:201, September 1993.

[157] I. Kayo, M. Takada, and B. Jain. Information content of weak lensing power spectrum and bispectrum: including the non-Gaussian error covariance matrix. MNRAS, 429:344–371, February 2013.

[158] S. Kazantzidis, A. R. Zentner, and A. V. Kravtsov. The Robustness of Dark Matter Density Profiles in Dissipationless Mergers. ApJ, 641:647–664, April 2006.

[159] T. Kitayama and Y. Suto. Formation rate of gravitational structures and the cosmic X-ray background radiation. MNRAS, 280:638, May 1996.

[160] A. Klypin, A. V. Kravtsov, J. S. Bullock, and J. R. Primack. Resolving the Structure of Cold Dark Matter Halos. ApJ, 554:903–915, June 2001.

[161] A. Klypin, G. Yepes, S. Gottlöber, F. Prada, and S. Heß. MultiDark simulations: the story of dark matter halo concentrations and density profiles. MNRAS, 457:4340–4359, April 2016.

[162] A. A. Klypin, S. Trujillo-Gomez, and J. Primack. Dark Matter Halos in the Standard Cosmological Model: Results from the Bolshoi Simulation. ApJ, 740:102, October 2011.

[163] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. R. Nolta, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. ApJS, 192:18, February 2011.

[164] G. V. Kulkarni, R. C. Nichol, R. K. Sheth, H.-J. Seo, D. J. Eisenstein, and A. Gray. The Three-Point Correlation Function of Luminous Red Galaxies in the Sloan Digital Sky Survey. MNRAS, 378:1196–1206, 2007.

[165] A. Kull. A Model for the Density Distribution of Virialized Cold Dark Matter Halos. ApJ, 516:L5–L8, May 1999.

[166] J. Kwan, S. Bhattacharya, K. Heitmann, and S. Habib. Cosmic Emulation: The Concentration-Mass Relation for wCDM Universes. ApJ, 768:123, May 2013.

[167] I. Lacerna and N. Padilla. The nature of assembly bias - I. Clues from a ΛCDM cosmology. MNRAS, 412:1283–1294, April 2011.

[168] I. Lacerna and N. Padilla. The nature of assembly bias - II. Halo spin. MNRAS, 426:L26–L30, October 2012.

[169] C. Lacey and S. Cole. Merger rates in hierarchical models of galaxy formation. MNRAS, 262:627–649, June 1993.

[170] C. G. Lacey and S. Cole. Merger rates in hierarchical models of galaxy formation. 2. Comparison with N body simulations. MNRAS, 271:676, 1994.

[171] E. Lawrence, K. Heitmann, J. Kwan, A. Upadhye, D. Bingham, S. Habib, D. Higdon, A. Pope, H. Finkel, and N. Frontiere. The Mira-Titan Universe. II. Matter Power Spectrum Emulation. ApJ, 847:50, September 2017.

[172] A. M. C. Le Brun, M. Arnaud, G. W. Pratt, and R. Teyssier. Internal dark matter structure of the most massive galaxy clusters. MNRAS, 473:L69–L73, January 2018.

[173] G. Lemson and G. Kauffmann. Environmental Influences on Dark Matter Halos and Consequences for the Galaxies Within Them. *ArXiv Astrophysics e-prints*, October 1997.

[174] M. Levi, C. Bebek, T. Beers, R. Blum, R. Cahn, D. Eisenstein, B. Flaugher, K. Honscheid, R. Kron, O. Lahav, P. McDonald, N. Roe, D. Schlegel, and representing the DESI collaboration. The DESI Experiment, a whitepaper for Snowmass 2013. *ArXiv e-prints*, August 2013.

[175] A. Lewis, A. Challinor, and A. Lasenby. Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. ApJ, 538:473–476, August 2000.

[176] N. Li, M. D. Gladders, K. Heitmann, E. M. Rangel, H. L. Child, M. K. Florian, L. E. Bleem, S. Habib, and H. J. Finkel. The importance of secondary halos for strong lensing in massive galaxy clusters across redshift. *ArXiv e-prints*, October 2018.

[177] Y. Li, H. J. Mo, and L. Gao. On Halo Formation Times and Assembly Bias. MNRAS, 389:1419, 2008.

[178] Y. Li, H. J. Mo, F. C. van den Bosch, and W. P. Lin. The Assembly History of Dark Matter Haloes. MNRAS, 379:689–701, 2007.

[179] W. P. Lin, Y. P. Jing, and L. Lin. Formation time-distribution of dark matter haloes: theories versus N-body simulations. MNRAS, 344:1327–1333, October 2003.

[180] E. V. Linder. Exploring the Expansion History of the Universe. Phys. Rev. Lett., 90(9):091301, March 2003.

[181] Y. Lithwick and N. Dalal. Self-similar Solutions of Triaxial Dark Matter Halos. ApJ, 734:100, June 2011.

[182] E. L. Łokas. Universal profile of dark matter haloes and the spherical infall model. MNRAS, 311:423–432, January 2000.

[183] LSST Science Collaboration, P. A. Abell, J. Allison, S. F. Anderson, J. R. Andrew, J. R. P. Angel, L. Armus, D. Arnett, S. J. Asztalos, T. S. Axelrod, et al. LSST Science Book, Version 2.0. *ArXiv e-prints*, December 2009.

[184] Y. Lu, H. J. Mo, N. Katz, and M. D. Weinberg. On the origin of cold dark matter halo density profiles. MNRAS, 368:1931–1940, 2006.

[185] A. D. Ludlow and R. E. Angulo. Einasto Profiles and the Dark Matter Power Spectrum. MNRAS, 465(1):L84–L88, 2017.

[186] A. D. Ludlow, M. Borzyszkowski, and C. Porciani. The formation of CDM haloes I: Collapse thresholds and the ellipsoidal collapse model. *ArXiv e-prints*, July 2011.

[187] A. D. Ludlow, S. Bose, R. E. Angulo, L. Wang, W. A. Hellwing, J. F. Navarro, S. Cole, and C. S. Frenk. The mass-concentration-redshift relation of cold and warm dark matter haloes. MNRAS, 460:1214–1232, August 2016.

[188] A. D. Ludlow, J. F. Navarro, R. E. Angulo, M. Boylan-Kolchin, V. Springel, C. Frenk, and S. D. M. White. The mass-concentration-redshift relation of cold dark matter haloes. MNRAS, 441:378–388, June 2014.

[189] A. D. Ludlow, J. F. Navarro, M. Boylan-Kolchin, P. E. Bett, R. E. Angulo, M. Li, S. D. M. White, C. Frenk, and V. Springel. The Mass Profile and Accretion History of Cold Dark Matter Halos. MNRAS, 432:1103, 2013.

[190] A. D. Ludlow, J. F. Navarro, M. Li, R. E. Angulo, M. Boylan-Kolchin, and P. E. Bett. The dynamical state and mass-concentration relation of galaxy clusters. MNRAS, 427:1322–1328, December 2012.

[191] Z. Lukić, D. Reed, S. Habib, and K. Heitmann. The Structure of Halos: Implications for Group and Cluster Cosmology. ApJ, 692:217–228, February 2009.

[192] A. V. Macciò, A. A. Dutton, and F. C. van den Bosch. Concentration, spin and shape of dark matter haloes as a function of the cosmological model: WMAP1, WMAP3 and WMAP5 results. MNRAS, 391:1940–1954, December 2008.

[193] A. V. Macciò, A. A. Dutton, F. C. van den Bosch, B. Moore, D. Potter, and J. Stadel. Concentration, spin and shape of dark matter haloes: scatter and the dependence on mass and environment. MNRAS, 378:55–71, June 2007.

[194] R. Mandelbaum, U. Seljak, R. J. Cool, M. Blanton, C. M. Hirata, and J. Brinkmann. Density profiles of galaxy groups and clusters from SDSS galaxy-galaxy weak lensing. MNRAS, 372:758–776, October 2006.

[195] R. Mandelbaum, U. Seljak, and C. M. Hirata. A halo mass–concentration relation from weak lensing. JCAP, 8:006, August 2008.

[196] A. Manrique, A. Raig, E. Salvador-Sole, T. Sanchis, and J. M. Solanes. On the origin of the inner structure of halos. ApJ, 593:26–37, 2003.

[197] Y.-Y. Mao, A. R. Zentner, and R. H. Wechsler. Beyond assembly bias: exploring secondary halo biases for cluster-size haloes. MNRAS, 474:5143–5157, March 2018.

[198] F. Marín. The Large-scale Three-point Correlation Function of Sloan Digital Sky Survey Luminous Red Galaxies. ApJ, 737:97, August 2011.

[199] F. A. Marin et al. The WiggleZ Dark Energy Survey: constraining galaxy bias and cosmic growth with 3-point correlation functions. MNRAS, 432:2654, 2013.

[200] S. Masaki, Y.-T. Lin, and N. Yoshida. Modelling colour-dependent galaxy clustering in cosmological simulations. MNRAS, 436:2286–2300, December 2013.

[201] C. K. McBride, A. J. Connolly, J. P. Gardner, R. Scranton, J. A. Newman, R. Scoccimarro, I. Zehavi, and D. P. Schneider. Three-point Correlation Functions of SDSS Galaxies: Luminosity and Color Dependence in Redshift and Projected Space. ApJ, 726:13, January 2011.

[202] C. K. McBride, A. J. Connolly, J. P. Gardner, R. Scranton, R. Scoccimarro, A. A. Berlind, F. Marín, and D. P. Schneider. Three-point Correlation Functions of SDSS Galaxies: Constraining Galaxy-mass Bias. ApJ, 739:85, October 2011.

[203] J. McBride, O. Fakhouri, and C.-P. Ma. Mass accretion rates and histories of dark matter haloes. MNRAS, 398:1858–1868, October 2009.

[204] M. Meneghetti, M. Bartelmann, K. Dolag, L. Moscardini, F. Perrotta, C. Baccigalupi, and G. Tormen. Strong lensing efficiency of galaxy clusters in dark energy cosmologies. A&A, 442:413–422, November 2005.

[205] M. Meneghetti and E. Rasia. Reconciling extremely different concentration-mass relations. *ArXiv e-prints*, March 2013.

[206] D. Merritt, A. W. Graham, B. Moore, J. Diemand, and B. Terzić. Empirical Models for Dark Matter Halos. I. Nonparametric Construction of Density Profiles and Comparison with Parametric Models. AJ, 132:2685–2700, December 2006.

[207] J. Merten, M. Meneghetti, M. Postman, K. Umetsu, A. Zitrin, E. Medezinski, M. Nonino, A. Koekemoer, P. Melchior, D. Gruen, L. A. Moustakas, M. Bartelmann, O. Host, M. Donahue, D. Coe, A. Molino, S. Jouvel, A. Monna, S. Seitz, N. Czakon, D. Lemze, J. Sayers, I. Balestra, P. Rosati, N. Benítez, A. Biviano, R. Bouwens, L. Bradley, T. Broadhurst, M. Carrasco, H. Ford, C. Grillo, L. Infante, D. Kelson, O. Lahav, R. Massey, J. Moustakas, E. Rasia, J. Rhodes, J. Vega, and W. Zheng. CLASH: The Concentration-Mass Relation of Galaxy Clusters. ApJ, 806:4, June 2015.

[208] B. Moore, S. Kazantzidis, J. Diemand, and J. Stadel. The origin and tidal evolution of cuspy triaxial haloes. MNRAS, 354:522–528, October 2004.

[209] B. Moore, T. Quinn, F. Governato, J. Stadel, and G. Lake. Cold collapse and the core catastrophe. MNRAS, 310:1147–1152, December 1999.

[210] A. Moradinezhad Dizgah, H. Lee, J. B. Muñoz, and C. Dvorkin. Galaxy Bispectrum from Massive Spinning Particles. JCAP, 1805(05):013, 2018.

[211] S. More, B. Diemer, and A. V. Kravtsov. The Splashback Radius as a Physical Halo Boundary and the Growth of Halo Mass. ApJ, 810:36, September 2015.

[212] B. P. Moster, R. S. Somerville, C. Maulbetsch, F. C. van den Bosch, A. V. Macciò, T. Naab, and L. Oser. Constraints on the Relationship between Stellar Mass and Halo Mass at Low and High Redshift. ApJ, 710:903–923, February 2010.

[213] R. Mostoghiu, A. Knebe, W. Cui, F. R. Pearce, G. Yepes, C. Power, R. Dave, and A. Arth. The Three Hundred Project: The evolution of galaxy cluster density profiles. MNRAS, 483:3390–3403, March 2019.

[214] J. C. Muñoz-Cuartas, A. V. Macciò, S. Gottlöber, and A. A. Dutton. The redshift evolution of Λ cold dark matter halo parameters: concentration, spin and shape. MNRAS, 411:584–594, February 2011.

[215] E. O. Nadler, A. Perko, and L. Senatore. On the Bispectra of Very Massive Tracers in the Effective Field Theory of Large-Scale Structure. JCAP, 1802(02):058, 2018.

[216] Y. Nan, K. Yamamoto, and C. Hikage. Higher multipoles of the galaxy bispectrum in redshift space. JCAP, 7:038, July 2018.

[217] J. F. Navarro, C. S. Frenk, and S. D. M. White. The Structure of Cold Dark Matter Halos. ApJ, 462:563, May 1996.

[218] J. F. Navarro, C. S. Frenk, and S. D. M. White. A Universal Density Profile from Hierarchical Clustering. ApJ, 490:493–508, December 1997.

[219] J. F. Navarro, E. Hayashi, C. Power, A. R. Jenkins, C. S. Frenk, S. D. M. White, V. Springel, J. Stadel, and T. R. Quinn. The inner structure of ΛCDM haloes - III. Universality and asymptotic slopes. MNRAS, 349:1039–1051, April 2004.

[220] E. Neistein and F. C. van den Bosch. Natural Downsizing in Hierarchical Galaxy Formation. MNRAS, 372:933–948, 2006.

[221] A. F. Neto, L. Gao, P. Bett, S. Cole, J. F. Navarro, C. S. Frenk, S. D. M. White, V. Springel, and A. Jenkins. The statistics of ΛCDM halo concentrations. MNRAS, 381:1450–1462, November 2007.

[222] A. B. Newman, T. Treu, R. S. Ellis, D. J. Sand, C. Nipoti, J. Richard, and E. Jullo. The Density Profiles of Massive, Relaxed Galaxy Clusters. I. The Total Density Over Three Decades in Radius. ApJ, 765:24, March 2013.

[223] R. C. Nichol et al. The effect of large-scale structure on the sdss galaxy three-point correlation function. MNRAS, 368:1507–1514, 2006.

[224] H. Niikura, M. Takada, N. Okabe, R. Martino, and R. Takahashi. Detection of universality of dark matter profile from Subaru weak lensing measurements of 50 massive clusters. PASJ, 67:103, December 2015.

[225] T. Nishimichi and A. Oka. Simulating the anisotropic clustering of luminous red galaxies with subhaloes: a direct confrontation with observation and cosmological implications. MNRAS, 444:1400–1418, October 2014.

[226] Y. Noh, M. White, and N. Padmanabhan. Reconstructing baryon oscillations. Phys. Rev. D, 80(12):123501, December 2009.

[227] A. Nusser and R. K. Sheth. Mass growth and density profiles of dark matter halos in hierarchical clustering. MNRAS, 303:685, 1999.

[228] G. Ogiya, D. Nagai, and T. Ishiyama. Dynamical evolution of primordial dark matter haloes through mergers. MNRAS, 461:3385–3396, September 2016.

[229] M. Oguri, M. B. Bayliss, H. Dahle, K. Sharon, M. D. Gladders, P. Natarajan, J. F. Hennawi, and B. P. Koester. Combined strong and weak lensing analysis of 28 clusters from the Sloan Giant Arcs Survey. MNRAS, 420:3213–3239, March 2012.

[230] N. Okabe and G. P. Smith. LoCuSS: weak-lensing mass calibration of galaxy clusters. MNRAS, 461:3794–3821, October 2016.

[231] N. Okabe, G. P. Smith, K. Umetsu, M. Takada, and T. Futamase. LoCuSS: The Mass Density Profile of Massive Galaxy Clusters at z = 0.2. ApJ, 769:L35, June 2013.

[232] N. Okabe, M. Takada, K. Umetsu, T. Futamase, and G. P. Smith. LoCuSS: Subaru Weak Lensing Study of 30 Galaxy Clusters. PASJ, 62:811–870, June 2010.

[233] C. Okoli and N. Afshordi. Concentration, ellipsoidal collapse, and the densest dark matter haloes. MNRAS, 456:3068–3078, March 2016.

[234] N. Padmanabhan and M. White. Calibrating the baryon oscillation ruler for matter and halos. Phys. Rev. D, 80(6):063508, September 2009.

[235] N. Padmanabhan, M. White, and J. D. Cohn. Reconstructing baryon oscillations: A Lagrangian theory perspective. Phys. Rev. D, 79(6):063523, March 2009.

[236] N. Padmanabhan, X. Xu, D. J. Eisenstein, R. Scalzo, A. J. Cuesta, K. T. Mehta, and E. Kazin. A 2 per cent distance to z = 0.35 by reconstructing baryon acoustic oscillations - I. Methods and application to the Sloan Digital Sky Survey. MNRAS, 427:2132–2145, December 2012.

[237] J. Pan and I. Szapudi. The Monopole moment of the three-point correlation function of the 2-degree Field Galaxy Redshift Survey. MNRAS, 362:1363, 2005.

[238] J. A. Peacock and R. E. Smith. Halo occupation numbers and galaxy bias. MNRAS, 318:1144–1156, November 2000.

[239] D. W. Pearson and L. Samushia. A Detection of the Baryon Acoustic Oscillation features in the SDSS BOSS DR12 Galaxy Bispectrum. MNRAS, 478:4500–4512, August 2018.

[240] P. J. E. Peebles. *The large-scale structure of the universe*. Princeton University Press, Princeton, NJ, 1980.

[241] W. J. Percival et al. The Clustering of Galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Including covariance matrix errors. MNRAS, 439(3):2531–2541, 2014.

[242] G. B. Poole, P. W. Angel, S. J. Mutch, C. Power, A. R. Duffy, P. M. Geil, A. Mesinger, and S. B. Wyithe. Dark-ages Reionization and Galaxy formation simulation - I. The dynamical lives of high-redshift galaxies. MNRAS, 459:3025–3039, July 2016.

[243] C. N. Poveda-Ruiz, J. E. Forero-Romero, and J. C. Muñoz-Cuartas. Quantifying and Controlling Biases in Estimates of Dark Matter Halo Concentration. ApJ, 832:169, December 2016.

[244] C. Power, A. Knebe, and S. R. Knollmann. The dynamical state of dark matter haloes in cosmological simulations - I. Correlations with mass assembly history. MNRAS, 419:1576–1587, January 2012.

[245] C. Power, J. F. Navarro, A. Jenkins, C. S. Frenk, S. D. M. White, V. Springel, J. Stadel, and T. Quinn. The inner structure of ΛCDM haloes - I. A numerical convergence study. MNRAS, 338:14–34, January 2003.

[246] F. Prada, A. A. Klypin, A. J. Cuesta, J. E. Betancort-Rijo, and J. Primack. Halo concentrations in the standard Λ cold dark matter cosmology. MNRAS, 423:3018–3030, July 2012.

[247] F. Prada, A. A. Klypin, E. Simonneau, J. Betancort-Rijo, S. Patiri, S. Gottlöber, and M. A. Sanchez-Conde. How Far Do They Go? The Outer Structure of Galactic Dark Matter Halos. ApJ, 645:1001–1011, July 2006.

[248] W. H. Press and P. Schechter. Formation of galaxies and clusters of galaxies by self-similar gravitational condensation. ApJ, 187:425–438, 1974.

[249] C. Ragone-Figueroa, M. Plionis, M. Merchán, S. Gottlöber, and G. Yepes. The relation between halo shape, velocity dispersion and formation time. MNRAS, 407:581–589, September 2010.

[250] A. Raig, G. Gonzalez-Casado, and E. Salvador-Sole. Scaling evolution of universal dark-matter halo density profiles. ApJ, 508:L129, 1998.

[251] C. Rampf and Y. Y. Y. Wong. Lagrangian perturbations and the matter bispectrum II: the resummed one-loop correction to the matter bispectrum. JCAP, 6:018, June 2012.

[252] E. Rangel, N. Frontiere, S. Habib, K. Heitmann, W. Liao, A. Agrawal, and A. Choudhary. Building halo merger trees from the q continuum simulation. In *2017 IEEE 24th International Conference on High Performance Computing (HiPC)*, pages 398–407. IEEE, 2017.

[253] D. Reed, F. Governato, L. Verde, J. Gardner, T. R. Quinn, J. Stadel, D. Merritt, and G. Lake. Evolution of the density profiles of dark matter halos. MNRAS, 357:82–96, 2005.

[254] D. S. Reed, F. Governato, T. R. Quinn, J. Stadel, and G. Lake. The age dependence of galaxy clustering. MNRAS, 378:777–784, 2007.

[255] D. S. Reed, S. M. Koushiappas, and L. Gao. Non-universality of halo profiles and implications for dark matter experiments. MNRAS, 415:3177–3188, August 2011.

[256] A. Refregier, A. Amara, T. D. Kitching, A. Rassat, R. Scaramella, J. Weller, and f. t. Euclid Imaging Consortium. Euclid Imaging Consortium Science Book. *ArXiv e-prints*, January 2010.

[257] M. Ricotti. Dependence of the inner dark matter profile on the halo mass. MNRAS, 344:1237–1249, October 2003.

[258] M. Ricotti and M. I. Wilkinson. On the origin of dark matter cores in dwarf galaxies. MNRAS, 353:867–873, September 2004.

[259] A. Rodríguez-Puebla, P. Behroozi, J. Primack, A. Klypin, C. Lee, and D. Hellinger. Halo and subhalo demographics with Planck cosmological parameters: Bolshoi-Planck and MultiDark-Planck simulations. MNRAS, 462:893–916, October 2016.

[260] E. Romano-Diaz, A. Faltenbacher, D. Jones, C. Heller, Y. Hoffman, and I. Shlosman. Constrained Cosmological Simulations of Dark Matter Halos. ApJ, 637:L93–L96, February 2006.

[261] E. Romano-Diaz, Y. Hoffman, C. Heller, A. Faltenbacher, D. Jones, and I. Shlosman. Evolution of Characteristic Quantities for Dark Matter Halo Density Profiles. ApJ, 657:56–70, 2007.

[262] A. J. Ross et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: Observational systematics and baryon acoustic oscillations in the correlation function. MNRAS, 464(1):1168–1191, 2017.

[263] N. Roth, A. Pontzen, and H. V. Peiris. Genetically modified haloes: towards controlled experiments in CDM galaxy formation. MNRAS, 455(1):974–986, 2016.

[264] A. N. Salcedo, A. H. Maller, A. A. Berlind, M. Sinha, C. K. McBride, P. S. Behroozi, R. H. Wechsler, and D. H. Weinberg. Spatial clustering of dark matter haloes: secondary bias, neighbour bias, and the influence of massive neighbours on halo properties. MNRAS, 475:4411–4423, April 2018.

[265] E. Salvador-Sole, J. M. Solanes, and A. Manrique. Merger versus accretion and the structure of dark matter halos. ApJ, 499:542, 1998.

[266] E. Salvador-Solé, J. Viñas, A. Manrique, and S. Serra. Theoretical dark matter halo density profile. MNRAS, 423:2190–2202, July 2012.

[267] G. Sato-Polito, A. D. Montero-Dorta, L. R. Abramo, F. Prada, and A. Klypin. The dependence of halo bias on age, concentration and spin. *arXiv e-prints*, October 2018.

[268] F. Schmidt. Effect of relative velocity and density perturbations between baryons and dark matter on the clustering of galaxies. Phys. Rev. D, 94(6):063508, September 2016.

[269] R. W. Schmidt and S. W. Allen. The dark matter haloes of massive, relaxed galaxy clusters observed with Chandra. MNRAS, 379:209–221, July 2007.

[270] M. Schmittfull, Y. Feng, F. Beutler, B. Sherwin, and M. Y. Chu. Eulerian BAO Reconstructions and N-Point Statistics. Phys. Rev. D, 92(12):123522, 2015.

[271] R. Scoccimarro. The Bispectrum: From Theory to Observations. ApJ, 544:597–615, December 2000.

[272] R. Scoccimarro. Fast estimators for redshift-space clustering. Phys. Rev. D, 92(8):083532, October 2015.

[273] R. Scoccimarro, S. Colombi, J. N. Fry, J. A. Frieman, E. Hivon, and A. Melott. Nonlinear evolution of the bispectrum of cosmological perturbations. ApJ, 496:586, 1998.

[274] E. Sefusatti, M. Crocce, and V. Desjacques. The matter bispectrum in N-body simulations with non-Gaussian initial conditions. MNRAS, 406:1014–1028, August 2010.

[275] E. Sefusatti, M. Crocce, S. Pueblas, and R. Scoccimarro. Cosmology and the bispectrum. Phys. Rev. D, 74(2):023522, July 2006.

[276] E. Sefusatti, M. Crocce, R. Scoccimarro, and H. M. P. Couchman. Accurate estimators of correlation functions in Fourier space. MNRAS, 460:3624–3636, August 2016.

[277] U. Seljak. Analytic model for galaxy and dark matter clustering. MNRAS, 318:203–213, October 2000.

[278] H.-J. Seo and D. J. Eisenstein. Probing Dark Energy with Baryonic Acoustic Oscillations from Future Large Galaxy Redshift Surveys. ApJ, 598:720–740, December 2003.

[279] M. Sereno and G. Covone. The mass-concentration relation in massive galaxy clusters at redshift ~ 1. MNRAS, 434:878–887, September 2013.

[280] P. R. Shapiro, I. T. Iliev, H. Martel, K. Ahn, and M. A. Alvarez. The Equilibrium Structure of CDM Halos. *arXiv e-prints*, September 2004.

[281] L. D. Shaw, J. Weller, J. P. Ostriker, and P. Bode. Statistics of Physical Properties of Dark Matter Clusters. ApJ, 646:815–833, August 2006.

[282] B. D. Sherwin and M. White. The Impact of Wrong Assumptions in BAO Reconstruction. *ArXiv e-prints*, August 2018.

[283] R. K. Sheth and G. Tormen. Formation times and masses of dark matter halos. MNRAS, 349:1464, 2004.

[284] R. K. Sheth and G. Tormen. On the environmental dependence of halo formation. MNRAS, 350:1385, 2004.

[285] J. Silk. Cosmic Black-Body Radiation and Galaxy Formation. ApJ, 151:459, February 1968.

[286] M. Simet, T. McClintock, R. Mandelbaum, E. Rozo, E. Rykoff, E. Sheldon, and R. H. Wechsler. Weak lensing measurement of the mass-richness relation of SDSS redMaPPer clusters. MNRAS, 466:3103–3118, April 2017.

[287] M. Sinha and K. Holley-Bockelmann. A First Look at Galaxy Flyby Interactions. I. Characterizing the Frequency of Flybys in a Cosmological Context. ApJ, 751:17, May 2012.

[288] Z. Slepian and D. J. Eisenstein. Computing the three-point correlation function of galaxies in $\mathcal{O}(N^2)$ time. MNRAS, 454(4):4142–4158, 2015.

[289] Z. Slepian and D. J. Eisenstein. Accelerating the two-point and three-point galaxy correlation functions using Fourier transforms. MNRAS, 455(1):L31–L35, 2016.

[290] Z. Slepian and D. J. Eisenstein. Modelling the large-scale redshift-space 3-point correlation function of galaxies. MNRAS, 469(2):2059–2076, 2017.

[291] Z. Slepian and D. J. Eisenstein. A practical computational method for the anisotropic redshift-space three-point correlation function. MNRAS, 478(2):1468–1483, 2018.

[292] Z. Slepian et al. Detection of baryon acoustic oscillation features in the large-scale three-point correlation function of SDSS BOSS DR12 CMASS galaxies. MNRAS, 469(2):1738–1751, 2017.

[293] Z. Slepian et al. The large-scale three-point correlation function of the SDSS BOSS DR12 CMASS galaxies. MNRAS, 468(1):1070–1083, 2017.

[294] Z. Slepian et al. Constraining the baryon-dark matter relative velocity with the large-scale three-point correlation function of the SDSS BOSS DR12 CMASS galaxies. MNRAS, 474(2):2109–2115, 2018.

[295] R. E. Smith, R. K. Sheth, and R. Scoccimarro. An analytic model for the bispectrum of galaxies in redshift space. Phys. Rev. D, 78:023523, 2008.

[296] R. S. Somerville and J. R. Primack. Semi-analytic modelling of galaxy formation: the local Universe. MNRAS, 310:1087–1110, December 1999.

[297] D. Spergel et al. Wide-Field InfraRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report. *ArXiv e-prints*, March 2015.

[298] K. Subramanian, R. Cen, and J. P. Ostriker. The Structure of Dark Matter Halos in Hierarchical Clustering Theories. ApJ, 538:528–542, August 2000.

[299] N. S. Sugiyama, S. Saito, F. Beutler, and H.-J. Seo. A complete FFT-based decomposition formalism for the redshift-space bispectrum. *ArXiv e-prints*, March 2018.

[300] T. Sunayama, A. P. Hearin, N. Padmanabhan, and A. Leauthaud. The Scale-Dependence of Halo Assembly Bias. MNRAS, 458(2):1510–1516, 2016.

[301] T. Sunayama, N. Padmanabhan, K. Heitmann, S. Habib, and E. Rangel. Efficient Construction of Mock Catalogs for Baryon Acoustic Oscillation Surveys. JCAP, 1605(05):051, 2016.

[302] D. Syer and S. D. M. White. Dark halo mergers and the formation of a universal profile. *arXiv e-prints*, pages astro–ph/9611065, November 1996.

[303] I. Szapudi. Three-Point Statistics from a New Perspective. ApJ, 605:L89–L92, April 2004.

[304] G. Taffoni, L. Mayer, M. Colpi, and F. Governato. On the life and death of satellite haloes. MNRAS, 341:434–448, May 2003.

[305] M. Takada et al. Extragalactic science, cosmology, and Galactic archaeology with the Subaru Prime Focus Spectrograph. PASJ, 66:R1, February 2014.

[306] A. Tasitsiomi, A. V. Kravtsov, S. Gottlober, and A. A. Klypin. Density profiles of lambda-CDM clusters. ApJ, 607:125–139, 2004.

[307] J. E. Taylor. Dark Matter Halos from the Inside Out. *Adv. Astron.*, 2011:604898, 2011.

[308] G. Tormen. The rise and fall of satellites in galaxy clusters. MNRAS, 290:411, 1997.

[309] G. Tormen. The assembly of matter in galaxy clusters. MNRAS, 297:648, 1998.

[310] G. Tormen, F. R. Bouchet, and S. D. M. White. The structure and dynamical evolution of dark matter haloes. MNRAS, 286:865–884, April 1997.

[311] D. Tseliakhovich and C. Hirata. Relative velocity of dark matter and baryonic fluids and the formation of the first structures. Phys. Rev. D, 82(8):083520, October 2010.

[312] S. M. Udrescu, A. A. Dutton, A. V. Macciò, and T. Buck. A deeper look into the structure of ΛCDM haloes: correlations between halo parameters from Einasto fits. MNRAS, November 2018.

[313] K. Umetsu, T. Broadhurst, A. Zitrin, E. Medezinski, D. Coe, and M. Postman. A Precise Cluster Mass Profile Averaged from the Highest-quality Lensing Data. ApJ, 738:41, September 2011.

[314] K. Umetsu and B. Diemer. Lensing Constraints on the Mass Profile Shape and the Splashback Radius of Galaxy Clusters. ApJ, 836:231, February 2017.

[315] K. Umetsu et al. CLASH: Weak-lensing Shear-and-magnification Analysis of 20 Galaxy Clusters. ApJ, 795:163, November 2014.

[316] K. Umetsu, A. Zitrin, D. Gruen, J. Merten, M. Donahue, and M. Postman. CLASH: Joint Analysis of Strong-lensing, Weak-lensing Shear, and Magnification Data for 20 Galaxy Clusters. ApJ, 821:116, April 2016.

[317] A. Vale and J. P. Ostriker. Linking halo mass to galaxy luminosity. MNRAS, 353:189–200, September 2004.

[318] M. Valluri, I. M. Vass, S. Kazantzidis, A. V. Kravtsov, and C. L. Bohn. On Relaxation Processes in Collisionless Mergers. ApJ, 658:731–747, April 2007.

[319] F. C. van den Bosch. The universal mass accretion history of cdm haloes. MNRAS, 331:98, 2002.

[320] F. C. van den Bosch, F. Jiang, A. Hearin, D. Campbell, D. Watson, and N. Padmanabhan. Coming of age in the dark sector: how dark matter haloes grow their gravitational potential wells. MNRAS, 445:1713–1730, December 2014.

[321] F. C. van den Bosch, G. Tormen, and C. Giocoli. The Mass function and average mass loss rate of dark matter subhaloes. MNRAS, 359:1029–1040, 2005.

[322] A. Vikhlinin, A. Kravtsov, W. Forman, C. Jones, M. Markevitch, S. S. Murray, and L. Van Speybroeck. Chandra Sample of Nearby Relaxed Galaxy Clusters: Mass, Gas Fraction, and Mass-Temperature Relation. ApJ, 640:691–709, April 2006.

[323] H. Wang, H. J. Mo, and Y. P. Jing. The distribution of ejected subhaloes and its implication for halo assembly bias. MNRAS, 396:2249–2256, July 2009.

[324] H.-Y. Wang, H. J. Mo, and Y. P. Jing. Environmental Dependence of Cold Dark Matter Halo Formation. MNRAS, 375:633–639, 2007.

[325] J. Wang, J. F. Navarro, C. S. Frenk, S. D. M. White, V. Springel, A. Jenkins, A. Helmi, A. Ludlow, and M. Vogelsberger. Assembly history and structure of galactic cold dark matter haloes. MNRAS, 413:1373–1382, May 2011.

[326] J. Wang and S. D. M. White. Are mergers responsible for universal halo properties? MNRAS, 396:709, 2009.

[327] Y. Wang, X. Yang, H. J. Mo, F. C. van den Bosch, and Y. Chu. The three-point correlation function of galaxies: comparing halo occupation models with observations. MNRAS, 353:287–300, September 2004.

[328] R. H. Wechsler, J. S. Bullock, J. R. Primack, A. V. Kravtsov, and A. Dekel. Concentrations of Dark Halos from Their Assembly Histories. ApJ, 568:52–70, March 2002.

[329] R. H. Wechsler, A. R. Zentner, J. S. Bullock, A. V. Kravtsov, and B. Allgood. The Dependence of Halo Clustering on Halo Formation History, Concentration, and Occupation. ApJ, 652:71–84, November 2006.

[330] A. R. Wetzel, J. D. Cohn, M. J. White, D. E. Holz, and M. S. Warren. The Clustering of Massive Halos. ApJ, 656:139–147, 2007.

[331] A. R. Wetzel and M. White. What determines satellite galaxy disruption? MNRAS, 403:1072–1088, April 2010.

[332] M. White. The Mass Function. ApJS, 143:241–255, December 2002.

[333] S. D. M. White and C. S. Frenk. Galaxy formation through hierarchical clustering. ApJ, 379:52–79, September 1991.

[334] L. L. R. Williams, A. Babul, and J. J. Dalcanton. Investigating the Origins of Dark Matter Halo Density Profiles. ApJ, 604:18–39, March 2004.

[335] A. W. C. Wong and J. E. Taylor. What Do Dark Matter Halo Properties Tell Us about Their Mass Assembly Histories? ApJ, 757:102, September 2012.

[336] H.-Y. Wu, O. Hahn, R. H. Wechsler, Y.-Y. Mao, and P. S. Behroozi. Rhapsody. I. Structural Properties and Formation History from a Statistical Sample of Re-simulated Cluster-size Halos. ApJ, 763:70, February 2013.

[337] X. Xu and Z. Zheng. Dependence of Halo Bias and Kinematics on Assembly Variables. MNRAS, 479(2):1579–1594, 2018.

[338] K. Yamamoto, Y. Nan, and C. Hikage. Analytic halo approach to the bispectrum of galaxies in redshift space. Phys. Rev. D, 95(4):043528, February 2017.

[339] V. Yankelevich and C. Porciani. Cosmological information in the redshift-space bispectrum. *ArXiv e-prints*, July 2018.

[340] J. Yoo, N. Dalal, and U. Seljak. Supersonic relative velocity effect on the baryonic acoustic oscillation measurements. JCAP, 7:018, July 2011.

[341] J. Yoo and U. Seljak. Signatures of first stars in galaxy surveys: Multitracer analysis of the supersonic relative velocity effect and the constraints from the BOSS power spectrum measurements. Phys. Rev. D, 88(10):103520, November 2013.

[342] P. Zarrouk et al. The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample: measurement of the growth rate of structure from the anisotropic correlation function between redshift 0.8 and 2.2. MNRAS, 477:1639–1663, June 2018.

[343] I. Zehavi, S. Contreras, N. Padilla, N. J. Smith, C. M. Baugh, and P. Norberg. The Impact of Assembly Bias on the Galaxy Content of Dark Matter Halos. ApJ, 853:84, January 2018.

[344] M. Zemp. On Physical Scales of Dark Matter Halos. ApJ, 792:124, September 2014.

[345] D. H. Zhao, Y. P. Jing, H. J. Mo, and G. Börner. Mass and Redshift Dependence of Dark Halo Structure. ApJ, 597:L9–L12, November 2003.

[346] D. H. Zhao, Y. P. Jing, H. J. Mo, and G. Börner. Accurate Universal Models for the Mass Accretion Histories and Concentrations of Dark Matter Halos. ApJ, 707:354–369, December 2009.

[347] D. H. Zhao, H. J. Mo, Y. P. Jing, and G. Börner. The growth and structure of dark matter haloes. MNRAS, 339:12–24, 2003.

[348] Z. Zheng, A. A. Berlind, D. H. Weinberg, A. J. Benson, C. M. Baugh, S. Cole, R. Davé, C. S. Frenk, N. Katz, and C. G. Lacey. Theoretical Models of the Halo Occupation Distribution: Separating Central and Satellite Galaxies. ApJ, 633:791–809, November 2005.

[349] G. Zhu, Z. Zheng, W. P. Lin, Y. P. Jing, X. Kang, and L. Gao. The Dependence of the Occupation of Galaxies on the Halo Formation Time. ApJ, 639:L5–L8, March 2006.