THE UNIVERSITY OF CHICAGO

FORWARD MODELING OF OPTICAL OBSERVABLES FOR GALAXY CLUSTER COSMOLOGY

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

DEPARTMENT OF ASTRONOMY AND ASTROPHYSICS

BY

ZHUOWEN ZHANG

CHICAGO, ILLINOIS

JUNE 2024

I dedicate this thesis to Walt Whitman.

Every moment of light and dark is a miracle.

— Walt Whitman, Miracles, *Leaves of Grass*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Earning my PhD was the hardest thing I've done in my life so far. I couldn't have done it without the support of many individuals, some of whom I want to acknowledge here. This cannot possibly be an exhaustive list so please forgive me for the inevitable omissions.

First and foremost I thank my advisor Joshua Frieman. Your patience and support throughout the years in guiding me to become to become both an independent and collaborative researcher is immensely appreciated. I additionally thank members of my thesis committee Chihway Chang, Irina Zhuravleva and Rich Kron for offering me valuable advice in my research direction.

I owe it to many colleagues who at various stages in my career offered me valuable support in my research and career. They include Yuanyuan Zhang, Haoyi (Heidi) Wu, Daisuke Nagai, Arya Farahi, Erwin Lau, Jim Annis, Maria Elidaiana da Silva Pereira, Constantin Payerne, Celine Combet, Tom Diehl, Andrew Hearin, Chunhao To, Matteo Costanzi, Sebastian Bocquet, Eduardo Rozo, Alex Drlica-Wagner, Leslie Rogers, Lindsey Bleem, Bradford Benson, John Carlstrom, Steve Padin and many others.

I thank the friends I made in our department and the wider astronomy community. They are Huanqing Chen, Rostom Mbarek, Emily Gilbert, Samantha Usman, Dhayaa Anbajagane, James Lasker, Ross Cawthon, Phil Mansfield, Tae-Hyeon Shin, Johnny Esteves, Abby Lee, Andresa Campos, Wei Quan, Andrea Bryant, Rebecca Diesing, Lucas Secco, Gabriela Marques, Anowar Shajib, Giulia Giannini, Jason Poh, Conghao Zhou, Matt Kwiecien, Manwei Chen, Nicholas Mehrle and many more.

Outside of astronomy I thank the friends and mentors I came across in different walks of life. Special shout out to members of the UChicago meditation group and Windmill drama club.

Lastly, I thank my parents for their unwavering support and unconditional love.

ABSTRACT

In my thesis, I use state-of-the-art simulations to model and quantitatively resolve several of the important systematics for cluster cosmology. First, I quantify the impact of triaxiality, that is, the ellipsoidal shapes and orientations of clusters, on cluster observables and explore the relation between triaxiality and other major systematics such as miscentering and projection effects. Second, I quantify and develop a model for the previously unexplored covariance between cluster observables. The covariance study divides into two parts, one focused on the intrinsic covariance of cluster signals due to the underlying formation physics, and the second that includes a realistic cluster finder to quantify the impact of extrinsic terms due to observational effects on the covariance. These results have been included in the analysis of current optical surveys as DES and will be useful for stage IV cluster surveys as LSST and DESI.

CHAPTER 1

INTRODUCTION

The leading paradigm of cosmology is the ΛCDM model, or the concordance model. The two main ingredients proposed by this model are cold dark matter and dark energy in the form of the uniform density vacuum energy Λ. During the 80s, observations of the clustering of matter in the universe made evident that the perturbations of the baryon-photon plasma at recombination was not large enough to seed the structure observed in the low redshift universe. Cold dark matter was introduced as a plausible scenario in which non-relativistic dark matter decoupled from radiation at early times and collapsed into deep enough potential wells for baryonic matter to trace, forming the structures we observe today (Davis et al., 1985; Peebles, 1982). Vacuum energy was introduced in the 90s when measurements for a CDM-only universe produced a matter spectrum inconsistent with large scale clustering of galaxies at late times and temperature anisotropies of the Cosmic Microwave Background at early times (Efstathiou et al., 1990; Kofman et al., 1993). The existence of a vacuum-energy like form of dark energy was confirmed when two teams independently demonstrated through the luminosity-distance relation extrapolated to high redshifts using Type IA Supernovae (SNe IA) that the universe was expanding at an accelerating rate (Riess et al., 1998; Perlmutter et al., 1999).

Galaxy clusters are powerful probes of cosmology. Arising from the peaks in the initial density fluctuations of the universe, galaxy clusters have collapsed since then to become the most massive gravitational bound structures in the universe. The number density of clusters is sensitive to the amplitude of the initial density fluctuations. Furthermore, the expansion rate of the universe not only dilutes their number density in the sky but their growth of structure. This double sensitivity to both the geometry of the universe and its growth of structure makes galaxy clusters a unique probe. Measurements of the cluster abundance as a function of cluster mass and redshift can be used to constrain Ω_m, the matter density of the

universe, and σ_8, the matter density fluctuation amplitude, provided that cluster masses can be determined with both high precision and accuracy. The constraints provided by galaxy clusters are nearly orthogonal to constraints by CMB anisotropy, making clusters not only a competitive probe in its own right but a complementary probe when combined with other measurements.

Within the ΛCDM paradigm, a series of theoretical uncertainties need to be modeled. Since the 90s, simulations have improved drastically in their ability to model the number density of halos, known as the halo mass function, as well as halo clustering properties over a wide mass and redshift range. The modeling of these properties are hampered by baryonic effects including radiative cooling, star formation, supernovae and active galactic nuclei (AGN) feedback that suppress the halo-matter power spectrum at small scales and alter the shape of the halo mass function.

Observationally, one cannot directly measure the masses of clusters, but must rather infer them using a mass proxy such as luminosity or the richness estimator in the optical regime, the integrated Compton-Y signal $Y_{\rm SZ}$ in the millimeter regime, or the luminosity (L_X) or temperature (T_X) in the X-ray regime. To make inferences on their masses one has to make certain assumptions on the morphology and dynamical state of clusters. One must also take into consideration the selection function of cluster finders. In the optical regime, some of the main observational systematics include photometric redshift uncertainty, projection effects, triaxiality, miscentering, and cluster galaxy member contamination. These effects have to be accurately modeled to yield accurate and precise cosmology constraints. Such systematics are at a level on par with statistical uncertainties in stage-III surveys as DES, HSC, KiDS-100 and will soon dominate in stage-IV surveys as DESI, LSST, Euclid, Roman, CMB-S4 and eROSITA.

This thesis uses state-of-the-art simulations to forward model systematic biases relating to the weak lensing profile and mass-estimator calibration in the optical regime. The results

of this work have been utilized by DES and LSST-DESC. The thesis is organized as follows:

Chapter 2 provides a basic overview of cosmology theory. Beginning with Einstein's theory of general relativity I introduce basic cosmology concepts and the statistical framework for describing density perturbations. I walk readers through the main eras of ΛCDM universe and offer a summary of the landscape of models beyond ΛCDM.

Chapter 3 offers an introduction to galaxy clusters. I present readers with a summary of the theoretical framework for depicting the formation of galaxy clusters as well as conventions used to describe their density profiles and mass definitions. I then lead readers through a summary of the modeling of the halo mass function and scaling relations which are crucial models for cluster cosmology. Lastly in this chapter, I describe how one can combine the ingredients as laid out in this chapter to yield cluster cosmology constraints.

Chapter 4 aims to quantify the effect of triaxiality bias on weak lensing analyses. We offer analytic templates for the triaxiality bias of observed-richness and lensing profiles. These templates are mapped as corrections to the richness-binned lensing profiles for redMaPPer clusters. The resulting mass bias confirms the DES Y1 finding that triaxiality is a leading source of bias in cluster cosmology. However, the richness-dependence of the bias confirms that triaxiality does not fully resolve the tension at low-richness between DES Y1 cluster cosmology and other probes. We also test if triaxiality is correlated with two other systematics — miscentering and projection effects. Our model can be used for quantifying the impact of triaxiality bias on cosmological constraints for upcoming weak lensing surveys of galaxy clusters.

Chapter 5 quantifies the intrinsic correlation of optical cluster observables. Specifically, we study the covariance between the weak lensing signal and the "true" cluster galaxy number count as measured within a spherical volume that is void of projection effects. By quantifying the impact of this covariance on mass calibration, this work reveals a significant source of systematic uncertainty. Our results reveal a negative covariance at small radial scales

($R \lesssim R_{200c}$) and a null covariance at large scales ($R \gtrsim R_{200c}$). We attribute the difference between our results and the positive bias seen in other works with (mock)-cluster finders to projection effects. By modeling the true richness and lensing signal as multi-(log)-linear equations of secondary halo properties, we provide a quantitative explanation for the physical origin of the negative covariance at small scales.

Chapter 6 uses a realistic mock-cluster finder on a mock LSST catalog to measure the impact of the total covariance that includes projection effects. Because of the ray tracing resolution of the simulation we only extract the lensing signals at large scales. We find that in most bins there exhibits a slight positive covariance between optical observables which is consistent with expectations from projection effects. Propagating the impact onto cluster scaling relations, we find that the large scale covariance can shift the scaling relation parameters by up to $\sim 1\sigma$, roughly the shift when adopting different concentration-mass relations. We find that the impact of the covariance is much greater than the bias introduced from the addition of photometric redshift errors for the source galaxy sample.

Chapter 7 offers final remarks.

CHAPTER 2

THEORETICAL FRAMEWORK OF COSMOLOGY

2.1 General Relativity

Our model of the universe at large scales is based on Einstein's theory of General Relativity. This theory provides a theoretical framework for describing the curvature of spacetime is a pseudo-Riemannian manifold and the the motion of particles within curved spacetime.

Under this framework, the metric tensor $g_{\mu\nu}$ is used to describe the spacetime interval between two (infinitesimal) distance squared between two points:

$$ds^2 = g_{\mu\nu}x^\mu x^\nu. \tag{2.1}$$

In Minkowski space the metric tensor is described as

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{2.2}$$

in which case the spacetime interval can be described by Euclidean geometry, i.e.

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2. \tag{2.3}$$

We will see later on the spacetime interval described in a pseudo-Riemannian manifold in the case of General Relativity as is the case with our universe at large scales. The spacetime interval is considered timelike for $ds^2 > 0$, spacelike for $ds^2 < 0$ and null for $ds^2 = 0$. The proper time τ which characterizes the time elapsed for an observer following the spacetime

trajectory is related to ds^2 by

$$\tau = \frac{1}{c}\int \sqrt{ds^2} = \frac{1}{c}\int \sqrt{g_{\mu\nu}\frac{dx^\mu(\lambda)}{d\lambda}\frac{dx^\nu(\lambda)}{d\lambda}}d\lambda, \tag{2.4}$$

where we have specified the four-vector coordinate x^μ by some parameter λ. To obtain the geodesic one can use calculus of variation methods to find the solution to Equation (2.4) that minimizes the proper time. This leads us to the geodesic equation that governs the trajectory for a photon or a massless particle traveling at the speed of light:

$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\lambda}\frac{dx^\sigma}{d\lambda} = 0, \tag{2.5}$$

where $\Gamma^\mu_{\rho\sigma}$ are the Christoffel symbols related to the metric tensor by the non-linear relation:

$$\Gamma^\sigma_{\mu\nu} = \frac{1}{2}g^{\sigma\rho}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu}). \tag{2.6}$$

We can contrast General Relativity with Newtonian physics. In the Newtonian paradigm, the motion of objects is governed by the Possion equation

$$\nabla^2\Phi = 4\pi G\rho, \tag{2.7}$$

with G denoting the gravitational constant, ρ the matter density and Φ the gravitational potential. In flat Euclidean space the components of the Christoffel symbol in Equation (2.6) vanish and we are left with a linear geodesic equation $\frac{d^2x^\mu}{d\lambda^2} = 0$. In curved spacetime the geodesic equation in (2.5) are complicated non-linear differential equations with respect to the metric tensor $g_{\mu\nu}$. The relation of matter and curved spacetime in General Relativity can be summarized by what's known as Einstein's field equations:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + g_{\mu\nu}\Lambda = \frac{8\pi G}{c^4}T_{\mu\nu}, \tag{2.8}$$

in which $R_{\mu\nu}$ is the rank-2 Ricci tensor, R the scalar curvature, $T_{\mu\nu}$ the rank-2 energy-momentum tensor and Λ the cosmological constant.

2.2 FLRW Universe

The cosmological principles states that at large enough scales ($\gtrsim 150$ Mpc) the universe looks homogeneous and isotropic. Under these two assumptions, an expanding/contracting curved universe can be described by the Friedman-Lemaître-Robertson-Walker (FLRW) metric:

$$ds^2 = -c^2dt^2 + a^2(t)\big[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\big], \tag{2.9}$$

where a(t) is the scale factor, the coordinates (r, θ, ϕ) are comoving coordinates with respect to the rest frame, and $k = -1, 0, 1$ describes the curvature of the universe. The case $k = -1$ describes an open universe characterized by a three-hyperboloid manifold, $k = 0$ a flat universe, and $k = 1$ a closed universe characterized by a three-sphere manifold.

In hyperspherical coordinates the FLRW metric can be written as:

$$ds^2 = -c^2dt^2 + a^2(t)\big[d\chi^2 + \chi^2(d\theta^2 + \sin^2\theta d\phi^2)\big], \tag{2.10}$$

where χ is:

$$\chi(r) = \begin{cases} \sqrt{k}^{-1}\sin(\sqrt{k}r) & \text{if } k = 1 \\ r & \text{if } k = 0 \\ \sqrt{k}^{-1}\sinh(\sqrt{k}r) & \text{if } k = -1 \end{cases} \tag{2.11}$$

We may describe the matter in the universe at these scales as a perfect fluid that is isotropic in its rest frame and has no heat flow or viscosity. The components of the energy-momentum tensor $T_{\mu\nu}$ of this fluid are fully specified by its density ρ and pressure p. In terms of its rest frame four-velocity $\mu^\mu = \{1, 0, 0, 0\}$ the energy-momentum tensor for a

perfect fluid is

$$T_{\mu\nu} = (p + \rho)\mu_\mu\mu_\nu + pg_{\mu\nu}. \tag{2.12}$$

Substituting Eq. (2.12) into the right hand side of Eq. (2.8) and the components of the FLRW metric into the left hand side of Eq. (2.8) we arrive at a set of two differential equations for the scale factor $a(t)$ known as the Friedman equations:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3} \tag{2.13}$$

$$\frac{\ddot{a}}{a} = \frac{-4\pi G}{3c^2}(\rho c^2 + 3p) + \frac{\Lambda}{c^2}. \tag{2.14}$$

The quantity $H \equiv \frac{\dot{a}}{a}$ is the Hubble parameter which governs the rate of the expansion of the universe. From measurements of the Cosmic Microwave Background (CMB) by Planck, the Hubble constant H_0 at present day is determined to be $H_0 = 67.4 \pm 0.5$ km s^{-1}Mpc^{-1} (Planck Collaboration et al., 2020).

2.3 Cosmological Redshift and Distances

The cosmological redshift can be directly derived from the FLRW metric. Along the photon path $ds = d\theta = d\phi = 0$, and hence for an incoming photon

$$\chi = \int_{t_E}^{t_R} \frac{cdt}{a(t)} = \int_0^{r_E} \frac{1}{(1 - kr^2)^{1/2}} dr. \tag{2.15}$$

This quantity in the metric is in comoving units that are independent of the overall scale $a(t)$ of the universe. If at a second (infinitesimally separated) instant δt_E the emitter seconds a

second light pulse received at $t_R + \delta t_R$ we have

$$\int_{t_E+\delta t_E}^{t_R+\delta t_R} \frac{cdt}{a(t)} = \int_0^{r_E} \frac{1}{(1-kr^2)^{1/2}} dr = \int_{t_E}^{t_R} \frac{cdt}{a(t)}. \tag{2.16}$$

If it follows that δt_E and δt_R are sufficiently small that $a(t)$ can be taken as constant, it follows from the equality that

$$1 + z = \frac{a(t_R)}{a(t_E)}. \tag{2.17}$$

Conventionally we define $a_0 = 1$ at the present day and hence $1 + z = 1/a$.

The luminosity distance is an operational definition that characterizes the relation between the absolute luminosity and flux received on a curved hyper-sphere. Analogous to Euclidean geometry its definition is given by

$$d_L = \Big(\frac{L}{4\pi F}\Big)^{1/2}. \tag{2.18}$$

To obtain the flux, we recognize that the proper area of a curved hyper-sphere is $A = 4\pi\chi^2(r)$. The flux of the photons is firstly diminished by one power of the scale factor by the redshift in photon frequency from the time of emission t_E and another by the reduction in photon arriving rate. Hence the observed flux is given by:

$$F = \frac{L}{4\pi\chi} a(t_E). \tag{2.19}$$

The luminosity distance is evaluated as:

$$d_L = \chi/a(t) \tag{2.20}$$

We then introduce the angular diameter distance as another operational definition. A source with proper diameter ℓ in Euclidean geometry would subtend an angle $\Delta\theta$ at an

distance d_A. Based on this analogy we can derive the angular diameter distance using the FLWR metric assuming r and ϕ are constant along the photon path in the transverse plane. From which the proper distance is $\ell = a(t)\chi\Delta\theta$. The angular diameter distance after dividing by $\Delta\theta$ is then:

$$d_A = a(t)\chi \tag{2.21}$$

2.4 Evolution of the Scale Factor

As a matter of convenience and convention it is desirable to re-scale the densities by the critical density

$$\rho_{\text{crit}} = \frac{3H^2}{8\pi G}. \tag{2.22}$$

Dividing by the critical density of the universe, one can now describe the densities by their densities parameters

$$\Omega_i(t) \equiv \frac{\rho_i}{\rho_{\text{crit}}} = \frac{8\pi G}{3H^2(t)}\rho_i(t). \tag{2.23}$$

In this way the first of the Friedman equations can be re-written in its more conventional form of

$$H^2(t) = \frac{8\pi G}{3}\sum \Omega_i(t) - \frac{c^2 k}{a^2} \tag{2.24}$$

The density contributions of the universe come from radiation which is a sum of photons and neutrinos ($\rho_r = \rho_\gamma + \rho_\nu$), matter as a sum of baryonic matter and dark matter ($\rho_m = \rho_b + \rho_{\text{dm}}$), vacuum density $\rho_\Lambda = \frac{\Lambda c^2}{8\pi G}$. In some instances one may characterize the curvature with the curvature density $\rho_k = \frac{-3kc^2}{8\pi G a^2}$ for conformity in notational convention though it is not strictly speaking a physical matter/energy density . Rescaling the densities in terms of their density parameters we find that

$$\Omega = \Omega_r + \Omega_m + \Omega_\Lambda = 1 - \Omega_k. \tag{2.25}$$

In most standard cosmology models we consider $\Omega_k = 0$. The flatness hypothesis comes from the cosmic Inflation theory (Guth, 1981) in which the universe at around 10^{-36} seconds after the Big Bang undergoes rapid exponential expansion such that regardless of the initial value of Ω it will soon reach a value close to 1, or in other words $\Omega_k \approx 0$. $\Omega = 1$ or when the density is equal to the critical density ρ_{crit} suggest a flat universe, $\Omega > 1$ is a closed universe, and $\Omega < 1$ an open universe.

We now derive the evolution of the different density components under this convention. To do so we impose energy-momentum conservation. This requires

$$\nabla_\mu T^{\mu\nu} = 0, \tag{2.26}$$

which leads to the continuity equation

$$\dot{\rho} + (\rho + \frac{p}{c^2})\frac{3\dot{a}}{a} = 0. \tag{2.27}$$

We can further remove a degree of freedom by imposing an equation of state that connects the pressure with density

$$p = \omega\rho c^2. \tag{2.28}$$

Substituting the equation of state equation into Eq. (2.27) and integrating with respect to the scale factor we obtain

$$\rho_i \propto a^{-3(1+\omega_i)}, \tag{2.29}$$

which by sustituting into Eq. (2.24) we find

$$\Omega_i = \Omega_{i,0}\Big(\frac{H_0}{H}\Big)^2 a^{-3(1+\omega_i)}, \tag{2.30}$$

where $\Omega_{i,0}$ and H_0 are respectively the density parameter and Hubble parameter at present

day, and by which we have generalized the density parameter to include the curvature term. The equation of state parameter is $\omega_r = 1/3$ for radiation, $\omega_m = 0$ for non-relativistic matter, $\omega_k = -1/3$, $\omega_\Lambda = -1$ for vacuum energy under the concordance model. The evolution of the scale factor can now be re-written in a more elegant form as

$$\left(\frac{da}{dt}\right)^2 = H_0^2\left(\Omega_{r,0}a^{-4} + \Omega_{m,0}a^{-3} + \Omega_{k,0}a^{-2} + \Omega_{\Lambda,0}\right). \tag{2.31}$$

The evolution of the scale factor may be solved numerically by integrating Eq. (2.31) with constraints on the present day energy budget constrained from experiments. At different eras in the evolution history of the universe one may make simplifying assumptions such that the evolution of the scale factor may be derived in an analytic form. We present the different eras in the section below.

2.5 Statistics of the Density Field

The FLRW metric derived under the cosmological principle is a good approximation to describing the evolution of the universe at very large scales when we can assume homogeneity. At small scales fluctuations in matter densities exist. They are seeded by small fluctuations that can be described using linear perturbation theory which grow over time. This section summarizes the statistical properties of the large scale structure of the universe.

We define the overdensity with the symbol δ, defined as

$$\delta(\mathbf{x}, t) = \frac{\rho(\mathbf{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}, \tag{2.32}$$

where $\rho(\mathbf{x}, t)$ is the mass density localized in time and space, and $\bar{\rho}(t)$ the mean matter density in all space at a given time. Under standard inflationary models (see Section 2.6) the primordial density fluctuations are Gaussian, meaning that δ satisfies the probability

distribution

$$P(\delta)d\delta = \frac{1}{\sqrt{2\pi}\sigma}\exp\frac{-\delta^2}{2\sigma^2}d\delta, \tag{2.33}$$

set initially around $\delta \approx 10^{-5}$, and remain Gaussian up until recent times as $\delta \gtrsim 1$. Here σ is the mean smoothed matter fluctuation which we will derive later in this section.

For a Gaussian density fluctuation, the matter distribution can be described entirely by the two point correlation function

$$\xi(r) = \langle\delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r})\rangle_{\mathbf{x}}, \tag{2.34}$$

with the subscript $\mathbf{x}$ denoting that the average is taken over all space. This two point correlation function describes the excess in matter density due to matter clustering at a separation of r. Because of the assumed isotropy of the universe we drop the angular dependence of r on the left hand side of the equation such that the dependence only depends on the distance $r = |\mathbf{r}|$.

We can transform real space coordinates to Fourier space wavenumbers $\vec{k}$

$$\delta_{\vec{k}} = \frac{1}{\sqrt{V}}\int \delta(\vec{r})e^{i\vec{k}\cdot\vec{r}}d^3r, \tag{2.35}$$

where V is the comoving volume. We describe the power spectra $P(k)$ as the Fourier transform of the two-point correlation function

$$P(k) = \langle\delta_{\vec{k}}\delta^*_{\vec{k}}\rangle \tag{2.36}$$

$$= \frac{1}{V}\int\int \xi(r_{12})e^{-i\vec{k}\cdot\vec{r_1}}e^{i\vec{k}\cdot\vec{r_2}}d^3\vec{r}_1 d^3\vec{r}_2, \tag{2.37}$$

which after trivially integrating over $\int d^3\vec{r}_2 = V$ and switching to polar coordinates for the

integration of the volume element we get

$$P(k) = \frac{4\pi}{k} \int_0^\infty \xi(r) r \sin(kr) dr, \tag{2.38}$$

where once again due to the isotropy of the universe $P(k)$ does not depend on the directionality of $\vec{k}$ but only its norm k.

We can inverse-Fourier transform $P(k)$ at the limit $r \to 0$ to get the zero-lag correlation which is simply the variance of the density fluctuations

$$\xi(0) = \lim_{r\to 0} = \frac{1}{2\pi^2} \int P(k) \lim_{r\to 0} \frac{\sin(kr)}{kr} k^2 dk \tag{2.39}$$

$$\equiv \int \Delta^2(k) d\ln k, \tag{2.40}$$

where we have defined the quantity

$$\Delta^2(k) \equiv \frac{k^3 P(k)}{2\pi^2} \tag{2.41}$$

as the unitless logarithmic band power, which gives the contribution to the variance in density per wavenumber. In practice this power spectrum is grainy at small scales, so we often apply a window function to compute a smoothed average of the power spectrum. A popular choice is the top-hat window function

$$W_{\mathrm{TH}}(r, R) = \frac{1}{(4\pi/3)R^3} H(R - r), \tag{2.42}$$

where $H(R - r)$ is the Heaviside step function for a smoothing radius of R. The window function needs to be convoled with the overdensity δ in real space. Fortunately, convolution in real space translates to multiplication in Fourier space. In Fourier space the window

function is

$$W_{\rm TH}(k,R) = 3\frac{j_1(k,R)}{kR}, \tag{2.43}$$

where j_1 is the Bessel function of order 1. The smoothed power spectrum is

$$P(k,R) = |W(k,R)|^2 P(k). \tag{2.44}$$

Applying this to the zero-lag two point correlation we arrive at the mean amplitude squared of the smoothed matter fluctuation

$$\sigma^2(R) = \int \Delta^2(k)\Big(\frac{j_1(k,R)}{kR}\Big)^2 d\ln k. \tag{2.45}$$

An important quantity in cosmology is the smoothed mean matter fluctuation at 8 Mpc h^{-1} for the present day:

$$\sigma_8 = \sigma(R = 8 \text{ Mpc } h^{-1}, z = 0). \tag{2.46}$$

The scale size is chosen as cosmologists first studied the clustering of galaxies around galaxy clusters which are around $5-10$ Mpc h^{-1} in physical scale. The quantity σ_8 can be thought of one way to normalize the amplitude of the power spectrum.

Having presented the formalism for describing matter density fluctuations, we can derive the initial power spectrum using Newtonian perturbation theory. To do so, we start with the three fundamental equations governing the evolution of fluids

$$\frac{D\rho}{Dt} + \rho(\nabla\cdot\boldsymbol{\mu}) = 0 \qquad \text{(Continuity equation)} \tag{2.47}$$

$$\nabla^2\Phi = 4\pi G\rho \qquad \text{(Poisson equation)} \tag{2.48}$$

$$\frac{D\boldsymbol{\mu}}{Dt} = -\frac{\nabla\rho}{\rho} - \nabla\Phi \qquad \text{(Euler equation)} \tag{2.49}$$

in which the symbol $\frac{D}{Dt}$ represents the Lagrangian derivative and μ the velocity of the fluid

in comoving coordinates. We perturb the steady state solution by switching from comoving to physical coordinates and by introducing the peculiar velocity of the object in relation to the bulk motion of the fluid. Invoking the first law of thermodynamics ($TdS = dU + pdV$) for isentropic (adiabatic) initial conditions ($\nabla S = 0$) we arrive at a 2nd order differential equation

$$\frac{\partial^2 \delta_k}{\partial t^2} + 2\frac{\dot{a}}{a}\frac{\partial \delta_k}{\partial t} = -\omega^2 \delta, \tag{2.50}$$

where

$$\omega^2 = (k^2 - k_J^2)\frac{c_s^2}{a^2}, \tag{2.51}$$

which defines the characteristic Jean's length

$$\lambda_J \equiv \frac{2\pi a}{k_J} = c_s \sqrt{\frac{\pi}{G\bar{\rho}}}, \tag{2.52}$$

where the speed of sound $c_s \equiv \left(\frac{\partial P}{\partial \rho}\right)^{1/2}$. The solution to this differential equation only permits oscillating modes for modes with $k > k_J$ from an effect known as Jean's instability and permits growing modes for $k < k_J$. Because the speed of sound is close to the speed of light ($c_s \approx c/\sqrt{3}$) before recombination (see different eras in Section 2.6) the Jean's length is too large for modes within the horizon to collapse. Only after radiation-matter equality does the speed of sound fall enough to allow growing modes for modes within the horizon distance. We characterize the growth of structure by the linear growth factor $D(a) = \frac{\delta(a)}{\delta(1)}$, which for a $\omega_\Lambda = -1$ universe takes on the form

$$D(a) \propto \frac{H(a)}{H_0} \int_0^a \frac{da'}{[\Omega_m/a' + \Omega_\Lambda a'^2 + (1 - \Omega_m - \Omega_\Lambda)]^{3/2}}. \tag{2.53}$$

The linear growth factor only grows logarithmically with time during the radiation dominated era and as $t^{3/2}$ during matter domination and as e^{2Ht} during dark energy dominated era.

The initial power spectrum $P_I(k)$ was speculated independently by Harrison, Zel'dovich,

and Peebles (Harrison, 1970; Zeldovich, 1972; Peebles & Yu, 1970) to be $P_I(k) \propto k^n$, with the spectral index $n \approx 1$. If n is much greater than 1 too many black holes would form too early, and n much less than 1 would predict the overabundance of superclusters and voids. From Inflation (see Section 2.6) this value is predicted to be $n = 1 - 6\epsilon + 2\eta + \mathcal{O}(\epsilon^2, \eta^2)$ to be close but slightly less than 1, where ϵ and η are called slow-roll inflationary parameters. The latest constraints from Planck yield a spectra index of $n = 0.9654 \pm 0.0042$ (Planck Collaboration et al., 2020).

The power spectrum evolves over time from the initial power spectrum as

$$P(k, a) = P_I(k) D^2(a) T^2(k, a). \tag{2.54}$$

The transfer function $T(k, a)$ is needed to modulate the power spectrum as different modes enter the horizon distance at different scales during different eras of the universe. Computing $T(k, a)$ requires solving the Boltzmann equation in a perturbed FLRW metric. Asymptotically speaking $T(k, a) \approx 1$ during the matter dominated era as for modes that entered the horizon scale after matter-radiation equality due to (near) scale invariance predicted by inflation. The characteristic horizon distance at the time of equality is $L_0 \approx 12(\Omega_m h^2)^{-1}$ Mpc ≈ 100 Mpc. Modes that entered the horizon during the radiation dominated era were stunted due to the logarithmic growth of the linear growth factor and the transfer function is suppressed by a factor of k^{-2}. Asymptotically then we have that $P(k) \propto k^n$ for $k \ll 1/L_0$ and $P(k) \propto k^{n-4}$ for $k \gg 1/L_0$.

2.6 ΛCDM Cosmology

The leading paradigm of cosmology is ΛCDM (Λ for dark energy, CDM for cold dark matter). In this model the universe started from the Big Bang at 13.8 billion years ago from a singularity at a very hot and dense state. As the name states, the two main assumption

of the ΛCDM model is the presence of cold dark matter and vacuum energy Λ, in addition to ordinary matter (e.g. leptons, quarks, hadrons) that in astronomy we denote as baryonic matter.

Without dark matter gravity from ordinary matter along would not be strong enough for the universe to form overdensities as galaxies and galaxy clusters. The first evidence of dark matter came from Fritz Zwicky who inferred the mass of the Coma cluster by the motion of its galaxy members through the virial theorem and contrasted it with the total mass from counting the number of galaxies (Zwicky, 1937) , where he concluded that the difference between these two mass estimates is from an unknown substance he coined dark matter; later Vera Rubin fitted observations of the rotation curves of galaxies with inferred mass profiles and found that dark matter should exceed baryonic mass by a factor of six to one (Rubin et al., 1980). Depending on the travel speed of dark matter when decoupled from baryonic matter we can categorize dark matter into hot dark matter (HDM), warm dark matter (WDM) and cold dark matter (CDM). Our measurements of the power spectrum prefers the CDM model. In the HDM model, dark matter traveling at relativistic speeds will quickly escape low mass density fluctuations. The only structures that survive are high mass objects with $M \gtrsim 10^{15} M_{\odot}$ and smaller structures forming inside it in a top-down hierarchical model; warm dark matter is an intermediate category that also suppresses the formation of structure at small scales; CDM models, because of the small travel speed of dark matter forms structure in a bottom-up hierarchy that is consistent with measurements of the power spectrum.

Dark energy is proposed as a vacuum energy with negative pressure to explain the accelerated expansion of the universe. It was first proposed by Einstein as an additional term to the Einstein field equations to maintain a static universe. As Hubble in 1929 discovered that the universe was expanding by showing that the redshift of galaxies was proportional to their distances away from Earth (Hubble, 1929). The idea of a static universe fell out of the favor

and Einstein remarked the cosmological constant his "greatest blunder." In the 90s, cosmologists revived the cosmological constant to explain the discrepancy between observations of large scale structure with an Einstein-de Sitter model $\Omega_m = 1$ with no dark energy as well as the uncomfortablely high observed expansion age $H_0 t_0$ compared to de-Sitter models due to cosmic acceleration. Two teams Riess and Perlmutter (Riess et al., 1998; Perlmutter et al., 1999) used high redshift type IA supernovae (SNe IA) to measure the Hubble diagram to much farther distances than was possible before. They showed that high redshift SNe IA were dimmer than was expected from a decelerating universe and yielded constraints on Ω_Λ consistent with an accelerating universe. A very large discrepancy still exists today known as the cosmological constant problem (Weinberg, 1989) is that were dark energy vacuum energy predicted from quantum field theory, the energy density inferred from cosmological probes and predicted from quantum field theory differ by 120 orders of magnitude. Invoking the hypothetical symmetry between fermions and bosons in a supersymmetric (SUSY) world would reduce this discrepancy to 60 orders of magnitude, one that is still embarrassingly large.

Here we summary some of the crucial periods during of the expansion history of the universe:

Inflationary era ($t = 10^{-36} - 10^{-32}$ s, $z =$?) At a period $10^{-36} - 10^{-32}$ seconds after the Big Bang the universe underwent a period of what's known as Inflation, during which the energy budget was dominated by a cosmological-constant type of vacuum energy. In a simplest single-scalar-field inflationary model the energy budget is described by a homogenuous scalar field ϕ. Comparing the scalar-field with the energy-momentum tensor of a perfect fluid its density and pressure can be described as

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi) \tag{2.55}$$

$$p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi), \tag{2.56}$$

where $\frac{1}{2}\dot{\phi}^2$ can be thought of as the effective kinetic energy term and $V(\phi)$ the potential term. Substituting the pressure and density into the Friedmann equations we arrive at the Euler-Lagrange equation for the scalar field:

$$\ddot{\phi} + 3H\dot{\phi} + V(\phi) = 0. \tag{2.57}$$

There is no fundamental theory that describes the form of the potential $V(\phi)$ so it is taken arbitrarily so long as it solves the three outstanding problems of the standard Big Bang model — the horizon problem, flatness problem and monopole problem, as we will discuss here. The horizon problems comes from the uniformity of the temperature in the CMB which suggests that particles came into causal contact to form a thermal bath. Under the standard Big Bang models, patches of sky above roughly 2 degrees apart should be larger than the horizon distance at the time of last scattering. The flatness problem states how the universe is fine tuned to have an energy density $\Omega \approx 1$ or in other words $\Omega_k \approx 0$ when $\Omega = 1$ is an unstable critical point, such that in most solutions to the Friedmann equation when Ω is not close enough to unity the universe will quickly collapse or rapidly expand and cool within the first second of the Big Bang. Lastly comes the monopole problem which is the lack of evidence for magnetic monopoles when many theories in particle physics predicts its existence. We will discuss how inflation can simultaneously solve all three problems.

One condition that the potential energy has to meet is to be flat enough for enough expansion to occur and has to reach a minimum that marks the end of inflationary period. The first of these conditions is called the slow-roll approximation. Using the slow-roll approximation, at the beginning of inflation $V(\phi) \gg \frac{1}{2}\dot{\phi}^2$, and the equation of state $\omega \equiv \frac{\rho_\phi}{p_\phi} \approx -1$. Inflation acts as an effective cosmological constant that drives the scale factor to expand as

$$a(t) \propto \exp(Ht), \tag{2.58}$$

causing the universe to expand by a factor of over 10^{30} over a very short amount of time.

This rapid expansion solves the horizon problem as pre-Inflation the typical comoving scales of the universe satisfy $\lambda < 1/\mathcal{H}$ for the universe to be in thermal equilibrium. The initial quantum fluctuations pre- and during Inflation seed the fluctuations in the temperature of the CMB and through perturbations under gravity as the universe grew seeded the structures of the universe today such as galaxy clusters. During inflation many modes quickly bacame super-horizon scale. At the end of inflation some of the modes re-entered horizon scales, forming the clustering of matter and radiation we see in the universe today.

Inflation solves the flatness problem as the rapid expansion during Inflation forces Ω to be extremely close to unity regardless of its value pre-Inflation.

The monopole problem is solved as when the universe reached the end of Inflation $V(\phi) = 0$, $\omega = 1$ and the potential energy is converted into quark-gluon plasma and leptons in a process called reheating. During this process the temperature does not get hot enough to generate unwanted thermal relics as magnetic monopoles, causing their densities to dilute to negligible levels.

Radiation Dominated Era (z = ? - 3500) Coming out of Inflation the universe was initially dominated by radiation. Due to an unknown process called "baryogenesis" there is a slight preference of particles over anti-particles formed, and this excess in abundance "freezes out" as the expansion rate of the universe exceeds the pair annihilation rate. Neutrinos were first to decouple from the thermal bath at around $t \sim 1$ s after the Big Bang, followed by electrons from their position pairs to form $e^- + e^+ \rightleftharpoons \gamma + \gamma$ at around $t \sim 10$ s; Big Bang necleosynthesis took place at around $t \sim 20$ minutes when proton-neutron interactions fell out of equilibrium and neutrons either decayed into protons or fused with them to form helium-4, deuterium and trace amounts of higher

elements.

The energy budget of the universe during this era can be approximated by $\Omega \approx \Omega_r$. Substituting this into the Friedmann equation (Eq. (2.24)) using the equation of state for radiation we find that at this era the scale factor grows as $a(t) \propto t^{1/2}$.

Matter Dominated Era (z=3500 - z=0.4) At matter domination $\Omega \approx \Omega_m$ the scale factor grows as $a(t) \propto t^{2/3}$. At around $z \sim 1400$ the temperature cooled down enough to slow down the reaction $e^- + p \rightleftharpoons H + \gamma$ for neutral hydrogen to form in a period known as recombination. At $z \sim 1100$ Thomson scattering $e^- + \gamma \rightleftharpoons e^- + \gamma$ becomes inefficient enough that the mean free path of the photons exceeds the size of cosmological scales in a period called last scattering that formed the Cosmic Microwave Background. Between last scattering to $z \sim 30$ in what is known as the "Dark Ages", structures seeded by quantum fluctuation during Inflation began to clump under the linear perturbation regime but they have not collapsed to form stars and galaxies. From $z \sim 30 - 10$ in a period known as reionization the first stars started to form, reionizing the intergalactic medium.

Relevant to this thesis work, the first proto-galaxy-clusters formed at around $z \sim 2$. Due to the bottom-up hierarchical formation of halos in the cold dark matter paradigm large structures as galaxy clusters were late to form. Proto-clusters have collapsed but are in evolutionary states that are not fully virialized until $z \sim 1$ for most clusters. Chapter 2 will give a more detailed overview of galaxy clusters.

Dark Energy Dominated Era (z = 0.4-0.0) At around z=0.4 the universe energy budget reached a point when dark energy started to dominate. The negative pressure of dark energy is responsible for the cosmic acceleration $\ddot{a} > 0$ we observe in the universe today. The scale factor grows as $a(t) \propto \exp\left(\sqrt{\Lambda/3}t\right)$ for a dark energy dominated universe. According to ΛCDM the universe expansion will continue to accelerate resulting in a

Big Freeze scenario. Other models of dark energy such as quintessence propose a scalar field that alters the equation of state ω of dark energy over time that could result in the universe collapsing on itself in a scenario known as the Big Crunch.

2.7 Beyond ΛCDM

The ΛCDM model has been the canonical model for cosmology in the past few decades. While successful in many regards, the model currently runs into a disprepancy between late-time and early-time probes in constraining H_0 and $S_8 \equiv \sqrt{\Omega_m/0.3}\sigma_8$.

The H_0 tension refers to the discrepancy in measured H_0 values from early-time and late-time probes. Early time probes as measurements of the CMB temperature anisotropy by Planck (Planck Collaboration et al., 2020) reports a $H_0 = 67.27 \pm 0.60$ km/s/Mpc. Baryonic Acoustic Oscillation (BAO) with prior on the baryonic density derived from deuterium measurements from Big Bang Nucleosynthesis (BBN) constraints when combined with galaxy clustering and weak lensing also yield low values of H_0 (Krause et al., 2017; Troxel et al., 2018; Abbott et al., 2018). In contrast, late time probes as the Cepheid distance scale from the SH0ES Collaboration report values of $H_0 = 74.03 \pm 1.42$ km/s/Mpc (Reid et al., 2019). The result is consistent with other late type probes as time delay from strong lensing (Wong et al., 2019). We observe a 4.4σ tension inferred by the SH0ES collaboration and Planck. This tension may be alleviated by using the Tip of the Red Giant (TRGB) distance scale (Freedman, 2021) which suggests that the H_0 tension may be due to systematics with local distance ladder measurements. The nature of this tension is not yet known and it is open to a new physics interpretation.

Adding to the confusion is the S_8 tension. Combined measurements from galaxy weak lensing measurements as Kids-1000 and DES Y3 yield $S_8 = 0.769 \pm 0.016$ (Abbott et al., 2023). The value obtained from Planck is $S_8 = 0.834 \pm 0.016$, at a 2.9σ tension with late-stage probes (Planck Collaboration et al., 2020). The eROSITA survey for galaxy clusters

yielded $S_8 = 0.86 \pm 0.01$ which is consistent with CMB measurements but at a 3σ tension with other late-time probes (Ghirardini et al., 2024). This hints that the S_8 tension may be in due to uncertainties in the intrinsic alignment of galaxies on shape measurements, uncertainties in the non-linear power spectrum and baryonic feedback processes that effect cosmic shear (Mandelbaum, 2018; Huterer, 2023) that do not require extensions to ΛCDM.

To resolve these tensions a wide range of new physics models beyond ΛCDM have been proposed. Many of these introduce Dark Energy set in at early epochs of the universe or vary the equation of state of Dark Energy from a constant $\omega = -1$ (Karwal & Kamionkowski, 2016; Guo et al., 2019; Lin et al., 2019; Valentino et al., 2020). Others have proposed increasing the relativistic degrees of freedom at recombination, parameterized by the number of equilavent light neutrino species $N_{\rm eff}$ (Mangano et al., 2005; de Salas & Pastor, 2016; Akita & Yamaguchi, 2020). Another solution is Modified Newtonian Dynamics (MOND) models that modify gravity at different redshifts or scales (Mangano et al., 2005; Rossi et al., 2019; Ballardini et al., 2020; Akita & Yamaguchi, 2020; Valentino et al., 2020). Modified recombination histories have been examined as a potential solution as an earlier recombination redshift would infer a smaller sound horizon that is compatible with a larger Hubble constant (Hart & Chluba, 2017; Chiang & Slosar, 2018; Jedamzik & Pogosian, 2020; Bose & Lombriser, 2021). Most of the extension to ΛCDM cannot simultaneously resolve the H_0 and S_8 tensions, leaving these discrepancies as open problems in cosmology. We refer to Di Valentino et al. (2021); Abdalla et al. (2022) for complete reviews of this problem.

CHAPTER 3

GALAXY CLUSTERS AND COSMOLOGY

3.1 Formation of Galaxy Clusters

Galaxy clusters are formed during the non-linear perturbation regime when $\delta \gtrsim 1$. The simplest of the non-linear collapse models is the top-hat spherical model in an Einstein-de Sitter universe ($\Omega_m = 1$). In this model the evolution of the collapse is fully specified by a top-hat radius $R(t)$ that satisfies the equation of motion for a shell at radius R free falling from the gravitational acceleration of a uniform sphere of mass $M = \frac{4\pi}{3}(1+\delta_i)\bar{\rho}R_i^3$:

$$\frac{d^2R}{dt^2} = -\frac{GM}{R^2} = -\frac{4\pi G}{3}\bar{\rho}(1+\bar{\delta})R. \tag{3.1}$$

By its parametric solution the equation can be solved analytically. Its solution includes a turnaround time $t_{\rm ta}$ when the perturbation is at its largest scale at the onset of collapse, and a collapse time $t_{\rm coll}$ which under the top-hat model is twice the turnaround time. Taken literally, due to the runaway gravitational collapse the radius at $t_{\rm coll}$ is $r \to 0$ which clearly is not the case. This unphysical instability is resolved due to a process of violent relaxation as described by Lynden-Bell in which dark matter inside the cluster fragmented into subunits and its tidal effects of neighboring perturbations resulted in the dynamic motion of mass creating a non-uniform density profile. Clusters in dynamical equilibrium satisfy the Virial theorem:

$$K + \frac{1}{2}V = 0, \tag{3.2}$$

where K is the kinetic energy and V the gravitational potential energy. The density contrast between the virialized cluster and the background density for an EdS universe $\Delta_{\rm vir} \equiv \rho_{\rm coll}/\rho_{\rm b} = 18\pi^2 \approx 177$. The density contrast for models with $\Omega_m < 1$ is larger as initial density perturbations δ_i of the same value have a larger initial radius and therefore

take longer to collapse. Simulations based studies (Bryan & Norman, 1998) have provided model fits of the density contrast for cosmologies with varying Ω_m and Ω_Λ.

We can use linear perturbation theory to extrapolate the expected overdensity at the time of collapse $\delta_c = \delta_{\rm ta} D_+(t_{coll})/D_+(t_{ta})$. The linear growth factor $D_+(a) \propto a \propto t^{2/3}$ during the matter dominated era, and from $t_{\rm coll} \approx 2t_{\rm ta}$ in the top-hat isotropic model, we have $\delta_c \approx 1.68$, a value that's independent of cosmology up at the $1-2\%$ level. This is because the cosmology dependence of $\delta_{\rm ta}$ and $D_+(a)$ roughly cancel out (Percival, 2005). Clusters clearly reside in the non-linear regime as their actual density contrast at the time of their collapse far exceeds that predicted by linear perturbation. This value δ_c is nonetheless important for determining the abundance of halos as only objects with density contrasts $\delta > \delta_c$ can collapse into gravitationally bound halos.

The peak height ν is defined as the critical overdensity over the density fluctuation σ at the radius of the cluster:

$$\nu = \frac{\delta_c}{\sigma(z, R)}, \tag{3.3}$$

in which $\sigma(z, R) = \sigma(z = 0, R) D_+(z)$, as opposed to $\delta_c(z)$ that does not vary with redshift, is very sensitive to the cluster redshift and evolves ν from 1 to increasing values with increasing redshift.

The top-hat model serves as a good starting point to intuitively grasp the physics of cluster formation. Notwithstanding, it brushes off many of the complexities in realistic scenarios that are usually calibrated using N-body simulations. Most notably, clusters do not have a sharp edge, but have a smooth density profile, and therefore do not have a distinct collapse time as different radii collapse at different epochs. The shapes of clusters are triaxial are are tidally coupled to the filamentary large scale structures connecting clusters. Finally, from the bottom-up hierarchical formation model for cold dark matter, the density field is not smooth but rather contains sub-structures lumped inside parent sub-structures. We refer readers to (Kravtsov & Borgani, 2012) for a review.

3.2 Density Profiles and Mass Definitions

Under hydrostatic equilibrium the inward gravitational force balances out the gas pressure exerted outwards $\nabla\phi = -\nabla p/\rho_g$. For a collisionless system such as dark matter the hydrostatic equilibrium is given by Jean's equation:

$$M_J(< r) = \frac{r\sigma_r^2}{G}\Big[\frac{d\ln\nu(r)}{d\ln r} + \frac{d\ln\sigma_r(r)^2}{d\ln r} + 2\beta(r)\Big], \tag{3.4}$$

in which $\beta = 1 - 2\frac{\sigma_t^2}{2\sigma_r^2}$ is the orbit anisotropy parameter defined by the tangential velocity dispersion $\sigma_t(r)$ and radial velocity dispersion $\sigma_r(r)$ (Binney & Tremaine, 2008).

Cluster masses are typically defined by enclosing a radius with average density contrast Δ. The radius of the cluster R_Δ is implied by solving the equation

$$M(< r) = \frac{4\pi}{3}\Delta\rho(z)R^3, \tag{3.5}$$

where $\rho(z)$ is the reference density at the redshift. This is usually taken to be the critical density $\rho_{\rm crit}$ or the background matter density ρ_b. The overdensity usually takes on the values $\Delta = 200$ for optical surveys as it resembles $\Delta_{\rm vir}$ for the top-hat model described above, and for X-ray and millimeter wave surveys it is usually taken to be $\Delta = 500, 2500$ as most of the signal comes from the inner region of the cluster at these wavelengths.

The density profile of cluster can take one on several parametric forms. The simplest of these models is the Singular Isothermal Sphere (SIS) profile

$$\rho(r) = \frac{\sigma_v^2}{2\pi G r^2}. \tag{3.6}$$

This comes from the assumption that particles inside the cluster are isothermal and exhibit a Maxwellian velocity distribution. As we will see the complexities inside the cluster cause the densities to deviate from $\rho \propto r^{-2}$.

The most commonly accepted form of parameterization is the Navarro-Frenk-White (NFW) (Navarro et al., 1997) profile:

$$\rho_{NFW}(r) = \frac{4\rho_s}{\left(\frac{r}{r_s}\right)\left(1+(\frac{r}{r_s})^2\right)}, \tag{3.7}$$

with r_s the scaled radius. Asymptotically it exhibits $\rho \propto r^{-1}$ for $r \ll r_s$ and $\rho \propto r^{-3}$ for $r \gg r_s$. This formalism allows us to introduce the concentration parameter $c_\Delta = R_\Delta/r_s$. The concentration exhibits a negative relation with mass in what is known as assembly bias from the top-down hierarchical formation of clusters. Many studies such as Bhattacharya et al. (2011); Ludlow et al. (2016); Diemer & Joyce (2019) have aimed to parameterize the concentration-mass relation. Lithwick & Dalal (2011) showed that shape of the profile is produced from adiabatic contraction from its initial shape.

3.3 Halo Abundances

We describe here the halo abundance model originally described by Press & Schechter (1974) that formed the basis for prescriptions of most halo abundance models to date. According to Press & Schechter (1974), only overdensities at a smoothing scale $\sigma(R)$ with $\delta_s > \delta_c$ can collapse into halos. We can interchange R with M by applying a top-hat filter to the smoothed overdensity with $M = 4\pi/3\bar{\rho}R^3$ inside the top-hat radius. The initial density perturbation follow Gaussian statistics, and hence the fraction of objects that will collapse into mass exceeding M follows

$$F(>M) = \frac{1}{\sqrt{2\pi}\sigma(M)} \int_{\delta_c}^{\infty} \exp\left[\frac{-\delta_s^2}{2\sigma(M)}\right] d\delta_s. \tag{3.8}$$

We expect that in the limit that $M \to 0$ one would recover $F(0) = 1$, recovering the whole content of the universe. Yet the Press-Schechter formalism underpredicts this ratio by a

factor of 1/2. This is explained by recognizing that half of the sub-halos are absorbed into larger halos, and a fudge factor of 2 is appended to the halo function.

Modeling of the halo-mass function $n(M)$ is an active area of research. Bond et al. (1991) provided a more rigorous exursion set model to account for the factor of 2 in the halo mass function. Sheth & Tormen (2002) extended the Press-Schechter ansatz to include ellipoisdal collapse. Many cosmological simulations have sought to parameterize the halo mass function by the universal variable $\nu(M, z)$ that accounts for both mass and redshift dependence (Jenkins et al., 2001; Evrard et al., 2002; Lukić et al., 2007; Bhattacharya et al., 2011). In particular, Tinker et al. (2008) captured the universality of halo-mass function up to 10% up to $z = 2.5$ using the formula:

$$n(M, z) = \frac{\rho}{M} f(M, z) \frac{d \ln \sigma_M^{-1}}{dM}, \tag{3.9}$$

with $f(\sigma(M, z))$ of the form

$$f(\sigma) = A\Big[\big(\frac{\sigma}{b}\big)^{-a} + 1\Big] \exp\left(-c/\sigma^2\right), \tag{3.10}$$

with (A, a, b, c) best-fit values provided for different overdensity mass definitions.

3.4 Scaling Relations

Cluster properties are expected to be self-similar across mass ranges. The Kaiser model (Kaiser, 1986) assumes that the initial power spectrum is scale invariant $P(k) \propto k^n$ and that the physics of cluster formation does not introduce new scale dependence. The model further assumes that under hydrostatic equilibrium (Eq. (3.4)) the logarithmic slope of temperature and gas density of the intracluster medium with respect to radius does not depend on the

mass. Under these assumptions X-ray temperature has the scaling relation

$$T_X \propto \frac{M}{R} \propto [E(z)M_\Delta]^{2/3}. \tag{3.11}$$

The X-ray luminosity in soft band regions with $T > 2$ keV is almost independent of temperature and has the scaling relation

$$L_{X,soft} \propto \rho_{\rm gas} r^3 \Lambda(T_X) \propto M_\Delta E(z)^2, \tag{3.12}$$

in which $\Lambda(T_X)$ is the radiative cooling function. The mass dependence of the X-ray luminosity in the soft band region $T = 0.5 - 2$ keV is harder to model. Blanchard et al. (1992) considered a model for which the radiative cooling function across the entire temperature range scales as $T_X^{1/2}$, such that the bolometric X-ray luminosity scales as

$$L_{X,bol} \propto M_\Delta^{4/3} E(z)^{7/3}. \tag{3.13}$$

The quantity $Y_X = M_{\rm gas} T$ is proportional to the global thermal energy of the ICM. As $T_X \propto M^{2/3}$ and $M_{\rm gas} \propto M_\Delta$ under the Kaiser model, Y_X scales with mass as

$$Y_X \propto M_\Delta^{5/3} E(z)^{2/3}. \tag{3.14}$$

Analogously in the millimeter regime one can derive an integrated $Y_{\rm SZ}$ proportional to the total energy of the Sunyaev-Zel'dovich distortion of the CMB around the cluster. The distortion of the temperature of the CMB due to the SZ effect is

$$\frac{\Delta T_{CMB}}{T_{CMB}} = g(\nu) y(\theta), \tag{3.15}$$

where $g(\nu)$ is the frequency dependence and $y(\theta)$ the local Compton-y parameter linearly

related to the total gas pressure along the line of sight. Assuming spherical symmetry we can integrate this signal towards the cluster edge to arrive at the total Compton-y parameter

$$Y_{SZ} \propto \int y(r) 2\pi r dr \propto M_{\Delta}^{5/3} E(z)^{2/3}. \tag{3.16}$$

In the optical regime red galaxies around the cluster center that fall into the red-sequence in the color-magnitude diagram are probabilistically weighed and summed into a quantity known as richness λ (Rozo et al., 2010; Rykoff et al., 2014). Assuming that galaxies on average have a stellar mass m_* and that the total stellar mass of the cluster is M_*, the richness scales with mass as

$$\lambda \propto \frac{M_*}{m_*} \propto M_{\Delta} \tag{3.17}$$

3.5 Cluster Cosmology

We can infer cosmology by inferring the best-fit cosmological parameter with cluster observables. Bayes theorem states that

$$P(\mathcal{M}|\mathcal{D}) = P(\mathcal{D}|\mathcal{M}) \frac{P(\mathcal{M})}{P(\mathcal{D})}, \tag{3.18}$$

where $\mathcal{M}$ and $\mathcal{D}$ are respectively short hand notations for model and data parameters. We denote $P(\mathcal{M}|\mathcal{D})$ as the posterior, $P(\mathcal{D}|\mathcal{M})$ the likelihood, $P(\mathcal{M})$ the prior and $P(\mathcal{D})$ the evidence. The ratio of evidence $P(\mathcal{D}_1)/P(\mathcal{D}_2)$ is called the Bayes factor and is used to compare the statistical evidence between two different models. When assessing parameters within a single model it is usually discarded as a constant term that does not influence the posterior distribution.

Bayes theorem suggests that data points are fixed and that the model follows a probability distribution. The posterior for one measurements using the Bayesian approach can be taken as the prior for a newer dataset, a process knonw as Bayesian updating. Contrast this with the

Table 3.1: Ingredients used to infer cosmology with galaxy clusters.

Parameter	**Description**
$\mathbf{p_c}$	Cosmological parameters
$\mathbf{p_o}$	Observable-mass relation parameters
$\mathbf{p} = \{\mathbf{p_c}, \mathbf{p_o}\}$	Combined parameters
$\mathcal{O}$	True observable quantities
$\hat{\mathcal{O}}$	Measured observable quantities
Φ	Selection function
$\frac{dV}{dzd\Omega_s}$	Volume element
Ω_s	Survey area

frequentist approach that treats the model as fixed and the measured data as drawing from a probability distribution when conducting every single experiment procedure. As we only have one universe the frequentist approach fell out of favor to pave the way for the Bayesian approach. The ergodicity hypothesis in cosmology states that this particular realization of our universe has the average properties of the ensemble of all hypothetical realizations under our cosmological model. As a reference, Efstathiou (2003) compares the results of the statistics of the CMB quadrupole moment using both the Bayesian and frequentist approach.

To infer cosmology we need to introduce the following ingredients in Table 3.1. We start with the differential number count of clusters given by

$$\frac{dN}{dmdzd\hat{\mathcal{O}}} = \Omega_s \frac{dV}{dzd\Omega_s} n(m,z) P(\hat{\mathcal{O}}|m,z,\mathbf{p}) \Phi(\hat{\mathcal{O}},m,z). \tag{3.19}$$

Due to the self-similarity of clusters $P(\hat{\mathcal{O}}_i|m,z,\mathbf{p})$ is usually assumed to form a power-law relation with mass. The selection function Φ depends on the survey depth and cluster observable. As the measured observable typically suffers from a wide range of systematic biases, one needs to convert the measured observables $\hat{\mathcal{O}}$ to the true observables $\mathcal{O}$ by forward

modeling the relation $P(\hat{\mathcal{O}}|\mathcal{O}, m, z, \mathbf{p})$. The volume element is described by

$$dV = D_H \frac{D_M^2(z)}{E(z)} d\Omega_s dz, \tag{3.20}$$

where $D_H = c/H_0$ is the Hubble distance and $D_M^2(z)$ the transverse comoving distance. Once we recover the "true" observables we need to model the joint distribution between observables. Most often they are taken to be multivariate correlated Gaussians of the form

$$P(\mathcal{O}|m, z, \mathbf{p}) = \frac{1}{(2\pi)^{n/2}|\det C|^{1/2}} \exp\Big[(\mathcal{O} - \langle\mathcal{O}\rangle)^T \mathcal{C}^{-1} (\mathcal{O} - \langle\mathcal{O}\rangle)\Big], \tag{3.21}$$

with $\mathcal{C}$ the covariance matrix between observables. The predicted number count given the measured observables is now

$$\frac{dN}{d\hat{\mathcal{O}}} = \Omega_s \int dm \int dz \int d\mathcal{O} \frac{dV}{dz d\Omega_s} \Phi(\hat{\mathcal{O}}, m, z) P(\hat{\mathcal{O}}|\mathcal{O}, m, z, \mathbf{p}) P(\mathcal{O}|m, z, \mathbf{p}) n(m, z) \tag{3.22}$$

The total expected number count is determined by marginalizing over the observable space

$$N_{\rm tot} = \int d\hat{\mathcal{O}} \frac{dN}{d\hat{\mathcal{O}}}|_{\mathbf{p}}. \tag{3.23}$$

When binned by the multivariate observable $\hat{\mathcal{O}}$ the number of clusters in each stacked bin is usually small enough that we can assign a Poisson likelihood to the measured rate of detection of clusters. The posterior distribution for $\mathbf{p}$ can be found by sampling with Monte-Carlo Markov Chains the log-likelihood

$$\ln \mathcal{L}(\mathbf{p}) = \sum_i \Big(\ln \frac{dN}{d\hat{\mathcal{O}}} \Big) \frac{dN}{d\hat{\mathcal{O}}_i} - N_{\rm tot} + \text{const}, \tag{3.24}$$

where $\ln \frac{dN}{d\hat{\mathcal{O}}}$ is the logarithm of the predicted number count in a given bin according to our model and $\frac{dN}{d\hat{\mathcal{O}}_i}$ the measured number count.

Clusters are powerful probes of cosmology as they are sensitive both to the geometry of the universe due to the cosmological dependence of dV and the growth of structure as parameterized by the halo mass function. We can use them to constraint the matter density of the universe, Ω_m, the amplitude of density fluctuations, σ_8, the equation of state parameter for dark energy, w, and the amplitude of primordial non-Gaussianity, e.g., f_{NL}.

CHAPTER 4

IMPACT OF CLUSTER TRIAXIALITY ON WEAK LENSING MEASUREMENTS

4.1 Introduction

The Dark Energy Survey (DES) used the 4-m Blanco Telescope and the Dark Energy Camera (Flaugher et al., 2015) to carry out a multi-band, 5,000 deg^2 survey over six years, with the primary goal of constraining cosmology and the nature of dark energy. Given its depth and wide-area coverage, DES observed ∼100,000 galaxy clusters up to redshift ~ 1 (Melchior et al., 2017). Initial cluster cosmology results, based on the first year of data (DES Y1), were published in Abbott et al. (2020). The cluster observable that DES Y1 employed as a mass proxy is a probabilistic cluster galaxy count called richness, computed with the redMaPPer algorithm (Rykoff et al., 2012).

Gravitational lensing, the shearing of galaxy images by foreground mass concentrations, is one of the most powerful methods for calibrating cluster mass-observable relations (Johnston et al., 2007; Gruen et al., 2014; Simet et al., 2017; McClintock et al., 2019). DES calibrates the cluster MOR through statistical weak lensing, in which shears from an ensemble of clusters are stacked to achieve high signal-to-noise (Bartelmann et al., 2001). In DES, stacked shear profiles are estimated for clusters binned in redMaPPer richness, enabling a determination of the mean halo mass as a function of richness (Melchior et al., 2017; McClintock et al., 2019).

Systematic effects in cluster selection or in calibration of the cluster MOR, if uncorrected for, can lead to biased cosmological inference from cluster abundance measurements. One such systematic arises from cluster triaxiality, the intrinsically elliptical shapes of galaxy clusters. N-body simulations indicate that dark halos can have major-to-minor axis ratios as high as 1.5 (Jing & Suto, 2002; Oguri et al., 2005), as confirmed observationally through

cluster weak lensing ellipticity measurements (Clampitt & Jain, 2016; Shin et al., 2018). Failing to account for cluster halo triaxiality may result in an overestimate of cluster mass by as much as 3-6% for stacked weak lensing measurements (Dietrich et al., 2014). Triaxiality was identified as one of the most important sources of systematic bias in the DES Y1 cluster lensing analysis, significant at the 2% level (McClintock et al., 2019). Recently Osato et al. (2018) showed that triaxiality not only biases the cluster surface mass density in the "one-halo" regime but also affects the surface density profile in the "two-halo" regime.

In this paper, we use redMaPPer cluster samples and associated halo catalogs in the Buzzard simulations to quantify cluster selection bias related to halo triaxiality properties such as orientation and ellipticity. We evaluate the impact of the triaxiality selection bias on 1) the richness–mass relation and 2) the excess surface mass density of individual halos (Osato et al., 2018). The stacked surface density profiles modeled with a triaxiality selection bias deviate from the isotropically stacked profiles; we find results comparable to those previously reported in the literature.

The paper is organized as follows. In Section 2, we describe the simulation data set used in the study and the halo–cluster matching algorithm. In Section 3 we examine the orientation and ellipticity distributions of triaxial halos associated with redMaPPer-selected clusters, quantifying the preference for halo orientation along the line of sight. In Section 4 we examine the boost in cluster richness for a given mass resulting from this orientation selection bias in the cluster sample. In Section 5 we test for correlation of halo triaxiality with other leading systematics, finding no evidence for such. In Section 6 we study halo surface mass densities as a function of orientation and the effect of orientation selection bias on stacked surface density measurements. We conclude in Section 7.

Throughout, we assume a flat ΛCDM cosmology with $\Omega_m = 0.283$, and $H_0 = 70$km s^{-1}Mpc^{-1}. Distances and masses, unless otherwise noted, are defined in units of h^{-1} Mpc and h^{-1} $M_\odot$.

4.2 The Simulation Data Set

4.2.1 Buzzard simulations

We make use of the N-body simulation catalogs from the suite of Buzzard simulations (DeRose et al., 2019) with the ΛCDM parameters given above. Detailed descriptions of the simulations can be found in MacCrann et al. (2018); DeRose et al. (2019); Wechsler et al. (2021); here we present a brief overview.

Halos are found by ROCKSTAR (Behroozi et al., 2013) with masses defined by M_{200b}, or more commonly referred to as M_{200m}, the mass enclosed in a radius within which the average matter density is 200 times the mean background matter density of the universe at the halo redshift. Galaxies are assigned to dark matter particles using ADDGALS, an empirical algorithm that places galaxies on dark matter particles based on a galaxy–dark matter relation learned from subhalo abundance matching catalogs and that is designed to accurately reproduce galaxy luminosities, colors, and spatial clustering over large volumes (DeRose et al., 2019). In particular, each massive halo is assigned a luminous, red galaxy at its center with the central galaxy's r-band absolute magnitude calibrated against the halo's virial mass (Wechsler et al., 2021).

The Buzzard flock is a set of 18 realizations of simulations that cover the DES Y1 footprint, each realization covering ∼1800 square degrees of the sky (Abbott et al., 2020; Drlica-Wagner et al., 2018). The galaxy catalog is complete toward a r-band magnitude of ~ 26.5 and z=2.35. By tuning the luminosity function of galaxies and their red fraction, the photometric redshift and errors follow the DES Y1 GOLD catalog, the the DES science-quality photometric catalog produced from Y1 data to enable cosmological analyses. To account for the masking of the DES Y1 footprint, Buzzard randomly downsampled galaxies by FRACGOOD, the percentage of un-masked pixels within a tile of the sky. As a second step, only galaxies that are brighter in the z-band than the local 10σ limiting magnitude are included

in the galaxy catalog.

The Buzzard simulations simultaneously achieve good spatial resolution and large volume by dividing the lightcone into three simulation boxes covering the redshift ranges $z \in [0.0, 0.34)$, $[0.34, 0.90)$, and $[0.90, 2.35)$, with respective minimally resolved dark matter particle masses of $2.7 \times 10^{10}\ h^{-1}\ \mathrm{M}_{\odot}$, $1.3 \times 10^{11}\ h^{-1}\ \mathrm{M}_{\odot}$, and $4.8 \times 10^{11}\ h^{-1}\ \mathrm{M}_{\odot}$. The increased resolution at low redshift captures non-linear structures at late times, while the lower resolution at high redshift enables the catalogs to encompass larger total volume. Particles are evolved using the L-Gadget2 code designed to efficiently run large-volume dark-matter only simulations (Springel et al., 2005).

4.2.2 redMaPPer cluster sample

With the advent of wide-field-imaging surveys, a plethora of optical cluster finding algorithms have emerged, such as those based on galaxy photometric redshifts, e.g. Kepner & Kim (2000), Soares-Santos et al. (2011), Wen et al. (2012), and Oguri (2014). In this paper, we study the cluster sample identified with the redMaPPer algorithm (Rykoff et al., 2014), which identifies cluster candidates as spatial over-densities of red-sequence galaxies. Clusters are assumed to be centered on a galaxy, with the central galaxy selected based on its luminosity and color (brightest central galaxy, or BCG). The algorithm also produces a richness estimate, λ, for each cluster candidate, a probabilistic count of cluster red-sequence galaxies above a luminosity threshold and inside a spatial aperture defined by $R_{\lambda} = 1h^{-1}\ \mathrm{Mpc}(\lambda/100)^{0.2}$ determined from iterative richness estimations.

The redMaPPer algorithm uses a sample of observed clusters with spectroscopic redshifts as a training set to build the initial redshift-dependent red-sequence model which cluster galaxies are fitted onto to determine the photometric redshift z_{λ}. The DES Y1 redMaPPer photometric redshifts are unbiased at the $|\Delta z| \leq 0.003$ level, and have a median photometric redshift scatter $\sigma z/(1+z) \approx 0.006$.

For DES Y1 cluster cosmology, redMaPPer clusters are taken from the GOLD galaxy catalog (Drlica-Wagner et al., 2018). The clusters are restricted to the redshift interval to $z \in [0.2, 0.65]$ and $\lambda > 20$, totaling 6504 clusters in the footprint. redMaPPer performance below redshift z = 0.2 is compromised by the lack of u-band data, while there are relatively few galaxy clusters in the catalog above redshift z = 0.65 (Abbott et al., 2020).

The redMaPPer cluster finder has been applied to the Buzzard catalogs to identify galaxy clusters. The Buzzard simulations come with a caveat that their richness-mass relation is biased low relative to the DES Y1 data which may is likely attributed to the spatial dependence of galaxy colors at small scales (DeRose et al., 2019). Nonetheless, we describe in Section 4.4 how we can use Buzzard to study the *relative* difference in richness-mass across orientation bins.

In this project, for sample completeness we make use of a redMaPPer sample with a maximum cluster redshift of $z < 0.90$ which is around the redshift detection limit of redMaPPer and the limit of the Buzzard light cone, and for sample purity we apply a richness cut of $\lambda > 20$ (Rykoff et al., 2016; McClintock et al., 2019). Halos are also cut at masses below $5 \times 10^{13}\ h^{-1} M_{\odot}$ which roughly corresponds to a richness of 20.

4.2.3 Cluster halo matching algorithm

Here we outline how redMaPPer clusters are matched to Buzzard halos. First, a cluster is labeled as centered or miscentered based on whether or not its redMaPPer BCG is a central galaxy in a Buzzard halo. Centered clusters have BCGs that share the same ID as that of the halo central galaxy; in this case, the cluster and halo central coordinates perfectly match. By this criterion, 63% of redMaPPer clusters are centered; the remaining were matched using the halo-cluster algorithm described below. A more detailed description of the centering properties of the redMaPPer catalogs can be found in Section 4.5.1.

The miscentered redMaPPer clusters were matched to Buzzard dark matter halos by

proximity. Halos were ranked by halo mass, and clusters were ranked by richness, both in descending order. We first search for halo-cluster pairs with redshift separation $\Delta z \leq 0.05$ between cluster photometric redshift and true halo redshift. This range of redshift separation is large compared to the typical photometric redshift error, $\Delta z \sim 0.005$, for redMaPPer-selected clusters. Then, for each halo, we identify those redMaPPer clusters with BCGs within a projected 2-D, comoving radius of 2 h^{-1} Mpc of the halo central galaxy. If there are multiple redMaPPer clusters satisfying these separation criteria, we match the halo to the richest such cluster that hasn't been previously matched. For each cluster, we repeat this matching process, selecting halos satisfying the redshift and projected distance criteria, and then choosing the most massive such halo still on the list as the one to be associated with that cluster. Clusters and halos that uniquely match with each other in both matching steps are considered valid matches.

Of the 24,243 initially identified redMaPPer cluster candidates in the suite of 18 catalogs, 23,658 or 97% are uniquely matched to a halo with the above prescription. We do not consider the non-uniquely matched clusters in this study.

This halo–cluster matching algorithm was cross-checked with an independent halo–cluster matching algorithm used in Farahi et al. (2016) that rank-orders halos and clusters by the number of galaxies they have in common. Using the Aardvark simulation, Farahi et al. (2016) uniquely matched 99% of redMaPPer clusters to halos, showing excellent agreement with this paper's algorithm on the completeness and uniqueness of cluster-to-halo matches. We cross checked our matching algorithm with that of Farahi et al. (2016) in a different version of Buzzard with a smaller patch of sky containing several hundred clusters and found almost identical halo-cluster pairings.

Due to the high number of particles per halo, Poisson noise plays a negligible role in our ellipticity measurements: at low redshift, with a mass resolution of 2.7×10^{10} h^{-1} $M_\odot$, a typical 3×10^{14} h^{-1} $M_\odot$-mass halo found through redMaPPer corresponding to a richness of

~ 40 will contain $\sim 10,000$ particles, and the same-mass halo at high redshift, with a poorer mass resolution of $1.3 \times 10^{11}\ \ h^{-1}\ \mathrm{M}_{\odot}$, contains $\sim 3,000$ particles. Simulations conducted by Jing & Suto (2002) demonstrated that these large numbers of particles per halo make Poisson noise negligible for our purposes. We do not consider halos with fewer than 100 particles with poor shape convergence, corresponding to group size objects with richnesses well below our $\lambda > 20$ cut.

4.3 Cluster Halo Triaxiality and Selection Bias

Previous studies have shown that optical cluster finders preferentially select halos with their major axes oriented along the line of sight (Corless & King, 2008; Dietrich et al., 2014). In this Section, we quantify this orientation bias of selected clusters using the redMaPPer catalogs and the Buzzard simulations. We also explore whether a cluster ellipticity selection effect exists, i.e., whether redMaPPer preferentially selects halos that are more or less elliptical than randomly selected halos.

4.3.1 Measurement of halo ellipticity and orientation

We make use of a quadrupole moment tensor method (Bett (2012) and references therein) to measure the shapes and orientations of halos. Many such algorithms solve for halo shapes by using particles inside a spherical envelope (Dietrich et al., 2014; Osato et al., 2018); this has the advantage of allowing easy comparison with other results, but it systematically underestimates the axial ratios for ellipsoidal profiles, an effect known as "edge bias." As described below, we correct for such an effect by using an iterative method to determine the shape of the enclosing envelope, in the vein of earlier works such as Dubinski & Carlberg (1991), Katz (1991) and Warren et al. (1992). To do so, we first measure the shape of the halo using particles inside a spherical envelope; once the axis ratios and the principal axes are found, the envelope adapts iteratively until both the axis ratios of the halo inside the

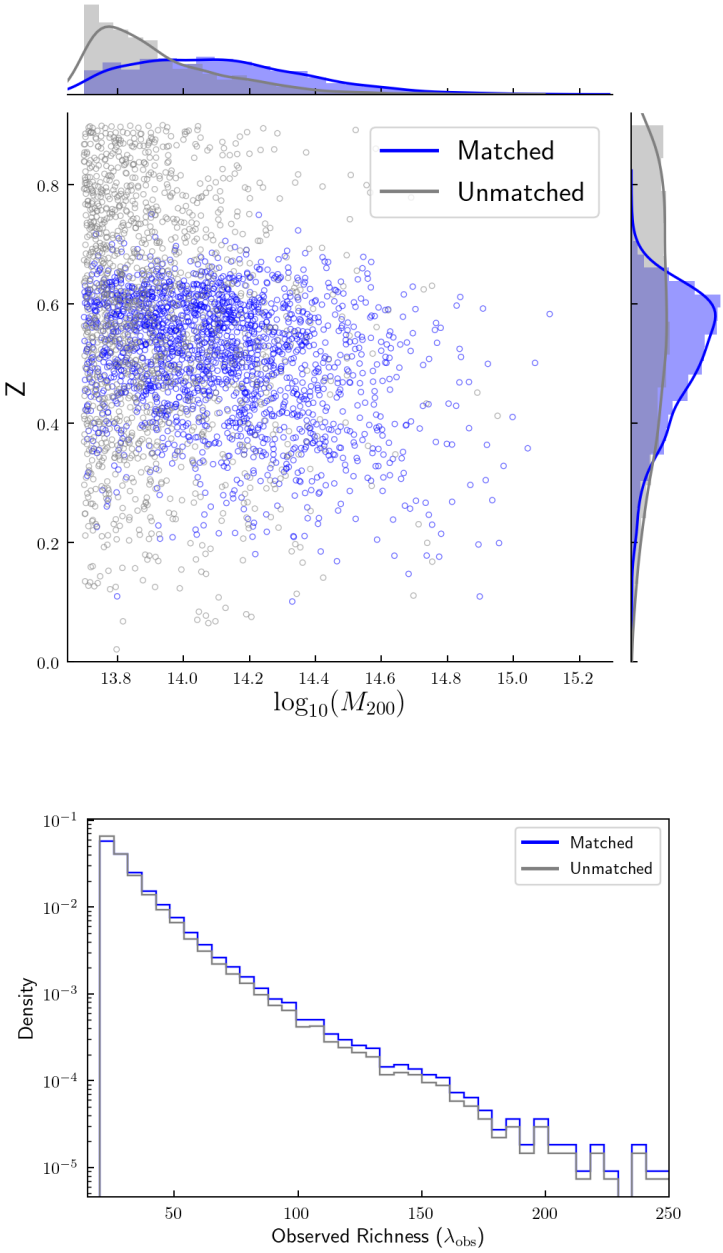


Figure 4.1: *Upper panel*: A 2-D distribution plot of the true M_{200m} and z of halos before and after matching with redMaPPer clusters. The halos are cut at $M_{200m} > 5\times10^{13}\ h^{-1}M_{\odot}$ and a redshift cut of $z < 0.90$ and are sparsely sampled for better visualization. *Lower panel*: The probability density function of the observed richness $\lambda_{\rm obs}$ before and after matching with halos. Because of the high match rate of redMaPPer clusters the two distributions are nearly identical.

envelope and the shape of the ellipsoidal envelope itself converge.

We now describe the halo ellipticity measurement algorithm in detail. It involves nested iteration of both the principal axes, as determined from the quadrupole moment tensor, and of the envelope shape. In the initial iteration, $l = 0$, of the envelope shape, the envelope is set to be a sphere centered on the halo center with a radius equal to the virial radius of the halo, $R_{\rm vir}$. The reduced quadrupole moment tensor is then calculated for the N_P dark matter particles inside the envelope. This tensor, with its principal-axis directions solved at the k-th iteration, is defined as:

$$\mathcal{M}_{ij}^{(k)} = \frac{1}{N_P^{(k)}} \sum_{p=1}^{N_p^{(k)}} \frac{R_{p,i}^{(k)} R_{p,j}^{(k)}}{\left(R_p^{(k)}\right)^2}, \tag{4.1}$$

where $R_{p,i}$ and $R_{p,j}$ are the distances from the center along Cartesian coordinate axes of the p-th particle and R_p^k is the triaxial radius, defined below, of the p-th particle solved at the k-th iteration.

We define a, b, and c as the major, intermediate, and minor axes lengths of a particle projected onto the unit sphere and $q \equiv \frac{c}{a}$ and $s \equiv \frac{b}{a}$ as the minor-major and intermediate-major axis ratios; the physical distances to the p-th particle along the minor, intermediate and major axes are denoted X_p, Y_p and Z_p. In this notation, the triaxial radius at the k-th iteration of the particle is expressed as:

$$R_p^{(k)} = \sqrt{\left(\frac{X_p}{q^{(k-1)}}\right)^2 + \left(\frac{Y_p}{s^{(k-1)}}\right)^2 + Z_p^2}. \tag{4.2}$$

The axis lengths projected onto the unit sphere are the square roots of the eigenvalues of the reduced tensor, and the axis directions are the corresponding eigenvectors. After each iteration, the principle axes are rotated by the rotation matrix $M^{(k)}$, where each row in the matrix is a principle axis found from the reduced tensor in the previous iteration. The

reduced tensor is computed again under the rotated coordinates. Starting from $q^{(k=0)} = 1$ and $s^{(k=0)} = 1$, the tensor is considered to have converged if

$$\left|1 - \frac{q^{(k)}}{q^{(k-1)}}\right| < 10^{-6} \quad \text{and} \quad \left|1 - \frac{s^{(k)}}{s^{(k-1)}}\right| < 10^{-6} \quad , \tag{4.3}$$

and is deemed divergent if convergence is not reached before the number of iterations k exceeds 100.

The total rotation matrix after n rotations is

$$\boldsymbol{M_{\rm tot} = M^{(n)} \dots M^{(k)} \dots M^{(1)}} \ , \tag{4.4}$$

where each row in $M_{\rm tot}$ gives the direction of the corresponding halo axis prior to rotation.

If after k iterations the axis ratios derived from the tensor converge, then the elliptical envelope of the particles is advanced from the previous $l-1$-th to the l-th (for $l > 0$) iteration, adapting its axis ratios and orientation to those of the halo as determined from the tensor with the previous envelope. Particles with elliptical distances of

$$R_p^{(l)} \equiv \sqrt{\left(\frac{X_p^{(l-1)}}{q^{(l-1)}}\right)^2 + \left(\frac{Y_p^{(l-1)}}{s^{(l-1)}}\right)^2 + (Z_p^{(l-1)})^2} < R_{\rm vir} \tag{4.5}$$

are selected. The sequence initializes at $q^{(l=0)} = s^{(l=0)} = 1$, and (X_p^0, Y_p^0, Z_p^0) along the original (x, y, z) axes of our coordinate system and converges using the same criteria as for the shape of the halo inside the envelope, Cf. equation 4.3. The shape of the halo is said to be convergent only if both the shape of the halo particles found inside the envelope and the shape of the envelope itself both converge.

We applied this technique to measure the shapes of simulated halos that are matched to the redMaPPer clusters; of the 23,658 matched redMaPPer clusters, the halo shape measurements converge by the above criteria for 22,790 of them. We use this sample in the following

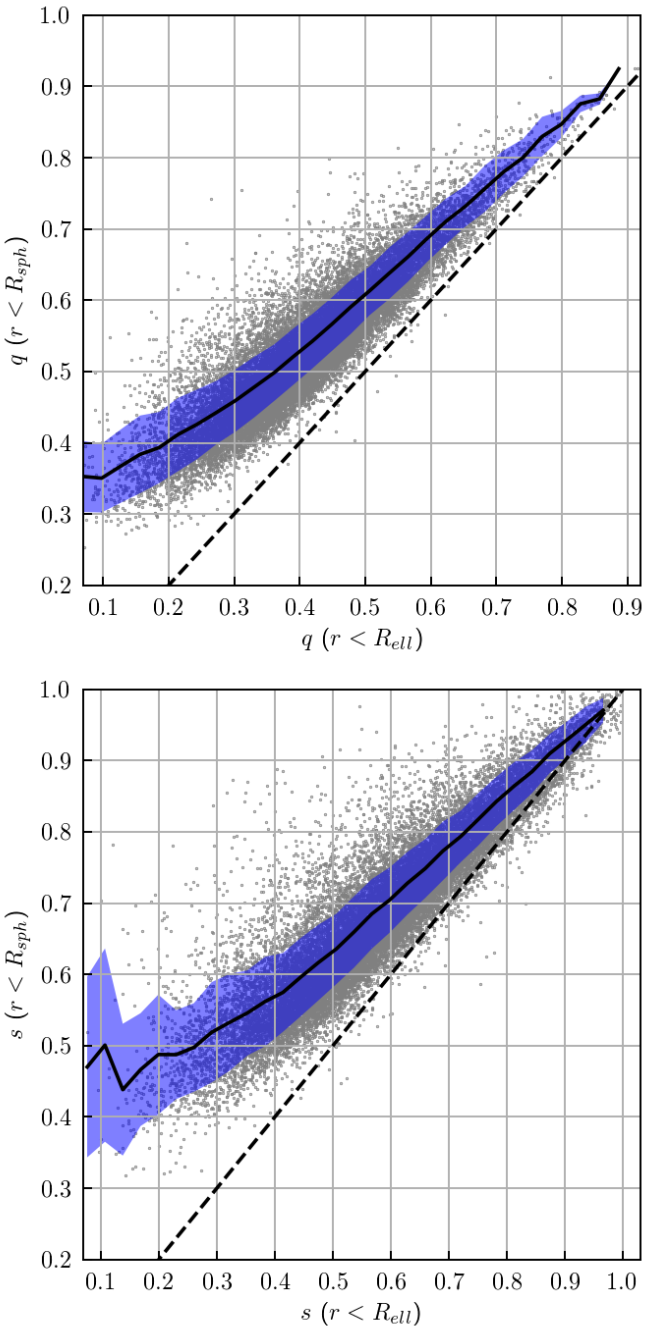


Figure 4.2: Axis ratios, q and s, for redMaPPer-matched halos measured with spherical vs. adaptive ellipsoidal envelopes. Solid black lines show the mean ratios in each axis-ratio bin, and the blue bands indicate the $1-\sigma$ scatter. Dashed lines would correspond to no difference in axis ratios between the two methods. The results demonstrate that edge bias reduces the measured ellipticities of halos from their true values, with larger bias at higher ellipticities (smaller q and s).

sections to explore orientation bias.

We can gauge the impact of the edge bias on halo shape measurement by comparing results with the adaptive ellipsoidal envelope to those using a fixed spherical envelope. In Fig. 4.2, we plot the halo axis ratios q and s found using spherical envelopes (ordinates) with those from the adaptive ellipsoidal envelopes (abscissas). We see clearly that the axis ratios are biased high (ellipticities biased low) when using spherical envelopes, with larger bias at higher ellipticities (lower values of the axis ratios). These results are in qualitative agreement with those of Shin et al. (2018), who studied 2-D projected ellipticities of observed galaxies in redMaPPer clusters. They found that the inferred 2-D ellipticity, $e \equiv (1+q)/(1-q)$ where q is the axis ratio for a 2-D ellipse, deviates by as much as 0.1 when using a circular aperture for the redMaPPer (Rykoff et al., 2014) cluster finder, $R_\lambda = 1h^{-1}$ Mpc$(\lambda/100)^{0.2}$, due to the cut-off of satellite galaxies along the major axis; they also found that the bias in ellipticity becomes worse at higher ellipticity (smaller q).

4.3.2 Distributions of cluster halo orientation and ellipticity

Armed with measurements of halo shapes for redMaPPer clusters, in this subsection we study the distributions of halo ellipticity and orientation. To test for redMaPPer-associated selection biases, we compare these distributions to those for a sample of 36,445 randomly selected halos with convergent shape measurements from the Buzzard catalog. The orientation of interest is the angle between the halo major axis and the line of sight, which we denote by i; a non-uniform distribution of i would signal the preferential selection of (prolate) clusters with these vectors aligned. For this analysis, we adopt the orientation bins $\cos(i) \in [0.0, 0.2),\ [0.2, 0.4),\ [0.4, 0.6),\ [0.6, 0.8),\ [0.8, 1.0)$.

The distributions of axis ratios for redMaPPer-matched halos and for randomly selected halos are shown for different orientation bins in the upper panels of Figure 4.3. Previous N-body studies found that more massive halos tend to be more elliptical (Kasun & Evrard,

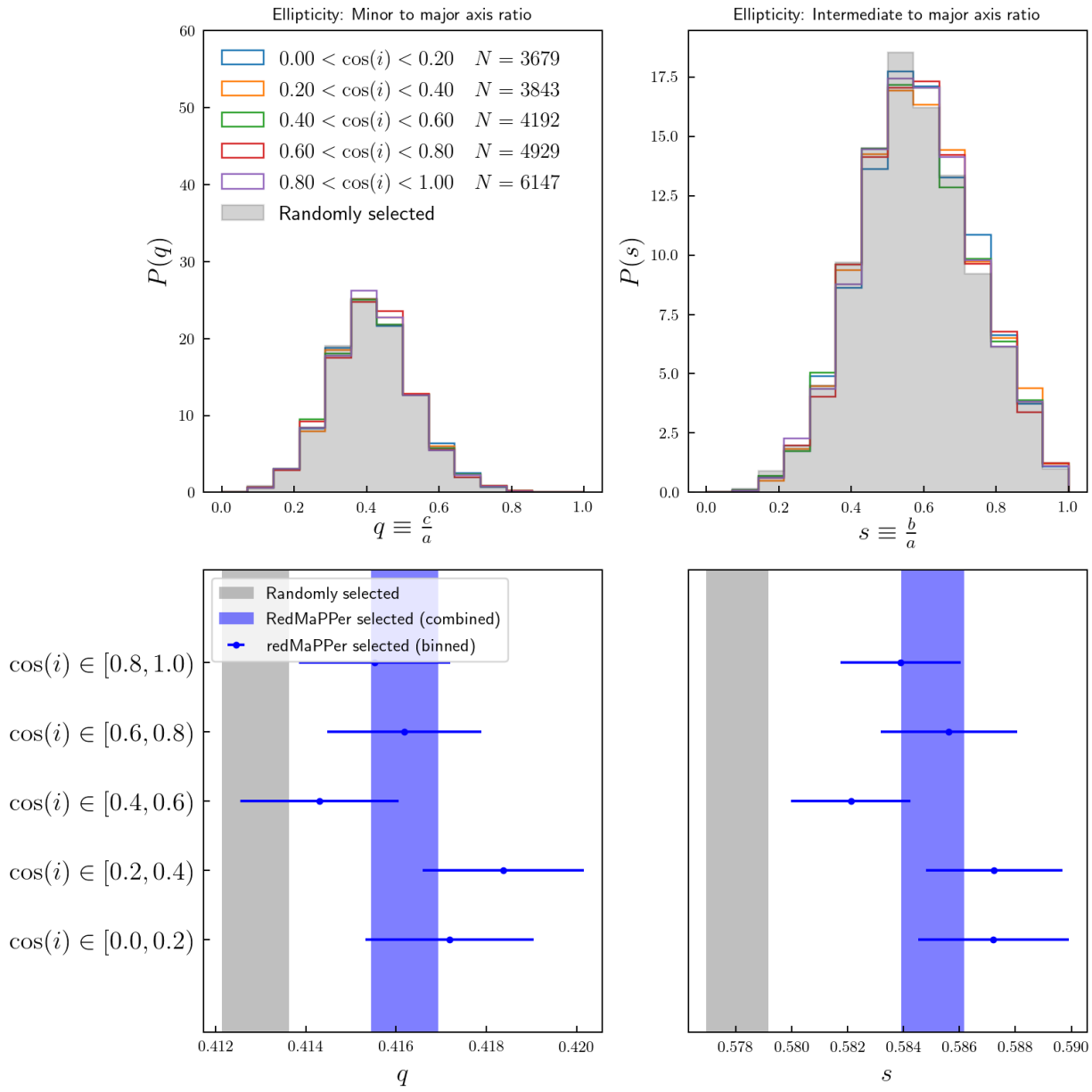


Figure 4.3: *Top panels*: Axis-ratio distributions for redMaPPer-matched clusters binned by orientation and for randomly selected halos from the Buzzard simulations. *Bottom panels*: Mean axis ratios with 1σ errors from jackknife resampling. Applying the 3σ significance cutoff rule, no significant shift is found in the shape parameters q and s for redMaPPer-matched and randomly selected halos. Also is the case that no statistically significant difference is found in the mean ellipticities across different orientation bins.

2005) as a result of tidal forces and mergers. To account for this effect, we resampled the randomly selected halos to match the halo mass function of the redMaPPer-matched halos. The upper panels of Fig. 4.3 indicate that the ellipticity distributions of the redMaPPer-matched halos are qualitatively very similar to those for the resampled random halos, with little dependence on orientation.

To quantify this comparison, in the bottom panels of Fig. 4.3 we show the mean axis-ratios for the redMaPPer-matched halos in different orientation bins (in blue), along with the means for the random halos (in grey). The errors on these measurements are estimated by jackknife resampling, with the simulated survey footprint split by the k-means algorithm *kmeans_radec*[1] into 40 non-overlapping patches—the error estimates come from the variance among the patches, each of them 37.5 square degrees. With this kind of spatial jackknife, the choice of the size of the jackknife patch is a compromise: for very large patch size, the number of patches (samples) would be too small to get a meaningful statistical sample; for very small patch size, large-scale structure would be highly correlated across adjacent patches, so they could not be treated as quasi-independent for error estimation.

The mean axis ratios differ by 0.7 and 1.2% for q and s respectively for redMaPPer vs. random halos. To determine if these differences are significant, we conduct a null-hypothesis test on q and s with their standard errors modeled as Student's t distributions. We find a 1.4σ difference in the minor-to-major axis ratio q for redMaPPer vs. randomly sampled halos and a 1.8σ difference in the intermediate-to-major axis ratio s. There are no statistically significant shifts in mean axis ratios for redMaPPer halos between different $\cos(i)$ bins. Thus, we do not find strong evidence of shifts in the ellipticity distributions.

Figure 4.4 (top panel) shows a similar analysis to that above, but now for the distribution of halo orientation in 3 different richness bins. In this case, there is a clear signal of orientation bias in the redMaPPer-matched clusters, with preferential selection of clusters

1. Code written by Erin Sheldon. Source: `https://github.com/esheldon/kmeans_radec`

with major axis oriented along the line of sight. The effect is more pronounced for clusters of higher richness: the lower panel shows an increase in the mean value of $\cos(i)$ with richness. Using the same method of null hypothesis testing, we find that the mean value of $\cos(i)$ for redMaPPer halos of 0.555 ± 0.002 is boosted compared to that for randomly selected halos with a 13.8σ significance. There is also a statistically significant shift in the mean value of $\cos(i)$ between richness bins: the mean $\cos(i)$ for $\lambda \in [30.0, 50.0)$ ($\lambda \in [50.0, 274.0)$) exceeds that for $\lambda \in [20.0, 30.0)$ at 3.7σ (4.8σ) significance. As a null test, we find that the randomly selected halos have a mean $\cos(i)$ consistent with 0.50.

In the next subsection, we will interpret the correlation of mean $\cos(i)$ with richness seen in Fig. 4.4 as due to the boosting of observed richness for clusters (of fixed mass) oriented along the line of sight.

4.4 Effect of Orientation on the Richness–Mass Relation

Since we have shown that the orientation distribution of redMaPPer-selected clusters is biased, it is important to understand how this may impact the observed cluster richness-mass relation, a key ingredient in cluster cosmology. In this Section, we explore how the cluster richness-mass relation varies with cluster orientation.

Figure 4.5 shows the empirical relation between Buzzard halo mass (defined by $M_{200\mathrm{m}}$) and observed richness for the redMaPPer-matched clusters. Following previous work (Saro et al., 2015; Simet et al., 2017; Melchior et al., 2017; McClintock et al., 2019), we model the relation between cluster mean richness $\mu(\lambda)$ and halo mass M as a linear relation between $\ln(\lambda)$ and $\ln(M)$, with a pivot point at 10^{14} $\mathrm{M}_{\odot}$:

$$\mu(\ln\lambda) = \ln(A) + B \times \left(\ln(M/M_{\odot}) - 14\ln(10)\right) \ . \tag{4.6}$$

We do not consider the redshift evolution of the richness-mass relation as results from pre-

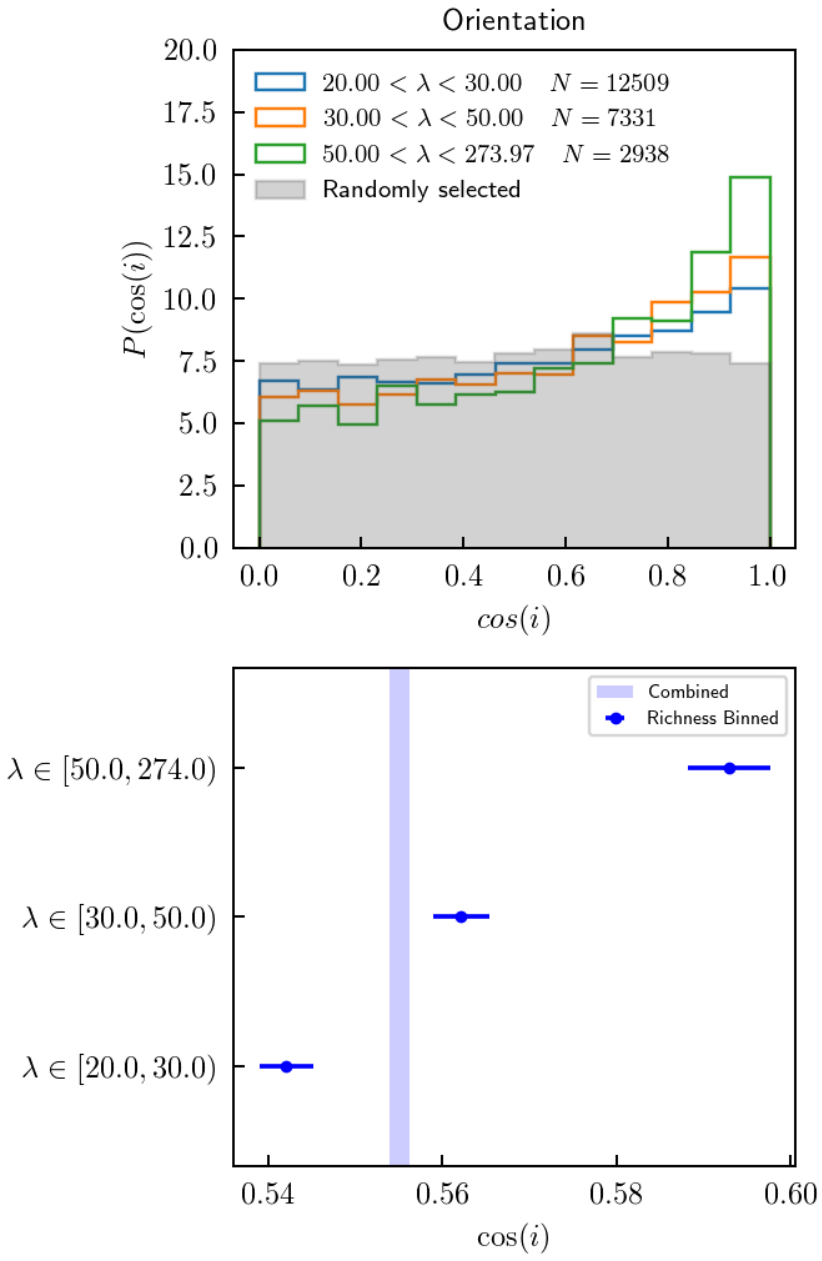


Figure 4.4: *Top panel*: Distribution of $\cos(i)$ for redMaPPer-matched halos in 3 richness bins and for randomly selected halos. *Bottom panel*: The mean $\cos(i)$ for redMaPPer-selected halos is boosted relative to that for randomly selected halos (0.50, not shown). The mean value of $\cos(i)$ also increases with redMaPPer richness. Errors are estimated from jackknife resampling.

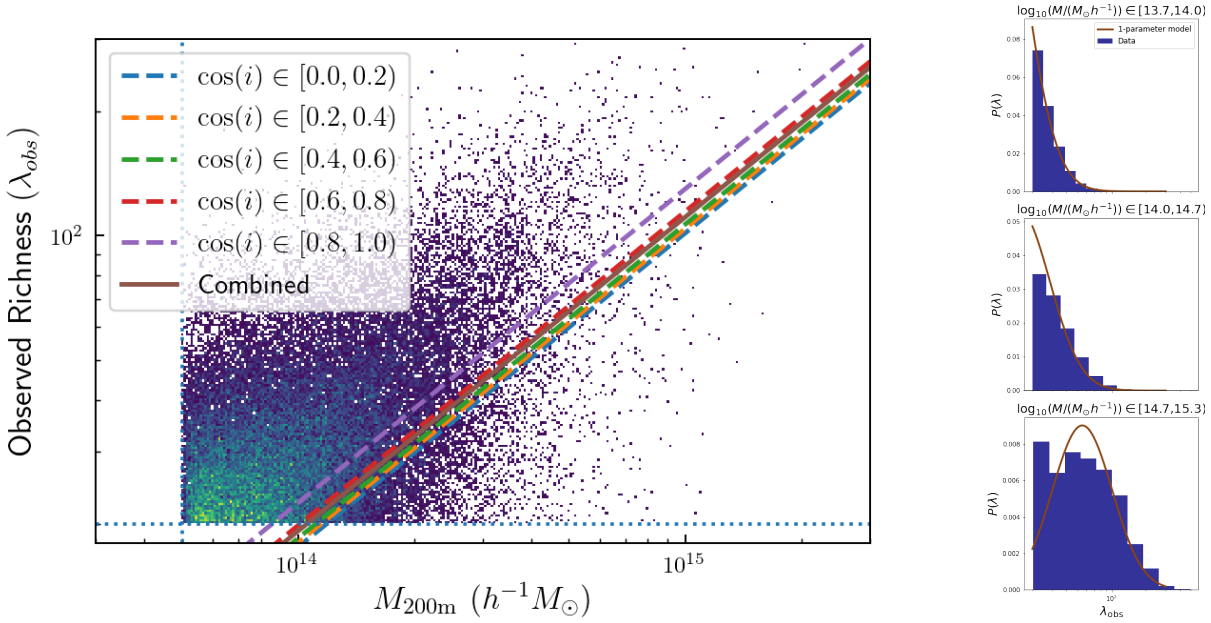


Figure 4.5: *Left panel:* Solid line labelled "Combined" shows the best-fit model to the full sample assuming a linear relationship between $\ln(\lambda)$ and $\ln(M)$. Dashed lines show best-fit models in each orientation bin, with the amplitude $\ln(A)$ allowed to vary from bin to bin. For halos of fixed mass, those oriented along the line of sight have larger observed redMaPPer richness. The dashed horizontal line indicates the richness cut at $\lambda > 20$ and dashed vertical line the mass cut at $M > 5 \times 10^{13}\ h^{-1}M_\odot$. Color coded is the density of the scatter points in the parameter space, with brighter colors indicating a higher density. *Right panel:* The richness distribution in mass bins for all data points overlaid with a truncated Gaussian fit using the best-fit parameters in the "Combined" 1-parameter model. In lower mass bins the best fit mean log-richness $\mu(\ln\lambda)$ is lower than the mean log-richness of the data points, as the peak of the truncated Gaussian fit lies below the $\lambda > 20$ cutoff.

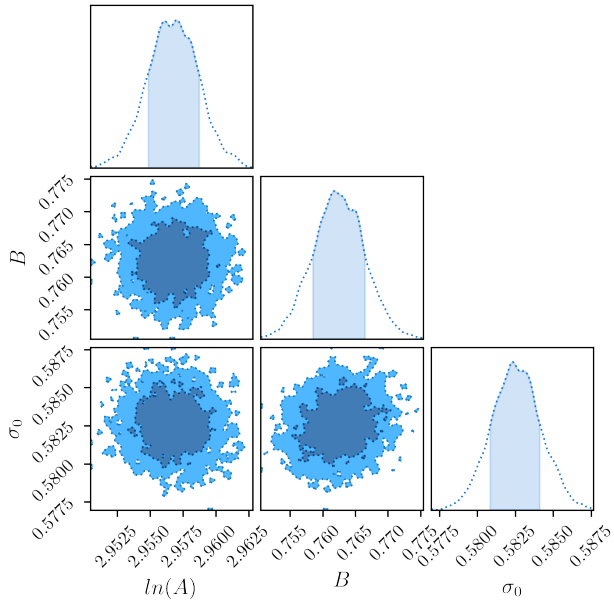


Figure 4.6: Posterior distributions of the richness–mass parameters derived using all redMaPPer-matched clusters. The shaded regions in the 2-D distributions show the 68 and 95% confidence regions; shaded regions in 1-D plots indicate the 68% confidence regions for the marginalized parameters. Posteriors for templates in different orientation bins share the same features.

vious multiwavelength scaling relations of galaxy clusters have prescribed a global redshift fit to the richness-mass relation (Simet et al., 2017) or those that do model the redshift dependence find it consistent with a null dependence (Saro et al., 2015; Melchior et al., 2017; McClintock et al., 2019; Bleem et al., 2020). In a recent work, To et al. (2021b) used Buzzard simulations to quantify the large scale bias of redMaPPer-redMaGic cross correlation that has a redshift dependence at $1-\sigma$ from null and that could be explained by the increase in observed richness at higher redshift from stronger projection effects.

We model the scatter of richness at fixed mass as truncated log-normal scatter that cuts off clusters with $\lambda < 20$:

$$P(\ln\lambda|\ln M) \propto \mathcal{N}(\mu(\ln\lambda), \sigma(\ln\lambda))H(\lambda - 20) \ , \tag{4.7}$$

where $H(x)$ is the Heaviside step function. The variance σ^2 is the sum of the intrinsic

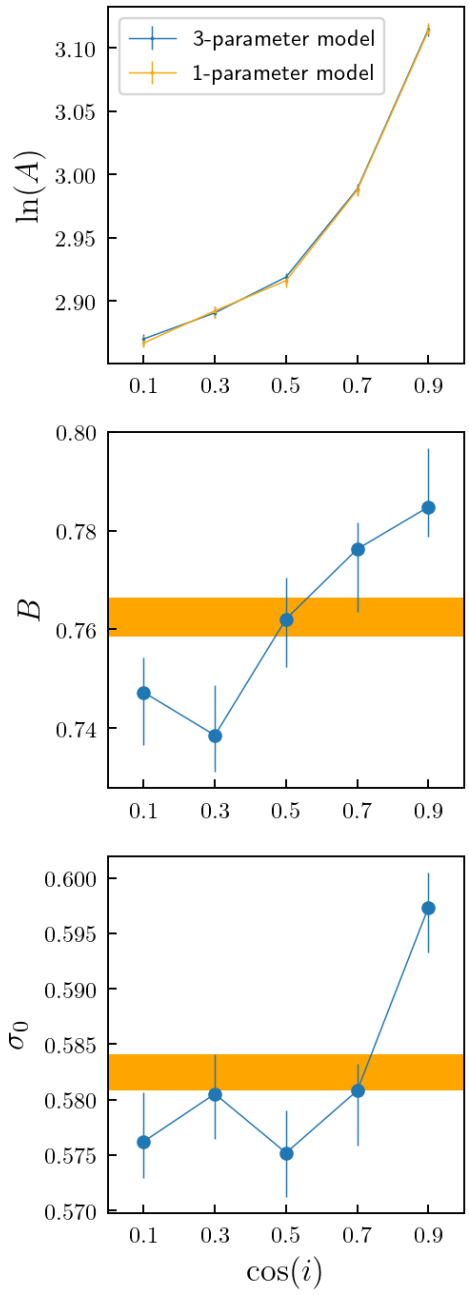


Figure 4.7: Dependence of redMaPPer richness–mass model parameters on halo orientation $\cos(i)$. Horizontal bands show the mean and 68% CL range for the global (full-sample) fit for each parameter. The top panel shows best-fit amplitude $\ln(A)$ vs. orientation when the other 2 parameters are allowed to vary with orientation (3-parameter model) and when they are fixed (1-parameter model), indicating little difference. The Bayesian Information Criterion (BIC) test favors the 1-parameter model.

variance σ_0^2 and a Poisson term due to finite richness,

$$\sigma^2(\ln\lambda) = \sigma_0^2 + \frac{\exp(\mu(\ln\lambda)) - 1}{\exp(2\mu(\ln\lambda))}. \tag{4.8}$$

According to Bayes' theorem, the posterior likelihood of the model parameters is given by

$$P(A, B, \sigma_0|\lambda, M) \propto P(\lambda, M|A, B, \sigma_0)P(A, B, \sigma_0), \tag{4.9}$$

where $P(A, B, \sigma_0)$ is the joint prior on the parameters which we set as non-informative uniform distributions.

The maximum likelihood estimates for the model parameters are found with a Markov Chain Monte-Carlo (MCMC) method implemented through the *pymc* module, assuming uniform priors for A, B, and σ_0. We run chains of 10^6 steps for each run, thin them by selecting every 200 steps, and remove the first 3000 steps (after thinning) as burn-in, yielding 2000 steps to sample the posterior distribution.

The solid line labelled "Combined" in Fig. 4.5 shows the best-fit model to the richness-mass relation for the full redMaPPer sample, with parameters given in the bottom line of Table 4.1. The posterior distributions for the "Combined" model parameters shown in Figure 4.6 show good convergence of the parameters and minimal correlation among them. The same trends are produced (but not shown) in the posterior distributions for different orientation bins. The reduced chi-square statistics shown in Table 4.1 show that the model is a good fit to the data.

Next, we assume that the richness-mass model of Eqns. (4.6-4.8) applies separately in each orientation bin. The 3-parameter model in each orientation bin is fit independently, with the results shown in Figure 4.7 and parameter values in the middle box of Table 4.1. We find that most of the dependence on orientation comes from the boosting of the amplitude parameter, $\ln(A)$, with $\cos(i)$. We therefore also consider a model in which only $\ln(A)$ varies

with orientation, with the other 2 parameters fixed to their global values. The top panel of Fig. 4.7 and Table 4.1 show that this 1-parameter model makes no appreciable change in the best-fit values of $\ln(A)$ in each bin. Moreover, reducing the number of parameters does not significantly compromise the goodness-of-fit of the MLE model relative to the number of extra parameters: as shown in Table 4.1, the reduced Bayesian Information Criterion (BIC) for the 1-parameter vs. the 3-parameter model marginally favors the simpler model.

The best-fit 1-parameter models in each orientation bin are indicated by the dashed lines in Fig. 4.5: the effect of orientation bias on the richness–mass relation is a boost in the amplitude, that is, in observed richness, at fixed halo mass, for halos with major axes aligned with the line of sight.

While the orientation-bias model studied here captures the behavior of redMaPPer-selected halos in the Buzzard simulations, a caveat is in order before applying the model to redMaPPer-selected clusters in the real universe. In particular, the redMaPPer richness at fixed halo mass in Buzzard has been found to be systematically lower at a 3σ level from that for redMaPPer clusters with weak-lensing calibrated masses in DES Y1 data (DeRose et al., 2019) which can be traced to the underestimation of the halo occupation distribution (HOD) of red galaxies identified by the red sequence in Buzzard. If this systematic is relatively independent of richness, we expect our model for the difference in richness amplitude with orientation, $\Delta \ln(A)$, to retain its validity, even if the central values of $\ln(A)$, B and σ_0 differ (note that the intrinsic scatter σ_0 is not constrained in the McClintock et al. (2019) weak lensing analysis of DES Y1 clusters). The dependence of the richness–mass relation on the HOD of red-sequence galaxies can be tested with studies using other simulations, such as the latest cosmoDC2 (Korytov et al., 2019), which populates halos with galaxies using a different set of semi-analytic and empirical methods from ADDGALS. Alternatively, one can construct and analyze new redMaPPer catalogs from the Buzzard simulations after injecting red-sequence galaxies to match the HOD of DES Y1 data.

Model parameters and BIC for richness–mass template									
	3-parameter model				1-parameter model				
$\cos i$	$\ln A$	B	σ_0	BIC	$\ln A$	B	σ_0	BIC	χ^2/ν
[0.0,0.2)	$2.869\pm^{0.004}_{0.006}$	$0.747\pm^{0.007}_{0.011}$	$0.576\pm^{0.004}_{0.003}$	8819	$2.866\pm^{0.005}_{0.003}$			8799	1.33
[0.2,0.4)	$2.890\pm^{0.005}_{0.004}$	$0.739\pm^{0.010}_{0.007}$	$0.581\pm^{0.004}_{0.004}$	8088	$2.892\pm^{0.003}_{0.006}$			8064	1.26
[0.4,0.6)	$2.919\pm^{0.003}_{0.006}$	$0.762\pm^{0.008}_{0.010}$	$0.575\pm^{0.004}_{0.004}$	8123	$2.916\pm^{0.004}_{0.005}$			8104	1.20
[0.6,0.8)	$2.988\pm^{0.004}_{0.005}$	$0.776\pm^{0.005}_{0.013}$	$0.581\pm^{0.002}_{0.005}$	6480	$2.986\pm^{0.004}_{0.005}$	$0.762\pm^{0.005}_{0.003}$	$0.582\pm^{0.002}_{0.002}$	6463	1.02
[0.8,1.0)	$3.115\pm^{0.003}_{0.005}$	$0.785\pm^{0.012}_{0.006}$	$0.597\pm^{0.003}_{0.004}$	2648	$3.114\pm^{0.006}_{0.003}$			2588	0.77
All	NA				$2.956\pm^{0.003}_{0.001}$			29807	1.09

Table 4.1: Maximum Likelihood estimates and 68% CL errors of richness-mass model parameters for redMaPPer clusters as a function of halo orientation $\cos(i)$ and for the full cluster sample ("All"). The middle box shows results when all 3 model parameters are allowed to vary with $\cos(i)$ (3-parameter model); right-most box shows results when only $\ln(A)$ is allowed to vary (1-parameter model). Also shown are the Bayesian Information Criterion (BIC) values for each case; the slightly lower values for the 1-parameter model indicate that it is marginally preferred. The reduced chi-square statistics $\chi^2/\nu \sim 1$ show that the 1-parameter model is a good fit to the data.

4.5 Correlation of Triaxiality with Other Systematics

Orientation bias is one significant systematic for the cluster richness-mass relation; miscentering and projection effects are two others. In modeling these systematics for cluster cosmology, it is important to know the degree to which they may be correlated. In this Section, we explore possible correlation of orientation bias with the other two.

4.5.1 Miscentering

As noted above in Section 4.2.3, in the simulated cluster catalog 37% of the matched clusters are miscentered in the sense that the galaxy identified by redMapper as the BCG is not the central galaxy in the corresponding Buzzard halo. In both the simulation and the real universe, miscentering can happen for a number of reasons. For example, a recent halo merger may result in two nearly-central galaxies of comparable luminosity, or a recent burst of star formation may move the central galaxy's color off the locus of the red sequence. (Cooke et al., 2019; Ragone-Figueroa et al., 2020; Zenteno et al., 2020). Alternatively, a red foreground galaxy along the line of sight to a cluster may be misidentified as the BCG, although Section 4.5.2 indicates that this is rare in the Buzzard simulations.

The miscentering distribution for redMapper clusters in DES Y1 data was estimated through comparison of redMaPPer BCG angular positions with the peaks of X-ray emission for a subsample of clusters with Chandra archival data (Zhang et al., 2019). A number of studies have indicated that X-ray peaks are accurate proxies for the centers of cluster potential wells, though they are subject to systematic errors as well (Lin & Mohr, 2003; Song et al., 2012; Stott et al., 2012; Mahdavi et al., 2013; Lauer et al., 2014). In Zhang et al. (2019), based on 144 redMaPPer clusters with X-ray data, $75 \pm 8\%$ of the redMapper clusters were found to be centered, i.e., they have very small projected separation between redMaPPer BCG and X-ray centroid. For the remainder, the distribution of radial separation between redMaPPer BCGs and X-ray peaks was modeled as a sum of a declining exponential

and a gamma function.

Here, we study the distribution of projected separation, $R_{\rm sep}$, between redMaPPer BCGs and Buzzard central galaxies for halo-matched clusters in the simulation. Since the separation is expected to scale with cluster size, we use the scaled separation, $R_{\rm sep}/R_\lambda$, where $R_\lambda = 1h^{-1}\ {\rm Mpc}(\lambda/100)^{0.2}$ is the characteristic circular aperture for the redMaPPer cluster finder.

We note here the difference in definition between centers. In real data the centering property for a single cluster is not known. Rather the separation distance between optical and X-ray center is modeled as a joint distribution for centered and miscentered clusters with the centered fraction as a model parameter with a maximum likelihood of $75\% \pm 8\%$. By contrast, Buzzard populates halo centers with galaxies using the ADDGALS algorithm, the centering of each individual cluster is a known quantity determined by whether the central galaxy determined by redMaPPer and the halo are one and the same. Among the 23658 halo-matched clusters, 14905 were correctly centered and 8753 are miscentered, the centered fraction being 63% which is within 2σ the centering fraction using X-ray follow-up (Zhang et al., 2019). We define the distance between the redMaPPer chosen BCG and the true halo center as the miscentering separation distance $R_{\rm sep}$.

The resulting separation distribution is shown in Fig. 4.8; the distribution is peaked at $R_{\rm sep} = 0.1R_\lambda$, with a tail that extends to $R_{\rm sep} \simeq R_\lambda$. The shape of the distribution is well-fit by a Γ distribution of functional form

$$P_{\rm miscent}(x|\tau) = \frac{x}{\tau^2}\exp\big(-x/\tau\big), \tag{4.10}$$

where $x \equiv R_{\rm sep}/R_\lambda$. Using methods of least squares, the best fit characteristic scale is found to be $\tau = 0.16$ which is well within the $1-\sigma$ range of the characteristic scale for Chandra to DES center offset found in Zhang et al. (2019). Using the Kolmogorov-Smirnov test, we find that the binned dataset is consistent with the best-fit Gamma distribution at a $\alpha = 0.05$ significance level.

We study differences in the properties of the centered and miscentered cluster populations in the simulation in Fig. 4.9. The upper panel shows that the probability distribution of cluster mass for the centered population is peaked at a slightly higher mass than for the miscentered population, that is, it is the lower-mass clusters that tend to be miscentered, which suggests that this may be a mass dependent bias more prone to low-mass and low-richness clusters. The same trend was not observed with X-ray luminosity and temperature, variables sensitive to the cluster mass with a sample size of only 144 redMaPPer SDSS clusters with X-ray follow up (Zhang et al., 2019). Near-future X-ray surveys as eRosita (Hofmann et al., 2017), which aims to detect 10^5 clusters with a lower mass limit of $\sim 10^{14}\ M_\odot$, will provide a much better handle on the mass distribution of centered and miscentered clusters. The lower panel of Fig. 4.9 shows that the normalized richness distribution of the centered clusters is higher than that of the miscentered ones at $\lambda > 60$, though the difference is marginal.

The centered fraction increases with increasing richness, from 63% for the full sample ($\lambda > 20$) to 60% for $\lambda > 40$, 67% for $\lambda > 60$ and 69% for $\lambda > 80$. This trend is qualitatively consistent with the consistency test carried out on data by Zhang et al. (2019): they compared redMaPPer BCG positions for DES and SDSS clusters where the two data sets overlap and found that for $\lambda > 40$ a large fraction of the BCG positions were within $0.05R_\lambda$ of each other. The archival data from XMM and Chandra has a sharp richness cutoff of $\lambda \gtrapprox 70$ (Farahi et al., 2019b), so any trend of miscentering of BCGs relative to X-ray centroids with richness is not yet detectable with current data.

To quantify the impact of miscentering on the redMaPPer richness estimate in the Buzzard simulations, we consider two approaches. The first method is to recalculate the observed richness by assigning the cluster center onto a different galaxy. It has the advantage that it can also be applied to cluster data but the disadvantage that it involves additional assumptions that have not been fully tested. For each cluster, the redMaPPer algorithm initially

identifies five galaxies as candidates for the BCG. At the end of its iterative procedure, it assigns a final probability of being the BCG to each of these five, produces richness estimates, λ_i, $i = 1, ..., 5$ assuming each of them is the BCG, and identifies the most probable as the BCG, with corresponding richness estimate λ_i. As the probability of it being the true center drops for each candidate, a comparison of richness for clusters targeted at different central candidates would yield information on the potential degree of miscentering for each cluster.

In this first approach, we can quantify the bias in miscentering by taking the ratio of the richness centered at the second most probable galaxy to the first most probable cluster central galaxy among the 5 candidates identified by redMaPPer. This ratio λ_2/λ_1 is an indication of the potential bias in observed richness that miscentering could play when choosing a different cluster center. Among the many selection effects of redMaPPer that come into play in the measurement of this quantity, it is primary a function of the separation distance between the two central candidates—λ_2/λ_1 shifts downward from unity with increasing separation distance $R_{\rm RM\_sep}$ between the cluster candidates, and also notably so does the dispersion increase with $R_{\rm RM\_sep}$. Here $R_{\rm RM\_sep}$ is the separation distance between the two redMaPPer central candidates which in some clusters could be the halo-cluster separation distance $R_{\rm sep}$ but is often not the case. As shown in the left panels of Fig. 4.10 and in Fig. 4.8, $R_{\rm sep}$ goes out to $\sim 1R_\lambda$ while $R_{\rm RM\_sep}$ can be extended to $\sim 2.5R_\lambda$.

The second method of quantifying the impact of miscentering on richness gives a "ground-truth" estimate of the richness bias, but it can only be estimated in the simulation, not from observations. There is a version of the redMapper catalog for the Buzzard simulation, called the *halorun* catalog, in which the redMaPPer BCG is constrained to be the halo central galaxy for each halo-matched cluster. By construction, correctly centered clusters in the *fullrun* redMaPPer catalog that we have been discussing so far have the same richness as those in the *halorun* catalog. On the other hand, for the miscentered *fullrun* clusters, there

is a bias in the estimated richness due to miscentering characterized by

$$\frac{\Delta\lambda}{\lambda} = \frac{\lambda_{\rm fullrun} - \lambda_{\rm halorun}}{\lambda_{\rm fullrun}} \ . \tag{4.11}$$

This fractional shift in richness is plotted as a function of the scaled miscentering separation in the lower left panel of Fig. 4.10.

It is apparent from visual inspection in the left panels of Fig. 4.10 that both methods of quantifying miscentering bias that richness bias increases in amplitude and dispersion with scaled separation as has been shown using DES Y1 clusters with X-ray follow-up data.

Having shown that the miscentering properties of the Buzzard redMaPPer catalog are consistent with those in DES Y1 data, we now turn to examining whether miscentering and triaxiality are correlated systematics. We do this by measuring the miscentering bias as a function of halo orientation, using both of the metrics described above. As the right panels of Fig. 4.10 show, we find that the mean values and dispersion of the two metrics have no systematic dependence on $\cos(i)$. Miscentering and triaxiality can thus be treated as independent systematics.

The fact that we find no correlation between these two systematics is useful for the modeling of systematics in future weak lensing studies but should not come as too unexpected in light of their different physical origins. Miscentering occurs when mergers introduce identical central galaxy candidates or from the star formation properties of the central galaxy that shifts its color out of the red-sequence (Cooke et al., 2019; Ragone-Figueroa et al., 2020; Zenteno et al., 2020), effects completely different from the geometric boosting in richness when clusters are oriented along the line of sight that induce triaxiality bias.

We also test if miscentering can be attributed to line of sight projections whose effect on clusters we describe in detail in subsection 4.5.2. If miscentering is due to projection effects then the BCG at the center of the matched-halo would be of a different redshift and not belong as a member of the matched redMaPPer cluster. Within the allowed $\Delta z \pm 0.05$ redshift

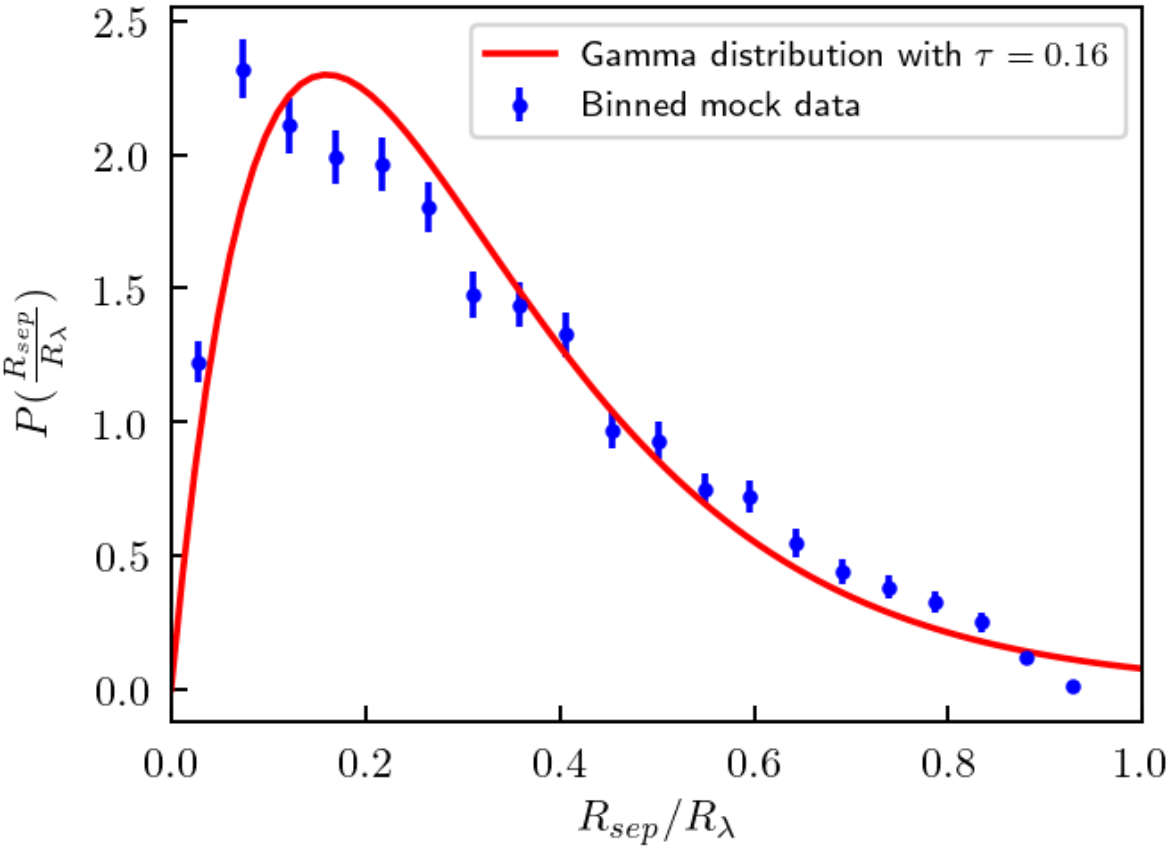


Figure 4.8: Probability distribution of the projected separation between Buzzard halo central galaxies and redMaPPer BCGs in the miscentered population. Scatter plots are the binned mock data points with Poisson error and the line is the best fit Gamma distribution. The two distributions are consistent according to the Kolmogorov-Smirnov test at a $\alpha = 0.05$ significance level.

separation between halo and cluster in our matching algorithm, all of the BCGs at the halo center belong as a member of the matched redMaPPer cluster. Additional tests beyond the scope of this paper need to be conducted to in order to conclude whether miscentering can be attributed to projection effects and if so to what degree, but simulations from Buzzard suggests that this may not be a strong effect.

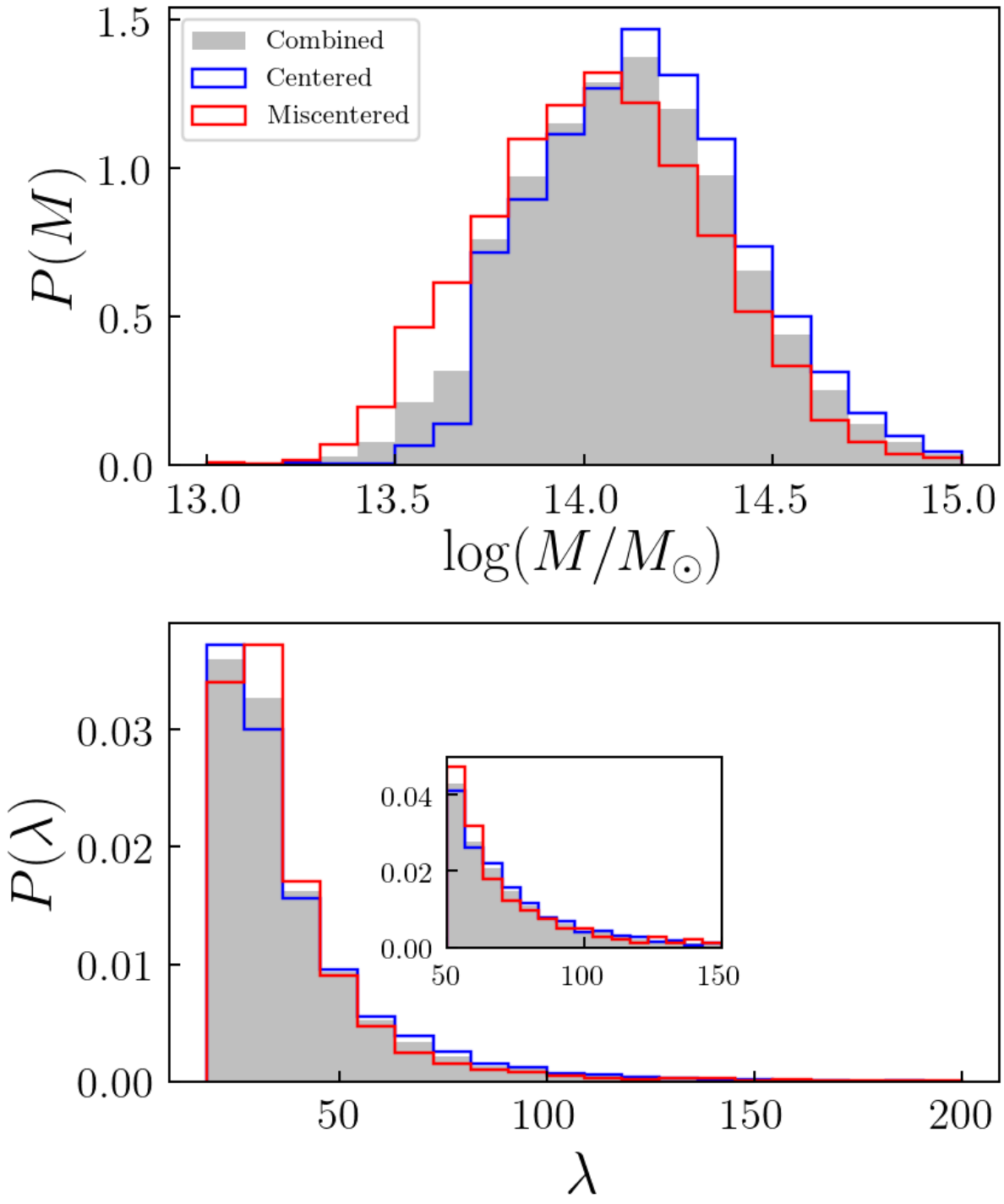


Figure 4.9: *Upper panel:* Mass distribution of the centered and miscentered redMaPPer clusters in the Buzzard simulations. The centered population is peaked at a higher mass. *Lower panel:* Richness distributions of centered and miscentered clusters and for the entire cluster sample. The inset plot shows a slightly higher fraction of centered clusters at high richness.

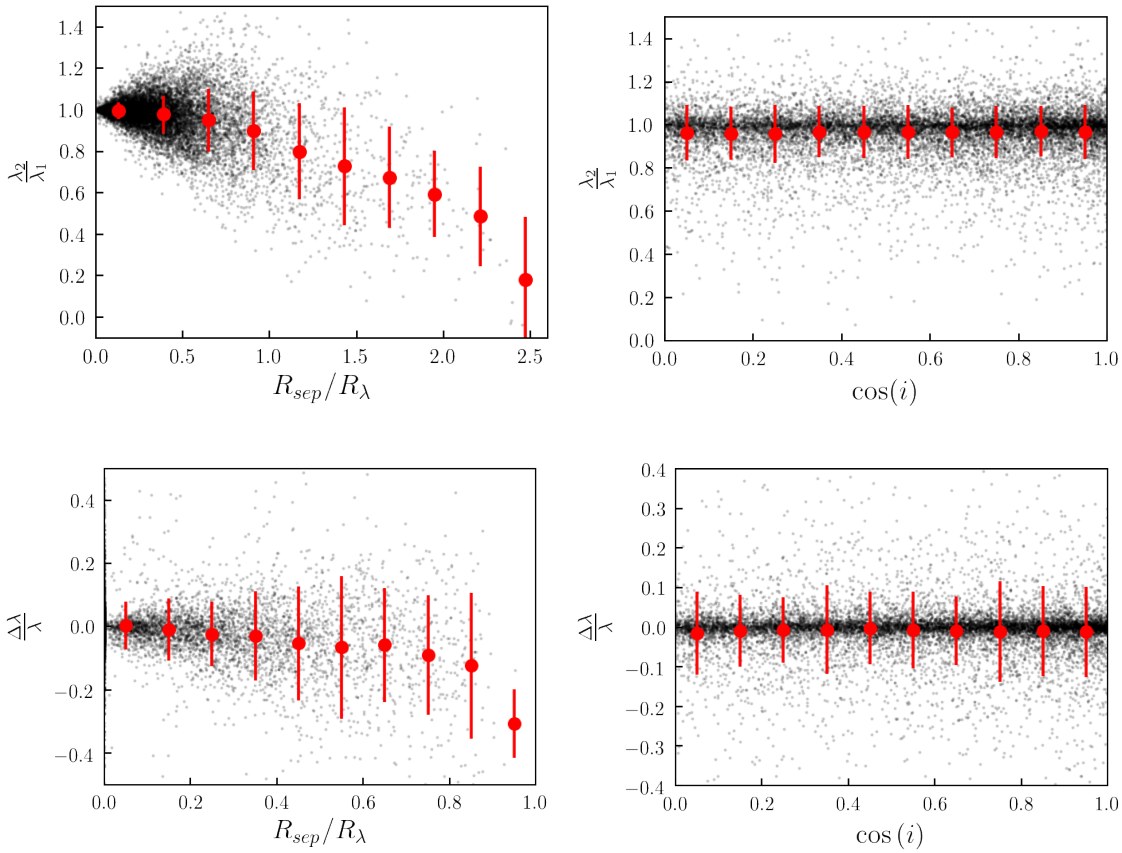


Figure 4.10: *Left panels:* Richness bias vs. miscentering separation for redMaPPer clusters in the Buzzard simulation. Both richness bias metrics λ_2/λ_1 and $\Delta\lambda/\lambda$ show larger bias and increased dispersion at larger miscentering distance. *Right panels:* Richness bias vs. orientation. The mean values of the richness bias metrics show no correlation with halo orientation angle, $\cos(i)$.

4.5.2 Projection

In this section we test for correlations between triaxiality and projection effects. Projections effects were modeled and quantified in Costanzi et al. (2019) using a different Buzzard halo catalog populated according to the assigned "true" richness–mass relation of Simet et al. (2017), and adopts an empirically calibrated back/foreground contamination to account for projection effects on the observed richness. We denote this catalog as the *C19 projection catalog*. Below we summarize the properties of projection effects and the quantities used in the C19 projection mock catalog for our analysis.

Cluster richness suffers from projection effects when non-member galaxies along the line of sight to a cluster are mistakenly classified as cluster members. These may be randomly located galaxies along the line of sight, galaxies spatially correlated with the cluster due to large-scale (e.g., filamentary) structure, or galaxies in a lower-richness cluster along the line of sight that "leak" into a larger one, a process in redMaPPer known as *percolation* (Costanzi et al., 2019). In combination, they bias the observed richness $\lambda^{\rm obs}$ away from the true richness $\lambda^{\rm true}$ by the amount:

$$\lambda^{\rm obs} - \lambda^{\rm true} = \Delta^{\rm bkg} + \Delta^{\rm prj}_{\rm non-cor} + \Delta^{\rm prj}_{\rm LSS} + \Delta^{\rm prc}, \tag{4.12}$$

where each term on the right hand side of the equation respectively denotes the background, non-correlated projection, large scale structure and percolation term.

Each component contributes to the observed richness in a different form. Background scatter, $\Delta^{\rm bkg}$, is assumed to be normally distributed around the true richness. The sum of the projection terms due to non-correlated clusters, $\Delta^{\rm prj}_{\rm non-cor}$, and correlated large-scale structure, $\Delta^{\rm prj}_{\rm LSS}$, are modeled as an exponential function with a cutoff at $\Delta^{\rm prj} \geqslant 0$, to ensure an upscatter of $\lambda_{\rm obs}$ as is physically motivated. The observed richness is painted on in the mock catalog by summing the richness of clusters along the light of sight weighted by the

redshift kernel $w(\Delta z, z)$:

$$\lambda_i^{\rm obs} = \lambda_i^{\rm true} + \Delta_i^{\rm prj} = \lambda_i^{\rm true} + \sum_{j\neq i}^{N} \lambda_j^{\rm true} f_{ij}^{\rm A} w(\Delta z_{ij}, z_j), \tag{4.13}$$

where f_{ij}^{A} is the geometric masking fraction of object j over i for an object j that's (partially) in the line of sight of i, and $w(\Delta z_{ij}, z_j)$ the redshift kernel which, as a function of redshift of i and the redshift separation between i and j is modeled as the functional form:

$$w(\Delta z|z_{\rm cl}) = \begin{cases} 1 - \frac{(\Delta z)^2}{\sigma_z(z_{\rm cl})^2}, & |\Delta z| < \sigma_z(z_{\rm cl}) \\ 0, & \text{otherwise,} \end{cases} \tag{4.14}$$

which can be interpreted intuitively as the diminishing strength of projection effects with redshift separation $|\Delta z|$ up to a maximum separation of $\sigma_z(z_{\rm cl})$.

For each cluster, its $\sigma_z^{\rm cl}(z)$ is fitted by sliding the redMaPPer redshift center away from the true cluster redshift so as to remove the excess richness $\Delta^{\rm prj}$ due to projection as a function of the redshift separation between assigned and true redMaPPer redshift. To recover the "leakage" function for clean line of sights, Costanzi et al. (2019) chooses the lower 5% of clusters in a given redshift as the leakage function. It is fit with a piecewise log-linear model with a transition at $z = 0.32$. Data from SDSS redMaPPer clusters (Costanzi et al., 2019) show that at $z \lesssim 0.3$ projections are from the width of the red-sequence and increase monotonically with increasing redshift from increasing photometric errors. At $z \gtrsim 0.3$ projection effects flatten out as the SDSS survey is no longer volume limited but magnitude limited, the faintest cluster galaxies residing near the magnitude limit of the survey at redshift above 0.3.

In this paper we introduce the derived quantity

$$\log\left(\sigma_z^{\rm proxy}(z_{\rm cl})\right) = \log\left(\sigma_z^{\rm cl}(z_{\rm cl})\right) - \log\left(\sigma_z^{5\%}(z_{\rm cl})\right) \tag{4.15}$$

as the difference between the log-scaled σ_z of an individual cluster and the lower 5% envelope of σ_z for all clusters at the redshift bin of the cluster. This quantity $\sigma_z^{\rm proxy}(z_{\rm cl})$ can be seen as the level of intrinsic excess projection after eliminating background noise and redshift-dependent observational biases.

Percolation is added into the full model of projection when clusters of lower richness are "absorbed" into one with higher richness. For each cluster j with richness smaller than that of i, the richness is taken from j to i by the amount

$$\Delta_{\rm i}^{\rm prc} = \sum_{j<i}^{N} \lambda_j^{\rm true}\big(1 - f_{ij}^{\rm A} w(\Delta z_{ij}, z_j)\big), \tag{4.16}$$

whose probability distribution $P(\Delta^{\rm prc}|\lambda^{\rm true}, z)$ is empirically determined to well resemble a boxcar function with $\Delta^{\rm prc} \in [-\lambda^{\rm true}, 0]$.

In the C19 projection catalog, each cluster is assigned a true richness using an empirically calibrated richness–mass relation from Simet et al. (2017) and given an observed richness using the projection effect algorithm described above by way of the redshift kernel $w(\Delta z|z_{cl})$. Hence the difference between the true and observed richness in this mock is due to projection effects alone. The probability distribution for $P(\Delta|\lambda^{\rm true}, z)$ for each component is then fit using this C19 projection mock, and upon convolution of the probability distributions for each individual component in Equation 4.12 we arrive at the final expression for $P(\lambda^{\rm obs}|\lambda^{\rm true}, z)$. We refer the reader to Costanzi et al. (2019) for the full expression and best-fit parameters. All halos in the mock projection catalog are artificially assigned an observed and true richness, whether or not such a halo could be detected and matched to a redMaPPer cluster. The observed richness is thus biased only from projection effects and does not suffer from all the other selection effects, including triaxiality and miscentering, that would exist had the halos undergone redMaPPer detection and cluster matching. This technique effectively isolates projection effects from potentially correlated systematics in the same vein that we used the

halorun catalog to isolate miscentering effects.

We find that projection effects are independent from triaxiality. Figure 4.11(a) shows that $\sigma^{\rm proxy}$, the strength of projection effects due to large-scale structure, is not correlated with $\cos(i)$. We further inspect the full scope of projection effects by studying the fractional difference between the observed and true richness in the projection mock catalog,

$$\frac{\Delta\lambda_{\rm prj}}{\lambda_{\rm prj}} = \frac{\lambda_{\rm prj}^{\rm obs} - \lambda^{\rm true}}{\lambda_{\rm prj}^{\rm obs}}, \tag{4.17}$$

which shows no correlation with $\cos(i)$, as shown in Figure 4.11(b). Finally, we run our fit to the richness–mass relation in the projection catalog of $\lambda_{\rm prj}^{\rm obs}$ binned in $\cos(i)$ and observe no difference in the observed richness–mass relation, shown in Figure 4.12. The $1-\sigma$ range of the best-fit parameters for the log-linear richness–mass template between different $\cos(i)$ bins all closely overlap with each other, with no clear trend.

The lack of correlation between projection and orientation may be puzzling at first in light of a common physical origin of these effects. The ΛCDM model of hierarchical structure formation facilitates the preferential gravitational collapse of dark matter halos that become galaxy clusters along the nodes of large-scale filaments. It is also widely understood that a halo's semi-major axis is preferentially aligned with the direction of the associated filament for halos residing in over-densities (e.g. Hahn et al. (2007), Forero-Romero et al. (2014)). It is thus sensible to expect a correlation between the strength of projection effects and halo orientation for halos residing in filaments.

The lack of correlation can be explained by the stochasticity of these effects along with the fact that not all halos share the same physical origin for this set of systematics. The boosting in richness from projection is from uncorrelated background noise and correlated large scale structure, the latter playing a much larger role. Adding the large scale structure into the modeling of the observed richness for projection boost the richness perturbation $\Delta^{\rm prj}$ by a factor of 2 and 4 in the $\lambda^{\rm true}$ range $20-100$ (Costanzi et al., 2019). It is also observed

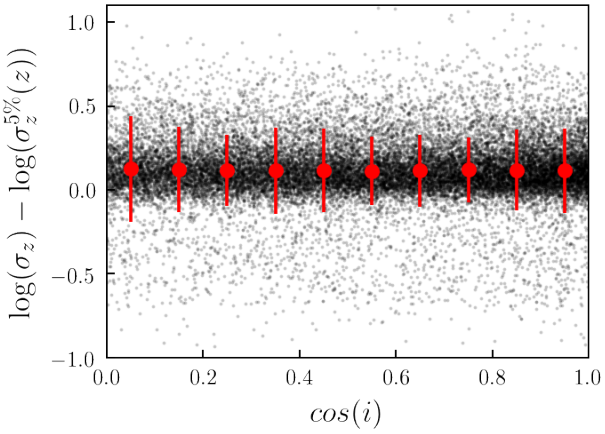
by N-body simulations from Sunayama et al. (2020) that a minority of clusters that reside in large scale filaments is responsible for the boosting of the stacked weak lensing signal of halos (see Section 4.6 on weak lensing) due to projection effects. This set of studies suggests that a small batch of clusters is responsible for the large degree of bias from projection effects.

Triaxiality bias, on the other hand, can occur whether halos reside in large scale filaments or in voids. That all halos, regardless of its external environment, is subject to the same degree of triaxiality bias while not the case for projection bias would explain the lack of correlation among an ensemble of stacked clusters. It would be interesting as a follow-up study to know if the correlation between projection and triaxiality can be detected for the minority of clusters residing in large scale structures that heavily boost the projection observable, but for the purposes of modeling redMaPPer selection effects, it is sufficient to know that for the entire sample of $\lambda^{\rm obs} > 20$ clusters detectable by redMaPPer, projection and triaxiality can be treated as separate systematics. A further study using spectroscopic redshift measurements of redMaPPer member galaxies from Magellan telescope data (Gruen D., in prep.) will provide the shape and orientation of clusters as well as test for non-member galaxies projected along the line of sight misidentified by redMaPPer, serving as a follow-up test of the correlation of these systematics using real data.

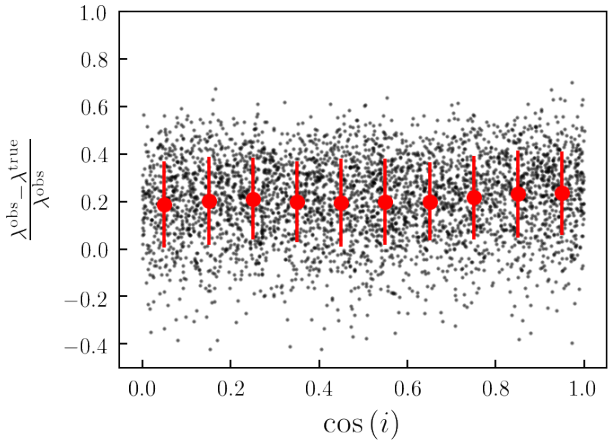
4.6 Effect of Halo Orientation on the Weak Lensing Profile

The effects of triaxiality on cluster optical detection are twofold—one through the boosting of the richness-mass relation as was covered in Section 4.4, the other through the boosting of radially dependent weak lensing signals.

This section quantifies the latter effect. It is split into three subsections—Section 4.6.1 models the boosting effect of the cluster weak lensing signal in the Buzzard simulations for individual halos before applying the redMaPPer cluster finder; Section 4.6.2 combines the result from Section 4.6.1 and our richness-mass model from Section 4.4 to predict the



(a) Correlation of σ_z^{proxy} with halo orientation.



(b) Correlation of $\frac{\Delta\lambda}{\lambda}$ with halo orientation.

Figure 4.11: Correlation of projection strengths and halo orientations measured in two mock catalogs. Top panel shows the measurement in the Buzzard simulations, where the σ_z^{proxy} (defined in Equation 4.15) is used to estimate the strength of projection effects. Bottom panel shows the measurement in the C19 projection mock, which is constructed using the same halo catalog as the Buzzard simulations. In the C19 catalog the galaxies are populated using a richness–mass relation and the observed richness is generated using a semi-analytic model (described in Section 4.5.2). In this mock, because we know the true galaxy content in each halo, we use the fractional difference between the observed richness and true richness (defined in Equation 4.17) as a proxy for projection. In both panels, we find there is no correlation between projection strengths and halo orientations.

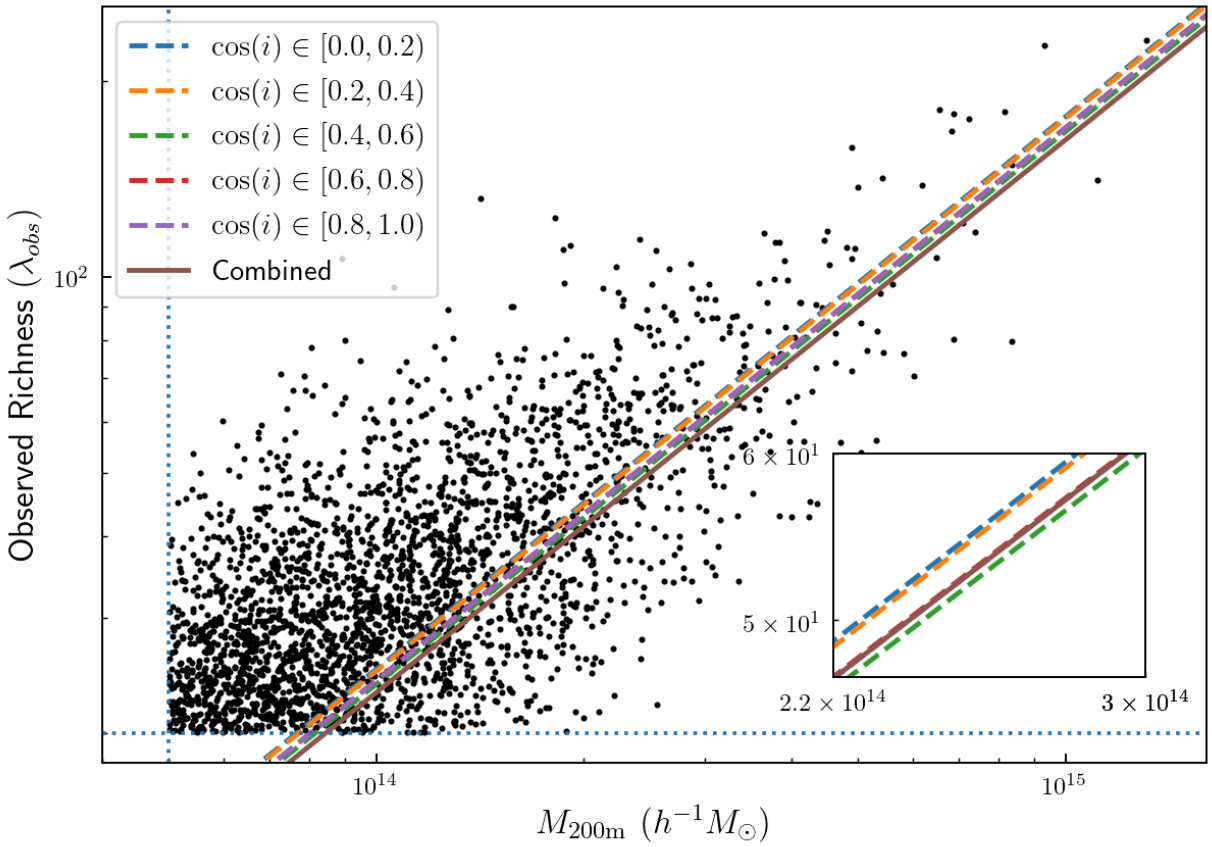


Figure 4.12: Observed richness–mass relation for different orientation bins in the projection mock catalog. No difference is observed in the observed richness–mass relation in the projection catalog with clusters of different orientation bins. The $1-\sigma$ contours for the best fit parameters $\ln(A)$, B and σ_0 (not shown) in all bins closely overlap with one another, indicating no correlation between the two systematics. The dashed horizontal line indicates the richness cut at $\lambda > 20$ and dashed vertical line the mass cut at $M_{200\mathrm{m}} > 5 \times 10^{13}\ h^{-1}M_\odot$.

observed boosting in stacked cluster lensing profiles at different richness bins after redMaPPer selection; Section 4.6.3 uses the result from Section 4.6.2 to conduct a Fisher matrix forecast on the mass bias of triaxiality for redMaPPer clusters stacked in different richness bins.

4.6.1 Modeling the effects of halo orientation on excess surface density before redMaPPer selection

In this section, we measure the excess surface densities of all halos with convergent shape measurements in a lightcone of $z < 0.90$. The masses of halos are binned in mass bins of $[5 \times 10^{13},\ 10^{14})$, $[10^{14},\ 2 \times 10^{14})$, $[2 \times 10^{14},\ 4 \times 10^{14})$ and $[4 \times 10^{14},\ \infty)$ $h^{-1}M_{\odot}$, and redshift bins of $[0,\ 0.34)$, $[0.34,\ 0.5)$, $[0.5,\ 0.7)$ and $[0.7,\ 0.9)$, for a total of 16 bins.

Another common expression for the density inside a halo is the halo–matter correlation $\xi_{hm}(r)$, which is related to the surface density Σ through the relation

$$\Sigma(R) = \overline{\rho}_m \int_{-\infty}^{+\infty} \left(1 + \xi_{hm}\Big(r = \sqrt{R^2 + z^2}\Big)\right) dz, \tag{4.18}$$

where $\overline{\rho}_m$ is the mean matter density at the redshift of the cluster, R is the projected radius in the plane of the sky, and z is the length along the line of sight.

In weak lensing, the tangential shear γ_t of the galaxies relative to the center of each foreground halo is related to the excess surface density by the relation

$$\Sigma_{\rm crit}\gamma_t = \overline{\Sigma}(< R) - \Sigma(R) \equiv \Delta\Sigma(R), \tag{4.19}$$

where the critical surface density $\Sigma_{\rm crit}$ defined as

$$\Delta\Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{D_s}{D_l D_{ls}}, \tag{4.20}$$

and where D_s, D_l and D_{ls} refer to the angular diameter distances to the source, to the lens,

and between the lens and source, respectively.

In this paper we measure $\Delta\Sigma(R)$, which has a one-to-one relationship with $\Sigma(R)$ and γ_t, all of which can be determined from the underlying halo–matter correlation $\xi_{hm}(r)$ and a fiducial cosmology for determining $\Sigma_{\rm crit}$. In the following sections, in order to reduce the clutter in the equations for modelling excess surface density as a function of orientation we use μ as a shorthand for $\cos(i)$.

When we measure $\Delta\Sigma(R)$ from the simulations, we use projected radii R extending from 0.1 h^{-1} Mpc to 100 h^{-1} Mpc in 30 equally log-spaced bins, and a projected distance symmetric about the halo of $\Delta D_p = 10,\ 50,\ 100,\ 200\ h^{-1}$ Mpc. For ease of visualization, the orientation dependence is plotted and fitted onto a template as the quantity

$$F(R,\mu) = \log \frac{\Delta\Sigma(R,\mu)}{\overline{\Delta\Sigma}(R)}, \tag{4.21}$$

where $\Delta\Sigma(R,\mu)$ is the average profile in an orientation bin for a given mass and redshift bin, and $\overline{\Delta\Sigma}(R)$ is the averaged profile across all orientation bins in the same mass and redshift bin.

The shapes of the profiles can be roughly divided into the "one-halo" regime ($R \lesssim R_{200m}$) and the "two-halo" regime ($R \gtrsim R_{200m}$) (Fig. 4.13). In the one-halo regime, halos with their major axes oriented towards the line of sight are boosted in their surface density relative to the mean, a result well explained by the triaxial halo model (Oguri et al., 2005; Corless & King, 2008). The transition between the one- and two-halo regimes produces a neck in the surface density, where the halo–matter correlation from neither regime dominates. In the two-halo regime, the trends of the lensing ratios in different orientation become inverted with respect to unity when increasing the projection depth from $\Delta D_p = 10h^{-1}$Mpc to $\Delta D_p = 200h^{-1}$Mpc. At $\Delta D_p = 10h^{-1}$Mpc, the ratio of excess surface densities in the two-halo regime of high $\cos(i)$ halos drop below the mean, which may be explained by an under-dense region surrounding the plane perpendicular to the major axes of the halos. As one

moves towards larger projection depths, halos with higher $\cos(i)$ exhibit boosted $\Delta\Sigma$ profiles in the two-halo regime relative to the mean as a result of the alignment of halos with their underlying large scale structure, i.e., the large projection depth captures much of the mass in the large-scale filaments for halos with $\cos(i) \sim 1$ (Hahn et al. (2007), Forero-Romero et al. (2014)). Because of the similarity of excess surface density profiles for $\Delta D_p = 100\ h^{-1}$Mpc and $\Delta D_p = 200 h^{-1}$Mpc, we deem the projection length $\Delta D_p = 100\ h^{-1}$Mpc as convergent. The excess surface density profiles in the one- and two-halo regimes and their dependence on projection depth agree well with Osato et al. (2018), who built profiles for a simulation of similar projections depths and with comparable mass resolution.

We model the log ratio of excess surface density, $F(R, \mu)$, in a $\mu \equiv \cos(i)$ bin relative to the mean with six free parameters given by the product of a multipole expansion over $\cos(i)$ and a Cauchy function:

$$\begin{aligned} F(R, \mu) &= A(\mu) f(R) \\ A(\mu) &= A_0 + A_1 \mu + A_2 \mu^2 + A_3 \mu^3 \\ f(x \equiv \ln(R)) &= 1 - \frac{1}{(x - x_0)^2 + \gamma}. \end{aligned} \tag{4.22}$$

The bottleneck shape of the $\Delta\Sigma$ profiles binned by $\cos(i)$ is well captured by the Cauchy function in most of the mass and redshift bins, with best-fit parameters and p-values listed in Table 4.2 and plotted in Figure 4.14. The parameters show no clear sign of monotonic evolution with mass or redshift that may hint at underlying physics, but they do differ in value from bin to bin, so for greater accuracy the templates are divided into different bins when estimating the stacked mass bias due to triaxiality as will be shown in subsection 4.6.3. The best-fit parameters are determined using a Nelder-Mead minimization method; with 10 log-spaced bins in each $\cos(i)$ binned $\Delta\Sigma$ profile and 5 $\cos(i)$ bins, the templates are fitted with 6 free parameters, totalling $5 \times 10 - 6 = 44$ degrees of freedom; the χ^2 and p-value

are calculated for each fit. Of the 16 fits, 8 have left- or right-handed p-values within 0.01, and 11 within 0.001. The over-fitted templates occur in high-mass or high-redshift bins, which suffer larger errors from the dearth of dark matter particle sampled in each bin, and the under-fitted ones result from a mismatch in the "two-halo regime" that exhibits more poorly constrained trends from bin to bin and the behavior of which is less well understood. Qualitatively, the fits preserve the basic underlying shape of the excess surface density ratios, as shown in Fig. 4.14.

The templates provided could be used as correction terms for Stage III and IV weak lensing cluster surveys such as in the comsoSIS pipeline (Zuntz et al. (2015)) for DES-Y3.

4.6.2 Modeling the effects of halo orientation on richness-binned excess surface density after redMaPPer selection

Table 4.2: Best fit parameters for Equation 4.22 across different mass and redshift bins.

$z_{\min}$	$z_{\max}$	$M_{\min}(M_\odot)$	$M_{\max}(M_\odot)$	A_0	A_1	A_2	A_3	x_0	γ	χ^2	Left-tail p-value	Right-tail p-value
0.00	0.34	5×10^{13}	1×10^{14}	-0.157	-0.001	0.091	0.485	1.346	0.378	85.204	0.9998	0.0002
0.00	0.34	1×10^{14}	2×10^{14}	-0.168	-0.0107	0.222	0.375	1.325	0.592	25.580	0.012	0.988
0.00	0.34	2×10^{14}	4×10^{14}	-0.197	0.373	-0.818	1.112	1.289	0.757	34.009	0.139	0.861
0.00	0.34	4×10^{14}	1×10^{16}	-0.190	-0.270	1.307	-0.457	1.245	1.504	18.551	0.9997	0.0003
0.34	0.50	5×10^{13}	1×10^{14}	-0.204	0.264	-0.489	0.909	1.320	0.403	65.605	0.981	0.020
0.34	0.50	1×10^{14}	2×10^{14}	-0.190	0.238	-0.472	0.888	1.261	0.782	23.623	0.995	0.005
0.34	0.50	2×10^{14}	4×10^{14}	-0.281	0.952	-2.056	1.913	1.342	1.141	25.567	0.988	0.012
0.34	0.50	4×10^{14}	1×10^{16}	-0.021	-0.268	-0.681	1.504	1.146	1.344	28.903	0.962	0.038
0.50	0.70	5×10^{13}	1×10^{14}	-0.212	0.190	-0.174	0.669	1.292	0.523	91.768	1.000	0.000
0.50	0.70	1×10^{14}	2×10^{14}	-0.203	0.103	0.017	0.547	1.307	0.784	66.490	0.9841	0.016
0.50	0.70	2×10^{14}	4×10^{14}	-0.214	0.095	0.213	0.350	1.228	1.126	48.257	0.305	0.695
0.50	0.70	4×10^{14}	1×10^{16}	-0.036	-0.996	2.188	-0.780	1.148	1.514	88.200	0.9999	0.0001
0.70	0.90	5×10^{13}	1×10^{14}	-0.208	0.209	-0.263	0.738	1.290	0.564	99.745	1.000	0.000
0.70	0.90	1×10^{14}	2×10^{14}	-0.213	0.305	-0.612	1.030	1.29	0.931	71.863	0.995	0.005
0.70	0.90	2×10^{14}	4×10^{14}	-0.287	0.655	-0.975	1.105	1.243	1.226	33.589	0.873	0.127
0.70	0.90	4×10^{14}	1×10^{16}	-0.158	-1.184	3.551	-1.839	1.260	1.904	21.298	0.9985	0.0015

Stacking refers to the process of building averaged excess surface density profiles of halos in different richness bins. This subsection describes the process of stacking used by the DES survey to calibrate the richness–mass relation and presents the effect of triaxiality on the stacked surface density.

The shapes of source galaxies behind a cluster along the line of sight will have small tangential distortions due to gravitational lensing. While individual distortions are small, this tangential shear can be measured at high signal to noise as a function of projected radial separation R in the stacked images of source galaxies around clusters binned, e.g., in richness and redshift. In the weak lensing regime, the tangential shear is related to the source-galaxy ellipticity by

$$\gamma_t \approx e^T + \text{noise}, \tag{4.23}$$

where e^T is the source ellipticity rotated to the tangential frame, and the noise is due to intrinsic ellipticities of the source galaxies (shape noise) and measurement uncertainty. The tangential shear, γ_t, as directly measured by observations can be converted to $\Delta\Sigma(R)$ through Equation 5.5. This paper directly measures $\Delta\Sigma(R)$ by computing the 2D dark-matter density along a cylinder of given projection depth centered around the cluster.

The model excess surface density is obtained by integrating the halo–matter correlation $\xi_{hm}(r)$ along the line of sight as in equation 5.2, and subtracting that from the mean surface density inside the projected radius as in equation 5.5. Typically, the halo–matter correlation in the "one-halo" regime is modeled as a spherical Navarro-Frenk-White (NFW) (Navarro et al., 1997) profile $\rho_{\rm NFW}(r|M)$,

$$\xi_{\rm 1h}(r|M) = \frac{\rho_{\rm NFW}(r|M)}{\rho_{m0}} - 1, \tag{4.24}$$

and the "two-halo" term as a linear matter correlation (Hayashi & White, 2008) scaled by

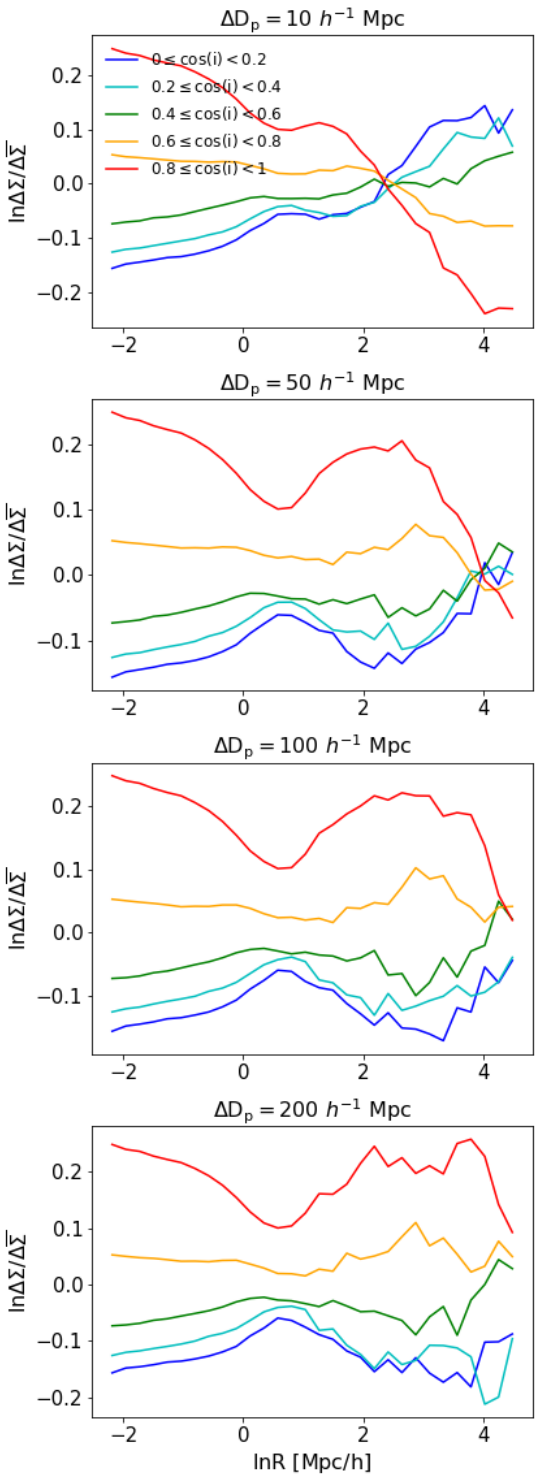


Figure 4.13: $\Delta\Sigma(R, \mu)$ for $M \in [10^{14}, 5 \times 10^{14})$ $M_\odot$ as a function of projection depth, ΔD_p. The lensing ratios in the "two-halo" regime reverses trends from low to high projection depth as a result of alignment of clusters with the large scale structure. The profiles with $\Delta D_p = 100$ h^{-1} Mpc are deemed convergent due to their similarity with the $\Delta D_p = 200$ h^{-1} Mpc profiles.

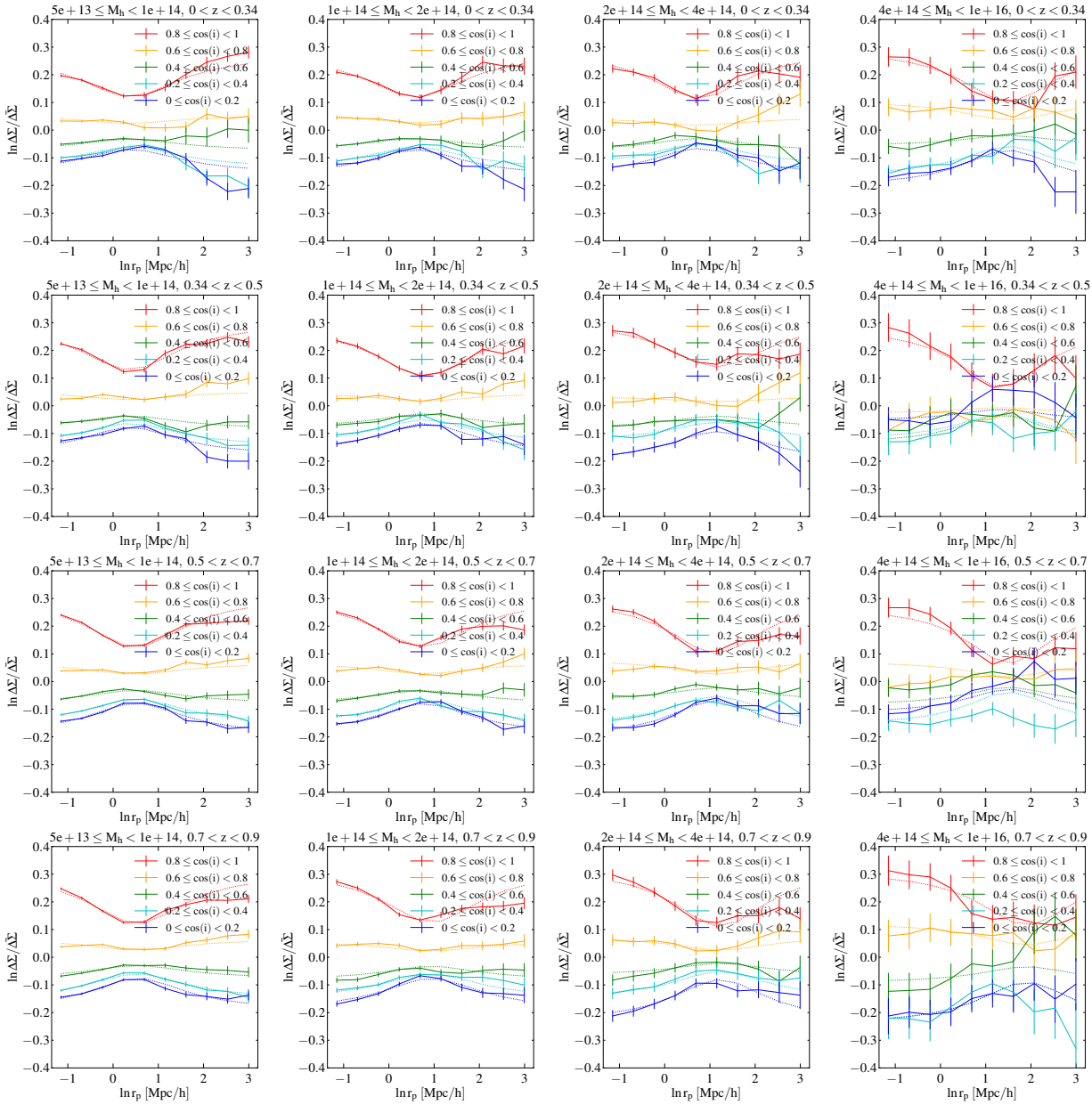


Figure 4.14: Stacked $\Delta\Sigma$ profiles in different orientations bins (solid lines) vs. Cauchy function fits (dashed lines) to the profiles governed by Equation 4.22 and with best fit parameters listed in Table 4.2. Error bars are the $1-\sigma$ deviations in measurements in a given orientation and radial bin.

the halo bias, (e.g. Tinker et al. (2010)):

$$\xi_{\rm 2h}(r|M) = b^2(M)\xi_{\rm lin}(r). \tag{4.25}$$

At the transition between the two regimes, DES Y1 follows Zu et al. (2014) in setting the halo–matter correlation to the maximum value of the two terms, i.e.,

$$\xi_{hm}(r|M) = \max\ \{\xi_{\rm 1h}({\rm r|M}),\ \xi_{\rm 2h}({\rm r|M})\} \tag{4.26}$$

In our analysis we reproduce the surface density templates from the procedures in the DES Y1 analysis using publicly available code—the linear power spectrum computed from CLASS (Lesgourgues (2011); Blas et al. (2011)) and the excess surface density computed from the *cluster_toolkit* module[2], which uses the spherical NFW profile for the "one-halo" term and refers to Tinker et al. (2010) for the halo bias—to generate isotropic profiles, which we denote $\overline{\Delta\Sigma}(R)$, calculated by integrating through Equation 5.2 ξ_{hm} in the form of Equation 5.7. In the "one-halo" regime we parametrize the NFW profile with a nominal concentration of $c_{200m} = 5$.

We investigate the difference in the stacked profile between the isotropic $\overline{\Delta\Sigma}(R)$ and $\overline{\Delta\Sigma}(R, M, \mu)$, the stacked profile as a function of orientation dependence.

The orientation dependence has two components—one is the scaling of individual lensing profiles by $\exp(F(R, \mu))$ as described in Section 4.6.1, and the other the effect of richness-mass, $P(\lambda|M, \mu)$, as modeled in Section 4.4, on the mass distribution of redMaPPer-selected clusters. The second component, $P(\lambda|M, \mu)$, biases the mass distribution of clusters in a

2. Code written by Tom McClintock. Source: http://cluster-toolkit.readthedocs.io/en/latest/index.html

richness bin $\widetilde{P}(M)$ through the form

$$\begin{aligned}\widetilde{P}(M) &= \int d\mu \int_{\lambda_1}^{\lambda_2} d\lambda P(M,\lambda,\mu) \\ &= \int d\mu \int_{\lambda_1}^{\lambda_2} d\lambda P(\lambda|M,\mu)P(\mu|M)P(M) \\ &\text{and safely assuming that } P(\mu|M) \text{ is constant,} \\ &\propto \int d\mu \int_{\lambda_1}^{\lambda_2} d\lambda P(\lambda|M,\mu)P(M), \end{aligned} \tag{4.27}$$

where $P(M)$ is the mass function of redMaPPer-selected clusters.

The conditional probability of richness, $P(\lambda|M,\mu)$, is log-normally distributed around a mean richness governed by Equation 4.6, and the standard deviation is given by equation 4.8. The equations are fit to the one-parameter model in which only $\log(A)$, the intercept of the $\log(\lambda)$-$\log(M)$ relation, is allowed to vary with orientation. We use a cubic spline to interpolate $\log(A)$ for $\mu \in [0,1)$. The halo-mass function of redMaPPer-selected clusters, $P(M)$, is constructed from a discrete histogram with 30 log-spaced mass bins in the mass range of the clusters.

Taking into account the two components for orientation dependence, the stacked surface density in a richness bin becomes

$$\begin{aligned}&\overline{\Delta\Sigma}(R,M,\mu) \text{ for } \lambda \in [\lambda_1,\lambda_2) \\ &= \int dM\ \Delta\Sigma(R,M,\mu)\widetilde{P}(M) \\ &= \int d\mu \int dM \int_{\lambda_1}^{\lambda_2} d\lambda\ \Delta\Sigma(R,M,\mu)P(\lambda|M,\mu)P(\mu|M)P(M) \\ &\propto \int d\mu \int dM \int_{\lambda_1}^{\lambda_2} d\lambda\ \Delta\Sigma(R,M,\mu)P(\lambda|M,\mu)P(M) \end{aligned} \tag{4.28}$$

The excess surface densities are computed for $\langle\Delta\Sigma(M,R,\mu)\rangle$ using equation 4.28 and $\overline{\Delta\Sigma}(M,R)$ using equations 5.3–5.7. We define the fractional difference with the shorthand

notation

$$\delta\langle\Delta\Sigma\rangle = \frac{\overline{\Delta\Sigma}(R,\mu) - \overline{\Delta\Sigma}(R)}{\overline{\Delta\Sigma}(R)}. \tag{4.29}$$

4.6.3 Mass bias estimation of stacked clusters

We are interested in estimating the effect of triaxiality on the mean weak lensing mass in clusters stacked in richness bins. The weak lensing mass is an observed quantity in weak lensing surveys derived by fitting the observed lensing profile to an analytic profile in a procedure akin to that in Section 4.6.2 and is used to constrain the mass-richness relation. We estimate the bias due to triaxiality on the weak lensing mass for stacked clusters by propagating the error on the lensing observable onto the mass model parameter using a Fisher matrix approximation.

In the most generic sense, the Fisher matrix F_{ij} in a given radial bin is defined as:

$$F_{ij}(R) = \frac{\partial\langle\Delta\Sigma\rangle(R)}{\partial p_i}\mathrm{Cov}(\langle\Delta\Sigma\rangle(\mathrm{R}))^{-1}\frac{\partial\langle\Delta\Sigma\rangle(R)}{\partial p_j}, \tag{4.30}$$

where the partial derivatives are of surface density profiles with respect to model parameters p_i of cluster mass M and concentration c, and the covariance matrix is that of surface density as a function of radius.

The mass-bias for stacked clusters due to triaxiality is given by the expression

$$\delta M_{\mathrm{binned}} = \sum_j (F^{-1})_{ij}\left[(\delta\langle\Delta\Sigma\rangle)\mathrm{Cov}(\langle\Delta\Sigma\rangle)^{-1}\frac{\partial\Delta\Sigma}{\partial p_j}\right], \tag{4.31}$$

estimated by inserting the fractional difference of stacked profiles, $\delta(\Delta\Sigma)$, into the bracketed expression and marginalizing over the concentration parameter. The total bias is then the

weighted sum of all mass and redshift bins marginalized over concentration and radius:

$$\delta M_{\rm total} = \sum_{M,z} P(M,z|\lambda) \left[\sum_{j,R} (F^{-1})_{ij}(R) \left(\delta\langle\Delta\Sigma\rangle {\rm Cov}(\langle\Delta\Sigma\rangle)^{-1} \frac{\partial\Delta\Sigma}{\partial p_j} \right) \right]. \tag{4.32}$$

The $\langle\Delta\Sigma\rangle$ profiles are binned in richness intervals of $\lambda \in [20,\ 30),\ [30,\ 50)$, and $[50,\ \infty)$, and are further divided into the same mass and redshift bins when computing individual $\Delta\Sigma(R)$ templates as described in Section 4.6.1. We make the simplifying assumption that the partial derivative of the bin-averaged surface density profile is well approximated by that for a numerical model for an individual halo, with M taken at the midpoint of the mass bin, and c derived from redshift and mass using the relation

$$c_{200m} = c_0 \Big(\frac{M}{M_0} \Big)^{-\beta}, \tag{4.33}$$

with functional form and best fit parameters of $c_0 = 4.6$ at $z = 0.22$ and $\beta = 0.13$ at a pivot mass of $M_0 = 10^{14} h^{-1} M_\odot$ (Mandelbaum et al., 2008), calculated at the midpoint value of said mass bin. The concentration-mass relation from Mandelbaum et al. (2008) is derived from a red-sequence finder in the SDSS survey in redshifts and mass ranges compatible with redMaPPer on DES Y1. We find that the impact on different concentration-mass relations e.g. (Oguri et al., 2012; Diemer & Joyce, 2019) has a sub 1% impact on the mass bias when folded into Eqn. 4.32. The approximation of $\langle\Delta\Sigma(R)\rangle$ profiles is computed using *cluster_toolkit* for the Buzzard cosmological parameters.

The covariance matrix for cluster weak lensing is taken from Wu et al. (2019), who calculated the matrices from a combination of analytic calculations and high-resolution N-body simulations for radii between 0.1 and 100 h^{-1} Mpc, discretized at 15 equally log-spaced bins. The covariance comes from a combination of shape noise, large scale structure and intrinsic noise. Modeled on a DES-like simulation with a galaxy density of $n_s \sim 10$/arcmin, the covariance is dominated by shape noise at projected radii $\lessapprox\ 5h^{-1}$ Mpc. The covariance

matrices are binned by mass in bins of $[10^{14},\ 2\times10^{14})$, $[2\times10^{14},\ 4\times10^{14})$ and $[4\times10^{14},\ \infty)$ $h^{-1}M_{\odot}$, and in lens/source redshift slices of $\{z_l=0.3, z_s=0.75\}$, $\{z_l=0.5, z_s=1.25\}$ and $\{z_l=0.7, z_s=1.75\}$, with z_l denoting the lens redshift and z_s the source redshift.

To address the different binning schemes used in the lensing covariance and stacked lensing profiles, we choose to evaluate the covariance at the central redshift slice of $\{z_l=0.5, z_s=1.25\}$, since the redshift dependence of the lensing covariance is weak. Because the covariance matrix is not applicable for masses below 10^{14} $h^{-1}M_{\odot}$, we ignore $\langle\Delta\Sigma(M,z)\rangle$ in the modeling for Equation 4.32 for the lowest mass bin of $[5\times10^{13},\ 1\times10^{14})$ $h^{-1}M_{\odot}$. Making this mass cut removes 35% of the redMaPPer clusters in total.

Using the covariance matrix from Wu et al. (2019) and the mass-concentration relation of Mandelbaum et al. (2008), we calculate the total mass bias through the propagation of bias from the lensing signal onto the mass model parameter through a Fisher matrix forecast. As shown in Figure 4.15, the mass is biased high at $1-5\%$, consistent with findings from **?** and Dietrich et al. (2014) and is highest at mid-richness ranges.

4.6.4 Comparison with DES Y1

Our weak lensing mass bias estimated from this paper is on the lower end but within 2σ of the bias estimated from the DES Y1 cluster cosmology paper (Abbott et al., 2020), which showed that the total bias for both triaxiality and projection effect is around $10\%-20\%$ depending on the richness and redshift bin.

The DES Y1 paper tested for systematics by controlling for variables that may introduce bias. The lensing profiles of two samples were compared—one selected by richness bins with its mass distribution left free to vary, and the second tracing the mass distribution of the richness-selected sample with its richness free to vary. The ratio of these profiles is an estimate of the total systematic bias due to redMaPPer selection in a given richness bin and radial range. The effects of triaxiality and projection effects can be teased out by re-sampling their

proxies $\cos(i)$ and $\sigma(z)$ in the richness-selected sample to match the mass-selected sample.

One notable finding in the DES Y1 cluster cosmology result is that known selection effects as orientation and projection resolve the weak mass discrepancy with other probes at $\lambda > 30$ but fail to expalin the discrepancy in the $\lambda \in [20, 30)$ range. This point was shown by comparing the weak lensing mass from the data with the inferred weak lensing mass using the cluster abundance information alone, combined with cosmological constraints derived from DES 3×2 point correlations (Abbott et al., 2018) (known as NC + 3×2). The comparison showed that the weak lensing mass after correcting for selection effects is consistent with NC + 3×2 at $\lambda > 30$ but the ratio is biased high at $\lambda \in [20, 30)$ when correcting for triaxiality and projection effects will only lower the inferred weak lensing mass.

We find in this paper that whe accounting for triaxiality biases the weak lensing mass will be lowered across all richnesses at a level consistent with findings in DES Y1. At $\lambda \in [20, 30)$ other unaccounted-for systematics must be at play that biases the weak lensing mass high compared to other probes.

4.7 Conclusion

The main findings of this work are as follows:

1. We find that the 3D axis ratios of redMaPPer-selected halos is consistent with the distribution of halos overall.

2. We find that the log-richness amplitude $\ln(A)$ of redMaPPer clusters for a given mass is boosted from the lowest to highest orientation bin with a significance of 14σ.

3. We find a null correlation between the bias in richness due to triaxiality and those for two other leading systematics in DES Y1 cluster cosmology—miscentering and projection—and offer explanations or follow-up studies for this result. The null correlation with projection effects was verified using both the Buzzard and C19 projection

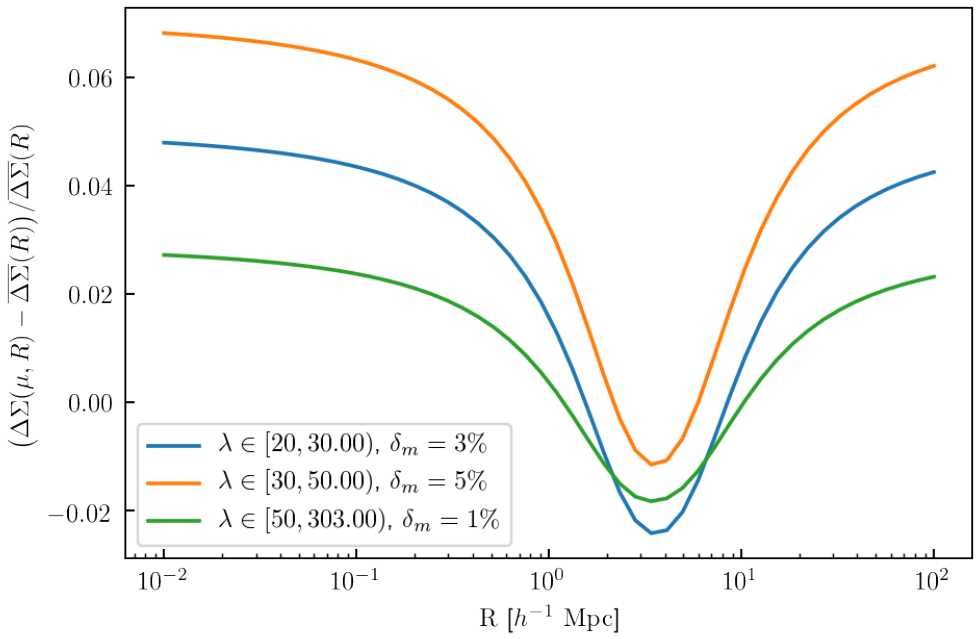


Figure 4.15: Fractional difference in lensing profiles $\delta\langle\Delta\Sigma\rangle$ for redMaPPer-selected clusters stacked in bins of richness. The total mass bias for each richness bin is measured by marginalizing $\delta\langle\Delta\Sigma\rangle$ as shown in plot through Equation 4.32 through propagating the errors of the lensing profile onto the mass model parameter using a Fisher forecast.

mock, catalogs with different galaxy-halo connection models.

4. We confirm the bottleneck shape in the transition between one- and two-halo regimes for halo lensing profiles first discovered by Osato et al. (2018) and fit it to redshift- and mass-dependent templates.

5. We quantify through items (ii) and (iv) the DES observable of richness-stacked redMaPPer cluster lensing profiles to predict a positive mass bias of $1-5\%$ due to triaxiality.

6. We find that the mean $P(\cos i)$ and the mass bias are both richness dependent and largest at mid-to-high richness, in accordance with the DES Y1 result that triaxiality does not fully resolve the tension in weak lensing mass at low richness.

Our findings are based on redMaPPer catalogs constructed using galaxies in the Buzzard simulations. The realistic red-sequence galaxy model in the Buzzard simulations allows us to run the redMaPPer algorithm in the same way as it was run on DES-Y1 data and hence enables us to quantify various selection effects introduced by the cluster finder. While this analysis provides evidence of redMaPPer selection effects and quantifies the relations between different systematics, we must acknowledge that there is one important caveat in this approach: the performance of the redMaPPer cluster finder depends on how galaxies are populated in the simulations, which might not precisely match the real universe. Since this analysis is only done on one specific simulation, the result in this paper can serve as a guidance for constructing a flexible enough model used in the analysis of real data.

These findings shed light on the impact of triaxiality on cluster selection, both their physical quantities and observed signals. Specifically, items (ii) and (iv) may be used as templates for current and near future weak lensing surveys as correction terms for this systematic. One important future work is to perform this analysis on different mock galaxy catalogs with different assumptions about the relations between galaxies and dark matter. Such an analysis will be essential to addressing the dependence of cluster finder performance

on galaxy population models.

CHAPTER 5

CORRELATED SCATTER FOR OPTICAL OBSERVABLES

5.1 Introduction

Cluster abundance and its evolution with redshift are linked to the constituents of the Universe through the growth of cosmic structure. (Allen et al., 2011, for a review). Cluster abundance measured in large-scale galaxy surveys offers power constraints on cosmological parameters (e.g., Vikhlinin et al., 2009; Mantz et al., 2015; de Haan et al., 2016; Mantz et al., 2016a; Dark Energy Survey Collaboration et al., 2016; Pierre et al., 2016; Costanzi et al., 2021). These constraints are based on accurate cluster mass measurements, which are not directly observable and must be inferred. Cluster mass calibration has been identified as one of the leading systematic uncertainties in cosmological constraints using galaxy cluster abundance (Mantz et al., 2010; Rozo et al., 2010; von der Linden et al., 2014; Applegate et al., 2014; Dodelson et al., 2016; Murata et al., 2019; Costanzi et al., 2021). Accurate mappings between a population of massive clusters and their observables are thus critical and essential in cluster cosmology. Considerable effort has been put into measuring the statistical relationships between masses and observable properties that reflect their baryon contents (see Giodini et al., 2013, for a review) and quantifying the sources of uncertainties.

The Dark Energy Survey (DES) cluster cosmology from the Year 1 dataset (Abbott et al., 2020) reported tension in Ω_m — the mean matter density of the universe — with the DES 3x2pt probe that utilizes three two-point functions from the DES galaxy survey (Abbott et al., 2018). The tension between these two probes that utilize the same underlying dataset may be attributed to systematics that bias the weak lensing mass of clusters low at the low mass end (To et al., 2021a; Costanzi et al., 2021). A possible origin for this discrepancy is that cluster masses are biased low due to systematics in cluster mass calibration. On the other hand, the tension can also originate from new physics that extends the Standard

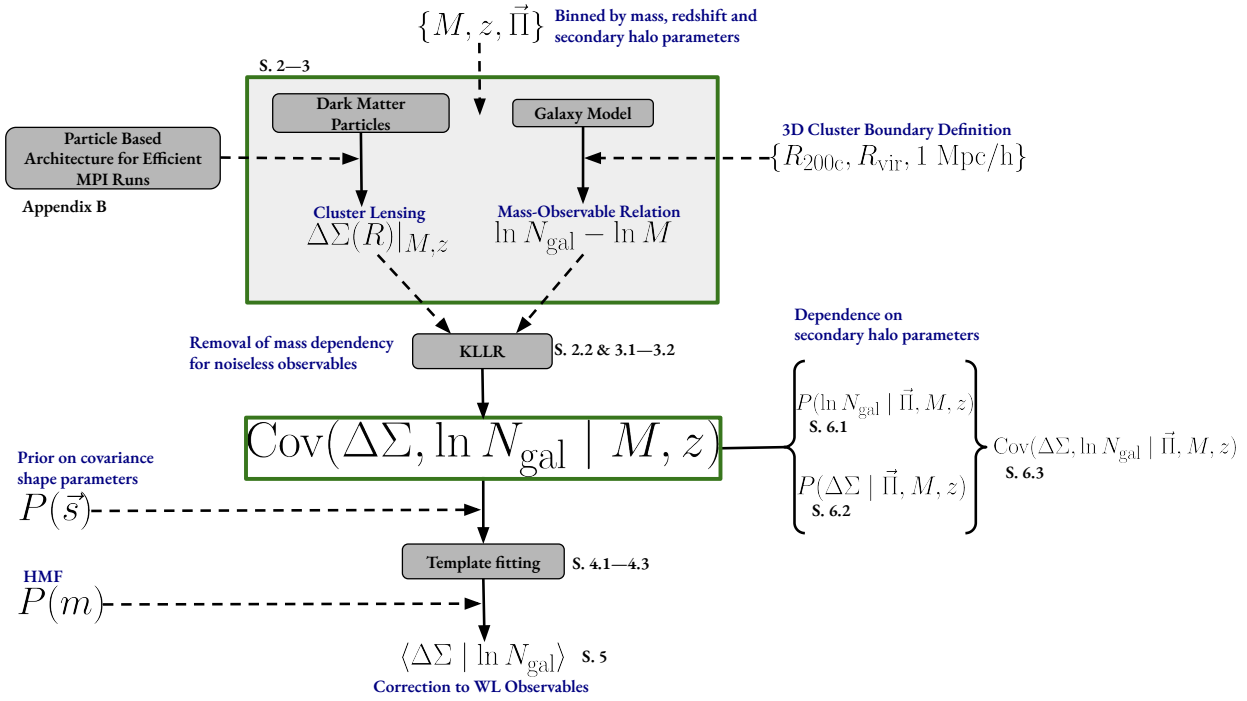


Figure 5.1: Graphic representation of the modeling of $\mathrm{Cov}(\Delta\Sigma, \ln N_{\mathrm{gal}} \mid M, z)$ — the covariance between the halo weak lensing signal $\Delta\Sigma(R)$ and log-richness $\ln N_{\mathrm{gal}}$ conditioned on mass and redshift — and its dependence with secondary halo parameters Π. The labels marked *S. XX* point to the location in the text. A full list of the notations used in this paper is introduced in Table 5.1 & 5.2.

Cosmological model. Thus, it is important to understand the systematics of cluster mass calibration. Cluster masses estimated from X-ray and SZ data are known to suffer from hydrostatic bias (Pratt et al., 2019). Conversely, cluster masses estimated from weak lensing have the potential to be more accurate compared to X-ray and SZ cluster masses. The systematics in the weak lensing mass calibration has just started to be explored recently (Applegate et al., 2014; Schrabback et al., 2018; McClintock et al., 2019; Kiiveri et al., 2021; Wu et al., 2022).

A relatively unexplored category of cluster systematics is the covariance between different cluster properties, including cluster observables and mass proxies. In cluster mass calibration, it is often assumed that this property covariance is negligible. However, as initially pointed out by Nord et al. (2008) and later shown in Evrard et al. (2014) and Farahi et al. (2018), non-zero property covariances between cluster observables can induce non-trivial, additive bias in cluster mass. As property covariance is additive, the systematic uncertainties that it induces will not be mitigated with the reduction of statistical errors as the sample size of the cluster increases. To achieve accurate cosmological constraints with the next generation of large-scale cluster surveys, it is imperative that systematic uncertainties in the property covariance be accurately and precisely quantified.

Although the property covariance linking mass to observable properties is becoming better understood and measured (Wu et al., 2015; Mantz et al., 2016a; Farahi et al., 2018, 2019a; Sereno et al., 2020), studies that specifically investigate weak lensing property covariance are scarce, which poses a challenge for upcoming lensing surveys of galaxy clusters such as the Rubin (Ivezić et al., 2019) observatories. To achieve the percentage-level lensing mass calibration goals for the upcoming observations, the property covariance of weak lensing must be quantified.

The physical origins of property covariance in lensing signals of galaxy clusters can be attributed to the halo formation history of the cluster and baryonic physics (Xhakaj et al.,

2022). Developing a first-principle physical model for the property covariance as a function of halo formation history and baryonic physics is a daunting task due to the highly non-linear and multi-scale physics involved in cluster formation. To make progress, in this paper, we adopt a simulation-based, data-driven approach whereby we develop semi-analytical parametric models of property covariance, which we then calibrate with cosmological simulations. We then apply our model to quantify the bias induced due to a non-zero property covariance in the expected weak lensing signal and the mass-observable scaling relation.

As will be presented in §5.3, a key element of this analysis is the estimation of true cluster *richness* by encircling clusters within a 3D radius within the physical vicinity of the halo center, as opposed to a 2D projected radius used by cluster finders as redMaPPer (Rykoff et al., 2014) by identifying galaxies within the red-sequence band in the color-magnitude space — the major difference being the removal of projection effects, or the mis-identification of non-cluster galaxies in the 2D projected radius from the photometric redshift uncertainty of the red-sequence when estimating the *true* richness from a gravitationally bound region around the halo. Furthermore, as this simulation-based study does not introduce other observational systematics as shape noise of galaxies, point spread function, miscentering, among others, this study can be used to explore the *intrinsic* covariance between observables prior to the addition of *extrinsic* systematics as projection effects. Our results will not only provide insight into the physical origin of the covariance, the difference between the *total* covariance as measured by observations and the *intrinsic* covariance will provide estimates on the amplitude of the *extrinsic* component.

The goals of this work are to (i) develop an analytical model that accounts for and quantifies the effect of non-zero covariance on cluster mass calibration, (ii) quantify this property covariance utilizing cosmological simulations, (iii) update uncertainties on inferred cluster mass estimates, and (iv) explain the physical origin of the covariance using secondary halo parameters. The rest of this paper is organized as follows. In §5.2, we present a population-

Table 5.1: Notations employed in our framework for the covariance in §5.2.3

Parameter	**Explanation**
$\Delta\Sigma$	Weak lensing signal
M	halo mass in $M_\odot h^{-1}$
$N_{\rm gal}$	optical richness enclosed inside 3D radius
z	redshift
r_p	projected and normalised radius

based analytical framework. In §5.3, we describe the simulations and data-vector employed in this work. In §5.4, we present our measurements and the covariance model. In §5.5, we present the impact of the covariance on weak lensing mass calibration. In §5.6, we quantify the physical origin of the covariance by parameterizing it using secondary halo parameters. In §5.7, we compare our work with those that employ realistic cluster finders. We conclude in §5.8.

5.2 Theoretical Framework

This section presents a theoretical framework that examines the impact of covariance on mass-observable scaling relations. In §5.2.1 we introduce the definitions of richness and weak lensing excess surface mass density and their scaling relations with cluster mass. We then describe the model of property covariance of richness and excess surface mass density in §5.2.2. In §5.2.3, we model the impact of covariance on stacked observable scaling relations. Finally, in §5.2.4, we develop a theoretical framework that explains the covariance based on a set of secondary halo parameters. A graphic representation of the outline of the paper is shown in Figure 5.1. The notations used in this section to describe the covariance are listed in Table 5.1 and notations for scaling relations are listed in Table 5.2.

Table 5.2: Scaling Relation Conventions.

Parameter	**Explanation**
π_a	normalization in scaling relation $\langle a \mid M \rangle$
α_a	slope in scaling relation $\langle a \mid M \rangle$
σ_a	scatter about $\langle a \mid M \rangle$
$r_{a,b}$	correlation between a and b at fixed M
$\pi_{a\|b}$	normalisation in scaling relation $\langle a \mid b \rangle$
$\alpha_{a\|b}$	slope in scaling relation $\langle a \mid b \rangle$
$\sigma_{a\|b}$	scatter about $\langle a \mid b \rangle$
a, b	$a, b \in \{\Delta\Sigma, \ln N_{\rm gal}\}$

5.2.1 Observable-mass relations

Excess Surface Mass Density $\Delta\Sigma$ from Weak Lensing

In weak lensing measurements of galaxy clusters, the key observable is the excess surface mass density, denoted $\Delta\Sigma$. The excess surface mass density is defined as

$$\Delta\Sigma(M, z, r_p) = \overline{\Sigma}(M, z, < r_p) - \Sigma(M, z, r_p), \tag{5.1}$$

where $\overline{\Sigma}(M, z, < r_p)$ denotes the average surface mass density within projected radius r_p, and $\Sigma(M, z, r_p)$ represents the average of the surface mass density at r_p. We model the average surface mass density Σ as

$$\Sigma(r_p) = \rho_m \int_{-\infty}^{+\infty} \left(1 + \xi_{hm}\Big(r = \sqrt{r_p^2 + \chi^2}\Big)\right) d\chi, \tag{5.2}$$

where ρ_m is the mean matter density at the redshift of the cluster, R is the projected radius in the plane of the sky, χ is the comoving distance along the line of sight centered around the cluster, and $\xi_{hm}(r)$ is the halo-matter correlation function which characterises the total mass density within a halo. Under the halo model, the halo-matter correlation function consists

of a "one-halo" term:

$$\xi_{\rm 1h}(r|M) = \frac{\rho_{\rm NFW}(r|M)}{\rho_{m0}} - 1, \tag{5.3}$$

and a "two-halo" term:

$$\xi_{\rm 2h}(r|M) = b(M)\xi_{\rm lin}(r), \tag{5.4}$$

where $\rho_{\rm NFW}$ is the Navarro-Frenk-White (NFW) density profile (Navarro et al., 1997), and $\xi_{\rm lin}$ is the linear matter correlation function, and b is the halo bias parameter.

In weak lensing, the excess surface density $\Delta\Sigma$ is tied to the tangential shear γ_t of the galaxies relative to the center of each foreground halo by the relation

$$\Sigma_{\rm crit}\gamma_t = \overline{\Sigma}(< R) - \Sigma(R) \equiv \Delta\Sigma(R), \tag{5.5}$$

where the critical surface density $\Sigma_{\rm crit}$ defined as

$$\Delta\Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{D_s}{D_l D_{ls}}, \tag{5.6}$$

and where D_s, D_l, and D_{ls} refer to the angular diameter distances to the source, the lens, and between the lens and source, respectively.

In this work, for each halo of mass M at redshift z, we compute the corresponding $\Delta\Sigma$ profile. We compare these measurements with theoretical predictions — in the one-halo regime we model the cluster overdensity as NFW profiles with their concentration determined by concentration-mass models of Prada et al. (2012), Ludlow et al. (2016) and Diemer & Joyce (2019), whereas in the two-halo term we adopt the linear matter correlation ξ_{mm} multiplied by halo biases using the Tinker et al. (2010), Pillepich et al. (2010) and Bhattacharya et al. (2011) models to derive the halo-matter correlation ξ_{hm}. At the transition radius between the one- and two-halo regimes, we follow SDSS (Zu et al., 2014) in setting the halo–matter

correlation to the maximum value of the two terms, i.e.,

$$\xi_{hm}(r|M) = \max\{\xi_{1\mathrm{h}}(r|M),\ \xi_{2\mathrm{h}}(r|M)\}. \tag{5.7}$$

In Fig. 5.2, the theoretical models described above are compared with our measurements of $\Delta\Sigma$ in cosmological simulations to validate our data product.

We model the mean $\langle \Delta\Sigma \mid M, z, r_p \rangle$ at fixed mass M, redshift z, and projected radius r_p as a log-linear relation given by

$$\langle \Delta\Sigma \mid M, z, r_p \rangle_1 = \pi_{\Delta\Sigma}(M, r_p, z) + \alpha_{\Delta\Sigma}(M, r_p, z) \ln M, \tag{5.8}$$

where $\alpha_{\Delta\Sigma}$ is the power-law slope of the relation and $\pi_{\Delta\Sigma}$ is a normalization that is a function of redshift and mass.

Optical Richness N_{gal}

Optical richness N_{gal} is an observable measure of the abundance of galaxies within a galaxy cluster. It is often defined as the number of detected member galaxies brighter than a certain luminosity threshold within a given aperture or radius around the cluster centre. Richness is often used as a proxy for cluster mass, as more massive clusters are expected to have more member galaxies (e.g., Rozo et al., 2014; Rykoff et al., 2014). The richness-mass scaling relation relates the richness of a galaxy cluster to its mass. In this work, we consider the mean N_{gal}-M_Δ scaling relation expressed as

$$\langle \ln N_{\mathrm{gal}} \mid M_\Delta, z \rangle_1 = \pi_{N_{\mathrm{gal}}}(M, z) + \alpha_{N_{\mathrm{gal}}}(M, z) \ln M_\Delta. \tag{5.9}$$

where M_Δ is the mass of the halo within a radius where the mean density is Δ times the critical density of the universe, $\alpha_{N_{\mathrm{gal}}}(M, z)$ is the power-law slope of the relation, $\pi_{N_{\mathrm{gal}}}(M, z)$

is a normalisation that is a function of redshift and mass.

Halo Mass and Radius Definitions

A common approach to defining a radial boundary of a galaxy cluster is such that the average matter density inside a given radius is the product of a reference overdensity $\Delta_{\rm ref}$ times the critical (ρ_c) or mean density (ρ_m) of the universe at that redshift. The critical density is defined as

$$\rho_c = \frac{3H_0^2}{8\pi G} E(z), \tag{5.10}$$

where $E(z)^2 = \Omega_{m,0}(1+z)^3 + \Omega_{\Lambda,0}$, $\Omega_{m,0}$ is the present day matter fraction of the universe, $\Omega_{\Lambda,0}$ is the dark energy fraction at the present age such that $\Omega_{m,0}+\Omega_{\Lambda,0} = 1$ for a flat universe ignoring the minimal contribution from the radiation fraction. The mean (background) density is defined as

$$\rho_b = \frac{3H_0^2}{8\pi G} (\Omega_{m,0}(1+z)^3). \tag{5.11}$$

The overdensity $\Delta_c = 200$ is commonly chosen as the reference overdensity in optical weak lensing studies and is closely related to the virial radius. Another radius definition is the virial radius $R_{\rm vir}$, with overdensity values calibrated from cosmological simulations (Bryan & Norman, 1998). In this work, we use $R_{\rm 200c}$ and $R_{\rm vir}$ to scale various observations, including the $\Delta\Sigma$ measurements and richness. Since the covariance is close to zero at the outskirts $R \gtrsim R_{\rm 200c}$ as shown in §5.4, we adopt $r_p = R/R_{\rm 200c}$ and $r_p = R/R_{\rm vir}$ as our normalised radii, as the cluster properties are more self-similar with respect to $\rho_c(z)$ compared to $\rho_b(z)$ (Diemer & Kravtsov, 2014; Lau et al., 2015). To test for the robustness of our covariance against different radii definitions, we also introduce a physical radius of a toy model of a constant $R = 1$ Mpc/h; here $h = 0.6777$ is the reduced Hubble constant used in this study.

5.2.2 *Covariance between $\Delta\Sigma$ and $N_{\rm gal}$*

In optical surveys, we cannot expect the *covariance* between richness $N_{\rm gal}$ and the excess surface mass density $\Delta\Sigma$ to be zero. Ignoring this covariance will lead to bias in cluster mass inferred from the excess surface mass density of the cluster selected based on richness. This work aims to quantify and analyse this covariance and its impact on the mass calibration relation. To achieve this objective, we must first specify the joint probability distribution of excess surface mass density and richness, $p(\Delta\Sigma, \ln N_{\rm gal} \mid M, z, r_p)$. In this work, we assume that the joint distribution follows a multivariate normal distribution (Stanek et al., 2010; Evrard et al., 2014; Mulroy et al., 2019; Miyatake et al., 2022), which is fully specified with two components, the mean vector and the property covariance. We have checked the goodness of this assumption in Appendix 5.9.4.

From the mean observable-mass scaling relations in Equation (5.8) and Equation (5.9), the scaling relation between these two observables can be modeled as a local linear relation given by

$$\langle \Delta\Sigma \mid N_{\rm gal}, z, r_p \rangle = \pi_{\Delta\Sigma|N_{\rm gal}}(N_{\rm gal}, z, r_p) + \alpha_{\Delta\Sigma|N_{\rm gal}}(N_{\rm gal}, z, r_p) \ln N_{\rm gal}, \tag{5.12}$$

where π and α are the normalization and slope of the model.

The property covariance matrix is a combination of scatter and correlation between the scatter of $\Delta\Sigma$ and $\ln N_{\rm gal}$ at a fixed halo mass, redshift, and projected radius. We use $\sigma_{N_{\rm gal}}(M, z)$ and $\sigma_{\Delta\Sigma}(M, z, r_p)$ to denote the scatter of the observable-mass relation for $\ln N_{\rm gal}$ and $\Delta\Sigma$, respectively, and use $r_{N_{\rm gal},\Delta\Sigma}(M, z, r_p)$ to denote the correlation between these scatters. The covariance matrix is then given by

$$\mathrm{Cov}_{i,j}(M, z, r_p) = r_{i,j}(M, z, r_p)\ \sigma_i(M, z, r_p)\ \sigma_j(M, z, r_p), \tag{5.13}$$

where i and $j \in \{\Delta\Sigma, \ln N_{\rm gal}\}$. Specifically, the covariance between $\Delta\Sigma$ and $N_{\rm gal}$ can be

expressed in terms of the residuals about the mean quantities

$$\mathrm{Cov}_{\Delta\Sigma,N_{\rm gal}}(M,z,r_p) = \mathrm{Cov}(\mathrm{res}_{\Delta\Sigma}(M,z,r_p), \mathrm{res}_{N_{\rm gal}}(M,z)), \tag{5.14}$$

where the residuals of the $\Delta\Sigma$ and $N_{\rm gal}$ are, respectively, are given by

$$\mathrm{res}_{\Delta\Sigma}(M,z,r_p) = \Delta\Sigma - \langle \Delta\Sigma \mid M,z,r_p\rangle, \tag{5.15}$$

$$\mathrm{res}_{N_{\rm gal}}(M,z) = \ln N_{\rm gal} - \langle \ln N_{\rm gal} \mid M,z\rangle. \tag{5.16}$$

To model the mass dependencies of the mean profiles of $\Delta\Sigma$ and $\ln N_{\rm gal}$, we employ the Kernel Localised Linear Regression (KLLR, Farahi et al., 2022a) method. This regression method fits a locally linear model while capturing globally non-linear trends in data and has shown to be effective in modeling halo mass dependencies in scaling relations (Farahi et al., 2018; Wu et al., 2022; Anbajagane et al., 2022). By developing a local-linear model of $\Delta\Sigma - \ln N_{\rm gal}$ with respect to the halo mass and computing the residuals about the mean relation, we remove the bias in the scatter due to the mass dependence and reduce the overall size of the scatter. As shown in Fig. 5.4 the $1-\sigma$ of the covariance is determined by bootstrap resampling.

5.2.3 Corrections to the $\Delta\Sigma - N_{\rm gal}$ relation due to Covariance

The shape of the halo mass function plays an important role in evaluating the conditional mean value of $\langle \Delta\Sigma \mid N_{\rm gal}, z, r_p\rangle$ where the scatter between two observables with a fixed halo mass is correlated. Ignoring the contribution from the correlated scatter, to the zeroth order, the expected $\Delta\Sigma$ evaluated at fixed richness is given by Equation (5.12). This is the model that has been used in mass calibration with stacked weak lensing profiles (Johnston et al., 2007; Kettula et al., 2015; McClintock et al., 2019; Chiu et al., 2020; Lesci et al., 2022).

The first- and second-order approximations of the scaling relation are given by

$$\langle \Delta\Sigma \mid N_{\rm gal}, z\rangle_1 = \langle \Delta\Sigma \mid N_{\rm gal}, z\rangle_{\rm fid} + \frac{\gamma_1}{\alpha_{N_{\rm gal}}} \times {\rm Cov}(\Delta\Sigma, \ln N_{\rm gal}), \tag{5.17}$$

and

$$\begin{aligned}\langle \Delta\Sigma | N_{\rm gal}, z\rangle_2 = \langle \Delta\Sigma \mid N_{\rm gal}, z\rangle_{\rm fid} + \\ {\rm Cov}(\Delta\Sigma, \ln N_{\rm gal}) \times \Big[\frac{x_s}{\alpha^2_{N_{\rm gal}}}(\alpha_{N_{\rm gal}}\gamma_1 + \gamma_2(\ln N_{\rm gal} - \pi_{N_{\rm gal}}))\Big],\end{aligned} \tag{5.18}$$

where $\langle \Delta\Sigma \mid N_{\rm gal}, z\rangle_{\rm fid}$ is the fiducial relation taking into account the curvature of the HMF but independent of the covariance; $x_s = (1 + \gamma_2\sigma^2_{M|N_{\rm gal},1})^{-1}$ is the compression factor due to curvature of the HMF, the subscript 1 denoting that the scatter is taken from the HMF expanded to first order; here we omit the (M, z) dependence of the covariance as a shorthand notation. These expansions around the pivot mass are for halos centered around a narrow enough mass bin. We show explicitly in Fig. 5.9 that the first-order expansion converges using our binning method. The derivations for the first and second-order expansion terms can be found in Evrard et al. (2014) and Farahi et al. (2018) and the derivation for this particular expression of the second-order term is shown in Appendix 5.9.3.

Here γ_1 and γ_2 are the parameters for the first and second-order approximations to the mass dependence of the halo mass function (e.g., Evrard et al., 2014):

$$\frac{{\rm d}n_{\rm hmf}(M, z)}{{\rm d}\ln M} \approx A(z) \exp\left[-\gamma_1(M, z)\ln M - \frac{1}{2}\gamma_2(M, z)(\ln M)^2\right]. \tag{5.19}$$

where $A(z)$ is the normalisation of the mass function due to the redshift alone. In deriving the above approximations, we have made use of the fact that ${\rm Cov}(\Delta\Sigma, \ln N_{\rm gal}|M, z) \equiv r_{\Delta\Sigma,N_{\rm gal}}\sigma_{\Delta\Sigma}\sigma_{N_{\rm gal}}$. The terms $\sigma_{\Delta\Sigma}$, $\sigma_{N_{\rm gal}}$, $r_{N_{\rm gal},\Delta\Sigma}$, γ_1 and γ_2 are evaluated at the mass implied by $\langle \ln M \mid N_{\rm gal}\rangle$.

Table 5.3: Notations employed in exploring the secondary halo parameter dependence

Parameter	**Explanation**
Π	Set of secondary halo parameters
$\Gamma_{\rm inst}$	instantaneous mass accretion rate (MAR)
$\Gamma_{\rm 100Myr}$	mean MAR over the past 100 Myr
$\Gamma_{\rm dyn}$	mean MAR over virial dynamical time
$\Gamma_{\rm 2dyn}$	mean MAR over two virial dynamical time
$\Gamma_{\rm peak}$	Growth rate of peak mass from current z to z+0.5
$a_{1/2}$	Half mass scale factor
$c_{\rm vir}$	$R_{\rm vir}$ concentration
T/|U|	Absolute value of the kinetic to potential energy ratio
$X_{\rm off}$	Offset of density peak from mean particle position (kpc h^{-1})

These property covariance correction terms are absent in the current literature. A key feature of this approximation method is that the second-order solution has better than percent-level accuracy when the halo mass function is known (Farahi et al., 2016). In Figure 5.9, we demonstrate that the statistical uncertainties for the first-order correction in Equation (6.18) is larger than the uncertainty in the halo mass function and the uncertainty due to the second-order halo mass approximation.

5.2.4 Secondary halo parameter dependence of $\mathrm{Cov}(\Delta\Sigma, \ln N_{\rm gal}|M, z)$

We elucidate the physical origin of the covariance between $\Delta\Sigma$ and $\ln N_{\rm gal}$ by developing a phenomenological model based on the secondary halo parameters listed in Table 5.3. These parameters are computed from the ROCKSTAR halo finder (Behroozi et al., 2013). They capture the halo's mass accretion history , which we hypothesise is the driving force behind the observed covariance. To incorporate these parameters into our model, we extend Equations (5.8) and (5.9) by introducing multi-linear terms that include the secondary halo

parameters denoted by the vector Π:

$$\begin{aligned}(\ln N_{\rm gal} \mid \Pi, M, z) =& \langle \ln N_{\rm gal} \mid M, z\rangle_1 + \vec{\beta}^{\intercal}_{N_{\rm gal}}(M,z)\cdot\Pi + \epsilon_{N_{\rm gal}} \\ =& \pi_{N_{\rm gal}}(M,z) + \alpha_{N_{\rm gal}}(M,z)\ln M + \\ & \vec{\beta}^{\intercal}_{N_{\rm gal}}(M,z)\cdot\Pi + \epsilon_{N_{\rm gal}}. \end{aligned} \tag{5.20}$$

$$\begin{aligned}(\Delta\Sigma \mid \Pi, M, z) =& \langle \Delta\Sigma \mid M, z\rangle_1 + \vec{\beta}^{\intercal}_{\Delta\Sigma}(M,z)\cdot\Pi + \epsilon_{\Delta\Sigma} \\ =& \pi_{\Delta\Sigma}(M,z) + \alpha_{\Delta\Sigma}(M,z)\ln M + \\ & \vec{\beta}^{\intercal}_{\Delta\Sigma}(M,z)\cdot\Pi + \epsilon_{\Delta\Sigma}, \end{aligned} \tag{5.21}$$

where Π is a vector of secondary halo parameters of potential interest listed in Table 5.3, and $\epsilon_{\Delta\Sigma}$ and $\epsilon_{N_{\rm gal}}$ are normally distributed intrinsic scatter terms with zero means and uncorrelated variances. In Appendix 5.9.5 we show that the residual conditioned on secondary halo parameters can largely be assumed to be Gaussian. Additionally, we assume $\langle \epsilon_{N_{\rm gal}}\epsilon_{\Delta\Sigma}\rangle = 0$, which implies that the scatter about the mean relations is uncorrelated after factoring in the secondary properties.

Due to the bilinearity and distributive properties of covariance, combining Equation (5.20) and Equation (5.21) yields:

$$\begin{aligned}\mathrm{Cov}(\Delta\Sigma, \ln N_{\rm gal} \mid M, z) =& \mathrm{Cov}(\langle\Delta\Sigma\rangle_1, \langle N_{\rm gal}\rangle_1) + \mathrm{Cov}(\langle\Delta\Sigma\rangle_1, \vec{\beta}^{\intercal}_{N_{\rm gal}}\cdot\Pi) \\ &+ \mathrm{Cov}(\langle\Delta\Sigma\rangle_1, \epsilon_{N_{\rm gal}}) + \mathrm{Cov}(\vec{\beta}^{\intercal}_{\Delta\Sigma}\cdot\Pi, \langle\ln N_{\rm gal}\rangle_1) \\ &+ \mathrm{Cov}(\vec{\beta}^{\intercal}_{\Delta\Sigma}\cdot\Pi, \vec{\beta}^{\intercal}_{N_{\rm gal}}\cdot\Pi) + \mathrm{Cov}(\vec{\beta}^{\intercal}_{\Delta\Sigma}\cdot\Pi, \epsilon_{N_{\rm gal}}) \\ &+ \mathrm{Cov}(\epsilon_{\Delta\Sigma}, \langle\ln N_{\rm gal}\rangle_1) + \mathrm{Cov}(\epsilon_{\Delta\Sigma}, \vec{\beta}^{\intercal}_{N_{\rm gal}}\cdot\Pi) \\ &+ \mathrm{Cov}(\epsilon_{N_{\rm gal}}, \epsilon_{\Delta\Sigma}), \end{aligned} \tag{5.22}$$

where we omit the explicit (M, z) dependence in $\langle\Delta\Sigma \mid M, z\rangle_1$, $\langle \ln N_{\rm gal} \mid M, z\rangle_1$, $\vec{\beta}^{\intercal}_{\Delta\Sigma}(M,z)$,

$\vec{\beta}^{\intercal}_{N_{\rm gal}}(M,z)$, $\epsilon_{N_{\rm gal}}(M,z)$ and $\epsilon_{\Delta\Sigma}(M,z)$ to simplify the notation. The KLLR method is utilised to estimate these mass-dependent parameters. All terms involving $\langle\Delta\Sigma\rangle$ and $\langle\ln N_{\rm gal}\rangle$ vanish, as these terms are independent of Π by definition. Terms involving $\epsilon_{N_{\rm gal}}$ and $\epsilon_{\Delta\Sigma}$ also go to zero, as they are uncorrelated Gaussian scatters. Only the term $\mathrm{Cov}(\vec{\beta}^{\intercal}_{\Delta\Sigma}\cdot\Pi, \vec{\beta}^{\intercal}_{N_{\rm gal}}\cdot\Pi)$ remains, and hence our final expression for the covariance is

$$\begin{aligned}\mathrm{Cov}(\Delta\Sigma, \ln N_{\rm gal} \mid M,z) &= \mathrm{Cov}(\vec{\beta}^{\intercal}_{\Delta\Sigma}\cdot\Pi, \vec{\beta}^{\intercal}_{N_{\rm gal}}\cdot\Pi) \\ &= \vec{\beta}^{\intercal}_{\Delta\Sigma}\mathrm{Cov}(\Pi,\Pi)\vec{\beta_{N_{\rm gal}}}.\end{aligned} \tag{5.23}$$

To estimate the error in the covariance due to each of the secondary halo parameters, we compute $\mathrm{Cov}(\Delta\Sigma, \Pi_i|M,z)$ for each secondary halo parameter i in each (r_p, M, z) bin and take their standard deviations as the error measurement. Modeling the richness-mass relation as in Equation (5.20) and using the same derivation as in Equation (5.22), we arrive at the expression

$$\mathrm{Cov}(\Delta\Sigma, \ln N_{\rm gal}|M,z) = \sum_i \beta_{N_{\rm gal},i}(M,z)\ \mathrm{Cov}(\Delta\Sigma, \Pi_i|M,z), \tag{5.24}$$

in which the error from each contributing term in Π is the standard deviation for $\mathrm{Cov}(\Delta\Sigma, \Pi_i|M,z)$ multiplied by the partial richness slope $\beta_{\rm N_{gal},i}$. The total variance of $\mathrm{Cov}(\Delta\Sigma, \ln N_{\rm gal}|M,z)$ are the errors of each term added in quadrature.

We test the validity of this model by checking how well the secondary halo parameters can explain covariance between lensing and richness in §5.6. After subtracting the covariance from each of the Π_i terms, the full covariance should be consistent with null, given the uncertainty. Our results confirm that the dependency of secondary halo parameters can indeed explain the covariance.

5.3 Dataset and Measurements

In this section, we describe the measurements on the individual ingredients that make up the covariance — $\Delta\Sigma$ the lensing signal in §5.3.1 and $\ln N_{\rm gal}$ the log-richness measurement in §5.3.2.

5.3.1 Measurements of $\Delta\Sigma$

We employ the MultiDark Planck 2 (MDPL2) cosmological simulation (Klypin et al., 2016) to measure halo properties. The MDPL2 is a gravity-only N-body simulation, consisting of 3840^3 particles in a periodic box with a side length of $L_{\rm box} = 1\ h^{-1}$Gpc, yielding a particle mass resolution of approximately $m_{\rm p} \approx 1.51 \times 10^9 h^{-1} M_\odot$. The simulation was conducted with a flat ΛCDM cosmology similar to Planck Collaboration et al. (2014), with the following parameters: $h = 0.6777$, $\Omega_{\rm m} = 0.307115$, $\Omega_\Lambda = 0.692885$, $\sigma_8 = 0.829$, and $n_{\rm s} = 0.96$. We use the surface over-density of down-sampled dark matter particles to measure the weak lensing signal. We selected cluster-sized halos using the ROCKSTAR (Behroozi et al., 2013) halo catalogue, which includes the primary halo property of mass and redshift and a set of secondary halo properties listed in Table 5.3 that we utilise in this analysis. To capture the contribution of both the one- and two-halo terms to ξ_{hm}, we use a projection depth of $D_p = 200$ Mpc h^{-1} to calculate $\Delta\Sigma$ (e.g., Costanzi et al., 2019; Sunayama et al., 2020). The MDPL2 data products are publicly available through the MultiDark Database (Riebe et al., 2013) and can be downloaded from the `CosmoSim` website[1].

The excess overdensity, $\Delta\Sigma$, is calculated by integrating the masses of the dark matter particles in annuli of increasing radius centred around the halo centre. However, since clusters do not have a well-defined boundary, we compare the results of two radial binning schemes. The first scheme uses 20 equally log-spaced ratios between 0.1 and 10 times $R_{\rm vir}$, while the second scheme spans 0.1 to 10 times $R_{\rm 200c}$. We consider the measurements binned at $R_{\rm 200c}$

1. `https://www.cosmosim.org`

as our final results to be consistent with the weak lensing literature. Figure 5.2 shows that our measurements are consistent with most models of the concentration-mass and halo-bias models at a 1σ level.

At a projection depth of $D_p = 200$ Mpc h^{-1}, the projection effects can be modeled as a multiplicative bias (Sunayama, 2023). In Sunayama (2023) the projection effects on $\Delta\Sigma$ are modelled as $\Delta\Sigma_{\rm obs} = (1+\alpha)\Delta\Sigma_{\rm true}$, where $\alpha = 18.4\pm8.6\%$. Although the multiplicative bias of projection effects may increase the amplitude of $\mathrm{Cov}(\Delta\Sigma, \ln N_{\rm gal})$ by a factor of $(1+\alpha)$, we argue that it does not introduce an additive bias into our model for $\mathrm{Cov}(\Delta\Sigma, \ln N_{\rm gal})$. This is because, under the richness model and $\Delta\Sigma$ in Equations (5.20) and (5.21), only terms in richness that are correlated with projection effects will contribute to the covariance. As demonstrated in §5.3.2, we enclose the halo within a 3D physical radius, so the number count $N_{\rm gal}$ of galaxies should not include projection effects. Therefore, projection effects should not introduce an additive bias to our covariance.

To remove the 2D integrated background density, we first computed the background density of the universe (ρ_b) at the cluster redshift using the cosmological parameters of the MDPL2 simulation. The integrated 2-D background density is given by $\Sigma_b = 2D_p\rho_b$, where factor 2 comes from the integration of the foreground and background densities.

5.3.2 Measurements of $N_{\rm gal}$

Dataset for $N_{\rm gal}$ — SAGE galaxy catalog

The Semi-Analytic GALAXY Evolution (SAGE) is a catalogue of galaxies within MDPL2, generated through a post-processing step that places galaxies onto N-body simulations. This approach, known as a semi-analytic model (SAM), is computationally efficient compared to hydrodynamical simulations with fully self-consistent baryonic physics. SAMs reduce the computational time required by two to three orders of magnitude, allowing us to populate the entire 1 $(\mathrm{Gpc/h})^3$ simulation volume with galaxies. SAGE's statistical power enables us

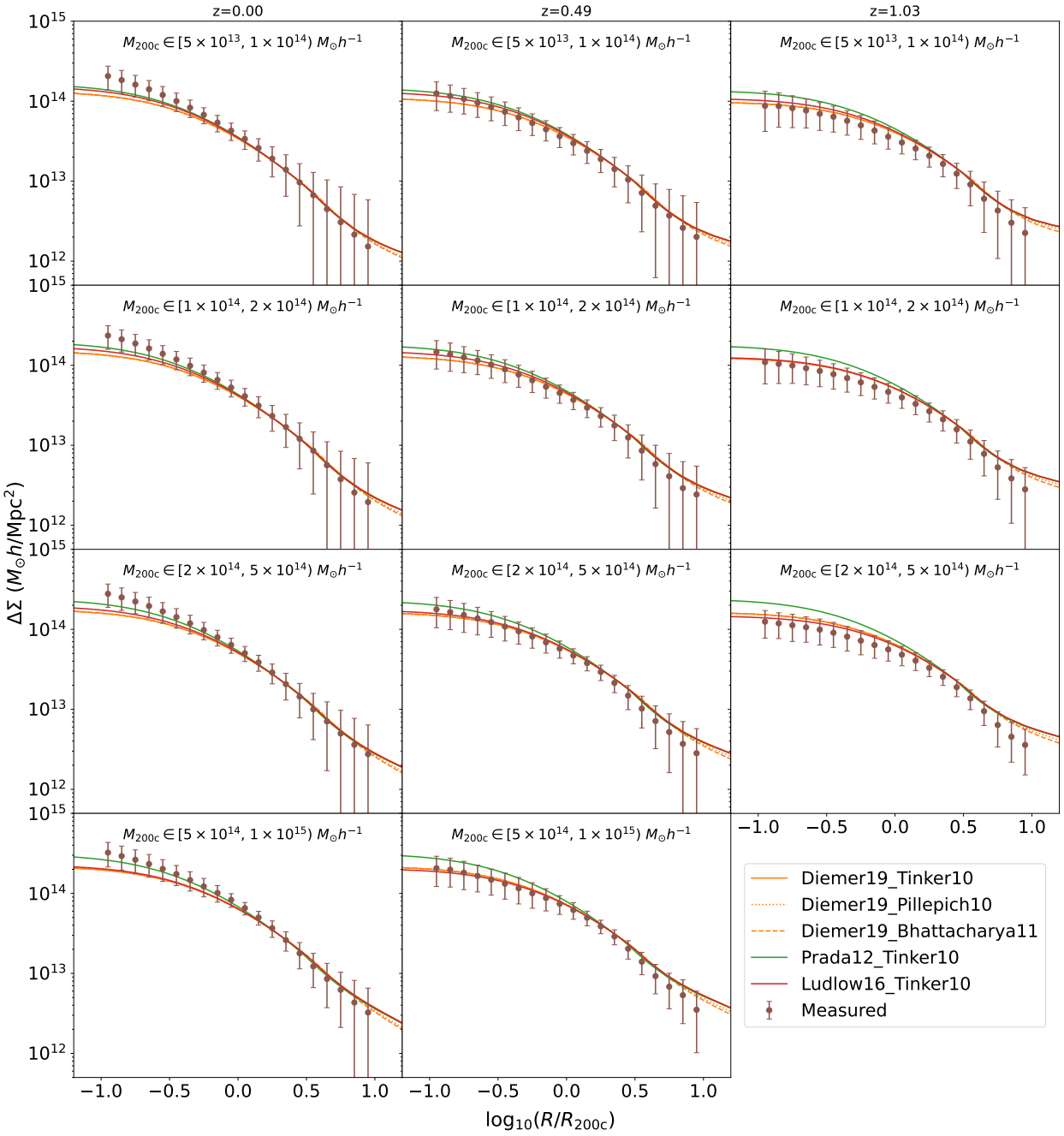


Figure 5.2: The measured $\Delta\Sigma$ profiles using downsampled particles for every 10 particles and theoretical $\Delta\Sigma$ as computed from the NFW profile using different concentration-mass relations (LHS in legend) in the one-halo regime and different halo-bias models (RHS in legend) in the two halo regimes, with errors taken to be $1-\sigma$ standard deviations; the measurements are consistent with theoretical predictions and the size of the errors is too large to distinguish between models. The same conclusion (not shown) holds for $\Delta\Sigma$ binned by R_{vir}.

to conduct stacked weak lensing analyses.

The baryonic prescription of SAGE is based on the work of Croton et al. (2016), which includes updated physics in baryonic transfer processes such as gas infall, cooling, heating, and reionization. It also includes an intra-cluster star component for central galaxies and addresses the orphan galaxy problem by adding the stellar mass of disrupted satellite galaxies as intra-"cluster" mass. SAGE's primary data constraint is the stellar mass function at $z = 0$. Secondary constraints include the star formation rate density history (Somerville et al., 2001), the Baryonic Tully-Fisher relation (Stark et al., 2009), the mass metallicity relation of galaxies (Tremonti et al., 2004), and the black hole–bulge mass relation (Stark et al., 2009).

Model for $N_{\rm gal}$

To determine the number of galaxies inside a cluster-sized halo, we utilise the SAGE semi-analytic model and compute the total number of galaxies within a 3D radius around the halo centre. We compare the true richness ($N_{\rm gal}$) to $M_{\rm 200c}$ scaling relations between different models and the observed richness-mass relations from Costanzi et al. (2021) using data from the DES Year-1 catalogue and mass-observable-relation from the South Pole Telescope (SPT) cluster catalogue (see Figure 5.3). The observed richness-mass relation is fitted as a log-linear model with $2-\sigma$ error bars that trace the posterior of the best-fit richness-mass model parameters. The $M_{\rm 500c}$ mass definition in the catalogue is converted to $M_{\rm 200c}$ using an NFW profile for the surface density of the cluster and adopting the Diemer & Joyce (2019) concentration-mass relation anchored at $z = 0.35$, which is roughly the median redshift of the cluster sample.

We use the KLLR method to determine the local linear fit for our $N_{\rm gal}$-mass model, which relaxes the assumption of global log-linearity (Anbajagane et al., 2020). Realistic cluster finders, such as redMaPPer (Rykoff et al., 2014), impose a colour-magnitude cut or

a stellar mass cut, which are highly dependent on the red-sequence model or the spectral energy density model. We found that imposing a stellar mass cut of $10.5\log(M/M_\odot)$ would correspond roughly to the bottom 5% percentile of SDSS detected galaxies (Maraston et al., 2013). However, this drastically decreases the number of galaxies in a halo, with most having $N_{\rm gal}$ in the single digits. As we are interested in the *intrinsic* covariance from the physical properties of the halo, we do not impose additional magnitude or stellar mass cuts. We confirm that, as described in Croton et al. (2016), the galaxy stellar mass distribution at $z = 0$ is consistent with the best-fit double Schechter function calibrated with low-redshift galaxies from the Galaxy and Mass Assembly (GAMA) (Baldry et al., 2012) down to stellar masses of $\mathcal{M} > 10^{8.5} M_\odot$.

Figure 5.3 illustrates that our $N_{\rm gal}$-mass models, which count the number of galaxies within a physical 3D radius and impose no colour-magnitude cut as redMaPPer does, resemble the general behavior of the observed richness-mass relations in terms of both slope and intercept. However, we acknowledge that redMaPPer may suffer from projection effects that artificially inflate the number count of red-sequence galaxies within its aperture because of line-of-sight structures. Additionally, the $N_{\rm gal}$ count within the $R_{\rm vir}$ radius exceeds that of $R_{\rm 200c}$ as $R_{\rm vir}$ is greater than $R_{\rm 200c}$. In the toy model scenario, where we use a constant 1 Mpc h^{-1} radius, the slope of the mass-richness relation starts to decrease as the mass increases due to the increasing physical size of the clusters, as expected. The diversity of cluster radii and the resulting variation in the local slope and intercept of the $N_{\rm gal}$-mass relations demonstrate the robustness of our covariance model. In Section 5.4, we show that different radii/mass definitions have little impact on the parameters of our covariance model, thus establishing its independence from different reference radii, the definitions of cluster edges, and the resulting richness-mass relations.

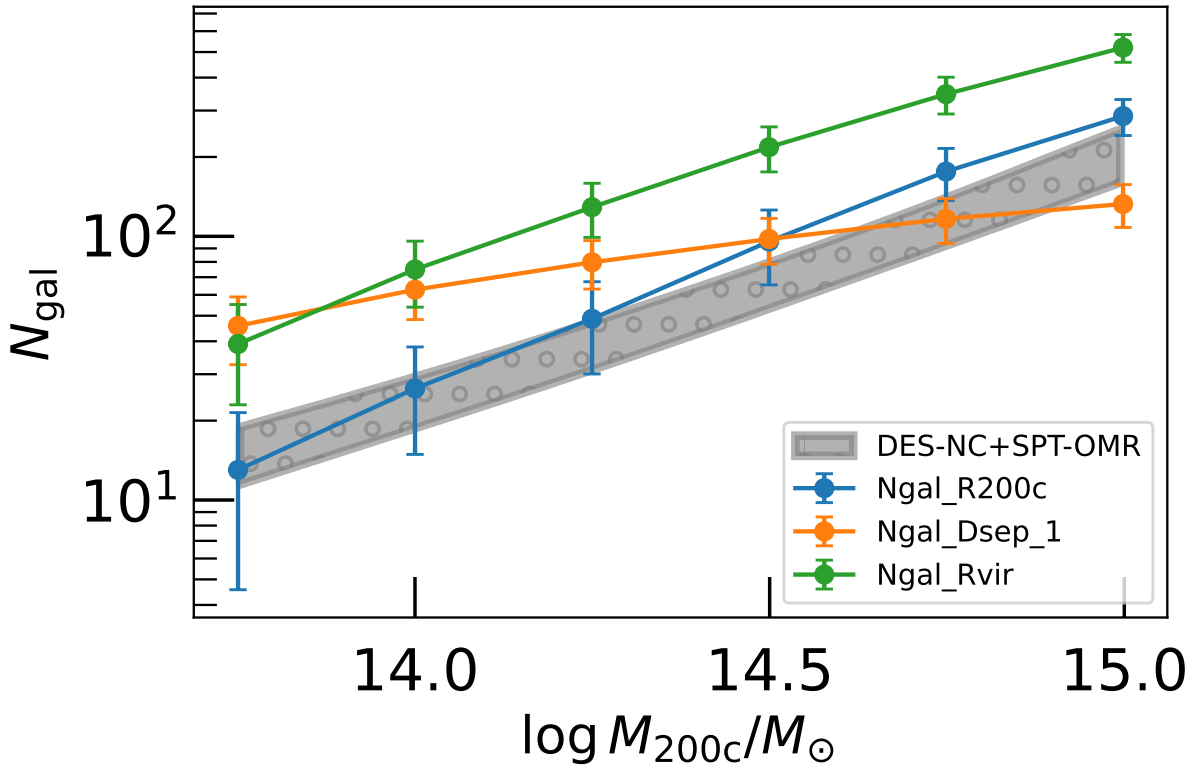


Figure 5.3: Using different prescriptions of the richness count, we compare with the SPT-DES data (Costanzi et al., 2021). The richness estimator, with no stellar or color-magnitude cut, shows a similar trend with the data. In §5.4, we show that the results are robust to changes in the definition of the number count estimator.

5.4 Results: Covariance Shape and Evolution

In this section, we report the measurements for our covariance. In Figure 5.4, we find an anti-correlation between $N_{\rm gal}$ and $\Delta\Sigma$ at small scales across most redshift and mass bins spanned by our dataset, which we fit with the best-fit "Sigmoid" functional form of the expression

$$\mathrm{Cov}(\tilde{\mathrm{x}}) = \mathrm{s}\Big(\mathrm{erf}(\frac{\sqrt{\pi}}{2}\tilde{\mathrm{x}}) + \mathrm{g}\Big), \tag{5.25}$$

with $x \equiv \log R/R_{\rm 200c}$ the log-radius and $\tilde{x} \equiv (x-\gamma)/\tau$ the scaled and offset log-radius. In Appendix 5.9.1, we offer statistical verification of the best-fit functional form.

We first describe the evolution of the covariance in §5.4.1 by binning across the (M, z) bins. Next, in §5.4.2, we present an alternative binning scheme based on halo peak height that can provide insight into the dependence of the time formation history of the covariance scale.

5.4.1 Binned in (M, z)

Our best-fit parameters in Table 5.6 indicate that in 9 out of 12 (M, z) bins, the $\mathrm{Cov}(\Delta\Sigma, \ln N_{\rm gal} \mid M, z)$ rejects the null-correlation hypothesis with high statistical significance ($p-\mathrm{value} < 0.01$). However, in two bins, specifically $M_{\rm 200c} \in [5\times10^{14}, 1\times10^{15})\ M_\odot h^{-1}$ at $z = 0.49$ and $M_{\rm 200c} \in [2\times10^{14}, 5\times10^{15})\ M_\odot h^{-1}$ at $z = 1.03$, the magnitude of the covariance is relatively small compared to the size of their errors. Consequently, it becomes challenging to constrain the shape parameters in these two bins, and the covariance is consistent with the null hypothesis. Furthermore, we exclude the bin $M_{\rm 200c} \in [5\times10^{14}, 1\times10^{15})\ M_\odot h^{-1}$, $z = 1.03$ due to the limited number of halos it contains.

Our results suggest that the shape of the covariances can be accurately described by the full error function. Additionally, for $R \geq R_{\rm vir}$ or $R \geq R_{\rm 200c}$, the covariance aligns with the null-correlation hypothesis. This alignment is reflected in the fact that all nine bins with

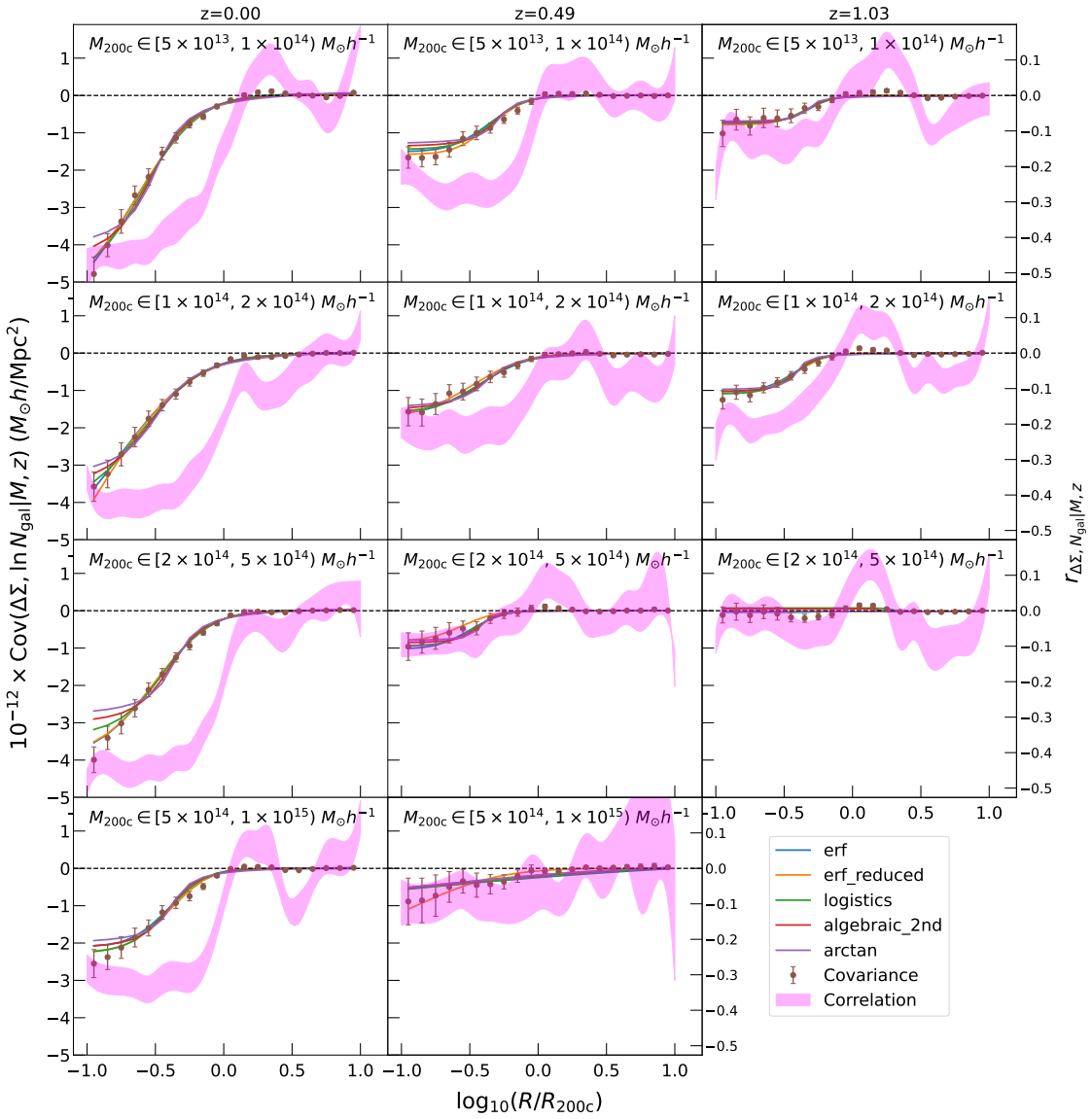


Figure 5.4: Measured against the left hand side y-axis are measurements of Cov($\Delta\Sigma$, $N_{\mathrm{gal}}|M,z$) with 1σ errors and different functional forms using the full model. The functions are classes of "Sigmoid" functions. In all bins, the error function outperforms other functional forms in their DIC parameters, providing good χ^2 values. For $M_{200c} \in [5 \times 10^{14}, 1 \times 10^{15})$ at $z = 0.49$ and $M_{200c} \in [2 \times 10^{14}, 5 \times 10^{15})$ at $z = 1.03$, the posteriors of the full models do not converge as the size of the covariance is too small. Measured against the right-hand side y-axis are the correlation coefficients $r_{\Delta\Sigma, N_{\mathrm{gal}}|M,z}$ with smoothed bands representing the $1-\sigma$ error. The errors are measured by bootstrap resampling.

constrained posterior shape have best-fit g values within 2σ of $g = -1$. Deviations from $g = -1$ can be interpreted as evidence of disagreements with the Press-Schechter formalism (Press & Schechter, 1974) of spherical collapse halos, which can be originated from the presence of anisotropic or non-Gaussian matter distribution around halos at large scales (Lokken et al., 2022), or it can be an indicator of an open-shell model of halos that allows for the bulk transfer of baryonic and dark matter in and out of the halo potential well during the non-linear collapse.

With $g = -1$ fixed, the reduced error function marginally improves the constraints in most bins. However, with the reduced model, we can provide posterior constraints for $M_{200c} \in [5 \times 10^{14}, 1 \times 10^{15}) M_\odot h^{-1}$ at $z = 0.49$ and $M_{200c} \in [2 \times 10^{14}, 5 \times 10^{15}) M_\odot h^{-1}$ at $z = 1.03$, which the full model failed to constrain but with very loose posterior constraints. The estimated parameters for both the full and reduced models are presented in Table 5.6 and Table 5.7, respectively.

To assess the impact of varying the definition of the halo radius on our measurements of the shape of the covariance, we considered two factors: the scale dependence of $\Delta\Sigma$ discussed in §5.2.1 and §5.3.1, and the alteration of the richness-mass relation as shown in Figure 5.3 in §5.3.2. Figure 5.5 demonstrates that there is no apparent evolution of the shape parameters $\boldsymbol{\theta} \in \{\tau, \gamma, g\}$ when altering the scale dependence for $\Delta\Sigma$ or the true richness count. However, we find marginal 3σ evidence of a difference in the amplitude parameter of the covariance s when changing the scale normalisation from $r_p = R/R_{200c}$ to $r_p = R/R_{\rm vir}$ while using the same true richness count. As halos exhibit more self-similarity in the inner regions when scaled by R_{200c} (Diemer & Kravtsov, 2014), we adopt this as our radius normalisation and use the number of galaxies enclosed within R_{200c} as our true richness count.

Subsequently, we explored the evolution of the shape parameters with respect to (M, z) and found no strong mass dependence. However, we observed a monotonically decreasing redshift dependence of the amplitude parameter s, as illustrated in Figure 5.6. To explain

both the halo mass and the redshift dependence, we used the peak height of the halo, $\nu(M, z)$.

5.4.2 Binned by peak height

An alternative binning scheme that encapsulates both the halo mass and redshift information is to bin halos by the peak height parameter, defined as

$$\nu = \frac{\delta_c}{\sigma(R, a)}, \tag{5.26}$$

where $\delta_c(z)$ is the collapse overdensity at which gravitational collapses enter the non-linear regime and $\sigma(R, a)$ is the smoothing scale seen in Equation (5.33) at the radius of the cluster. For an Einstein-de Sitter universe ($\Omega_m = 1$, $\Omega_\Lambda = 0$) $\delta_c \approx 1.686$ at the epoch of collapse and is weakly dependent on cosmology and redshift (Percival, 2005). $\sigma(R, a)$ scales as $\sigma(M, a) = \sigma(M, a = 1)D_{+0}(a)$ at the linear collapse regime, where $D_{+0}(a) \equiv D_+(a)/D_+(a = 1)$. Here $D_+(a)$ is the linear growth factor defined as

$$D_+(a) = \frac{5\Omega_M}{2} E(a) \int_0^a \frac{da'}{[a'E(a')]^3}, \tag{5.27}$$

for a ΛCDM cosmology, where $E(a) \equiv H(a)/H_0$ is the normalised Hubble parameter. $\sigma(R, z)$ depends strongly on redshift, and hence, the peak height ν strongly depends on the halo radius and the redshift of non-linear collapse.

The peak height has been adopted to simplify the mass and redshift dependence in various halo properties, such as halo concentration (Prada et al., 2012) and halo triaxiality (Allgood et al., 2006). Here, we explore whether the peak height can serve as a universal parameter to explain the scale and shape of $\mathrm{Cov}(\Delta\Sigma, \ln N_{\mathrm{gal}} \mid M, z)$. We bin the halos into deciles of ν and set posterior constraints on the shape of the covariance using our *erf* model in the full model case. In the highest decile (90%-100% percentile), we reject the null-correlation hypothesis at the $p = 0.01$ level, but due to the size of the error bars, the shape of the

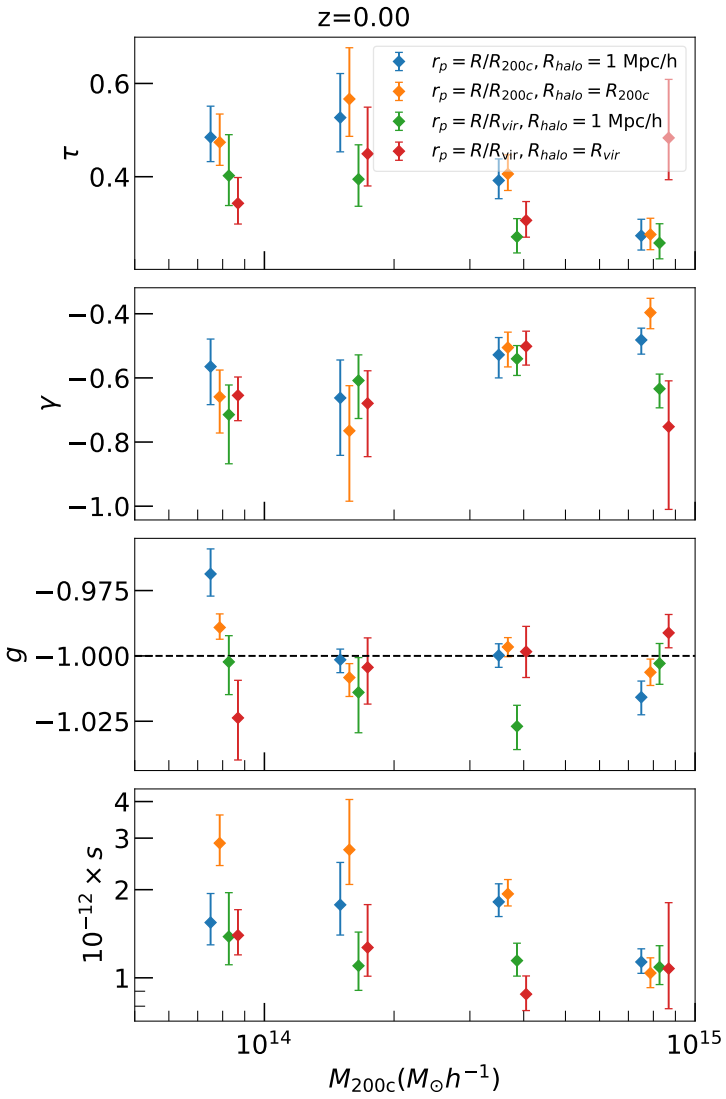


Figure 5.5: Evolution of $\mathrm{Cov}(\Delta\Sigma,\ N_{\rm gal} \mid M, z)$ shape parameters with respect to mass and radial binning schemes and $N_{\rm gal}$ definition at fixed redshift at z=0. $\Delta\Sigma$ is binned in equal log-space radial bins in $R/R_{\rm 200c}$ or $R/R_{\rm vir}$; for each radial binning, the number count of galaxies inside the cluster is given by a constant radius of 1 Mpch^{-1} or $R_{\rm 200c}$ when binned by $R_{\rm 200c}$ and $R_{\rm vir}$ when binned by $R_{\rm vir}$. We find no strong evolution in the shape or scale of the covariance under different binning schemes or $N_{\rm gal}$ definitions. The trend is consistent across different redshift bins and demonstrates the robustness of the covariance under different true richness definitions. The error bars indicate the $1-\sigma$ distribution of the posteriors.

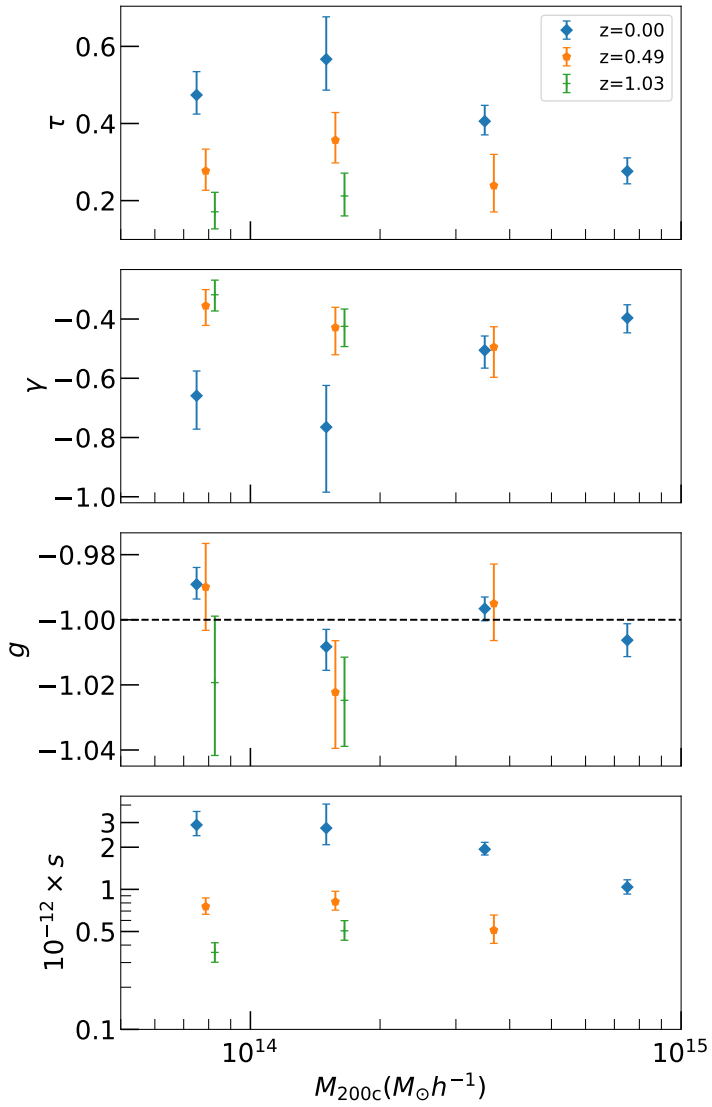


Figure 5.6: Evolution of $\mathrm{Cov}(\Delta\Sigma,\ N_{\mathrm{gal}} \mid M, z)$ shape parameters of the error function with respect to mass and redshift, binned in units of R_{200c} and with N_{gal} taken to be the number of clusters inside the R_{200c} radius of the cluster. There is no strong dependence of τ, γ, and g with respect to mass and redshift and a strong monotonically decreasing s with respect to redshift. At $M_{200c} \in [5 \times 10^{14}, 1 \times 10^{15})$, $z = 0.49$ and $M_{200c} \in [2 \times 10^{14}, 5 \times 10^{15})$, $z = 1.03$ the covariance is consistent with null at $p = 0.01$ and $p = 0.05$ levels, respectively. The error bars indicate the $1 - \sigma$ distribution of the posteriors.

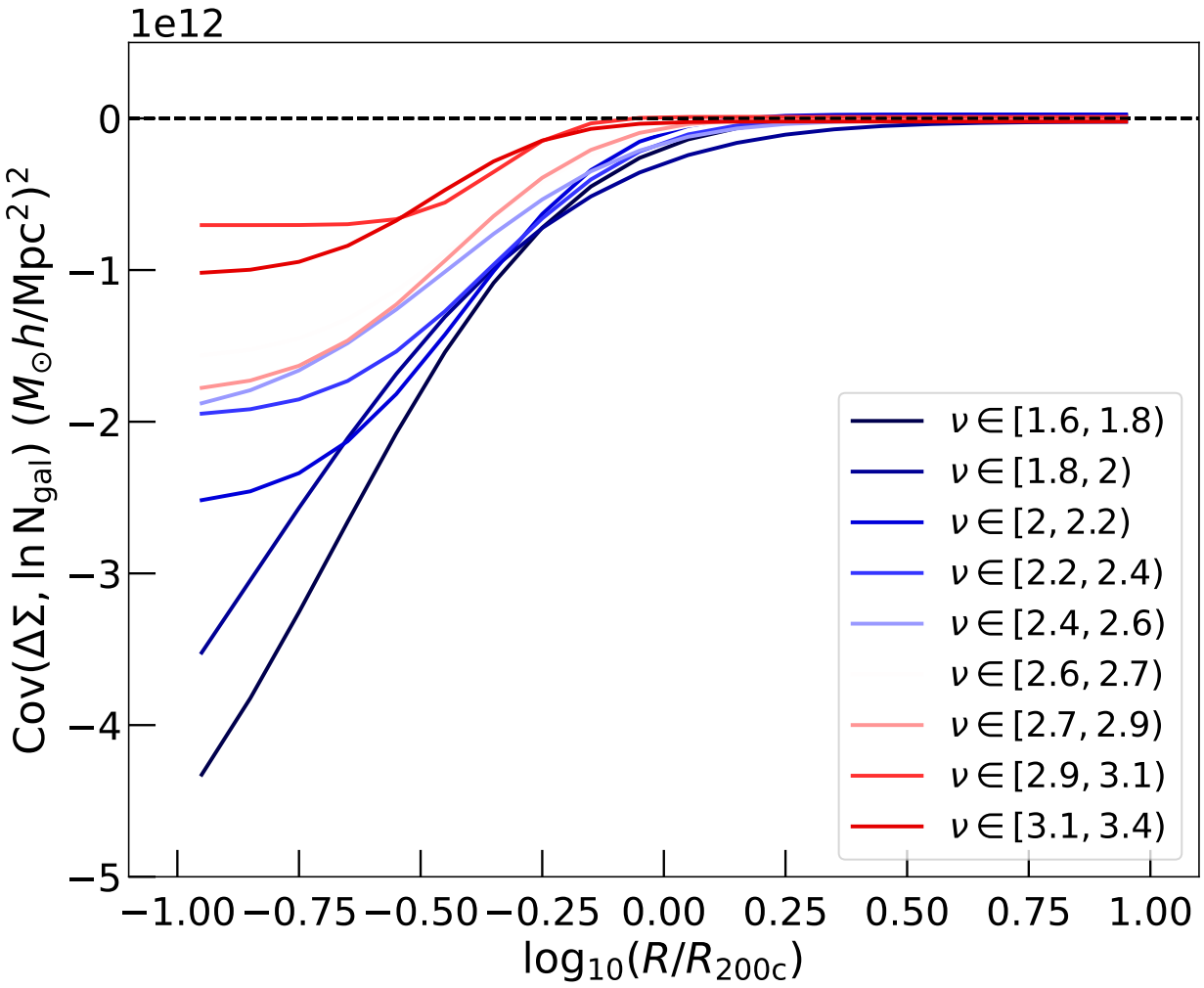


Figure 5.7: Best-fit "full" error function model for $\mathrm{Cov}(\Delta\Sigma, \ln N_{\mathrm{gal}} \mid M, z)$ when binned in deciles of halo peak height. The first nine bins reject the null hypothesis at a $p < 0.01$ level, and the highest decile rejects the null hypothesis at a $p < 0.05$ level. We can provide posterior constraints for all bins in peak height except for the one with the highest peak height value.

parameters τ, γ, and g and largely unconstrained and $s = 10^{12} \times 0.16^{+0.34}_{-0.12}$. Due to the large degeneracy, we exclude the highest decile from our dataset and limit the range of our model to $\nu \in [1.57, 3.40)$, which spans 0% to 90% of our sample set. The large error bars may be due to the fact that the halo abundance as a function of ν falls precipitously around $\nu \sim 4$, so the highest decile spans a wide tail of high $\nu \in [3.4, 4.6)$. The plots in Figure 5.7 are the best-fit templates when binned by peak height, and Figure 5.8 shows the best-fit parameters as a function of peak height. We do not see a strong dependence on the peak height for τ, γ and g. For s, its dependence on ν can be modeled as a log-linear relation of the form

$$\log_{10}(s) = C_s + \alpha\nu, \tag{5.28}$$

with $C_s = 13.07^{+0.26}_{-0.26}$ and $\alpha = -0.44^{+0.11}_{-0.11}$. At the highest decile, the $s = 10^{12} \times 0.16^{+0.34}_{-0.12}$ falls within the 1σ confidence band of the log-linear fit. Compared to the first nine deciles, the fit yields a χ^2 p-value of 0.73. The negative slope between s and ν indicates that more massive halos at the cosmic era of their formation exhibit a lesser anti-correlation between $\Delta\Sigma$ and $\ln N_{\rm gal}$.

5.5 Impact of $\mathrm{Cov}(\Delta\Sigma, \ln N_{\rm gal} \mid M, z)$ on Weak Lensing Measurements

To assess the impact of $\mathrm{Cov}(\Delta\Sigma, \ln N_{\rm gal} \mid M, z)$ on the scaling relation $\langle \Delta\Sigma \mid N_{\rm gal}, z \rangle$, we utilise Equation (6.18) for the first-order correction and Equation (5.18) for the 2nd order. The mean mass of the halos in each (M, z) bin is chosen as the pivot mass around which the HMF is Taylor expanded, and the intercept $\pi_{N_{\rm gal}}$ and slope $\alpha_{N_{\rm gal}}$ for the richness-mass scaling relation shown in Equation (5.9) are computed locally at the pivot point in each bin (M, z) bin. Our mock data, binned in R_{200c}, yields results that are consistent with the global richness-mass relation found in the literature (Bocquet et al., 2016; Costanzi et al., 2021; To et al., 2021a), as shown in Figure 5.3.

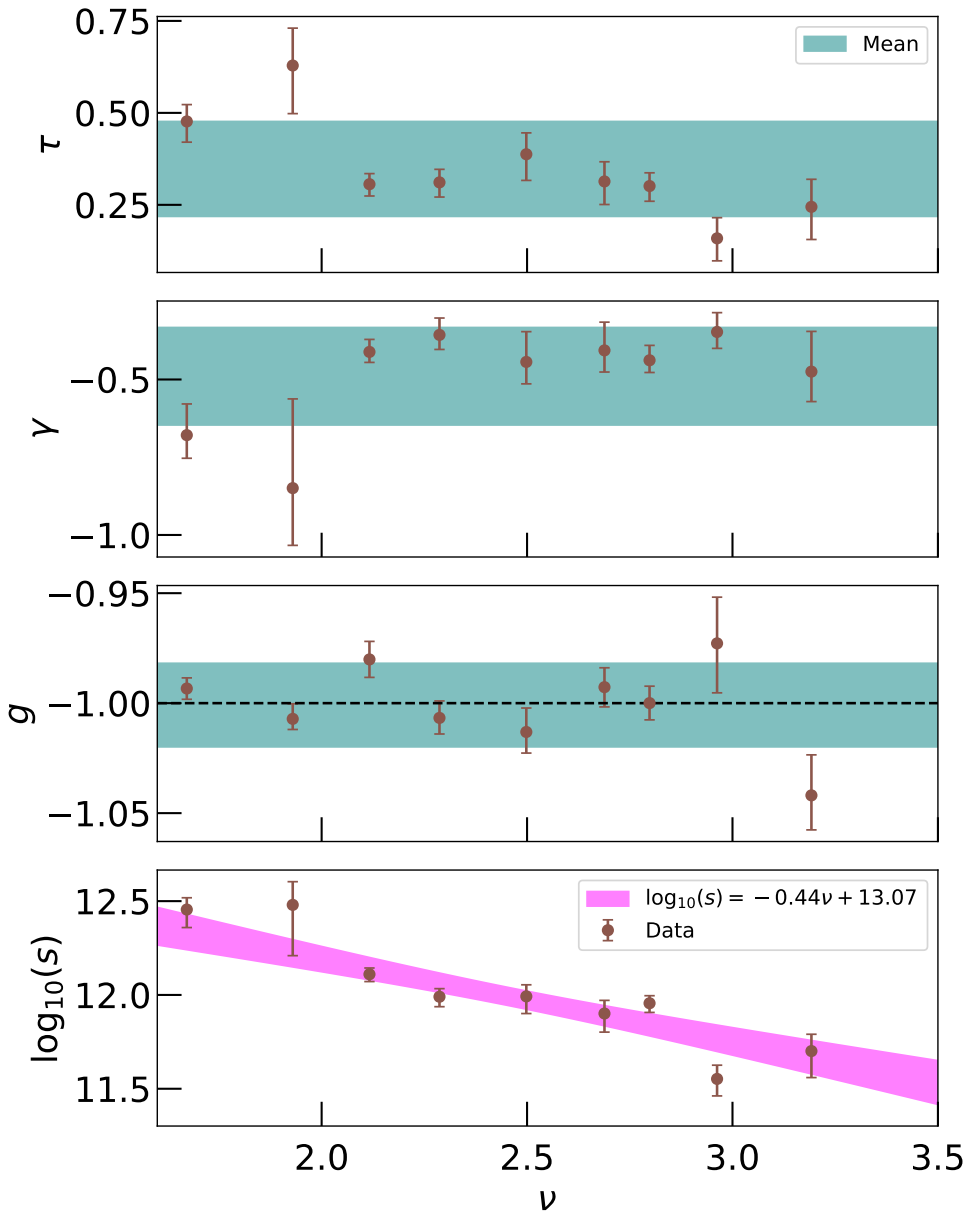


Figure 5.8: The evolution of shape parameters for $\mathrm{Cov}(\Delta\Sigma, \ln N_{\mathrm{gal}} \mid M, z)$ binned in deciles of the peak height ν, excluding the highest decile. The parameters τ, γ, and g show little dependency with ν while the amplitude s exhibits a log-linear relationship with ν of the form shown in Equation (5.28). The mean g is consistent with -1. The horizontal teal bands fill the 1σ range around the mean, and the pink line is the best log-linear fit between s and ν with 1σ confidence bands.

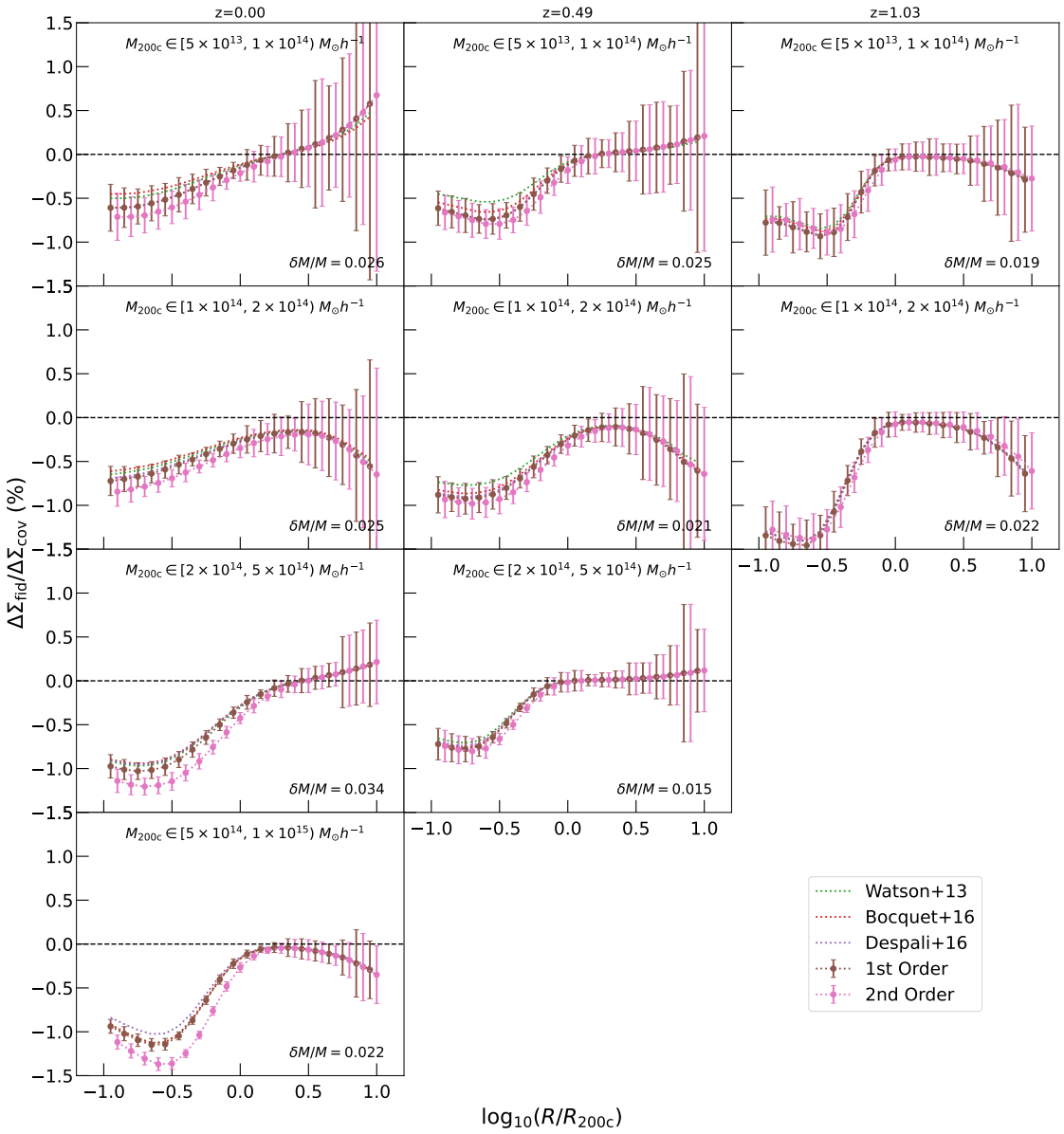


Figure 5.9: The percent level change in stacked $\Delta\Sigma$ measurements after including the covariance terms in Equations (6.18) and (5.18) as denoted by $\Delta\Sigma_{\rm cov}$ and without applying corrections as denoted by $\Delta\Sigma_{\rm fid}$. The slope and curvature of the halo mass function are calculated numerically from the Tinker et al. (2008) halo mass function in our nominal correction. The errors are taken from bootstrapped errors of the covariance. We compare the results with first-order corrections from other halo mass functions using Watson et al. (2013); Bocquet et al. (2016); Despali et al. (2016). We find that the percentile difference in $\Delta\Sigma$ far exceeds the uncertainty in the choice of halo mass functions, and that second-order corrections are subdominant to the first-order correction itself, which is at a $\sim 1\%$ level at small scales for $\Delta\Sigma$ and propagates into an *upward correction* of stacked halo mass of $\delta M/M \sim 2-3\%$ for most bins *after* applying the correction.

For the correction terms, we adopt the Tinker et al. (2008) halo mass function as our nominal model and compute the numeric log-derivative for values of γ_1 and γ_2, the log-slope and curvature of the halo mass function around the pivot mass. We compare the Tinker mass function results with others, including Watson et al. (2013); Bocquet et al. (2016); Despali et al. (2016), and find that the difference is subdominant to the first-order correction, which is at a $\sim 1\%$ level at small scales, as shown in Figure 5.9.

To estimate the mass bias in each bin, we stack $\langle \Delta\Sigma \mid M, z \rangle$ and model the profiles as if they were individual halos with a mean mass, redshift and concentration as described in Equations (5.2)-(5.7). We assume NFW profile using the concentration-mass model of Diemer & Joyce (2019) in the one-halo regime. The two-halo regime should not be affected, as the covariance is consistent with zero at $R \gtrsim R_{200c}$. We convert the 3D overdensity of the modeled halo ξ_{hm} to $\Delta\Sigma$ using Equations (5.2) and (5.5), and then apply the first-order correction in Equation (6.18). Using a Monte Carlo method we obtain the expected mass with and without this correction and report the change in the mean halo mass with this correction for each (M, z) bin. As shown in Figure 5.9, we find that adding the correction leads to an *upward correction* of the stacked halo mass of approximately $\delta M/M \sim 2 - 3\%$ for most (M, z) bins.

5.6 Explaining the Covariance

5.6.1 Secondary halo parameter dependence of $\ln N_{\rm gal}$

We employed a multi-variable linear regression model to determine the best-fit when incorporating secondary properties in the regression. Initially, considering the full set of parameters listed in Table 5.3, we applied a backward modeling scheme to identify the relevant parameters of interest. Details of this process can be found in Appendix 5.9.5, which led to the selection of the following secondary halo parameters for our model: $\Pi \subset$

$\{\Gamma_{\rm 2dyn}, a_{1/2}, c_{\rm vir}, T/|U|, X_{\rm off}\}$. The resulting model demonstrated good explanatory power, as indicated by a high R^2 coefficient. Additionally, the model passed various tests, including variance inflation, global F-statistic, partial F-statistic, T-statistic, scatter heteroscedasticity, and scatter normality in most cases. Specifically, through a comparison of $F_{\rm partial}$ values, we found that richness could be modeled by a multi-linear equation involving all secondary halo parameters. Further information can be found in Table 5.8, where the F-statistic demonstrates that all parameters are statistically significant. Only when considered collectively can they accurately reflect the dependence of richness on halo formation history.

To establish informative priors for upcoming weak lensing surveys such as HSC and LSST, we examined whether the dependence of $N_{\rm gal}$ on secondary halo properties, as inferred from the slope $\beta_{N_{\rm gal}}$, aligns with arguments based on halo formation physics. We expected that $\beta_{N_{\rm gal},c_{\rm vir}}$ resulting from the formation of satellite galaxies (equivalent to $N_{\rm gal}-1$ in the presence of a central galaxy) within halos would exhibit a negative relation, i.e., $\beta_{N_{\rm gal},c_{\rm vir}} < 0$. Simulation-based studies have suggested that early-forming halos possess higher concentrations (Wechsler et al., 2002), and correspondingly, high-concentration halos (which form early) have fewer satellite galaxies due to galaxy mergers within the halos (Zentner et al., 2005). This effect is known as galaxy assembly bias (Wechsler & Tinker, 2018) — the change in galaxy properties inside a halo at fixed mass due to the halo formation history. There is marginal evidence of the existence of assembly bias from recent observations using galaxy clustering techniques (Zentner et al., 2019; Wang et al., 2022), as well as measurements of the magnitude gap between the brightest central galaxy (BCG) and a neighboring galaxy as a proxy for formation time (Hearin et al., 2013; Golden-Marx & Miller, 2018; Farahi et al., 2020).

As noted in Table 5.8, the signs of $\beta_{N_{\rm gal},i}$ for the remaining parameters $i \in \{a_{1/2}, T/|U|, \Gamma_{\rm 2dyn}, X_{\rm off}\}$ align with our expectations of assembly bias in most bins — late-forming clusters undergo more rapid mass accretion (higher $\Gamma_{\rm 2dyn}$) and are less virialized

(higher T/|U|), and because they also from the galaxy assembly bias mentioned above are richer in galaxy number counts when conditioned on the mass, we expect a positive partial slope $\beta_{N_{\rm gal},\Gamma_{\rm 2dyn}}$ and $\beta_{N_{\rm gal},{\rm T/|U|}}$. The case for $a_{1/2}$ and $X_{\rm off}$ is more complicated. Under the isolated formation of halos $a_{1/2}$ and $X_{\rm off}$ would be smaller for earlier forming halos due to the monotonic mass accretion and relaxation of halos over long time scales. However, as halos undergo mergers and tidal stripping the monotonicity of the parameters over time is not guaranteed. Therefore, we see a mixture of positive and negative partial slopes $\beta_{N_{\rm gal},{\rm a}_{1/2}}$ and $\beta_{N_{\rm gal},{\rm X_{off}}}$ in these cases. To describe the physical mechanisms on a case-by-case basis would require that we probe into the halo merger tree history of individual halos.

In this paper, we take a closer look at the sign of $\beta_{N_{\rm gal},c_{\rm vir}}$ and observe that while the partial scope matches our expectations in most bins, in some mass bins at medium and high redshifts it changes signs from negative at lower redshifts to positive at higher ones. While we observe a diminishing impact of secondary halo properties on richness (indicated by a smaller absolute value for $\beta_{N_{\rm gal},c_{\rm vir}}$), the reversal of the coefficient's sign cannot be solely attributed to statistical fluctuations around zero, as some values are inconsistent with zero at levels exceeding 3σ.

This issue can be attributed to the effect of major mergers on concentration. Recent studies (Ludlow et al., 2012; Wang et al., 2022; Lee et al., 2023) have shown that halos, during major merger events, experience a transient fluctuation in concentration before returning to the mean relation over a time period slightly less than the dynamical time of the halo. The measured concentration spike during major mergers, particularly prominent at higher redshifts, could explain a positive $\beta_{\{N_{\rm gal},c_{\rm vir}\}}$.

To test this hypothesis, we employ a toy model that divides halos in each (M, z) bin based on the median $\Gamma_{\rm 1dyn}$ into low-$\Gamma_{\rm 1dyn}$ and high-$\Gamma_{\rm 1dyn}$ subsamples. Given the timescale of mergers to be roughly the dynamical time of the halo, we choose $\Gamma_{\rm 1dyn}$ as a good proxy for potential merger events even though this parameter is excluded in the final linear regression

model due to multicollinearity (see Appendix 5.9.5.)

Figure 5.10 displays the halo concentration plotted against the richness residuals, separated by $\Gamma_{\rm 1dyn}$, at benchmark bins of $M_{\rm 200c} \in [5 \times 10^{13}, 1 \times 10^{14})\ M_{\odot} h^{-1}$ at three different redshift snapshots of $z = 0, 0.49, 1.03$. At $z = 0$, we observe a negative slope as expected from halo formation physics for both low-$\Gamma_{\rm 1dyn}$ and high-$\Gamma_{\rm 1dyn}$ subsamples, as well as for the overall sample. Furthermore, we observe a change in the slope between the low-$\Gamma_{\rm 1dyn}$ and high-$\Gamma_{\rm 1dyn}$ sub-samples, which can be explained by the gradual increase (or decrease) in concentration ($\Gamma_{\rm 1dyn}$) over time, even without major merger events (Wechsler et al., 2002; Zhao et al., 2003; Lu et al., 2006). At redshifts of $z = 0.49, 1.03$, we observe a positive slope in the overall and/or high-$\Gamma_{\rm 1dyn}$ samples, which contradicts the scaling relations between HOD and concentration in models that track their gradual evolution over $T \gg T_{\rm 1dyn}$. However, in the presence of major mergers, when $\Gamma_{\rm 1dyn}$ is significantly enhanced, the halo concentration may also experience a transient spike after the merger. The deviation from hydrostatic equilibrium provides a *plausible* explanation for a positive $\beta\{N_{\rm gal}, c_{\rm vir}\}$, which could be fully tested on MDPL2 through the reconstruction of halo merger trees, an analysis beyond the scope of this paper.

5.6.2 Secondary halo parameter dependence of $\Delta\Sigma$

In this section, we employ a multi-linear regression approach to model the lensing signal, similar to the methodology described in §5.6.1. We extend this approach to the model $P(\Delta\Sigma|N_{\rm gal}(\Pi), M, z)$ as a linear function of Π. Upon analysing different (M, z, r_p) bins, we observe that the reduced parameters $\Pi \subset \{a_{1/2}, c_{\rm vir}, T/U, \Gamma_{\rm 2dyn}, X_{\rm off}\}$ pass the variance inflation factor (VIF) test for multi-collinearity, or in other words we showed that the variance is not inflated and thereby made less reliable in the case that the secondary halo parameters in the full model are highly correlated. As with the case for the lensing signal, this indicates that $\ln N_{\rm gal}$ can be described without redundancy by a linear decomposition of these reduced

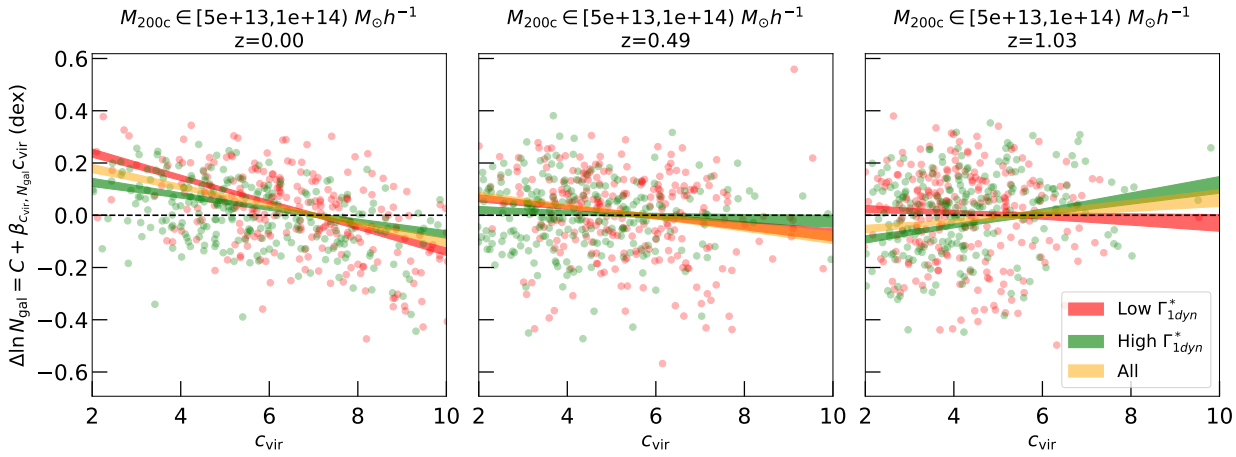


Figure 5.10: Residual log-richness versus concentration relation in subsets of halo mass accretion rate (MAR). The figure consists of three panels — left, middle, and right panels corresponding to $z = 0$, $z = 0.49$, and $z = 1.03$, respectively, all on a benchmark mass bin of $M_{200c} \in [5{\times}10^{13}, 1{\times}10^{14}) M_\odot h^{-1}$. The sample is split into low and high Γ_{1dyn} based on their median values. The scatter plot illustrates the data points, while the shaded regions show the best-fit linear fit with 1σ confidence interval for the main sample and each sub-sample. At $z = 0$, the richness-concentration relation exhibits a negative slope, consistent with our expectations of halo formation physics. The slopes for the low and high Γ_{1dyn} subsamples diverge due to the negative correlation between concentration and MAR. However, at $z = 0.49$ and $z = 1.03$, the slopes for the entire sample and/or the high Γ_{1dyn} subsample become positive, contrary to our observations of the richness-concentration relation. In contrast, the low Γ_{1dyn} subsample still shows a negative slope. These findings suggest that at medium to high redshifts, a subset of unrelaxed and recently merged halos with high MAR could elevate the concentration from its expected value at hydrostatic equilibrium.

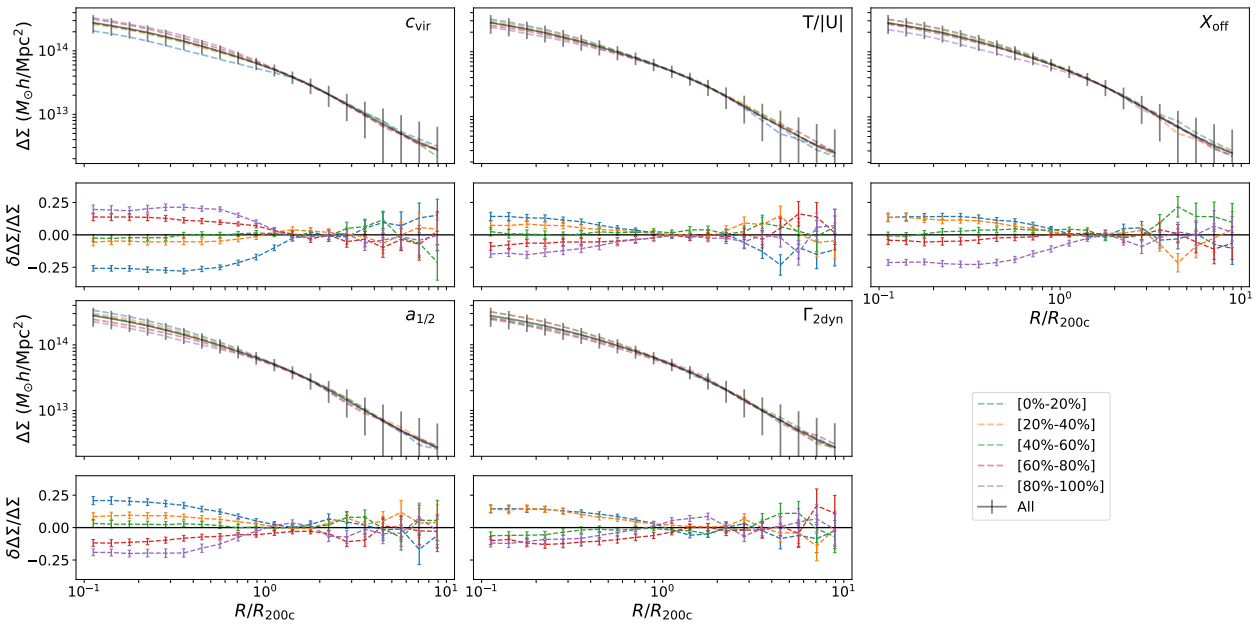


Figure 5.11: The dependence of $\Delta\Sigma$ on accretion history parameters in $M_{\rm 200c} \in [2\times10^{14}, 5\times10^{14})M_\odot h^{-2}$, $z = 0$. In each of the 5 panels is plotted the $\Delta\Sigma$ in different quintiles of the accretion history parameter $\Pi \in \{a_{1/2}, c_{\rm vir}, T/|U|, \Gamma_{\rm 2dyn}, X_{\rm off}\}$ compared to the mean $\Delta\Sigma$. For $\{a_{1/2}, T/|U|, \Gamma_{\rm 2dyn}, X_{\rm off}\}$ there is a strong negative correlation at small scales at $R \lesssim R_{\rm 200c}$ and for $c_{\rm vir}$ we find a strong positive correlation at $R \lesssim R_{\rm 200c}$. Comparing the width of $\Delta\Sigma$ binned at different quintiles of Π with the standard deviation of the profiles, we find that accretion history parameters account for much of the variance at small scales and play a negligible role at large scales.

parameters. Furthermore, most bins exhibit homoscedasticity, as confirmed by passing the Breusch-Pagan Lagrange multiplier test. This implies that the scatter terms $\sigma_{\Delta\Sigma}$ and σ_{Π_i} remain constant within each bin, with a few exceptions. Lastly, the scatter $\sigma_{\Delta\Sigma|N_{\rm gal}}$ in most bins (with a few exceptions) meets the criteria of the Shapiro-Wilk test for Gaussianity, suggesting that the distribution closely resembles a Gaussian distribution.

The multi-linear regression is a good fit to the conditioned lensing signal if we assume that $P(\ln N_{\rm gal}|M,z)$ and $P(\Delta\Sigma|M,z)$ can be modeled with a normal distribution (e.g., Anbajagane et al., 2020; Costanzi et al., 2021; To et al., 2021a). In this case $P(\Delta\Sigma \mid N_{\rm gal}, M, z)$ is a multi-linear equation with respect to the secondary halo parameters with mean

$$\begin{aligned}\langle\Delta\Sigma|N_{\rm gal},M,z\rangle \ &= \langle\Delta\Sigma|N_{\rm gal},M,z\rangle \\ &+C_1\sigma_{\Delta\Sigma}\Big(\sum_i \frac{\beta_{N_{\rm gal},i}}{\sigma_{\Pi_i}}\rho_{\Delta\Sigma-\Pi_i}\times(\Pi_i-\langle\Pi_i|M,z\rangle)\Big)\end{aligned} \tag{5.29}$$

and is normally distributed around the mean with variance

$$\begin{aligned}\sigma^2_{\Delta\Sigma|\ln N_{\rm gal}} \ &= \sigma_0^2 + C_2\sum_i \beta^2_{N_{\rm gal},i}\sigma^2_{\Pi_i}(1-\rho^2_{\Delta\Sigma-\Pi_i})+ \\ &C_3\sum_{i,j}^{j\neq i}\rho_{\Pi_i-\Pi_j}\sigma_{\Pi_i}\sigma_{\Pi_j}.\end{aligned} \tag{5.30}$$

The parameters $C1$, $C2$, $C3$, and σ_0 can be explicitly derived where $P(\ln N_{\rm gal}|M,z)$ and $P(\Delta\Sigma|M,z)$ are known, but the exact values are not essential for this paper. We refer the reader to Appendix 5.9.6 for derivations of Equations (5.29) & (5.30).

We note that only in bins of $R \lesssim R_{\rm 200c}$ do the multilinear regression models pass the global F-statistic test and the T-statistic test for each parameter. This result suggests that, at $R \gtrsim R_{\rm 200c}$, we find little correlation between $\Delta\Sigma$ and Π_i. Because the scatter still passes the Shapiro-Walk test for Gaussanity, the conditional probability $P(\Delta\Sigma|N_{\rm gal},M,z)$ at large

scales is still normally distributed, but with $\rho_{\Delta\Sigma-\Pi_i} = 0$. By setting $\rho_{\Delta\Sigma-\Pi_i} = 0$ the variance can be reduced to

$$\sigma^2_{\Delta\Sigma|N_{\rm gal}} = \sigma^2_0 + \sum_i \beta^2_{N_{\rm gal},i}\sigma^2_{\Pi_i} + \sum_{i,j}^{j\neq i} \rho_{\Pi_i-\Pi_j}\sigma_{\Pi_i}\sigma_{\Pi_j}. \tag{5.31}$$

We visualise the dependence of $\rho_{\Delta\Sigma-\Pi_i}$ on $R/R_{\rm 200c}$ in Figure 5.11. By dividing $\Delta\Sigma$ into quintiles of Π_i we find a strong correlation for all parameters at $R \lesssim R_{\rm 200c}$ and a null correlation at $R \gtrsim R_{\rm 200c}$. On small scales, our results show a positive correlation for concentration and a negative correlation for $\{a_{1/2}, T/U, \Gamma_{\rm 2dyn}, X_{\rm off}\}$. We observe that this trend holds for all (M, z) bins plotted for a benchmark bin of $M_{\rm 200c} \in [2 \times 10^{14}, 5 \times 10^{14}) M_\odot h^{-1}$ at $z = 0$.

The dependence of $\Delta\Sigma$ on secondary halo parameters qualitatively agrees with Xhakaj et al. (2022) wherein they targeted a narrow mass bin, with residual mass dependency inside the bin resampled so that mass follows the same distribution. In our work, we remove the mass dependency with the KLLR method (Farahi et al., 2022a), which achieves the same effect. We extend their results to mass and redshift bins probed by the optical surveys and quantitatively show that the dependence of $\Delta\Sigma$ on Π can be modeled as a multi-linear equation.

5.6.3 Results: secondary halo parameter dependence of $\mathrm{Cov}(\Delta\Sigma, \ln N_{\rm gal} \mid M, z)$

In Figure 5.12, we observe that the total covariance $\mathrm{Cov}(\Delta\Sigma, \ln N_{\rm gal}|M, z)$, which remains after removing the contribution of each secondary halo parameter $\beta_{N_{\rm gal},i}\mathrm{Cov}(\Delta\Sigma, \Pi_i|M, z)$, is consistent with zero at a significance threshold of 0.05 in all bins. The errors on the total covariance and individual contributions are computed using bootstrapping, and the errors on the remaining term are determined by adding the errors of the total and individual terms

in quadrature.

Based on our hypothesis in Equation (5.24), we conclude that the set of secondary halo parameters Π, which are related to the formation time and the mass accretion history of the halos, can fully explain the joint distribution of $\Delta\Sigma$ and $\ln N_{\rm gal}$ given the precision allowed by current errors, limited by the resolution limit (see Appendix 5.9.2 for information on particle resolution and measurement errors).

Since the joint distribution of $\Delta\Sigma$ and $\ln N_{\rm gal}$ follows a multivariate normal distribution, $P(\Delta\Sigma, \ln N_{\rm gal}|M, z)$ is completely characterised by its mean relation and $\mathrm{Cov}(\Delta\Sigma, \ln N_{\rm gal} \mid M, z)$. It should be noted that the contribution of each individual parameter to the total covariance, $\beta_{N_{\rm gal},i}\mathrm{Cov}(\Delta\Sigma, \Pi_i \mid M, z)$, is determined by the richness dependency captured by the slope $\beta_{N_{\rm gal},i}$ and the $\Delta\Sigma$ dependency represented by $\mathrm{Cov}(\Delta\Sigma, \Pi_i \mid M, z)$. Qualitatively, individual contributions to total covariance maintain their sign when both $\Delta\Sigma$ and $N_{\rm gal}$ contributions preserve their sign. Consistent with the arguments of halo formation, Π correlates with $\Delta\Sigma$ at small scales, as demonstrated in Figure 5.11 across all (M, z) bins. In most cases, the dependence of the richness on secondary halo parameters also maintains its sign across the (M, z) bins. In instances where we encounter a sign reversal in the $\ln N_{\rm gal} - c_{\rm vir}$ relation, we speculate that it is due to a transient increase in concentration following a major merger.

Furthermore, the total and individual contributions to the covariance tend to decrease in magnitude at smaller scales with increasing redshift. This decrease in covariance can be attributed to two factors: the decreasing explanatory power of Π on richness, as indicated by the decreasing values of R^2 and $F_{\rm partial}$ values in Table 5.8, and the decreasing absolute value of $\mathrm{Cov}(\Delta\Sigma, \Pi_i \mid M, z)$. This trend aligns with the idea that as halos have more time to form, the secondary halo properties related to the mass accretion history become more significant both in richness and in $\Delta\Sigma$. As discussed in §5.4, the dependence of the mass and redshift on covariance can be explained by the halo peak height, $\nu(M, z)$.

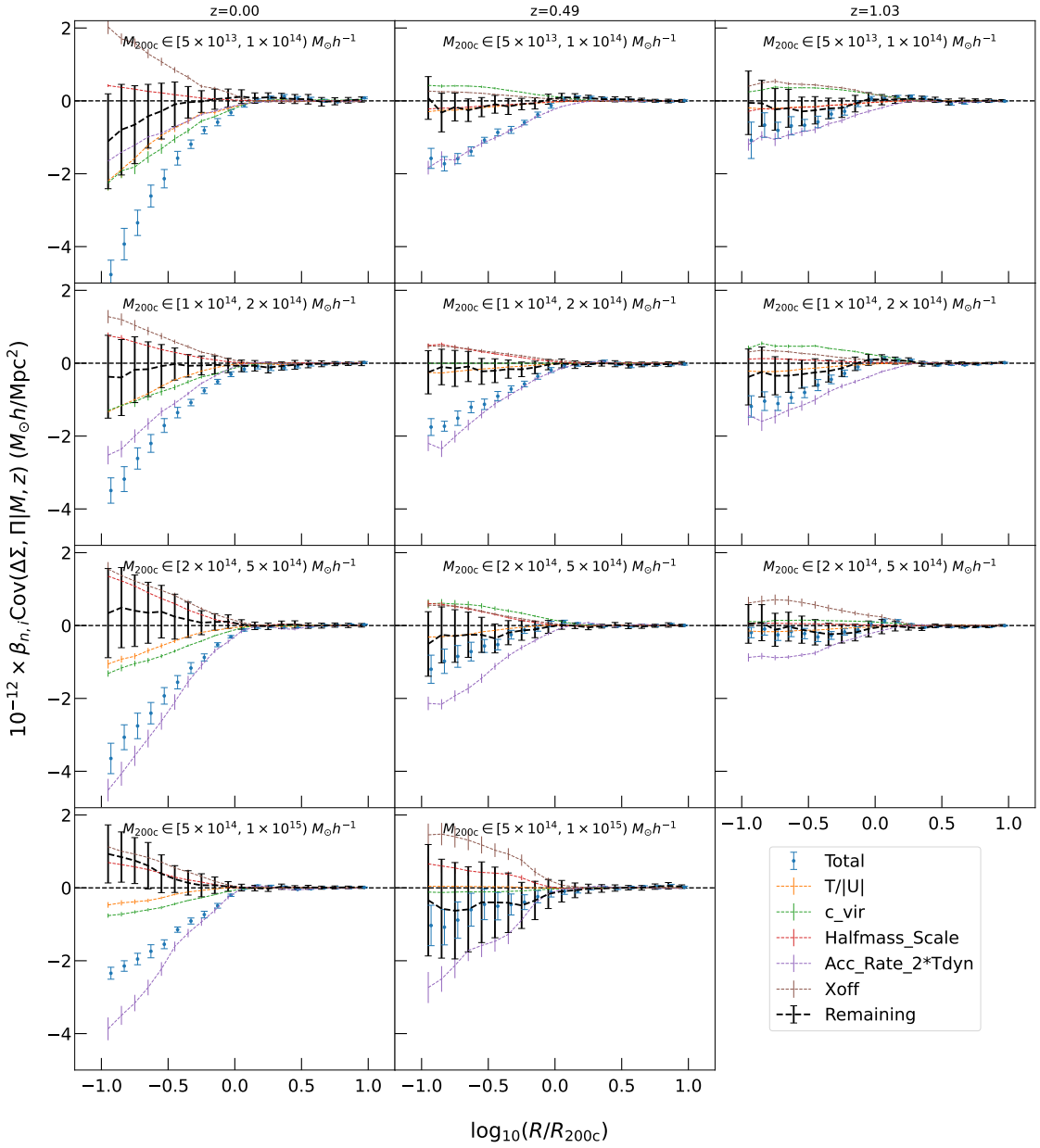


Figure 5.12: The dependence of $\mathrm{Cov}(\Delta\Sigma, \ln N_{\mathrm{gal}})$ on secondary halo parameters Π. The solid blue line is the total covariance. The dashed lines of the lines represent the covariance contribution coming from each of the secondary halo parameters modeled and governed by Equation (5.24), where $\mathrm{Cov}(\Delta\Sigma, \Pi|M,z)$ comes from the dependence of $\Delta\Sigma$ and the slope β_i is the dependency of richness. The thick black dashed line is the remaining covariance after removing the contribution from each Π_i term; the errors are obtained by adding the total and individual errors in quadrature without considering the correlations between terms. In agreement with our hypothesis in Equation (5.24), the remaining term is consistent with null at a $p < 0.01$ level for all bins.

5.7 Discussions

Intrinsic vs. Extrinsic Covariance. Our study distinguishes between the intrinsic covariance investigated here and that observed in empirical cluster data sets. This distinction arises from systematic biases introduced by the cluster-finding algorithm (extrinsic component) and the underlying physics governing halo formation (intrinsic component). Specifically, our analysis involves counting galaxies in 3D physical space. In contrast, a realistic cluster finder like redMaPPer (Rykoff et al., 2014) employs a probabilistic assignment of galaxies to halos in 2D physical space, considering projected radii and redshift through color matching onto the red sequence. Our study does not account for the observational systematics in redMaPPer associated with uncertainties in photometric redshifts and projection effects (Rozo et al., 2014; Farahi et al., 2016).

We find that the fractional amplitude of the bias and the scale dependence on the lensing signal observed in our results (Fig. 5.9) are comparable to those reported by Farahi et al. (2022b), who measured the covariance between the dark matter density and the galaxy number count enclosed inside a halo after applying a realistic stellar mass cut. The enclosed mass within a 3D radius using the IllustrisTNG100 simulation is anchored at $z = 0.24$. By comparing our findings to those obtained using a realistic cluster finder such as redMaPPer, we can unravel intrinsic and extrinsic contributions to the covariance between weak lensing observables (Wu et al., 2022). This work provides more profound insights into the distinct effects originating from the underlying physics and the methodology employed in cluster-finding algorithms (Euclid Collaboration et al., 2019).

Projection Effects. A noteworthy distinction arises regarding the covariance observed in our study compared to others examining clusters in projected space. Projection effects can potentially introduce a sign flip in covariance, as they can positively bias both $\Delta\Sigma$ and richness (Costanzi et al., 2019; Wu et al., 2022; Zhang & Annis, 2022; Zhang et al., 2023).

Particularly, Zhang et al. (2023) showed that the lensing signal can be affected both at large and small scales from the preferential alignment of halo orientation with the underlying large-scale structure filament. Wu et al. (2022) detected a positive correlation between $\Delta\Sigma$ and $\ln N_{\rm gal}$ employing Buzzard simulations, where $\Delta\Sigma$ were measured using dark matter particles and galaxy counts were performed within a cylindrical region of depth 60 Mpc/h. Their investigation revealed that the positive correlation primarily stems from galaxy number counts beyond the halo's virial radius and within 60 Mpc/h. On the other hand, using the Dark Quest emulator and HOD-based galaxy catalogues (Nishimichi et al., 2019), Sunayama et al. (2020) found negligible deviations from the mean relation at small scales and an overall reduction in selection bias at large scales, approximately halving the effect observed by Wu et al. (2022). Furthermore, Huang et al. (2022), using data from the Subaru Hyper Suprime-Cam (HSC) survey, observed that the selection bias is most prominent in the vicinity of the transition from the one-halo to the two-halo regime, as evidenced by the comparison between the outer stellar mass proxy and richness. In a study based on the IllustrisTNG300 simulation, Zhang & Annis (2022) discovered a net positive correlation between the fitted weak lensing mass and the projected 2D number count of the halo when conditioned on the halo mass.

These results suggest that projection effects can potentially introduce a positively correlated bias to both $\Delta\Sigma$ and $N_{\rm gal}$. We can estimate the impact of projection effects by comparing the intrinsic covariance measured in our study with the total covariance observed in the projected space. Consequently, our results serve two essential purposes: (i) elucidating the physical origins of the negative covariance and (ii) discerning intrinsic and extrinsic components to determine the covariance attributable to projection effects accurately.

Radial Dependence. There is a notable difference in the reported amplitude and scale dependence of covariance, which can be attributed to discrepancies in the employed halo occupation density models. Notably, a distinctive scale dependence discrepancy exists be-

tween simulation-based investigations of projection effects (Sunayama et al., 2020; Wu et al., 2022; Salcedo et al., 2020) and observational data from the HSC (Huang et al., 2022). In particular, the analysis of observational data reveals a prominent 1 Mpc bump, which could be explained by uncertainties inherent in observations, such as miscentering effects. It is crucial to gain insight into the sensitivity of the covariance with respect to the model parameters and the influence of selection effects. Understanding these factors is essential to comprehensively interpret and account for the observed covariance in galaxy cluster survey studies.

Accuracy vs. Precision. The statistical power of current and future surveys enables us to determine the normalisation and slope of the mass–observable relations at a few percent levels (e.g., **?**Mantz et al., 2016b; Mulroy et al., 2019; To et al., 2021a). However, these estimates are susceptible to known and unknown sources of systematic errors that inflate the uncertainties. These uncertainties introduce biases and degrade the accuracy of the results. Therefore, it is essential to carefully identify, quantify, and account for these systematic effects to ensure robust and reliable measurements. In this work, we focus on studying one of these sources of systematic uncertainty that was not considered previously.

5.8 Summary

This work reveals insights into the scale-dependent covariance between weak lensing observables and the physical properties of the halo. Using the MDPL2 N-body simulation with galaxies painted using the SAGE semi-analytic model, we present several key findings:

- We observe that the intrinsic covariance between $\Delta\Sigma$ and $\ln N_{\rm gal}$ enclosed within a 3D radius is negative at small scales and null at large scales in $(\ln M, z)$ ranges that cover optical surveys.

- We model the shape of the covariance across all bins using an error function that is

insensitive to the radius definition used to define halo boundaries.

- We find that the magnitude of the covariance is relatively insensitive to mass and decreases considerably with increasing redshift. The (M, z) dependence of the shape of the covariance can be encapsulated by the peak height parameter $\nu(M, z)$, which suggests that the scale of the covariance is related to the formation history of halos.

- We show that incorporating the covariance into $\langle \Delta\Sigma | N_{\rm gal}, z, r_p \rangle$ using the first-order expansion of the halo mass function yields about $> 1\%$ bias on $\langle \Delta\Sigma | N_{\rm gal}, z, r_p \rangle$ at small scales, which implies a mass bias of $> 2\%$ in the halo mass estimates in most bins.

- Our analysis reveals that the covariance between $\ln N_{\rm gal}$ and $\Delta\Sigma$ can be fully explained by secondary halo parameters related to the history of the halo assembly. This finding provides strong evidence that the non-zero covariance results from the variation in the formation history of dark matter halos.

5.9 Appendix

5.9.1 Functional form

This section aims to characterize the shape of the covariance across mass and redshift bins by fitting a template curve. The process involves several transformations and adjustments. First, a logarithmic transformation is applied to the radial bins, denoted as $x = \log_{10}(R/R_{\rm 200c})$. Then, a horizontal offset is introduced using a parameter γ, and scaling is applied using a parameter τ. This results in a transformed variable $\tilde{x} = (x - \gamma)/\tau$.

To analyze the transformed data vector $f(\tilde{x})$, we test a set of functional forms presented in Table 5.5. The normalization factors and coefficients associated with these functions are chosen in such a way that $f(\tilde{x})$ approaches 1 at large scales, -1 at small scales, $f'(0) = 1$ and

Parameters	Priors	
	Full	Reduced
τ	Uniform (0, 10)	Uniform (0, 10)
γ	Uniform (-5, 5)	Uniform (-5, 5)
g	Uniform (-2, 1)	Fixed at $g = -1$
$10^{-12} \times s$	Log-uniform (0.01, 10)	Log-uniform (0.01, 10)

Table 5.4: Priors for the model. We introduce two sets of priors. In the "full" models, the parameters are given physical (i.e., $\tau > 0$, $s > 0$) but non-informative uniform or log-uniform priors. In the "reduced" case, assuming that $\mathrm{Cov}(\Delta\Sigma, \ln N_{\rm gal} \mid M, z) = 0$ at large scales, we restrict $g = -1$ while assigning the same set of priors to all other parameters.

$f(0) = 0$.

A linear transformation of $f(\tilde{x})$ is then performed, given by $s\big(f(\tilde{x})+g\big)$, where g represents a vertical shift and s represents a scaling factor. The magnitude of s is comparable to the magnitude of $\mathrm{Cov}(\Delta\Sigma, \ln N_{\rm gal} \mid M, z)$, while the parameters γ, τ, g are of the order of unity. These parameters, along with s, form the set of parameters denoted as $\boldsymbol{\theta} \in \{\tau, \gamma, g, s\}$, which define our best-fit model. If $g = -1$, it implies a zero covariance at large scales. We fit two models: a full model with all parameters free, and a reduced model with $g = -1$, and the rest of the parameters free. The priors for the parameters are specified in Table 5.4.

For the full model, we choose the error function as our fiducial functional form. The estimated parameters for both the full and reduced models are presented in Table 5.6 and Table 5.7, respectively. Appendix 5.9.4 provides robustness testing to determine the best-fit model for our covariance. In Table 5.6, we compare the model parameters, χ^2 p-value, and the difference in Deviance Information Criterion (DIC) with our fiducial model using the candidate functions listed in Table 5.5. The error function generally outperforms other models when all parameters $\boldsymbol{\theta} \in \{\tau, \gamma, g, s\}$ are allowed to vary. Table 5.7 shows that the DIC of the reduced error function ($g = -1$) marginally outperforms the full error function in most cases, along with the posterior constraints of parameters shown in Figure 5.14.

Error function (fiducial)	$s\Big(\mathrm{erf}(\frac{\sqrt{\pi}}{2}\tilde{x}) + g\Big)$
Logistics function	$s\Big(\frac{2}{1+e^{\tilde{x}}} - 1 + g\Big)$
Inverse tangent	$s\Big(\frac{2}{\pi}\arctan\big(\frac{\pi}{2}\tilde{x}\big) + g\Big)$
Algebraic second order	$s(\tilde{x}/(1+\tilde{x}^2)^{1/2} + g)$

Table 5.5: Functional forms to model $\mathrm{Cov}(\Delta\Sigma, \ln N_{\rm gal})$. The radius in log-space x is transformed to $\tilde{x} \equiv (x-\gamma)/\tau$ by a horizontal offset γ and a characteristic scale τ. The functions $f(\tilde{x})$ are normalized so that $f(\tilde{x})$ asymptotically goes to 1 at $+\infty$, -1 at $-\infty$, $f(0) = 0$ and $f'(0) = 1$. Finally, we wrap $f(\tilde{x})$ by the function $p(f(\tilde{x})) \equiv s(f(\tilde{x}+g)$ to include a vertical offset g and amplitude parameter s. Together $\theta \in \{\tau, \gamma, g, s\}$ form the set of model parameters that allow us to make apple-to-apple comparisons between models.

5.9.2 Particle resolution and its impact on measurement errors

Using the 300 Cori Haswell node hours allocated by the NERSC to this project, we measured $\Delta\Sigma$ for $\sim$ 5000 clusters using dark matter particles downsampled by a factor of 10, in 20 log-spaced radial bins, at a projection depth of 200 h^{-1}Mpc.

At a downsampling rate of 10, our effective dark matter particle resolution is $M_{\rm p} \approx 1.51 \times 10^{10} h^{-1} M_\odot$. The error in $\Delta\Sigma$ comes from three sources: (i) cosmic variance, (ii) Poisson noise, and (iii) the intrinsic diversity of halos accounted for by secondary halo properties. In §5.6.2, we presented the contribution to $\Delta\Sigma$ scatter from secondary halo properties. Here, we compare the Poisson noise to the cosmic variance floor.

The cosmic variance introduces fluctuations in a 2D surface density fluctuation, given by

$$\delta\Delta\Sigma(R) = D_p \rho_m \sigma(R), \tag{5.32}$$

where $D_p = 200h^{-1}$Mpc is the projection depth, $\rho_m(z)$ is the mean density of the universe at that redshift, and $\sigma(R)$ is the root mean squared matter density fluctuation, given by

$$\sigma^2(R) = \int \Delta^2(k) \Big(\frac{3j_1(kR)}{kR}\Big)^2 d\ln k, \tag{5.33}$$

which is smoothed over an area of $A = 4\pi R^2$, $\Delta^2(k)$ is the matter power spectrum for a wavenumber k, and j_1 is the Bessel function of the first order.

Figure 5.13 shows the standard error of $\Delta\Sigma$ at a benchmark bin $M_{200c} \in [1 \times 10^{14}, 2 \times 10^{14}) M_\odot h^{-1}$ at z=0.00. At particle downsampling factors of 200, 100, and 10, the reduction in error is consistent with the Poisson term of $\sqrt{N}$, indicating that at these sampling rates, Poisson noise dominates. At our current downsampling rate of nth=10, the standard error is just above the cosmic variance floor at small scales and drops below the cosmic variance floor at large scales. In the ideal case that Poisson noise accounts for all the standard error, fully sampling all particles (nth=1, red dotted line) will reduce the standard error by a factor of $\sqrt{10}$, rendering it just below the cosmic variance floor at small scales. In the realistic case that the standard error for $\Delta\Sigma$ contributes from both Poisson noise and the intrinsic diversity of halo profiles, the fully sampled standard error should be on par with the cosmic variance at small scales. A future study with fully sampled particles should yield greater statistical constraints.

5.9.3 Derivation of second order expansion around the HMF

Following the formalism from Evrard et al. (2014) we derive Equation 5.18. The mean observable-mass scaling relation is given by the expression

$$\langle s_i \mid \ln M_0 \rangle = \alpha_i \ln M_0 + \pi_i, \tag{5.34}$$

for $s_i \in \{\Delta\Sigma, \ln N_{\rm gal}\}$ and pivot mass M_0. We now denote the deviation from the mean relation as δ_i, which from rearranging the terms in Equation (5.34) is given by $\delta_i = \frac{(s_i - \pi_i)}{\alpha_i} - \ln M_0$. From Evrard et al. (2014) the expression for the observable scaling relation for generic observables $\{a, b\}$ that follow a log-linear scaling relation as in Equation (5.34) is given by

$$\langle \delta_b \mid s_a \rangle = x_a \big[\langle \ln M \mid s_a \rangle + (\gamma_1 + \gamma_2 \delta_a) r_{ab} \sigma_{\ln M|a,1} \sigma_{\ln M|b,1} \big], \tag{5.35}$$

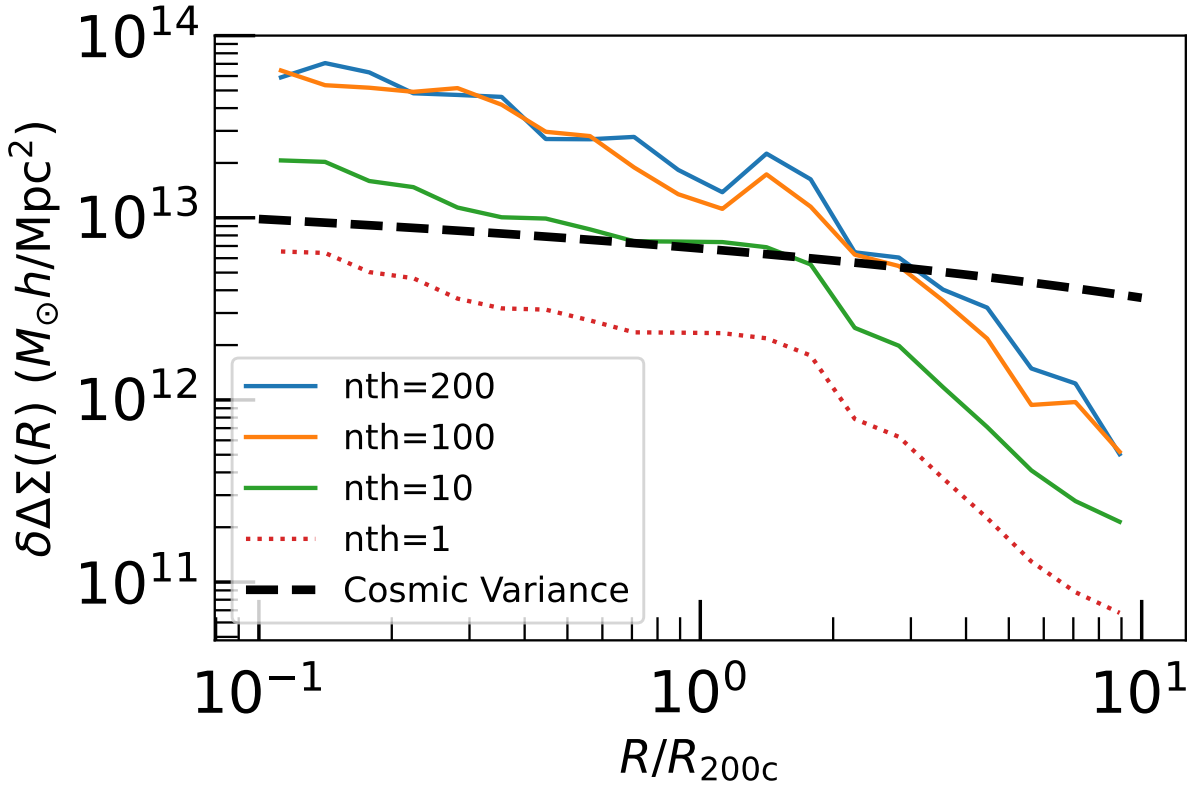


Figure 5.13: The standard error of $\Delta\Sigma$ measurements tested on a benchmark bin of $M_{200c} \in [1 \times 10^{14}, 2 \times 10^{14}) M_\odot h^{-1}$ at z=0. The standard error is estimated using the bootstrap method for the $N = 500$ clusters with dark matter particles downsampled by a factor of 200, 100, and 10 (solid lines). At our current resolution (nth=10, solid green line), the standard error is just above the cosmic variance at small scales and drops below the cosmic variance at large scales. The solid black line is density fluctuation estimated from the cosmic variance floor, as described in Equations (5.33) and (5.32). In the ideal case that Poisson noise accounts for all the standard error, fully sampling all particles (nth $= 1$, red dotted line) will reduce the standard error by a factor of $\sqrt{10}$, rendering it just below the cosmic variance floor at small scales. In the realistic case that the standard error for $\Delta\Sigma$ contributes from both Poisson noise and the intrinsic diversity of halo profiles, the fully sampled standard error should be on par with the cosmic variance at small scales.

where γ_1 and γ_2 are the first and second order coefficients of the Taylor expansion of the HMF around the pivot mass M_0 and $x_a = (1 + \gamma_2 \sigma^2_{\ln M|a,1})^{-1}$ the curvature term. The subscript 1 denotes the scatter for the HMF expanded to first order.

We now convert the left hand side of Equation (5.35) from the deviation from the mean scaling relation, δ_b, to the observable s_b to arrive at the expression

$$\begin{aligned}\langle s_b \mid s_a \rangle =& \big[\alpha_b x_a \langle \ln M \mid s_a \rangle + \ln M_0 - \pi_b\big] + \\ & \big[\alpha_b x_a (\gamma_1 + \gamma_2 \delta_a) r_{ab}\big] \sigma_{\ln M|a,1} \sigma_{\ln M|b,1} \\ =& \langle b \mid a, z \rangle_{\rm fid} + \\ & \mathrm{Cov}(a,b) \times \Big[\frac{x_a}{\alpha_a^2}(\alpha_a \gamma_1 + \gamma_2 (s_a - \pi_a))\Big], \end{aligned} \tag{5.36}$$

where we made use of the fact $\mathrm{Cov}(a, b|M, z) = r_{ab}\sigma_{a|M}\sigma_{b|M}$ and that the mass scatter conditioned on the observable to first order is related to the observable scatter by $\sigma_{\ln M|a,1} = \sigma_{a|M}/\alpha_a$, as shown in Equation 4 in Evrard et al. (2014) for the multivariate case. Substituting $\ln N_{\rm gal}$ for a and $\Delta\Sigma$ for b yields the expression for Equation (5.18).

5.9.4 Robustness testing of covariance modeling

The shape posterior is sampled by a Monto Carlo Markov Chain (MCMC) using the emcee package (Foreman-Mackey et al., 2013). We test for convergence by ensuring that the number of steps exceeds $100 t_{\rm auto}$ for all parameters, where $t_{\rm auto}$ is the integrated autocorrelation time as defined by Goodman & Weare (2010) and by ensuring that the convergence diagnostic denoted with R (Gelman & Rubin, 1992) across all walkers satisfy $R < 1.05$.

The posterior distribution according to Bayes theorem is given as:

$$p(\boldsymbol{\theta}|\{y_i\}) \propto p(\{y_i\}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{5.37}$$

$$= \prod_i p(y_i|\boldsymbol{\theta})p(\boldsymbol{\theta}), \tag{5.38}$$

where the second line assumes independent and identical distribution (i.i.d) for the data vectors. We set uniform priors $p(\boldsymbol{\theta})$ shown in Table 5.4 with signs and ranges motivated by the shape of the covariance (i.e., a negative γ and positive τ to offset $f(\tilde{x})$ to the left and a positive s and negative g shifts the fitted curve downwards).

We measure the goodness of fit using the left-tail p-value for the χ^2 with $N_{\rm data} - N_{\rm dim} = 20 - 4 = 16$ degrees of freedom. We compare between models by the Deviance Information Criterion defined as

$$\mathrm{DIC} = 2\overline{D(\boldsymbol{\theta})} - D(\overline{\boldsymbol{\theta}}), \tag{5.39}$$

where $\overline{\theta}$ is the best-fit parameters, and $D(\boldsymbol{\theta})$ is defined as

$$D(\boldsymbol{\theta}) \equiv -2\log\left(P(\{x_i\}|\boldsymbol{\theta})\right). \tag{5.40}$$

The performance between different functional forms (Table 5.5) is reported in Table 5.4.

The summary statistics for the posterior distribution of the covariance models are listed in Table 5.6 as plotted against measurements in Figure 5.4. Among the functions, the error function has either a better or comparable fit to all other functions in all other bins, as indicated by their DIC parameters. In two bins $M_{\rm 200c} \in [5\times10^{14}, 1\times10^{15})$ at $z = 0.49$ and $M_{\rm 200c} \in [2\times10^{14}, 5\times10^{15})$ at $z = 1.03$, the amplitude of the covariance is too small relative to their errors for shape parameters to be well-constrained. The right tail p-value for χ^2 is $p > 0.05$ for all but one bin. For this reason, we take the full error function as the nominal functional form.

For $R \geq R_{\rm vir}$ or $R \geq R_{\rm 200c}$, we find the covariance to be null at p-values > 0.01. A zero covariance at large scales implies $g = -1$ which coincides with the reduced model. We compare the results of the error function of the reduced model to the full model and find their performance varies from bin to bin as indicated by the DIC (Table 5.7). The posteriors of the reduced model provide marginally tighter constraints than the full model (Figure 5.14).

Mass & redshift	Error function (full): $s\left(\mathrm{erf}(\frac{\sqrt{\pi}}{2}\tilde{x}) + g\right)$							
	τ	γ	g	$10^{12} \times s$	$\Delta\mathrm{DIC}_{\mathrm{log}}$	$\Delta\mathrm{DIC}_{\mathrm{alg}}$	$\Delta\mathrm{DIC}_{\mathrm{arctan}}$	p-value
$[5 \times 10^{13}, 1 \times 10^{14})$, $z = 0.00$	$0.47^{+0.06}_{-0.05}$	$-0.66^{+0.08}_{-0.11}$	$-0.99^{+0.005}_{+0.004}$	$2.88^{+0.71}_{-0.47}$	4.8	17.8	31.5	0.27
$[1 \times 10^{14}, 2 \times 10^{14})$, $z = 0.00$	$0.57^{+0.11}_{-0.08}$	$-0.76^{+0.14}_{-0.22}$	$-1.008^{+0.0005}_{-0.0007}$	$2.73^{+0.13}_{-0.65}$	-2.9	-1.7	2.8	0.88
$[2 \times 10^{14}, 5 \times 10^{14})$, $z = 0.00$	$0.41^{+0.04}_{-0.04}$	$-0.51^{+0.05}_{-0.06}$	$-0.997^{+0.004}_{-0.0004}$	$1.93^{+0.24}_{-0.18}$	4.4	17.4	35.0	0.79
$[5 \times 10^{14}, 1 \times 10^{15})$, $z = 0.00$	$0.27^{+0.03}_{-0.03}$	$-0.40^{+0.04}_{-0.05}$	$-1.006^{+0.005}_{-0.0005}$	$1.03^{+0.13}_{-0.11}$	-0.27	14.3	29.1	0.0009
$[5 \times 10^{13}, 1 \times 10^{14})$, $z = 0.49$	$0.28^{+0.6}_{-0.05}$	$-0.36^{+0.05}_{-0.07}$	$-0.990^{+0.013}_{-0.013}$	$0.75^{+0.11}_{-0.09}$	2.7	9.0	15.1	0.72
$[1 \times 10^{14}, 2 \times 10^{14})$, $z = 0.49$	$0.35^{+0.07}_{-0.06}$	$-0.43^{+0.07}_{-0.09}$	$-1.022^{+0.016}_{-0.017}$	$0.81^{+0.15}_{-0.10}$	1.7	5.4	9.3	0.97
$[2 \times 10^{14}, 5 \times 10^{14})$, $z = 0.49$	$0.24^{+0.08}_{-0.07}$	$-0.49^{+0.07}_{-0.10}$	$-0.995^{+0.012}_{-0.011}$	$0.51^{+0.14}_{-0.01}$	0.6	2.4	4.0	0.63
$[5 \times 10^{14}, 1 \times 10^{15})$, $z = 0.49$				NA				
$[5 \times 10^{13}, 1 \times 10^{14})$, $z = 1.03$	$0.17^{+0.05}_{-0.04}$	$-0.31^{+0.05}_{-0.05}$	$-1.019^{+0.020}_{-0.022}$	$0.35^{+0.06}_{-0.05}$	4.3	6.5	8.3	0.05
$[1 \times 10^{14}, 2 \times 10^{14})$, $z = 1.03$	$0.21^{+0.06}_{-0.05}$	$-0.42^{+0.05}_{-0.06}$	$-1.024^{+0.013}_{-0.0014}$	$0.51^{+0.09}_{-0.07}$	-0.3	3.3	6.1	0.01
$[2 \times 10^{14}, 5 \times 10^{14})$, $z = 1.03$				NA				

Table 5.6: Summary statistics for $\mathrm{Cov}(\Delta\Sigma,\ N_{\mathrm{gal}} \mid M, z)$ binned by $R_{200\mathrm{c}}$ and M_{200c}, with N_{gal} defined inside the halo $R_{200\mathrm{c}}$. Columns 2-5 are the best-fit parameters for the nominal error function and their 1σ ranges. Columns 6-8 are the difference between the DIC of the logistics, algebraic, and inverse tangent models with the nominal error function, respectively. Column 9 is the right-tail p-value as measured by the χ^2 statistic with 20-4 = 16 degrees of freedom. Across all bins with applicable posterior constraints, the error function out-performs or is comparable to alternative functional forms as indicated by the difference in DIC, and has $p \geq 0.01$ in all but one bin. In two bins $M_{200\mathrm{c}} \in [5 \times 10^{14}, 1 \times 10^{15})$ at $z = 0.49$ and $M_{200\mathrm{c}} \in [2 \times 10^{14}, 5 \times 10^{15})$ at $z = 1.03$ the size of the covariance is too small relative to the size of their errors for shape parameters to be constrained. The covariance in these two bins is consistent with null at $p = 0.01$ and $p = 0.05$ levels, respectively.

5.9.5 Modeling secondary properties

We describe the linear regression model for richness used in §5.6.1. The same methodology is applied to $\Delta\Sigma$ in §5.6.2.

To model the expected natural logarithm of galaxy count ($\ln N_{\rm gal}$), we decompose it linearly using secondary halo parameters listed in Table 5.3, as shown in Equation (5.20). We employ the least squares method for linear regression and examine parameter redundancies. For over half of the bins, the parameters $\Gamma_{\rm inst}$, $\Gamma_{\rm 100Myr}$, $\Gamma_{\rm 2dyn}$, and $\Gamma_{\rm peak}$ exhibit collinearity, with Variance Inflation Factors (VIF) exceeding 5. This outcome is expected, as these quantities represent the same physical quantities smoothed over different time scales. As for the reduced set of non-collinear parameters, their correlation coefficient are quantified in Shin & Diemer (2023) using the Erebos simulation suite. To determine which parameters to retain, we utilize the partial F-statistic.

Table 5.8 demonstrates the diminishing explanatory power of Π on the richness as seen by the diminishing R^2 and $F_{\rm partial}$. We consider the partial F-statistic serves as a heuristic measure for the explanatory power of a variable, defined as:

$$F_{\rm partial} = \frac{(\mathrm{RSSE}_{\rm reduced} - \mathrm{RSSE}_{\rm full})/p}{\mathrm{RSSE}_{\rm full}/(n-k)}, \tag{5.41}$$

where RSSE is the residual sum of squared errors for the reduced model after removing the parameter in question and the full model containing $\Pi \subset \{a_{1/2}, c_{\rm vir}, T/U, \Gamma_{\rm 2dyn}, X_{\rm off}\}$, p is the number of parameters removed from the full model which in our case is by construction set to $p = 1$, n the number of data points, and k is the number of parameters in the full backward model. This statistic can be shown to be proportional to the contribution to the total R^2 *uniquely* explained by this parameter alone.

A partial F-statistic test reveals that $\Gamma_{\rm 2dyn}$ exhibits the highest partial F-statistic of all accretion rate parameters. Therefore, we retain this parameter in the reduced dimensional

Mass & redshift	Error function (reduced): $s\Big(\mathrm{erf}(\frac{\sqrt{\pi}}{2}\tilde{x}) - 1\Big)$				
	τ	γ	$10^{12} \times s$	$\Delta\mathrm{DIC}_{\mathrm{erf-full}}$	p-value
$[5 \times 10^{13}, 1 \times 10^{14})$, $z = 0.00$	$0.44^{+0.04}_{-0.04}$	$-0.63^{+0.07}_{-0.09}$	$2.66^{+0.56}_{-0.39}$	-5.2	0.06
$[1 \times 10^{14}, 2 \times 10^{14})$, $z = 0.00$	$0.69^{+0.11}_{-0.09}$	$-0.98^{+0.20}_{-0.27}$	$4.10^{2.31}_{1.19}$	4.5	0.96
$[2 \times 10^{14}, 5 \times 10^{14})$, $z = 0.00$	$0.40^{+0.04}_{-0.03}$	$-0.50^{+0.05}_{-0.06}$	$1.90^{+0.26}_{-0.20}$	0.2	0.78
$[5 \times 10^{14}, 1 \times 10^{15})$, $z = 0.00$	$0.29^{+0.04}_{-0.04}$	$-0.40^{+0.04}_{-0.04}$	$1.12^{+0.13}_{-0.11}$	6.1	0.005
$[5 \times 10^{13}, 1 \times 10^{14})$, $z = 0.49$	$0.25^{+0.04}_{-0.03}$	$-0.35^{+0.04}_{-0.04}$	$0.79^{+0.08}_{-0.07}$	6.2	0.57
$[1 \times 10^{14}, 2 \times 10^{14})$, $z = 0.49$	$0.38^{+0.14}_{-0.09}$	$-0.48^{+0.12}_{-0.24}$	$0.85^{+0.46}_{-0.19}$	-6.9	0.81
$[2 \times 10^{14}, 5 \times 10^{14})$, $z = 0.49$	$0.25^{+0.21}_{-0.12}$	$-0.55^{+0.14}_{-0.42}$	$0.40^{+0.69}_{-0.13}$	NA	0.48
$[5 \times 10^{14}, 1 \times 10^{15})$, $z = 0.49$	$0.61^{+0.27}_{-0.24}$	$-0.90^{+0.49}_{-0.76}$	$1.01^{+3.05}_{-0.60}$	-5.7	0.99
$[5 \times 10^{13}, 1 \times 10^{14})$, $z = 1.03$	$0.20^{+0.06}_{-0.05}$	$-0.35^{+0.06}_{-0.08}$	$0.38^{+0.07}_{-0.06}$	-0.75	0.04
$[1 \times 10^{14}, 2 \times 10^{14})$, $z = 1.03$	$0.22^{+0.06}_{-0.05}$	$-0.42^{+0.06}_{-0.8}$	$0.54^{+0.11}_{-0.09}$	3.9	0.3
$[2 \times 10^{14}, 5 \times 10^{14})$, $z = 1.03$	$5.97^{+2.76}_{-3.01}$	$-0.55^{+3.30}_{-2.99}$	$0.02^{+0.02}_{-0.008}$	NA	0.15

Table 5.7: Summary statistics for $\mathrm{Cov}(\Delta\Sigma, N_{\mathrm{gal}} \mid M, z)$ binned by $R_{200\mathrm{c}}$ and M_{200c} with N_{gal} defined inside the halo $R_{200\mathrm{c}}$ for the reduced error function model. Compared with the full error function model, the performance of the reduced model varies from bin to bin – using $\Delta\mathrm{DIC} > 3$ as a statistically significant result, it outperforms the full model in 5/9 overlapping bins, under-performs in 3/9 bins, and is comparable in 2 bins. The reduced model is able to yield convergent chains for $M_{200\mathrm{c}} \in [5 \times 10^{14}, 1 \times 10^{15})$ at $z = 0.49$ and $M_{200\mathrm{c}} \in [2 \times 10^{14}, 5 \times 10^{15})$ at $z = 1.03$ but with poor constraints on the parameters.

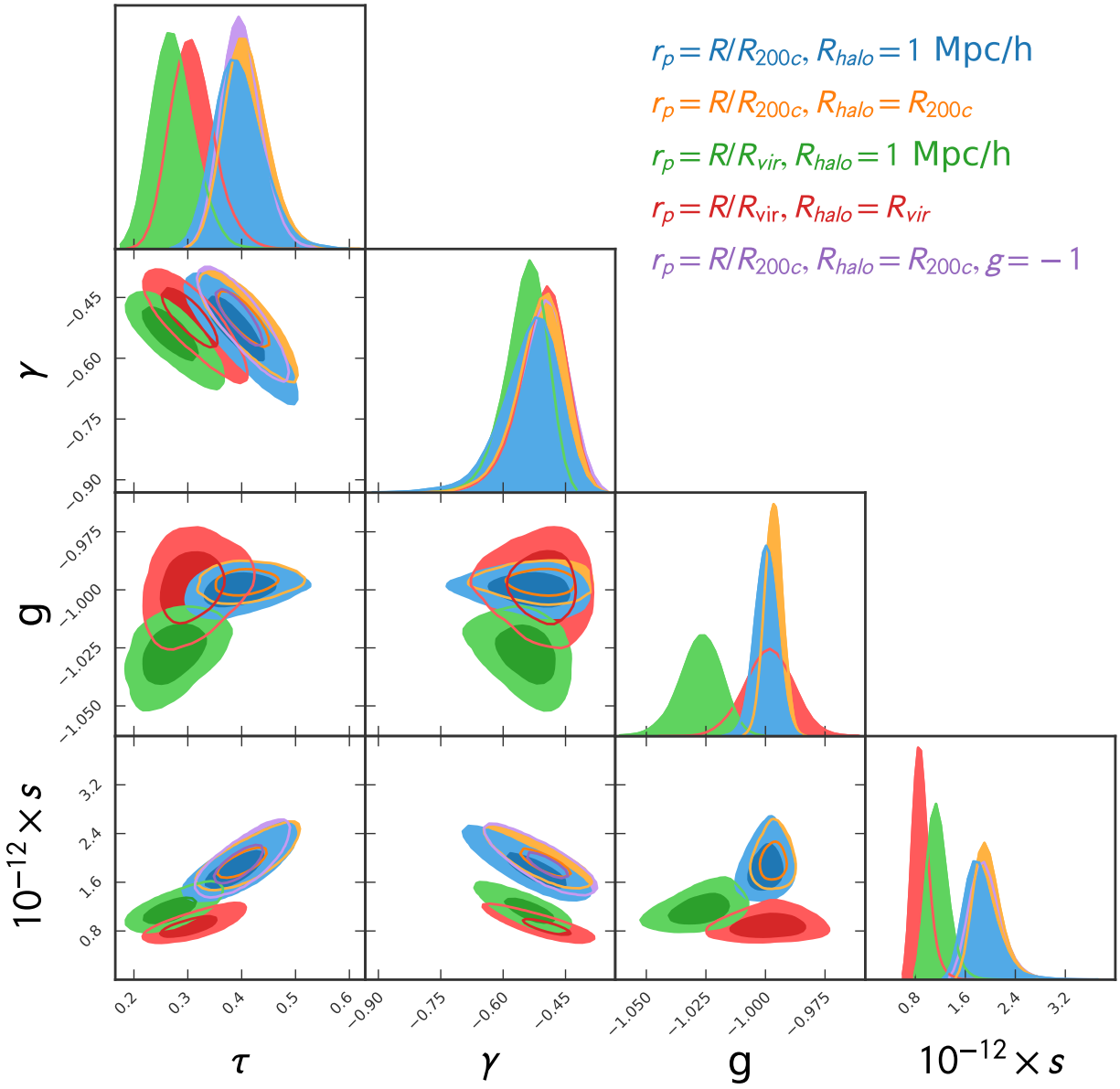


Figure 5.14: Posterior distribution of shape parameters in a benchmark bin of $M_{200c}/M_{vir} \in [2 \times 10^{14}, 5 \times 10^{14})$ at $z = 0.00$ under different binning schemes r_p and N_{gal} models with different halo boundaries R_{halo}. The marginalized parameter constraints for the full model closely overlap one another, and using the reduced model with $g = -1$ marginally improves the posterior constraints. The plot was generated using pygtc (Bocquet & Carter, 2016).

linear regression. The final model includes the following parameters $\Pi \subset \{a_{1/2}, c_{\rm vir}, T/U, \Gamma_{\rm 2dyn}, X_{\rm off}\}$. To ensure the robustness of the linear model across all bins, we perform the following tests:

- Variance inflation factor (VIF) test with a cutoff of 5 to detect multicollinearity.
- Global F-statistic for the entire model with a significance level of 0.05 to examine the correlation between the dependent variable and all parameters.
- Partial F-statistic for the entire parameter set to compare the relative importance of each parameter. The Partial F-statistic measures the additional contribution of each parameter to the multi-linear fit by estimating its corresponding R^2 value.
- T-statistic for each parameter to verify that the coefficients significantly deviate from zero at a significance level of 0.05.
- Breusch-Pagan Lagrange Multiplier test (Breusch & Pagan, 1979) at a significance level of 0.05 to assess heteroscedasticity.
- Shapiro-Wilk test (Shapiro & Wilk, 1965) at a significance level of 0.05 to evaluate the Gaussianity of residuals.

Across all bins, the reduced model successfully satisfies the first four tests. However, some bins fail the Shapiro-Wilk test due to a negative skew and positive kurtosis. Nonetheless, a visual examination of Q-Q plots indicates that the residuals predominantly follow a Gaussian distribution, except for deviations at the tail ends. Q-Q plots, quantile-quantile plots, are visualization tools used to compare the quantiles of a dataset to the quantiles of a theoretical distribution, typically a normal distribution. They provide a visual assessment of how well the data aligns with the assumed distribution.

Redshift & M_{200c} ($M_\odot h^{-1}$)	R^2	Const	c_{vir}		T/|U|		$a_{1/2}$		Γ^*_{2dyn}		X_{off}	
			$\beta_{N_{gal}}$	F	$\beta_{N_{gal}}$,	F	$\beta_{N_{gal}}$	F	$10^6 \times \beta_{N_{gal}}$	F	$10^3 \times \beta_{N_{gal}}$	F
$z = 0.00$												
$[5 \times 10^{13}, 1 \times 10^{14})$	0.45	-0.69	-0.035(0.004)	171	1.448(0.135)	185	0.103(0.084)	166	2.743(1.68)	134	-1.62(0.23)	69
$[1 \times 10^{14}, 2 \times 10^{14})$	0.48	-0.46	-0.033(0.003)	217	1.086(0.120)	179	-0.198(0.071)	170	2.668(0.628)	206	-0.76(0.15)	86
$[2 \times 10^{14}, 5 \times 10^{14})$	0.49	-0.41	-0.029(0.003)	196	-0.794(0.086)	145	0.292(0.007)	213	-0.722(0.223)	167	0.57(0.11)	67
$[5 \times 10^{14}, 1 \times 10^{15})$	0.46	-0.32	-0.019(0.002)	119	1.503(0.068)	111	0.148(0.054)	201	0.483(0.096)	226	-0.25(0.06)	51
$z = 0.49$												
$[5 \times 10^{13}, 1 \times 10^{14})$	0.25	-0.71	0.006(0.005)	24	0.325(0.125)	69	0.188(0.101)	142	3.90(1.64)	102	-0.3(0.19)	35
$[1 \times 10^{14}, 2 \times 10^{14})$	0.21	-0.45	-0.450(0.082)	27	0.314(0.135)	46	-0.353(0.158)	107	1.80(0.71)	89	-0.37(0.13)	14
$[2 \times 10^{14}, 5 \times 10^{14})$	0.13	-0.56	-0.003(0.004)	0	0.323(0.110)	26	0.670(0.155)	51	0.388(0.108)	26	-0.27(0.11)	2
$[5 \times 10^{14}, 1 \times 10^{15})$	0.13	-0.32	-0.010(.008)	0	0.007(0.286)	0	0.697(.309)	5	-0.538(0.331)	3	-0.41(.14)	3
$z = 1.03$												
$[5 \times 10^{13}, 1 \times 10^{14})$	0.27	-1.12	0.015(0.005)	6	0.471(0.136)	21	0.569(0.269)	116	2.7(1.2)	63	-0.81(0.18)	0
$[1 \times 10^{14}, 2 \times 10^{14})$	0.22	-0.84	0.016(0.004)	10	0.374(0.135)	36	-0.233(0.234)	83	1.076(0.478)	36	-0.35(0.12)	2
$[2 \times 10^{14}, 5 \times 10^{14})$	0.20	-0.6	0.0015(0.005)	10	0.284(0.185)	1	-0.157(0.266)	15	0.160(0.213)	6	-0.55(0.11)	13

Table 5.8: Best-fit parameters, global R^2, and explanatory power indicators for log-richness modeled in Equation (5.20). Values in parentheses represent 1σ confidence intervals to the partial slopes β. The partial F-statistic, defined in Equation (5.41), is used to quantify the explanatory power of each variable. A higher partial F-statistic indicates a greater amount of predictive power uniquely attributed to that variable. Statistical significance is determined by an F-statistic of $F > 10$.

5.9.6 *Derivation of* $P(\Delta\Sigma \mid N_{\rm gal}, m, z)$

We demonstrate that $P(\Delta\Sigma \mid N_{\rm gal}, M, z)$ can be modeled as a multi-linear relation of secondary halo parameters of mean and variance given by Equations (5.29) and (5.30).

From our assertion that $P(\Delta\Sigma, \vec{\Pi} \mid M, z)$ is a bivariate normal, the conditional probability $P(\Pi_i \mid \Delta\Sigma, M, z)$ in each radial bin can be expressed as a normal distribution with mean

$$\langle \Pi_i | \Delta\Sigma, M, z \rangle = \langle \Pi_i | M, z \rangle + \rho_{\Pi_i - \Delta\Sigma} \frac{\sigma_{\Pi_i}}{\sigma_{\Delta\Sigma}} (\Delta\Sigma - \langle \Delta\Sigma | M, z \rangle), \tag{5.42}$$

and variance

$$\sigma^2_{\Pi_i | \Delta\Sigma} = \sigma^2_{\Pi_i} (1 - \rho^2_{\Pi_i - \Delta\Sigma}). \tag{5.43}$$

Here, we omit the radial dependence $R/R_{\rm 200c}$ for all variables and the conditional dependence on (M, z) in the subscripts for ρ and σ (i.e., $\sigma_{\Delta\Sigma}$ should be explicitly written as $\sigma_{\Delta\Sigma|M,z}(R)$).

For an independent random variable $Z = X + Y$ with X and Y uncorrelated independent random variables with distributions of the form $\sim \mathcal{N}(u, \sigma^2)$ and $\sim \mathcal{N}(\nu, \tau^2)$, respectively, Z is another Gaussian of $\mathcal{N}(u + v, \sigma^2 + \tau^2)$. In the case that X and Y are correlated, we must introduce a cross term in the variance of probability distribution Z, namely $P(Z) \sim \mathcal{N}(u + v, \sigma^2 + 2\rho_{X-Y}\sigma\tau + \tau^2)$.

In our specific case, we want to model the distribution $P(\sum_i \beta_{s,i} \Pi_i \mid \Delta\Sigma, M, z)$ from $P(\Pi_i \mid \Delta\Sigma, M, z)$ and the correlation here refers to the correction between secondary halo parameter ρ_{Π_i, Π_j}.[2] From the convolution theorem and the expressions for $P(\Pi_i \mid \Delta\Sigma, M, z)$ in Equations (5.42) and (5.43), we obtain the expression for $P(\ln N_{\rm gal} \mid \Delta\Sigma, M, z)$ as a

2. In §5.6.2, we show through the Variance Inflation test that set of reduced model parameters $\Pi \in \{a_{1/2}, c_{\rm vir}, T/U, \Gamma^*_{\rm 2dyn}, X_{\rm off}\}$ are not multi-collinear.

normal distribution with mean:

$$\langle \ln N_{\rm gal}|\Delta\Sigma, M, z\rangle \;= \langle \ln N_{\rm gal,0}|\Delta\Sigma, M, z\rangle + \sigma_{\Delta\Sigma}\Big(\sum_i \frac{\beta_{N_{\rm gal},i}}{\sigma_{\Pi_i}}\rho_{\Delta\Sigma-\Pi_i}\times(\Pi_i - \langle\Pi_i|M,z\rangle)\Big), \tag{5.44}$$

and variance

$$\sigma^2_{\ln N_{\rm gal}|\Delta\Sigma} \;= \sigma^2_{N_{\rm gal,0}} + \sum_i \beta^2_{N_{\rm gal},i}\sigma^2_{\Pi_i}(1-\rho^2_{\Delta\Sigma-\Pi_i}) + \sum_{i,j}^{j\neq i}\rho_{\Pi_i-\Pi_j}\sigma_{\Pi_i}\sigma_{\Pi_j}. \tag{5.45}$$

We now want to derive the scaling relations for $P(\Delta\Sigma \mid N_{\rm gal}, M, z)$. From the Bayes theorem,

$$P(\Delta\Sigma \mid N_{\rm gal}, M, z) = P(\ln N_{\rm gal} \mid \Delta\Sigma, M, z)\frac{P(\Delta\Sigma \mid M, z)}{P(\ln N_{\rm gal} \mid M, z)}, \tag{5.46}$$

wherein we assume that $P(\ln N_{\rm gal}|M,z)$ and $P(\Delta\Sigma|M,z)$ can be modeled as normal distributions, then $P(\Delta\Sigma \mid N_{\rm gal}, M, z)$ is another normal distribution with mean

$$\langle\Delta\Sigma|N_{\rm gal}, M, z\rangle \;= \langle\Delta\Sigma|N_{\rm gal,0}, M, z\rangle + C_1\sigma_{\Delta\Sigma}\Big(\sum_i \frac{\beta_{N_{\rm gal},i}}{\sigma_{\Pi_i}}\rho_{\Delta\Sigma-\Pi_i}\times(\Pi_i - \langle\Pi_i|M,z\rangle)\Big) \tag{5.47}$$

and variance

$$\sigma^2_{\Delta\Sigma|\ln N_{\rm gal}} \;= \sigma^2_0 + C_2\sum_i \beta^2_{N_{\rm gal},i}\sigma^2_{\Pi_i}(1-\rho^2_{\Delta\Sigma-\Pi_i}) + C_3\sum_{i,j}^{j\neq i}\rho_{\Pi_i-\Pi_j}\sigma_{\Pi_i}\sigma_{\Pi_j}. \tag{5.48}$$

The parameter C_1 for the mean relation can be explicitly derived if we know the posterior distribution of $P(\ln N_{\rm gal}|M,z)$ and $P(\Delta\Sigma|M,z)$ by the exercise of completing the squares inside the exponents, i.e. by matching the quadratic, linear and constant terms inside the exponents of normal distributions on the left and right hand sides of Equation (5.46).

The parameters C_2, C_3, and σ_0 can be explicitly derived first by using the variance of product law of correlated Gaussians, i.e. $\mathrm{Var}(Z \equiv XY) = 1 + \rho^2$ after transforming X and Y into unit variance, zero mean Gaussians with correlation efficient ρ, and then by using the variance of quotient approximation

$$\mathrm{Var(X/Y)} = \big(\frac{\mu_{\rm X}}{\mu_{\rm Y}}\big)^2\Big[\frac{\sigma^2_{\rm R}}{\mu^2_{\rm R}} - 2\frac{\mathrm{Cov(X,Y)}}{\mu_{\rm X}\mu_{\rm Y}} + \frac{\sigma^2_{\rm X}}{\mu^2_{\rm Y}}\Big], \tag{5.49}$$

where μ_X, σ^2_X are the mean and variances of $P(\Delta\Sigma \mid M,z)$ and μ_Y, σ^2_Y the mean and variances of $P(\ln N_{\rm gal} \mid M,z)$ in our specific case.

The exact values of $C1$, $C2$, $C3$, and σ_0 are not essential for this paper as we aim to derive a general expression for $P(\Delta\Sigma \mid N_{\rm gal}, M, z)$ as a function of secondary halo parameters.

CHAPTER 6

TOTAL COVARIANCE AND IMPACT ON OPTICAL SCALING RELATIONS

6.1 Introduction

The previous chapter (Ch.5) measured the *intrinsic* covariance between cluster optical observables due to their formation history, parameterized by the secondary halo parameters such as the mass accretion rate and halo concentration.

Comparing those results with previous works might suggest an incompatibility in the sign of the covariance. Wu et al. (2022) measured the ratio of $\Delta\Sigma$ between an "observed" and "expected" sample and found a ratio above unity at large scales using a semi-realistic cluster finder. The study paints galaxies onto N-body dark matter halos using a semi-analytic model, and identifies as cluster members galaxies within a cylinder centered around the halo center. The height of the cylinder is extended to 60 Mpc to mimic projection effects of realistic cluster finders that misidentify galaxies along the line of sight not gravitational bound to the cluster. The "expected" sample conditions on the clusters true mass and samples their observed richness, taken as the number count of galaxies within the cylinder. The "observed" sample conditions on the cluster richness to match the "expected" data set. In the case that richness is a non-biased proxy of mass the ratios of lensing signals from these samples would be consistent with unity. A ratio above unity suggests a positive correlation between $\Delta\Sigma$ and λ when conditioned on the true mass.

Other studies have likewise found a positive bias. Sunayama et al. (2020), using a different simulation that paints galaxies using a halo occupation density (HOD) approach likewise found a positive bias between the "expected" and "observed" lensing signal at large scales when counting galaxies inside a cylinder. Comparing with Wu et al. (2022) the covariance they found at small scales is close to null and the amplitude of the bias at large scales is

around a factor of 2 smaller. Using real data collected by HSC, Huang et al. (2022) took the ratio of lensing signals between those selected by redMaPPer (Rykoff et al., 2014) and another proxy that measures the total stellar mass at an annulus of 50-100 kpc around the brightest central galaxy. The latter, because of its smaller radii and the steep fall off of the galaxy luminosity with radius, e.g as modeled by a Sersic profile (Sérsic, 1963), should not be contaminated by projection effects. Huang et al. (2022) found an increase in the lensing ratio at around 1 Mpc. These studies that take into account the projection effects of realistic clusters yield a positive sign. Different among them are the radial dependence and amplitude of the positive bias, which remains a open problem in the field.

The difference in signs between the work in the previous chapter and other studies that measure the bias using a (semi)-realistic cluster finder suggests that the total covariance is a sum of an *intrinsic* term due to the formation physics of cluster and *extrinsic* terms that take into account observational biases as projection effects, miscentering, and photometric redshift errors.

This work examines the impact of the total covariance and its impact on cluster scaling relations using a mock catalog that mimics realistic cluster finders for the Vera Rubin Legacy Survey of Space and Time (LSST) (Ivezić et al., 2019). It additionally explores the impact that photometric redshift errors has on the posterior distribution of the model parameters and compares the magnitude of observational systematics with different choices in modeling the concentration-mass relation.

LSST is a stage-IV photometric survey that will cover 18,000 square degrees of the sky and will reach redshifts up to $z = 3$. For cluster science it will detect over 1,000,000 group sized or more massive objects in its 10 year run. The number count of cluster sized objects is more than an order of magnitude more than Stage-III surveys as DES and the fainter magnitude limits will push our sample to lower mass and higher redshift thresholds, allowing us to construct a more complete and pure dataset. For this work we use the Dark Energy

Science Collaboration Data Challenge 2 (DESC DC2) mock catalogs (Abolfathi et al., 2021) accompanied with this survey to model the impact of observational systematics on cluster scaling relations.

6.2 Theoretical Framework

This section introduces the modeling of the mass-richness relation, the cluster abundance and the cluster lensing profiles.

6.2.1 Mass-richness relation

We forward model the conditional probability $P(\lambda|m,z)$ as

$$P(\lambda|m,z) \propto \frac{1}{\lambda}\exp\left\{-\frac{[\ln\lambda - \langle\ln\lambda|m,z\rangle]^2}{\sigma^2_{\ln\lambda|m,z}}\right\}, \tag{6.1}$$

in which the mean richness is modeled as

$$\langle\ln\lambda|m,z\rangle = \ln\lambda_0 + \mu_z\log\left(\frac{1+z}{1+z_0}\right) + \mu_m\log_{10}\left(\frac{m}{m_0}\right), \tag{6.2}$$

and the conditional scatter is given by

$$\sigma_{\ln\lambda|m,z} = \sigma_{\ln\lambda_0} + \sigma_z\log\left(\frac{1+z}{1+z_0}\right) + \sigma_m\log_{10}\left(\frac{m}{m_0}\right). \tag{6.3}$$

The parameters $\ln\lambda_0$, μ_z, μ_m, $\sigma_{\ln\lambda_0}$, σ_z and σ_m are taken as free parameters, and m_0 and z_0 are fixed pivot values.

6.2.2 Cluster number count

We will observe a set of cluster counts $\{\widehat{N}_{1\leq k\leq n_b}\}$ in n_b redshift-richness bins. For richness-detected clusters, the redshift-richness cluster number density is given by the following formula

$$\frac{d^2N(\lambda,z)}{dzd\lambda} = \Omega \int_{m_{\rm min}}^{+\infty} dm \ \frac{dn(m,z)}{dm}\frac{d^2V(z)}{dzd\Omega}\Phi(\lambda,m,z)P(\lambda|m,z), \tag{6.4}$$

where Ω is the survey sky area, $\Phi(\lambda,m,z)$ is the redMaPPer selection function and $P(\lambda|m,z)$ corresponds to the forward modeling of the mass-richness relation. The mass $m_{\rm min}$ correspond to the limit mass accessible in the survey, generally set by the selection function. The observable in the DC2 simulations is the cluster counts, so the predicted cluster count in the i-th richness and j-th redshift bin is given by

$$N_{ij} = \int_{z_i}^{z_{i+1}} dz \int_{\lambda_j}^{\lambda_{j+1}} d\lambda \frac{d^2N(\lambda,z)}{dzd\lambda}. \tag{6.5}$$

The selection function Φ associated to the cluster finder algorithm and survey strategy. It reflects the fact that the cluster finder algorithm may miss a non-negligible fraction of true clusters of galaxies that are in reality associated to halos, as well as detect "false" clusters that are not related to underlying collapsed dark matter structures. In this work, we follow the selection function parametrization from Aguena & Lima (2018). The selection function is then given by

$$\Phi(\lambda,m,z,z_{\rm obs}) = \frac{c(m,z)}{p(\lambda,z_{\rm obs})}, \tag{6.6}$$

where $c(m,z)$ is the completeness and $p(\lambda,z_{\rm obs})$ is the purity. The selection function usually accounts for the true halo redshift z and the observed redshift $z_{\rm obs}$ of the detected galaxy cluster, that may differ from the true halo redshift. Here, we consider $z = z_{\rm obs}$ for simplicity otherwise it should account for the conditional distribution $P(z_{\rm obs}|z)$. We see that completeness and purity answer two different questions about cluster detection: (i) contam-

ination: does the cluster finder miss some halos, and if yes how many? (ii) purity: does the cluster finder detect some "false" clusters, and if yes how many? The selection function can be accessed by running the cluster finder on simulations, by performing a geometrical match between the detected clusters and the true dark matter halos. It then allows to disentangle between the effects of purity and completeness. Here, because we are already using simulations, we can directly measure these quantities.

6.2.3 Cluster weak lensing

The matter content around the cluster arise from its single contribution (1-halo term), but also from the matter density from its neighboring halos (2-halo term). Thus, the full $\Delta\Sigma$ profile is given by the sum of two contributions $\Delta\Sigma(R) = \Delta\Sigma_{1h}(R) + \Delta\Sigma_{2h}(R)$. The 1-halo term $\Delta\Sigma_{1h}$ is computed using a given expression of single halo matter density field. The 2-halo term $\Delta\Sigma_{2h}(R)$ denotes the contribution of the surrounding halos to the dark matter density field, and writes (Oguri & Takada, 2011)

$$\Delta\Sigma_{2h}(R) = \frac{\rho_m(z) b(M,z)}{(1+z)^3 D_A(z)^2} \int \frac{ldl}{(2\pi)} P_{\rm mm}(k_l, z) J_2(l\theta), \tag{6.7}$$

where $P_{\rm mm}(k,z)$ is the linear matter power spectrum at the halo redshift z, θ is the separation angle given by $R/D_A(z)$ and J_2 is the second order Bessel function of the first kind. $\rho_m(z)$ is the physical matter density at redshift z, and $b(M,z)$ is the halo bias at mass M and redshift z. The scaled wave vector k_l is given by $k_l = l/(D_A(z)(1+z))$.

For a given sample of clusters within the i-redshift bin and the j-redshift bin, the total predicted excess surface density profile is given by

$$\begin{aligned}\Delta\Sigma_{ij} = \frac{1}{N_{ij}} \int_{z_i}^{z_{i+1}} dz \int_{\lambda_j}^{\lambda_{j+1}} d\lambda \int_{m_{\rm min}}^{+\infty} dm \times \\ \frac{d^2N(m,z)}{dzdm} c(m,z) P(\lambda|m,z) \Delta\Sigma(R|m,z).\end{aligned} \tag{6.8}$$

In this paper, we estimate the excess surface density signal around an ensemble of clusters rather than for individual clusters. This stacking strategy increases the signal-to-noise ratio and is particularly interesting for low mass clusters, for which the strength of gravitational lensing is weak. Moreover, the stacking strategy allows to average the shear signal over the intrinsic triaxiality of individual dark matter halos within the stack, and to recover the hypothesis of an effective spherical symmetry of the average cluster, since the halo properties are fitted using a spherical model.

The maximum likelihood estimator of $\Delta\Sigma$ considering a stack of N_l clusters where each one has $N_{ls}(R)$ background sources in the radial bin $[R, R + \Delta R]$ can be written (Shirasaki & Takada, 2018; Sheldon et al., 2004)

$$\widehat{\Delta\Sigma}_+(R) = \sum_{l=1}^{N_l} \sum_{s=1}^{N_{ls}(R)} \tilde{w}_{ls} \widehat{\Sigma_{\rm crit}}(z_s, z_l) \epsilon_+^{l,s} \tag{6.9}$$

where the sum runs over all lens-source (l, s) pairs, located at the physical projected radius interval $[R, R+\Delta R[$ from the lens l. Here, $\epsilon_+^{l,s} = -[\epsilon_1 \cos(2\phi_{\rm s}) + \epsilon_2 \sin(2\phi_{\rm s})]$ is the tangential ellipticity of the galaxy s relative to the lens l, and $\phi_{\rm s}$ is the polar angle of the galaxy relative to its lens. The quantity $\widehat{\Sigma_{\rm crit}}(z_s, z_l)$ is the effective critical surface mass density of the lens-source system, averaged over the photometric redshift probability density function $p(z_s)$ of the galaxy with index s, such as

$$\widehat{\Sigma_{\rm crit}}(z_s, z_l)^{-1} = \int_{z_l}^{+\infty} dz_s \; p(z_s) \Sigma_{\rm crit}(z_s, z_l)^{-1}. \tag{6.10}$$

The weights w_{ls} maximize the signal-to-noise ratio for this estimator (Sheldon et al., 2004)

and can be written as the product $w_{ls} = w_{ls}^{\rm geo} w_{ls}^{\rm shape}$ such as

$$w_{ls}^{\rm geo} = \langle \widehat{\Sigma_{\rm crit}}(z_s, z_l) \rangle^{-2}, \tag{6.11}$$

$$w_{ls}^{\rm shape} = \frac{1}{\sigma_{\rm rms}^2(\epsilon_s^+) + \sigma_{\rm meas}^2(\epsilon_s^+)}. \tag{6.12}$$

The quantity $\sigma_{\rm rms}(\epsilon_s^+)$ is the standard deviation of the tangential component of the ellipticity, whereas $\sigma_{\rm meas}(\epsilon_s^+)$ denotes the error on shape measurement. In the ideal case where galaxy redshifts and shapes are perfectly known, these weights reduce to $w_{ls}^{\rm geo} = (\Sigma_{\rm crit}(z_s, z_l))^{-2}$ and $w_{ls}^{\rm shape} = 1/\sigma_{\rm rms}^2(\epsilon_s^+)$. The weights $\tilde{w}_{ls}$ are defined with respect to w_{ls} such that they are normalized to 1 in each radial bin.

6.3 Covariance between Optical Observables

One relatively unexplored category of cluster systematics is the covariance between different cluster properties. In the regime of weak lensing for optically selected clusters, this would be the covariance between the two cluster observables — tangential shear and richness — left unconstrained in previous cluster cosmology analyses as McClintock et al. (2019); Abbott et al. (2020).

As first pointed out by Nord et al. (2008) and later shown in Evrard et al. (2014) and Farahi et al. (2018), the property covariance can induce additive biases that cannot be mitigated with increased sample size and reduced shape noise (Wu et al., 2019). To achieve percent-level mass calibration of clusters, this new category of systematics must be accurately and precisely quantified (Rozo et al., 2014). Here, we follow the prescription of Zhang et al. (2024) in modeling and quantifying the systematic uncertainty induced by property covariance of optically selected clusters.

We make a note of the difference between intrinsic and extrinsic covariance. In Zhang et al. (2024) the authors measured the intrinsic covariance of cluster weak lensing observables

by encircling galaxies within a 3D physical radius around the halo to measure the "true" richnesses and directly measured the integrated dark matter density around the halo. This "ground-truth" measured of the covariance is caused by the galaxy assembly bias, the degree of which can be quantified by secondary halo properties (e.g. concentration, mass accretion rate, kinetic-to-potential energy ratio) that relate to the formation history of the cluster-sized halo. In this work we measure the total covariance (intrinsic + extrinsic) using the mock redMaPPer cluster finder algorithm to determine the observed richness and the galaxy shear as a proxy for the surface density. This realistic mock catalog introduces observational systematic as blending of galaxies (Nourbakhsh et al., 2022), photometric redshift uncertainty (Graham et al., 2017) and shape noise (Wu et al., 2019) for the galaxy shear signal and projection and percolation effects (Costanzi et al., 2019), triaxiality bias (Zhang et al., 2023) and a set of other observational and modeling systematics (McClintock et al., 2019) related to a realistic cluster finder as redMaPPer.

Our goal is to model and corrected for the systematic uncertainty of $\langle \Delta\Sigma(R) \rangle_m$ selected in richness-redshift stacks of m as described in (6.9) and a log-richness-mass mean relation given by (6.2). Generally speaking, the observable-mass relation can be modeled according to Evrard et al. (2014) as

$$\boldsymbol{s} = \boldsymbol{\pi}(\boldsymbol{z}) + \boldsymbol{\alpha} \ln M, \tag{6.13}$$

where $\boldsymbol{s}$ is the vector of observables, $\boldsymbol{\pi}(\boldsymbol{z})$ a constant that captures the normalization and redshift dependence, and $\boldsymbol{\alpha}$ the partial slope with respect to the log-mass. In the case that $\boldsymbol{s}$ is reduced to $\ln \lambda$ given the parameterization in Eqn. (6.2) we have

$$\pi(z)_{\ln \lambda} = -\left(\mu_m \ln M_0 + \mu_z \ln \frac{1+z}{1+z_0} \right) \tag{6.14}$$

$$\alpha_{\ln \lambda} = \mu_m. \tag{6.15}$$

Similarly, we model $\langle \Delta\Sigma | M, z \rangle$ as a log-linear function of mass with a different different set

of coefficients $\pi(z)_{\Delta\Sigma}$ and $\alpha_{\Delta\Sigma}$. We model the correlated scatter between these observables as normally distributed with a covariance matrix

$$\mathbf{C} = \begin{pmatrix} \sigma^2_{\ln\lambda|M,z} & r\sigma_{\ln\lambda|M,z}\sigma_{\ln\gamma|M,z} \\ r\sigma_{\ln\lambda|M,z}\sigma_{\ln\gamma|M,z} & \sigma^2_{\ln\gamma|M,z}, \end{pmatrix}, \tag{6.16}$$

To obtain the correlation term to the observable scaling relation we expand the halo mass function (e.g. Tinker et al. (2008)) to first order:

$$\frac{\mathrm{d}n_{\mathrm{hmf}}(M,z)}{\mathrm{d}\ln M} \approx A(z)\exp\left[-\beta_1(M,z)\ln M\right]. \tag{6.17}$$

From Bayes theorem we can derive an observable-scaling relation $\langle\Delta\Sigma|\lambda,z\rangle$ by marginalizing over $n(\ln M)$ (see Evrard et al. (2014); Zhang et al. (2024) for derivation) to arrive at the expression

$$\langle\Delta\Sigma \mid \lambda,z\rangle_1 = \langle\Delta\Sigma \mid \lambda,z\rangle_0 + \frac{\beta_1}{\alpha_{\ln\lambda}} \times \mathrm{Cov}(\Delta\Sigma, \ln\lambda|M,z), \tag{6.18}$$

with $\alpha_{\ln\lambda} = \mu_m$ the partial slope the of the richness-mass relation with respect to log-mass.

Incidentally, the conditional richness scatter $\sigma^2_{\ln\lambda|M,z}$ is related to the conditional mass scatter $\sigma^2_{\ln M|\lambda,1}$ by the expression $\sigma^2_{\ln M|\lambda,1} = (\sigma_{\ln\lambda|M}/\alpha_{\ln\lambda})^2$ which can be derived from Bayes theorem when expanding the halo mass function to first order. We then have

$$\begin{aligned} \sigma^2_{\ln M|\lambda,1} &= \mu_m^{-2}\sigma^2_{\ln\lambda|M} \\ &= \mu_m^{-2}\left(\sigma_{\ln\lambda_0} + \sigma_z\log\left(\frac{1+z}{1+z_0}\right) + \sigma_m\log_{10}\left(\frac{m}{m_0}\right)\right)^2 \end{aligned} \tag{6.19}$$

where we made use of (6.3) for the expression for $\sigma^2_{\ln\lambda|M}$.

6.4 Dataset and Observables

The Data Challenge 2 (DC2), is a vast simulated astronomical dataset covering 440 $\deg^2$, that is designed to help develop and test the pipeline and analysis tools of DESC for interpreting the LSST data (see full details in Abolfathi et al. (2021)).

The workflow of the production in the Data Challenge 2 goes from the cosmological N-body simulations up to the processing of simulated images with the LSST Science pipelines. The production of this dataset represented a major effort of the DESC collaboration. Here, we will not enter into the details but simply give an overview of the main steps. The workflow is separated into four main components.

The first step contains the run of a cosmological N-body simulation, the modeling of the galaxy population and the generation of the cosmoDC2 extra-galactic catalog. The OuterRim N-body (gravity-only) simulation (Heitmann et al., 2019) is used as a starting point of the Data Challenge 2, where past light-cones have been created from simulation snapshots. The dark matter halos have then been identified by a friend-of-friend (FoF) halo finder with linking length $b = 0.168$. Each halo has been assigned with a mass $M_{\rm FoF}$ (the sum of the individual dark matter particles associated between them) and a spherical overdensity mass $M_{\rm 200c}$, that are provided in the cosmoDC2 extension called Skysim5000 (covering 5000 $\deg^2$, with an improved ray-tracing resolution by a factor of two in angular scale) obtained by fitting a NFW profile to each dark matter particle distribution. The correspondence between $M_{\rm FoF}$ and $M_{\rm 200c}$ for cosmoDC2 halos is explored in Kovacs et al. (2022). To maintain a high degree of physical realism, the several key properties of the galaxies were drawn from the Galacticus semi-analytic model of galaxy formation (Benson, 2012), and were painted onto dark matter halos using GalSampler (Hearin et al., 2020). The derived galaxy properties include stellar mass, morphology, spectral energy distributions, broadband filter magnitudes and host halo information. The weak lensing shears and convergences at each galaxy position were estimated by a ray-tracing algorithm applied to the past light-cone

particles in the simulation.

We now present the different catalogs that will be used in this paper, that differ by their level of complexity in accounting for observational effects or by taking the inventory of truth galaxies that are simulated on top of the N-body simulation. All the galaxies in these catalogs will serve as sources to estimate the lensing signal around lenses. Then we will present the lens catalogs, i.e., catalogs of galaxy clusters that were detected by different cluster finding algorithm.

6.4.1 Source galaxies: cosmoDC2 extra-galactic catalog

The cosmoDC2 extra-galactic catalog (detailed in Korytov et al. (2019), tested and validated in Kovacs et al. (2022)) contains $\sim$ 2.26 billions galaxies and takes the inventory of $\sim$ 550 properties of the "true" galaxies (i.e. true magnitudes in the six LSST bands, true redshift, true shapes, etc.) as well as the ray-tracing quantities per galaxy (shear and convergence) up to a magnitude depth of 28 in the r-band and to redshift $z \sim 3$. In that sense, it represents an ideal LSST dataset (there is no extinction from dust, stars, etc., only galaxies). A first level of complexity is added by computing the photometric redshifts of cosmoDC2 galaxies, using existing photoz codes. Two photometric redshift codes have been run and the corresponding estimated galaxy redshifts are stored in two add-on catalogs. The first is FlexZBoost[1] (Izbicki & Lee, 2017). FlexZBoost is an empirical technique that uses machine learning and learns the mapping between the galaxy's observed colors and the true cosmoDC2 redshift by using a training dataset (for real data, it is done by matching a spectroscopic reference sample with the photometric dataset, the former redshifts being much more precise). In cosmoDC2, FlexZBoost was trained with a complete subsample of galaxies that extended to $i < 25$. The second is a template-based algorithm called BPZ[2] (Bayesian Photometric Redshifts, Benítez

1. The code is available here `https://github.com/rizbicki/FlexCoDE`.

2. The code is available here `https://www.stsci.edu/~dcoe/BPZ/`.

(2011)). The BPZ method is a template-fitting technique that formulates a likelihood of the galaxy's observed colors from a set of Spectral Energy Distribution (SED) models.

6.4.2 Lenses: redMaPPer cluster catalog

Galaxy clusters are the observational counterpart of dark matter halos and can be identified by running cluster finder algorithms on the DC2 catalogs. The redMaPPer cluster finder (Rykoff et al., 2014), which identifies galaxy clusters through the presence of red-sequence galaxies, has been run on cosmoDC2. It has already been widely used on SDSS (Abdullah et al., 2020) and DES (Abbott et al., 2020) data and as such it is one of the best-studied cluster finders. The resulting cluster catalog provides the cluster positions, redshifts and richnesses (calculated as the sum of membership probabilities of galaxies around the cluster) along with the list of potential member galaxies. In addition to redMaPPer, other detection methods have been applied to find clusters of galaxies in the cosmoDC2 data (WaZP Aguena et al. (2021), AMICO Bellagamba et al. (2018)). A third catalog was built by geometrically matching the redMaPPer detected clusters to the cosmoDC2 dark matter halo catalog. Then, each redMaPPer cluster with richness λ_k is assigned with a "true" spherical overdensity mass M_k. To perform this match between the two catalogs, we used the DESC code ClEvaR[3] (Aguena et al., in prep.) that provides a user-friendly software to perform this task. The catalog was obtained considering $M_{200c} > 10^{13} M_\odot$ for the dark matter halo catalog and $\lambda > 5$ for the redMaPPer cluster catalog.

6.4.3 The cosmoDC2 ray-tracing resolution

This attenuation has already been observed in the galaxy-galaxy lensing in DC2 (Korytov et al., 2019) and is associated to the limited resolution of the ray tracing to compute the lensing shear and convergence at each galaxy positions, as a consequence, we chose to use

3. The code is publicly available here `https://github.com/LSSTDESC/ClEvaR`.

only the $R > 1$ Mpc region for each stack in the analysis. This is a conservative choice that appears valid over the full mass and redshift range. However, this also means we cannot use the innermost region that has the largest SNR values. DC2 is an impressive achievement for the collaboration but this intrinsic limitation will limit the cluster-related forecasts of what could be achieved with the LSST data. As we neglect the inner part of the profile (corresponding to $R < 1$ Mpc), we do not account for the effect of miscentering and non-weak lensing corrections (Mandelbaum et al., 2006).

6.5 Methodology

For the work presented in this section, we have considered the redshift bin edges [0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 1] and the richness bin edges [20, 35, 70, 100, 200]. For each cluster in the redMaPPer catalog, we extract the source catalog in a circular aperture of $R = 10$ using the cosmoDC2 extra-galactic catalog (version 1.1.4) f Mpc. We first apply a the cut $r < 28$ and $i < 25$ so that the effective number density of galaxies $n_{gal} \approx 25$ arcmin^{-2} to be comparable to that will be used in the context of LSST after 5 years of data (Chang et al., 2013). when considering true source redshift, background source selection is made by considering $z_z > z_l + 0.1$. When using photometric redshift, we use simultaneously two cuts baed on the photometric PDF, namely $\langle z \rangle_s > z_l + 1$ and $P(z_s > z_l) > 0.9$. For each stack of N_l clusters, we consider the 10 log-spaced radial bins from 0.5 Mpc to 10 Mpc, and we estimate the stacked lensing profile in Eq. (6.9). Similarly, we measure the abundance of redMaPPer clusters in each bins. We discuss in the new section how do we compute the errors on the cluster lensing profiles and cluster counts. To infer the scaling relation parameters from the cluster observables, we draw the posterior distribution given by the Bayes theorem

$$\mathcal{P}(\theta|\text{data}) = \frac{\mathcal{L}(\text{data}|\theta)|\pi(\theta)}{\mathcal{L}(\text{data})} \tag{6.20}$$

where we consider the joint likelihood

$$\mathcal{L}(\widehat{N}, \widehat{\Delta\Sigma}) = \mathcal{L}_{\rm N}(\widehat{N}|\vec{\theta})\mathcal{L}_{\rm WL}(\widehat{\Delta\Sigma}|\vec{\theta}) \tag{6.21}$$

For the cluster count likelihood $\mathcal{L}_{\rm N}$, we consider the multivariate Gaussian probability distribution with mean corresponding to the predicted cluster number counts, and a count covariance matrix. The covariance accounts for the standard Poisson shot noise, associated to the intrinsic variance of a counting experiment but also the Super-Sample Covariance (SSC), denoting the contribution from the fluctuation of the matter density within and beyond the survey volume (Hu & Kravtsov, 2003). The count covariance can be estimated via resampling techniques directly on the data (see e.g. Escoffier et al. (2016)), that allows to capture all the statistical features in the observable; However, since the DC2 cluster sample is relatively small, the noise in the measurement of the cluster count matrix is too important, especially due to the effects of super-sample covariance on cluster counts. We rather model the cluster count covariance as

$$\Sigma_{ij} = N + \langle b \rangle_i \langle b \rangle_j N_i N_j S_{ij} \tag{6.22}$$

where $\langle b \rangle_i$ is the average halo bias in the i-th redshift-richness bin, N_i is the predicted cluster count, and S_{ij} is the matrix of the amplitudes of matter fluctuation in the respective redshift bins (Lacasa et al., 2018).

$\mathcal{L}_{\rm WL}(\widehat{\Delta\Sigma}|\vec{\theta})$ denotes the stacked lensing profiles likelihood (refereed as $\mathcal{L}_{\rm WLp}$) as inspired from Park et al. (2023); Murata et al. (2019) and more recently in Sunayama et al. (2023). Generally, the cluster lensing likelihood is taken to be a multivariate Gaussian probability distribution. The covariance of lensing profiles originate from a variety of phenomena (Wu et al., 2019; Gruen et al., 2015; McClintock et al., 2019) such as the intrinsic scatter in the shape of background galaxies as well as the limited sample of clusters and background

galaxies as the main source of scatter in the measure of $\Delta\Sigma$. The more clusters there are in the stack, the more background galaxies are used, and the less the variance of the stacked profile will be. All the other contributions introduce more variance and correlations between two radial bins, that drop rapidly by approximately an order of magnitude for large-scale separations and slower for small-scale separations (Wu et al., 2019). First, contribution from the uncorrelated large-scale structures, i.e., random structures along the line-of-sight (in addition to the detected cluster) introduce a scatter in the measured convergence (then the shear) when it is measured in angular bins (see Hoekstra (2003)). Another contribution quantifies the impact of the stochastic variation of correlated halos around a cluster, associated to the variation of the 2-halo term, and depends on both matter power spectrum and halo bias. Another contribution is associated to the level of scattering of the lensing profile due to the variation in intrinsic halo properties, i.e. the variation of concentration at fixed mass, the halo ellipticity and orientation. It may also denote the scatter in individual masses of richness-selected clusters within the stack. We will consider bootstrap resampling as used in Simet et al. (2017) and we will keep only the diagonal terms since the off-diagonal terms are too noisy due to the low count statistics in each bins.

On the other side, we will also use the stacked lensing mass likelihood (refereed as $\mathcal{L}_{\rm WLm}$), by using inferred stacked masses from the stacked lensing profiles in each richness-redshift bins. Doing so, we can test the impact using a two-step procedure (i.e. masses then scaling parameters) for the inference of the scaling relation compared to using lensing profiles directly. This method was used in Abbott et al. (2020). For each stack, we will infer the corresponding *effective* mass of a halo whose lensing profile matches the stacked lensing profile in the redshift-richness bin. We will use a Gaussian likelihood with covariance obtained from bootstrap resampling on stacked lensing profiles. Then, we consider the stacked lensing masses and error-bars $\{\log_{10} M_{\rm WL} + \sigma(\log_{10} M_{\rm WL})\}_i$ (the error-bars are the dispersion of the mass posteriors) in each redshift-richness bin in a joint likelihood, with prediction $\log_{10} M$

where M is given by

$$M_{ij} = \frac{1}{N_{ij}} \int_{z_i}^{z_{i+1}} dz \int_{\lambda_j}^{\lambda_{j+1}} d\lambda \int_{m_{\rm min}}^{+\infty} dm \times \frac{d^2N(m,z)}{dzdm} c(m,z) P(\lambda|m,z) m. \tag{6.23}$$

Then, we combine the weak lensing likelihood (either $\mathcal{L}_{\rm WLm}$ or $\mathcal{L}_{\rm WLp}$) with the count likelihood, to constrain the scaling relation parameters.

For the estimation of the stacked lensing profiles, we use the CLMM[4] code that provides various tools for the halo modeling with respect to the spherical overdensity mass definition. We use the Core-Cosmology Library (CCL, Chisari et al. (2019)) for the prediction of the halo mass function, the partial comoving volume and the halo bias (in this work, we consider their implementation of the Tinker et al. (2010) halo bias). To draw the posterior, we use the implementation of the Monte Carlo Markov chains in the `emcee`(Foreman-Mackey et al., 2013). The combination of all these codes was performed in the DESC CLCosmoSim repository[5]. S_{ij} quantities are computed using the `PySSC` package (Lacasa et al., 2018; Gouyou Beauchamps et al., 2022). The code is available at `https://github.com/fabienlacasa/PySSC`.

6.6 Results

This section test the impact of the covariance between weak lensing observables $\Delta\Sigma$ and $\ln\lambda$ on the mass-observable scaling relation and inferred cosmology. While Zhang et al. (2024) observed a radially dependent covariance at small scales $R < R_{200c}$ due to cluster formation physics, because of the attenuation of the cluster lensing signal at scales due to the ray tracing resolution we focus on the impact of the covariance at radial scales of $R > 0.5$ Mpc/h.

4. The GitHub repository is available at `https://github.com/LSSTDESC/CLMM`

5. The GitHub repository is available at `https://github.com/LSSTDESC/CLCosmo_Sim`

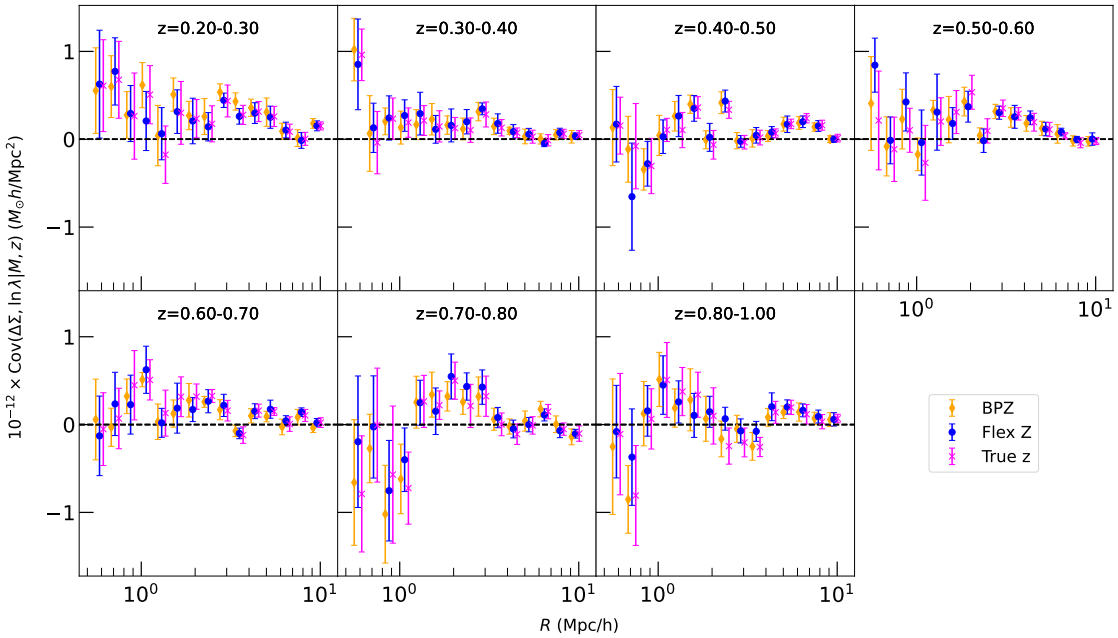


Figure 6.1: Measurements of covariance binned in redshifts and combined in mass bins. We perform the covariance measurements using true source galaxy redshifts marked in magenta, BPZ (Benítez, 2011) marked in yellow and FlexZBoost (Izbicki & Lee, 2017) marked in blue.

To effectively condition on the mass we apply a Kernel Local Linear Regression (KLLR) (Farahi et al., 2022a) to find the best fit local mean with mass as the independent variable and $\Delta\Sigma(R)$ and $\ln\lambda$ as the dependent variables. Specifically, the residuals of $\Delta\Sigma$ and $\ln\lambda$ are taken around their their local mean quantities $\langle\Delta\Sigma \mid M, z, R\rangle$ and $\langle\ln\lambda \mid M, z\rangle$

$$\mathrm{res}_{\Delta\Sigma}(M, z, R) = \Delta\Sigma - \langle\Delta\Sigma \mid M, z, R\rangle, \tag{6.24}$$

$$\mathrm{res}_{\ln\lambda}(M, z) = \ln\lambda - \langle\ln\lambda \mid M, z\rangle \tag{6.25}$$

and the covariance is measured around the local residuals, i.e.

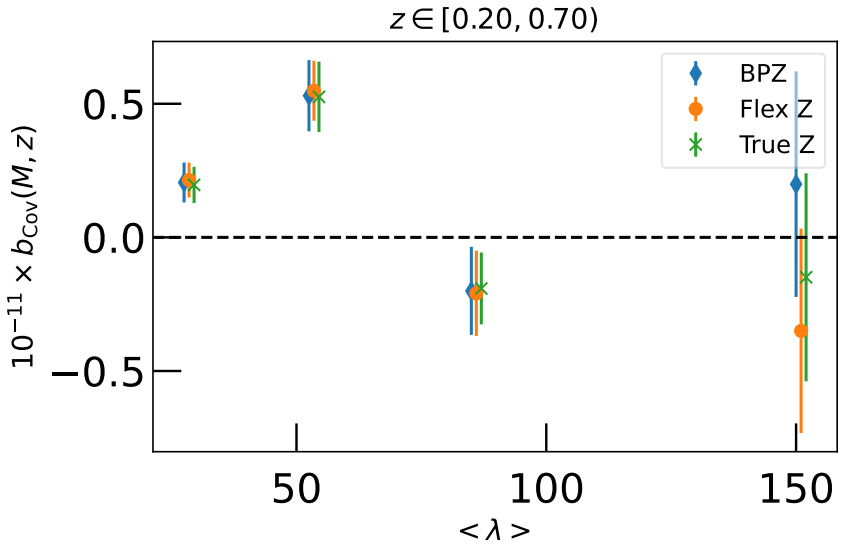
$$\mathrm{Cov}_{\Delta\Sigma,\ln\lambda}(M, z, R) = \mathrm{Cov}(\mathrm{res}_{\Delta\Sigma}(M, z, R), \mathrm{res}_{\ln\lambda}(M, z)). \tag{6.26}$$

It is important to note that as we remove the residual dependence of the mean mass on the covariance, the size of the covariance itself can still nonetheless be a function of (M,z),

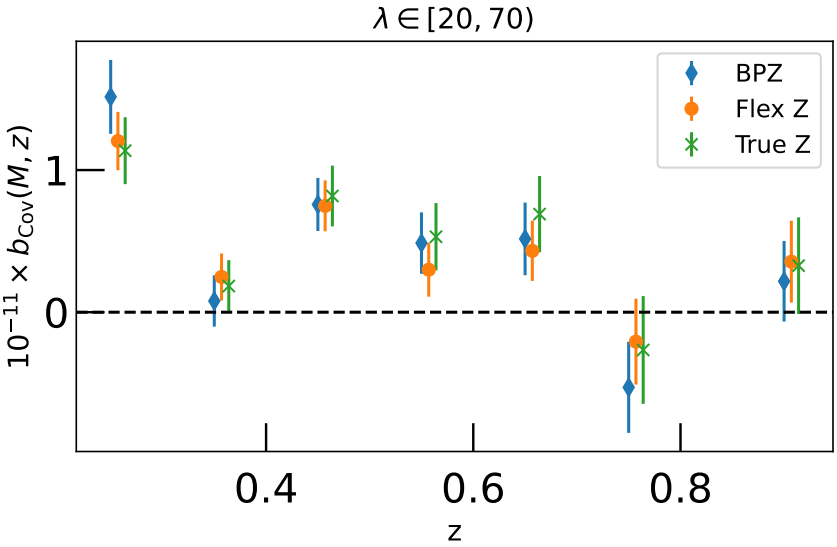
i.e. the scatter and correlation are modeled as $\sigma_{\ln\lambda|M,z}(M,z)$ and $\sigma_{\Delta\Sigma|M,z}(M,z)$ and $r_{\ln\lambda-\Delta\Sigma|M,z}(M,z)$. For this reason we test the covariance on small enough (M,z) bins that the scatter can be considered *homoskedastic*, or in order words approximated as a constant.

In our benchmark tests we choose to bin the covariance either in mass or redshift but not by both so as to retain enough statistical constraint especially for high mass, low redshift bins with a low cluster count. We plot the covariance merged by mass and binned in redshift in Fig. 6.1. Visually we recognize a slight positive covariance in most bins. The same trend (not shown) can be said of a benchmark test for the covariance binned by richness and merged between $z \in [0.2, 0.7)$ in which we discard the redshifts between $z \in [0.7, 1.0)$ as it can be shown that the covariances at these redshifts are null. To explicitly demonstrate the mass and redshift dependence we model the covariance as a constant bias across radii which we denote $b_{\rm Cov}(M,z)$. The covariance trend with respect to mass and redshift is plotted in Fig. 6.2. We demonstrate that in lower richness and redshift bins a slightly positive covariance between cluster observables at large radial smalls, consistent with expectations from projection effects (Sunayama et al., 2020; Wu et al., 2022; Huang et al., 2022).

The impact that the covariance has on the posterior constraints on the scaling relations can be seen in Fig. 6.3. The fiducial scaling relations in the absence of the covariance term is given by Eq. (6.2) & (6.3) using the true M_{200c} of halos. Adding different levels of complexity, we first introduce the addition of the covariance term (Eq. (6.18)) using the true source galaxy redshift, and then test the effects of including both the covariance term and photometric redshift errors using the BPZ (Benítez, 2011) and FlexZBoost (Izbicki & Lee, 2017) models. We then compare the impact that the covariance term has with different concentration-mass relations that do not include the covariance term. These include models from Diemer & Kravtsov (2014); Bhattacharya et al. (2013); Prada et al. (2012); Duffy et al. (2008); Bocquet et al. (2016). We find that including the covariance can shift the posterior constraints of scaling relation parameters by as much as $\sim 1\sigma$ which is at



(a) Covariance conditioned on richnesses, merged in redshift.



(b) Covariance conditioned on redshifts, merged in richnesses.

Figure 6.2: The covariance (Eq. (6.26)) modeled as a constant bias term $b_{\rm Cov}(M, z)$ across radius when binned by richness or redshift. The positive covariance at large scales at lower redshift and richness bins is consistent with expectations from projection effects.

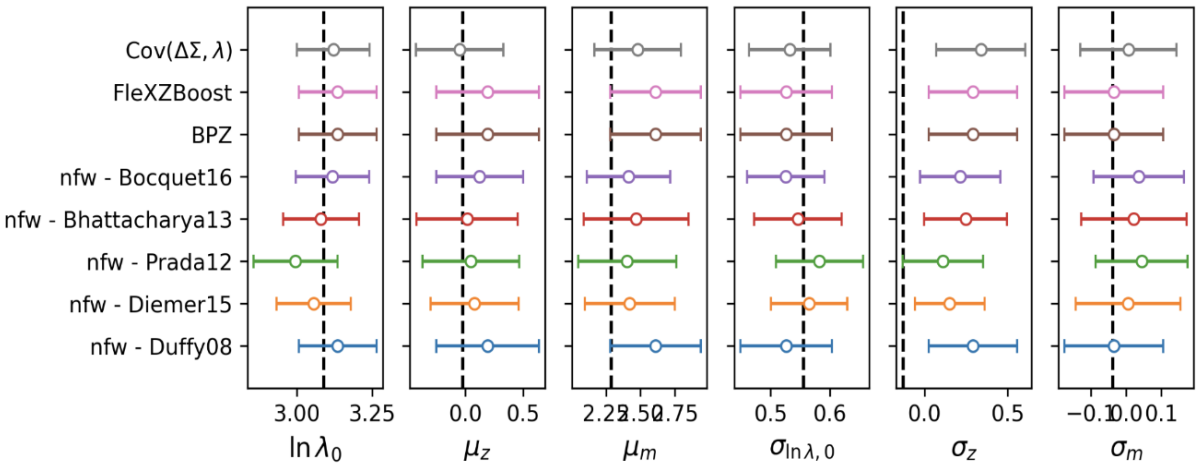


Figure 6.3: Impact of modeling choices on scaling relations. The vertical black dashed lines represent the best-fit parameters using fiducial scaling relations in the absence of the covariance term and using the true masses of halos, given given by Eq. (6.2) & (6.3). Adding different levels of complexity, in the first column we first introduce the addition of the covariance term (Eq. (6.18)) using the true source galaxy redshift, and then test the effects of including both the covariance term and photometric redshift errors using the FlexZBoost (Izbicki & Lee, 2017) and BPZ (Benítez, 2011) and models in respectively the second and third columns. In the remaining columns we compare the impact that the covariance term has with different concentration-mass relations that do not include the covariance term. These include models from Diemer & Kravtsov (2014); Bhattacharya et al. (2013); Prada et al. (2012); Duffy et al. (2008); Bocquet et al. (2016). We find that introducing the covariance term can have a $\sim 1\sigma$ on posterior constraints, at the same level of different choices of concentration-mass relations, and that the photometric redshift error for source galaxies makes a minimal impact on the scaling relations.

the same level of different choices of concentration-mass relations. We also observe that introducing photometric redshift uncertainty to the source galaxy redshift minimally impacts the posterior constraints.

CHAPTER 7

FINAL REMARKS

This thesis work touches on multiple aspects of cluster cosmology in the optical regime. We quantified several important systematics and explained their physical origins. The results of these work not only are useful in understanding the formation, properties and correlations between properties of clusters but the templates provided has been or will be directly applied to current and next generation cluster surveys to achieve a $< 1\%$ accuracy in cluster mass estimates. We summarize the main results of this thesis as follows:

- We demonstrate that cluster triaxiality can both impact the cluster lensing and richness-mass relation. This bias is propagated as a $1 - 5\%$ mass bias in the stacked lensing signal. We find that cluster triaxiality and projection effects resolve the tension between DES Y1 inferred weak lensing mass at richnesses of $\lambda > 30$ but does not resolve the "lensing is low" problem for $\lambda \in [20, 30)$.

- We measure the correlation between the strength of triaxiality bias and those of two other major systematic biases — miscentering and projection effects. Using multiple proxies to characterize the strength of these systematics we find a null correlation in all cases. This result allows us to construct triaxiality as an independent systematic. We offer explanations as to why there is no correlation.

- We quantify the intrinsic covariance between the cluster lensing signal and "true" richness and find a negative covariance at small radial scales. We show that from a negative log-linear relation between the scale of the covariance and the cluster peak height that the intrinsic covariance is related to the formation history of the cluster.

- We explicitly demonstrate the physical origin of the intrinsic covariance by modeling $\Delta\Sigma$ and $N_{\rm gal}$ as multi-linear relations with respect to secondary halo parameters re-

lating to the cluster's assembly history. Our model shows that the secondary halo parameters fully account for the negative covariance at large scales.

- We model the total covariance of cluster optical observables in a LSST mock catalog that includes systematics as projection effects. We find a slight positive covariance at large radial scales consistent with expectations from projection effects. The results of this work will be folded into the LSST DESC cluster cosmology pipeline.

REFERENCES

Abbott T., et al., 2018, Physical Review D, 98

Abbott T., et al., 2020, Physical Review D, 102

Abbott T. M. C., et al., 2023, Open J. Astrophys., 6, 2305.17173

Abdalla E., et al., 2022, Journal of High Energy Astrophysics, 34, 49–211

Abdullah M. H., Klypin A., Wilson G., 2020, , 901, 90

Abolfathi B., et al., 2021, , 253, 31

Aguena M., Lima M., 2018, , 98, 123529

Aguena M., et al., 2021, , 502, 4435

Akita K., Yamaguchi M., 2020, Journal of Cosmology and Astroparticle Physics, 2020, 012–012

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

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

Anbajagane D., Evrard A. E., Farahi A., Barnes D. J., Dolag K., McCarthy I. G., Nelson D., Pillepich A., 2020, , 495, 686

Anbajagane D., Evrard A. E., Farahi A., 2022, , 509, 3441

Applegate D. E., et al., 2014, , 439, 48

Baldry I. K., et al., 2012, Monthly Notices of the Royal Astronomical Society, pp no–no

Ballardini M., Braglia M., Finelli F., Paoletti D., Starobinsky A. A., Umiltà C., 2020, Journal of Cosmology and Astroparticle Physics, 2020, 044–044

Bartelmann M., King L. J., Schneider P., 2001, , 378, 361

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

Bellagamba F., Roncarelli M., Maturi M., Moscardini L., 2018, , 473, 5221

Benítez N., 2011, ASCL, p. ascl:1108.011

Benson A. J., 2012, , 17, 175

Bett P., 2012, , 420, 3303

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

Bhattacharya S., Habib S., Heitmann K., Vikhlinin A., 2013, , 766, 32

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

Blanchard A., Wachter K., Evrard A. E., Silk J., 1992, , 391, 1

Blas D., Lesgourgues J., Tram T., 2011, , 7, 034

Bleem L. E., et al., 2020, , 247, 25

Bocquet S., Carter F. W., 2016, The Journal of Open Source Software, 1

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

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

Bose B., Lombriser L., 2021, Physical Review D, 103

Breusch T., Pagan A., 1979, Econometrica, 47, 1287

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

Chang C., et al., 2013, , 434, 2121–2135

Chiang C.-T., Slosar A., 2018, Inferences of H_0 in presence of a non-standard recombination (arXiv:1811.03624)

Chisari N. E., et al., 2019, , 242, 2

Chiu I. N., Umetsu K., Murata R., Medezinski E., Oguri M., 2020, , 495, 428

Clampitt J., Jain B., 2016, , 457, 4135

Cooke K. C., Kartaltepe J., Tyler K., Darvish B., Fogarty K., O'Dea C. P., 2019, in American Astronomical Society Meeting Abstracts #233. p. 176.05

Corless V. L., King L. J., 2008, , 390, 997

Costanzi M., et al., 2019, , 482, 490

Costanzi M., et al., 2021, , 103, 043522

Croton D. J., et al., 2016, , 222, 22

Dark Energy Survey Collaboration et al., 2016, , 460, 1270

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

DeRose J., et al., 2019, The Buzzard Flock: Dark Energy Survey Synthetic Sky Catalogs (arXiv:1901.02401)

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

Di Valentino E., et al., 2021, Classical and Quantum Gravity, 38, 153001

Diemer B., Joyce M., 2019, The Astrophysical Journal, 871, 168

Diemer B., Kravtsov A. V., 2014, ApJ, 789, 1

Dietrich J. P., et al., 2014, , 443, 1713

Dodelson S., Heitmann K., Hirata C., Honscheid K., Roodman A., Seljak U., Slosar A., Trodden M., 2016, arXiv e-prints, p. arXiv:1604.07626

Drlica-Wagner A., et al., 2018, The Astrophysical Journal Supplement Series, 235, 33

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

Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, , 390, L64

Efstathiou G., 2003, Monthly Notices of the Royal Astronomical Society, 346, L26–L30

Efstathiou G., Sutherland W. J., Maddox S. J., 1990, , 348, 705

Escoffier S., et al., 2016, arXiv e-prints, p. arXiv:1606.00233

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

Evrard A. E., et al., 2002, The Astrophysical Journal, 573, 7–36

Evrard A. E., Arnault P., Huterer D., Farahi A., 2014, , 441, 3562

Farahi A., Evrard A. E., Rozo E., Rykoff E. S., Wechsler R. H., 2016, , 460, 3900

Farahi A., Evrard A. E., McCarthy I., Barnes D. J., Kay S. T., 2018, , 478, 2618

Farahi A., et al., 2019a, Nature Communications, 10, 2504

Farahi A., et al., 2019b, , 490, 3341

Farahi A., Ho M., Trac H., 2020, , 493, 1361

Farahi A., Anbajagane D., Evrard A. E., 2022a, The Astrophysical Journal, 931, 166

Farahi A., Nagai D., Anbajagane D., 2022b, The Astrophysical Journal, 933, 48

Flaugher B., et al., 2015, , 150, 150

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, , 125, 306

Forero-Romero J. E., Contreras S., Padilla N., 2014, , 443, 1090

Freedman W. L., 2021, The Astrophysical Journal, 919, 16

Gelman A., Rubin D. B., 1992, Statistical Science, 7, 457

Ghirardini V., et al., 2024, The SRG/eROSITA All-Sky Survey: Cosmology Constraints from Cluster Abundances in the Western Galactic Hemisphere (`arXiv:2402.08458`)

Giodini S., Lovisari L., Pointecouteau E., Ettori S., Reiprich T. H., Hoekstra H., 2013, Space Science Reviews, 177, 247

Golden-Marx J. B., Miller C. J., 2018, ApJ, 860, 2

Goodman J., Weare J., 2010, Communications in Applied Mathematics and Computational Science, 5, 65

Gouyou Beauchamps S., Lacasa F., Tutusaus I., Aubert M., Baratta P., Gorce A., Sakr Z., 2022, , 659, A128

Graham M. L., Connolly A. J., Ivezić Z., Schmidt S. J., Jones R. L., Jurić M., Daniel S. F., Yoachim P., 2017, The Astronomical Journal, 155, 1

Gruen D., et al., 2014, , 442, 1507

Gruen D., Seitz S., Becker M. R., Friedrich O., Mana A., 2015, , 449, 4264

Guo R.-Y., Zhang J.-F., Zhang X., 2019, Journal of Cosmology and Astroparticle Physics, 2019, 054–054

Guth A. H., 1981, Phys. Rev. D, 23, 347

Hahn O., Carollo C. M., Porciani C., Dekel A., 2007, , 381, 41

Harrison E. R., 1970, Phys. Rev. D, 1, 2726

Hart L., Chluba J., 2017, Monthly Notices of the Royal Astronomical Society, 474, 1850–1861

Hayashi E., White S. D. M., 2008, , 388, 2

Hearin A. P., Zentner A. R., Newman J. A., Berlind A. A., 2013, MNRAS, 430, 1238

Hearin A., Korytov D., Kovacs E., Benson A., Aung H., Bradshaw C., Campbell D., LSST Dark Energy Science Collaboration 2020, , 495, 5040

Heitmann K., et al., 2019, , 245, 16

Hoekstra H., 2003, , 339, 1155

Hofmann F., et al., 2017, , 606, A118

Hu W., Kravtsov A. V., 2003, The Astrophysical Journal, 584, 702–715

Huang S., et al., 2022, MNRAS, 515, 4722

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Huterer D., 2023, The Astronomy and Astrophysics Review, 31

Ivezić Ž., et al., 2019, , 873, 111

Izbicki R., Lee A. B., 2017, arXiv e-prints, p. arXiv:1704.08095

Jedamzik K., Pogosian L., 2020, Physical Review Letters, 125

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

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

Johnston D. E., et al., 2007, preprint, (arXiv:0709.1159)

Kaiser N., 1986, , 222, 323

Karwal T., Kamionkowski M., 2016, Physical Review D, 94

Kasun S. F., Evrard A. E., 2005, , 629, 781

Katz N., 1991, , 368, 325

Kepner J., Kim R., 2000, ArXiv Astrophysics e-prints,

Kettula K., et al., 2015, , 451, 1460

Kiiveri K., et al., 2021, , 502, 1494

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

Kofman L. A., Gnedin N. Y., Bahcall N. A., 1993, , 413, 1

Korytov D., et al., 2019, , 245, 26

Kovacs E., et al., 2022, OJAp, 5, 1

Krause E., et al., 2017

Kravtsov A. V., Borgani S., 2012, , 50, 353

Lacasa F., Lima M., Aguena M., 2018, , 611, A83

Lau E. T., Nagai D., Avestruz C., Nelson K., Vikhlinin A., 2015, , 806, 68

Lauer T. R., Postman M., Strauss M. A., Graves G. J., Chisari N. E., 2014, , 797, 82

Lee W., et al., 2023, , 945, 71

Lesci G. F., et al., 2022, , 665, A100

Lesgourgues J., 2011, preprint, (arXiv:1104.2932)

Lin Y.-T., Mohr J. J., 2003, , 582, 574

Lin M.-X., Raveri M., Hu W., 2019, Physical Review D, 99

Lithwick Y., Dalal N., 2011, The Astrophysical Journal, 734, 100

Lokken M., et al., 2022, ApJ, 933, 134

Lu Y., Mo H. J., Katz N., Weinberg M. D., 2006, MNRAS, 368, 1931

Ludlow A. D., Navarro J. F., Li M., Angulo R. E., Boylan-Kolchin M., Bett P. E., 2012, MNRAS, 427, 1322

Ludlow A. D., Bose S., Angulo R. E., Wang L., Hellwing W. A., Navarro J. F., Cole S., Frenk C. S., 2016, MNRAS, 460, 1214

Lukić Z., Heitmann K., Habib S., Bashinsky S., Ricker P. M., 2007, , 671, 1160

MacCrann N., et al., 2018, preprint, (arXiv:1803.09795)

Mahdavi A., Hoekstra H., Babul A., Bildfell C., Jeltema T., Henry J. P., 2013, , 767, 116

Mandelbaum R., 2018, Annual Review of Astronomy and Astrophysics, 56, 393–433

Mandelbaum R., Seljak U., Cool R. J., Blanton M., Hirata C. M., Brinkmann J., 2006, , 372, 758

Mandelbaum R., Seljak U., Hirata C. M., 2008, , 2008, 006

Mangano G., Miele G., Pastor S., Pinto T., Pisanti O., Serpico P. D., 2005, Nuclear Physics B, 729, 221–234

Mantz A., Allen S. W., Ebeling H., Rapetti D., Drlica-Wagner A., 2010, , 406, 1773

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

Mantz A. B., Allen S. W., Morris R. G., Schmidt R. W., 2016a, , 456, 4020

Mantz A. B., et al., 2016b, , 463, 3582

Maraston C., et al., 2013, MNRAS, 435, 2764

McClintock T., et al., 2019, , 482, 1352

Melchior P., et al., 2017, , 469, 4899

Miyatake H., et al., 2022, , 106, 083519

Mulroy S. L., et al., 2019, , 484, 60

Murata R., et al., 2019, , 71, 107

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

Nishimichi T., et al., 2019, , 884, 29

Nord B., Stanek R., Rasia E., Evrard A. E., 2008, , 383, L10

Nourbakhsh E., Tyson J. A., Schmidt S. J., Armstrong B., Burchat P., Sánchez J., 2022, Monthly Notices of the Royal Astronomical Society, 514, 5905–5926

Oguri M., 2014, , 444, 147

Oguri M., Takada M., 2011, , 83

Oguri M., Takada M., Umetsu K., Broadhurst T., 2005, , 632, 841

Oguri M., Bayliss M. B., Dahle H., Sharon K., Gladders M. D., Natarajan P., Hennawi J. F., Koester B. P., 2012, Monthly Notices of the Royal Astronomical Society, 420, 3213

Osato K., Nishimichi T., Oguri M., Takada M., Okumura T., 2018, , 477, 2141

Park Y., Sunayama T., Takada M., Kobayashi Y., Miyatake H., More S., Nishimichi T., Sugiyama S., 2023, , 518, 5171

Peebles P. J. E., 1982, , 263, L1

Peebles P. J. E., Yu J. T., 1970, , 162, 815

Percival W. J., 2005, , 443, 819

Perlmutter S., et al., 1999, The Astrophysical Journal, 517, 565–586

Pierre M., et al., 2016, , 592, A1

Pillepich A., Porciani C., Hahn O., 2010, , 402, 191

Planck Collaboration et al., 2014, , 571, A16

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

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

Pratt G. W., Arnaud M., Biviano A., Eckert D., Ettori S., Nagai D., Okabe N., Reiprich T. H., 2019, , 215, 25

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

Ragone-Figueroa C., Granato G. L., Borgani S., De Propris R., García Lambas D., Murante G., Rasia E., West M., 2020, , 495, 2436

Reid M. J., Pesce D. W., Riess A. G., 2019, The Astrophysical Journal Letters, 886, L27

Riebe K., et al., 2013, Astronomische Nachrichten, 334, 691

Riess A. G., et al., 1998, The Astronomical Journal, 116, 1009–1038

Rossi M., Ballardini M., Braglia M., Finelli F., Paoletti D., Starobinsky A. A., Umiltà C., 2019, Phys. Rev. D, 100, 103524

Rozo E., et al., 2010, , 708, 645

Rozo E., Evrard A. E., Rykoff E. S., Bartlett J. G., 2014, , 438, 62

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

Rykoff E. S., et al., 2012, , 746, 178

Rykoff E. S., et al., 2014, ApJ, 785, 104

Rykoff E. S., et al., 2016, , 224, 1

Salcedo A. N., Wibking B. D., Weinberg D. H., Wu H.-Y., Ferrer D., Eisenstein D., Pinto P., 2020, , 491, 3061

Saro A., et al., 2015, , 454, 2305

Schrabback T., et al., 2018, , 474, 2635

Sereno M., et al., 2020, , 492, 4528

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

Shapiro S. S., Wilk M. B., 1965, Biometrika, 52, 591

Sheldon E. S., et al., 2004, , 127, 2544

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

Shin T., Diemer B., 2023, MNRAS, 521, 5570

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

Shirasaki M., Takada M., 2018, , 478, 4277

Simet M., McClintock T., Mandelbaum R., Rozo E., Rykoff E., Sheldon E., Wechsler R. H., 2017, , 466, 3103

Soares-Santos M., et al., 2011, , 727, 45

Somerville R. S., Primack J. R., Faber S. M., 2001, , 320, 504

Song J., et al., 2012, , 761, 22

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

Stanek R., Rasia E., Evrard A. E., Pearce F., Gazzola L., 2010, , 715, 1508

Stark D. V., McGaugh S. S., Swaters R. A., 2009, , 138, 392

Stott J. P., et al., 2012, , 422, 2213

Sunayama T., 2023, MNRAS, 521, 5064

Sunayama T., et al., 2020, MNRAS, 496, 4468

Sunayama T., et al., 2023, arXiv e-prints, p. arXiv:2309.13025

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

Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, , 724, 878

To C., et al., 2021a, , 126, 141301

To C.-H., et al., 2021b, , 502, 4093

Tremonti C. A., et al., 2004, ApJ, 613, 898

Troxel M., et al., 2018, Physical Review D, 98

Valentino E. D., Melchiorri A., Silk J., 2020, Journal of Cosmology and Astroparticle Physics, 2020, 013–013

Vikhlinin A., et al., 2009, , 692, 1060

Wang K., Mao Y.-Y., Zentner A. R., Guo H., Lange J. U., van den Bosch F. C., Mezini L., 2022, MNRAS, 516, 4003

Warren M. S., Quinn P. J., Salmon J. K., Zurek W. H., 1992, , 399, 405

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

Wechsler R. H., Tinker J. L., 2018, , 56, 435

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

Wechsler R. H., DeRose J., Busha M. T., Becker M. R., Rykoff E., Evrard A., 2021, arXiv e-prints, p. arXiv:2105.12105

Weinberg S., 1989, Rev. Mod. Phys., 61, 1

Wen Z. L., Han J. L., Liu F. S., 2012, , 199, 34

Wong K. C., et al., 2019, Monthly Notices of the Royal Astronomical Society, 498, 1420–1439

Wu H.-Y., Evrard A. E., Hahn O., Martizzi D., Teyssier R., Wechsler R. H., 2015, , 452, 1982

Wu H.-Y., Weinberg D. H., Salcedo A. N., Wibking B. D., Zu Y., 2019, , 490, 2606

Wu H.-Y., et al., 2022, , 515, 4471

Xhakaj E., Leauthaud A., Lange J., Hearin A., Diemer B., Dalal N., 2022, , 514, 2876

Zeldovich Y. B., 1972, , 160, 1P

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

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

Zentner A. R., Hearin A., van den Bosch F. C., Lange J. U., Villarreal A. S., 2019, MNRAS, 485, 1196

Zhang Y., Annis J., 2022, , 511, L30

Zhang Y., et al., 2019, , 487, 2578

Zhang Z., et al., 2023, Monthly Notices of the Royal Astronomical Society, 523, 1994–2013

Zhang Z., et al., 2024, Monthly Notices of the Royal Astronomical Society, 530, 3127

Zhao D. H., Jing Y. P., Mo H. J., Brner G., 2003, ApJ, 597, L9

Zu Y., Weinberg D. H., Rozo E., Sheldon E. S., Tinker J. L., Becker M. R., 2014, , 439, 1628

Zuntz J., et al., 2015, Astronomy and Computing, 12, 45

Zwicky F., 1937, , 86, 217

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

de Salas P. F., Pastor S., 2016, Journal of Cosmology and Astroparticle Physics, 2016, 051–051

von der Linden A., et al., 2014, , 439, 2