Multi-Wavelength Facets of Galaxy Clusters

by

Arya Farahi

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics and Scientific Computing)
in The University of Michigan
2018

Doctoral Committee:

Professor August E. Evrard, Chair
Assistant Professor Emanuel Gull
Professor Dragan Huterer
Associate Professor Christopher Miller

Arya Farahi

aryaf@umich.edu

ORCID iD: 0000-0003-0777-4618

ACKNOWLEDGEMENTS

I am incredibly grateful to everyone who has made this exceptional journey possible through their mentorship, support, friendship, and love. First and foremost, my sincerest thanks are expressed to August Evrard, who has been a valued advisor, a close collaborator, and a good friend throughout the last six years. I would like to acknowledge the major role you played in this research and giving me freedom to pursue many research directions, in addition to being the best advisor I could possibly have hoped for. The breadth of your knowledge was a source of motivation for me to work harder, and furthermore you taught me how to become an original thinker and scientist. You inspire me with genuine enthusiasm and dedication, intellectual support, and confidence in my work. I hope to carry those lessons with me throughout my career and life.

No achievement is possible without support and great mentors. I want to thank all the people whom I had the privilege to collaborate with and mentored by, especially Andrew Benson, Leopoldo Pando Zayas, Christopher Miller, Elena Rasia, Kathy Romer, Jacob Abernethy, Eric Schwartz, and Danai Koutra. I truly appreciate your effort and thank you so much for being an excellent mentor and a great inspiration. I appreciate all the great feedback during my graduate career. I would especially like to thank Dragan Huterer for your support, guidance, discussion, ideas, and feedback which all have been absolutely invaluable. I have sincerely enjoyed the time that I have spent talking with you. Your feedback and criticism profoundly improved the quality of my work and have had an impact on my professional viewpoints and

development.

I would like to thank Christopher Miller and Elena Rasia for funding my research each for one semester. Additionally, I would like to thank Dragan Huterer, Christopher Miller, and Emanuel Gull for serving on my thesis committee. I would like to especially thank my classmates and friends, Anthony Kremin, Jessica Muir, Noah Weaverdyck, Hui Li, Jonathan Stroud, Yuanyuan Zhang, and Brian Nord, who made my time at the University of Michigan unforgettable.

I am highly grateful and indebted to Andrew Benson, without whom starting this journey was a mere dream. Thank you for replying to a random email you received and playing a wild card. I never forget the first question that I asked, "what does redshift mean?." Now that I am thinking of the past, I am surprised that you patiently replied to my email and did not ask me to take astronomy 101. You taught me not to judge people on the first impression and be patient, which I hope to carry with me throughout my life.

A huge thanks to Leo Pando Zayas for giving me an opportunity to learn about strings world and teaching me all I know about the string theory. Since my undergraduate I was wondering if there is a way to use computational model to study a string theory model. You made this wonder happen, which led to the first empirical evidence for the last conjecture that Stephen Hawking came up with. Leo, I hope to carry on the enthusiasm and passion that you have for research.

I am inspired by Jacob Abernethy, who shaped the way I think about outreach and education by introducing the Michigan Data Science Team (MDST). Your enthusiasm, hard work, and grit shaped the team and made it a great success. Thank you Jake and Eric for your dedication, support, and being such a great faculty advisors for the team. Moreover, I am grateful of all members who twice elected me

to serve them on the leadership board of MDST, it was a real privilege and honor to serve the MDST community. I am so proud of MDST's accomplishments in the past two years, which was not possible without contribution of every single member of the team, especially Jonathan Stroud. I wish the team a continuing success in the coming years. Data science for social good is something that this world needs more. I also wish to thank the Michigan Institute for Data Science (MIDAS) for their generous support and their belief in MDST's mission.

Last, but certainly not least, I want to thank my friends and family both here and afar for all their love over these past six years. I need to thank Sussan Bano and Kambouzia Farahi, my parents, for constantly believing in me and raising me to be the man I am today. Finally, I must express my gratitude to Negar Farzaneh, my wife, for her continued support and encouragement. Most of all, thank you for coming with me halfway across the world for more than half a decade that I can study the night sky. I was continually amazed by your willingness to read and correct countless pages of my emails, proposals, and talk scripts, and by your patience during all of the ups and downs of my research and life. I hope I continue to make you all proud.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

Cluster cosmology, as investigated by the number counts method, is deeply linked to the constituent properties of our Universe and small-scale astrophysical phenomena. In the number counts method, a key challenge is relating observations of cluster galaxy members or the gas component to the total mass of the system. This dissertation aims to address this challenge by developing a better understanding of mass–observables relation, with a subsequent goal of enhancing the interpretation of cluster samples that have emerged from large-scale multi-wavelength surveys. These surveys include the XMM-XXL project, the Local Cluster Substructure Survey (LoCuSS), and eventually the Dark Energy Survey data (DES). The results of this work support the science goal of understanding the content and evolution of the Universe's most massive systems, thereby improving cosmological constraints leading to a better understanding of the constituents of our Universe.

In this dissertation, I propose a novel method for cluster mass estimation based on member galaxy kinematics. I demonstrate a percent-level accuracy for the expected conditional log-mass, which implies that this algorithm is one of the most accurate algorithms available in the literature. The accuracy of this algorithm is extensively evaluated on a set of large-scale simulations. Next, all key systematics are identified and calibrated. With this method, we then estimate dynamical masses of a large, optically-selected cluster sample derived from the Sloan Digital Sky Survey (SDSS) and an X-ray-selected cluster sample derived from the XXL Survey.

The multi-wavelength scaling behavior of cluster observables is driven by the astrophysical evolution of the baryonic components within the potential well of massive halos. To facilitate the multi-wavelength scaling modeling, I study the stellar and gas content of dark matter halos extracted from the BAHAMAS simulations, a set of large-scale, full-physics hydrodynamical simulations. The results verify the popular log-normal model of the halo population, but deviate from the power-law approximation. With these simulations, I establish a new set of predictions, most importantly an intrinsic anti-correlation between gas mass and stellar content of these systems. This anti-correlation is a key prediction that we continue to strive to confirm through a subset of the LoCuSS cluster sample.

I implement a robust hierarchical Bayesian inference algorithm, which models the effects of sample selection and the measurement error covariance, to examine the gas and stellar contents of the underlying dark matter halos. To study the relation between the mass of dark matter halos and the multi-wavelength cluster observables, I apply this model to a subset of the LoCuSS cluster sample. Most importantly, this model enables us to examine the predicted anti-correlation between gas and stellar content of these systems. Finally, the results of this study establish the first empirical evidence for this anti-correlation, which has a profound implication for how the Universe's most massive structures formed and evolved.

CHAPTER I

Introduction

The never-ending quest for understanding the fundamental laws of our Universe has inspired many generations of scientists, philosophers, and mathematicians. Humankind's mental engagement with the Universe certainly predates civilization and the record of history. The onset of this journey began with pondering the natural world and the Universe, which goes back to ancient history. Philosophers are the ones who inaugurated this inquiry, searching for the fundamental laws of the Universe, with the natural philosophy movement, i.e. the philosophical study of nature and the physical universe. This expedition later branched out into numerous intellectual arenas under the umbrella of natural science. A major branch of natural science, which studies nature on the grand scale, is cosmology, and another branch is astronomy, which studies the physical phenomena occurring in the cosmos.

1.1 Cosmology and Astronomy as an Empirical Science

Cosmology and astronomy are two overlapping branches of natural science. Similar to other branches, these disciplines progress with the growth of empirical data. The first human who first gazed up at the dark sky obtained the first data which tells a story about our Universe. Today, gigantic data collection instruments are instead taking these data for us. Theoretical and experimental scientists have been working

together to provide a new interpretation of the accumulated data, to confirm existing theories, or rule them out. Sporadically, a new theory is developed to explain observational data which could not be explained with the past theories.

Obtaining reliable and interpretable observational data requires precise and calibrated instruments and appropriate measurement techniques. Such instruments emerged in the early seventeenth century with the invention of the optical telescope. This new instrument revolutionized the entire field of astronomy by providing a new and precise means of measuring and observing astronomical objects and events. The data obtained through this new instrument began the shift in our understanding of the Universe, which led to the modern astronomy as we know it today.

1.2 Scientific Discoveries of Early Days

Before the invention of the optical telescope by Europeans, early astronomers made a number of significant observational discoveries and established some dazzling theoretical models. These pioneering observational discoveries and theoretical feats played a key role in boosting scientific discovery. These discoveries aided astronomers of the seventeenth century and later in establishing the modern view of astronomy. In the following, a few of these remarkable, but mostly under-appreciated, discoveries, which are mostly ignored in typical western historical anecdotes, are pointed out (Ragep 2007).

One noteworthy example is a Persian astronomer, Abd al-Rahman al-Sufi (Azophi), who made the first documented observations of the Large Magellanic Cloud and the Andromeda Galaxy (Hafez et al. 2011). Around the year 964, Abd al-Rahman al-Sufi, in his book of Fixed Stars, described these new systems as a "Little Cloud". Figure 1.1 is taken from a copy of the Al-Sufi's manuscript in which shows the Andromeda

Figure 1.1: The Andromeda galaxy from the earliest copy of the Al-Sufi's manuscript dated 1009-10. The drawing of the Andromeda galaxy, which is described in the main text as a "Little Cloud", lying near the mouth of the Big Fish. This drawing illustrates the Andromeda galaxy with a small cloud of dots, which is the first documented observation of another galaxy. Source: Oxford, Bodleian Library MS. Marsh 144, page 167. `http://bodley30.bodley.ox.ac.uk:8180/luna/servlet/s/fj1968`

Galaxy with a small cloud of dots. These data were the gold standard in astronomy at the time when the naked eye still was the only observational instrument. About 700 years later a European astronomer, Simon Marius, rediscovered the Andromeda Galaxy with observations made through telescopes (Bond 1848).

Ghiyath ad-Din Abul-Fath Umar ibn Ibrahim al-Khayyam Nishapuri (Umar Khayyam), a prominent mathematician, astronomer, poet, and an atheist of the eleventh century, established and directed an observatory at Isfahan to reform the Persian calendar. Under his direction, a team of scholars recalibrated the Persian calendar by fixing the first day of the year, beginning of spring or Nowruz, at the exact moment of the passing of the Sun's center across vernal equinox. This work led to the Jalali calendar, which is claimed to be the most accurate calendar in use today (Akrami 2011; Aminrazavi 2013).

In the early thirteenth century, Muhammad ibn Muhammad ibn al-Hasan al-Tusi (Nasir al-Din Tusi), who is praised as the founder of trigonometry as a mathematical discipline (Lindberg & Shank 2013), established the Maragha observatory. The Maragha observatory is probably the first major international institution where scientists of many disciplines gathered to teach and conduct research (Ballay 1990). The observations he made at this observatory led to accurate tables of planetary movements, which is illustrated in his book "*Ziji ilkhani*". However, he is best known for the development of a mathematical method to describe the plants' motion. Ptolemy initially proposed his model Equant to describe the motion of planets; however, Tusi developed the "Tusi-couple" as an alternative to Ptolemy's model. This method models the harmonic motion via a circle of radius R rolling inside a circle of radius $2R$ (see Figure 1.2). A property of the Tusi-couple is that points on the inner circle that are not on the circumference trace ellipses. This method is

Figure 1.2: A sketch by Nasir al-Din Tusi in his book which shows his mathematical model of planets motion on the sky. Source: Biblioteca Apostolica Vaticana, Vat. Arabic ms 319, fol. 28 verso. `http://digi.vatlib.it/view/MSS_Vat.ar.319/0062`.

further developed by other members of the Maragha School, including Ibn al-Shatir, and eventually appeared in Nicolaus Copernicus' famous book the "*De revolutionibus orbium coelestium*" (On the Revolutions of the Heavenly Spheres)[1] to explain the planets' motion on the sky (Kennedy 1966; Neugebauer 2012).

Computational techniques are not unique to the modern era. These techniques had been used to find the solution of mathematical equations or the motion of the planets in the sky for many centuries before computers were invented. Ghiyath al-Din Jamshid Masud al-Kashi (al-Kashani), a Persian astronomer and mathematician, performed a numerical method to estimate the numerical value of π, and successfully performed this calculation to 16 decimal digits (Azarian et al. 2010). More importantly, Al-Kashani was one of the pre-modern astronomers who pioneered the field of computational astronomy. In the early fifteenth century, Al-Kashi invented two

[1]Kennedy (1966), who coined the term Tusi-couple, noticed that there is a striking similarity between models in Copernicus's "*De revolutionibus orbium coelestium*", including his Mercury and lunar models, and Ibn al-Shatir's models.

mechanical computing machines which could determine the time of an astronomical event or the location of the planetary system on the sky based on the historical observational data (Kennedy 1947, 1950, 1951, 1952). The first one is the Plate of Zones, which was used to determine the longitude of the Sun, Moon, and the other planets. The second is the Plate of Conjunctions, which was used to ascertain the time of day at which planetary conjunctions.

1.3 Emergence of New Paradigms in Scientific Discovery

Even though scholars in the pre-nineteenth century developed or performed computational techniques, these techniques were not widely used by the scientific community. Theory and observation (or experiment) have been at the forefront of and constituted the two primary pillars of the scientific discovery for many centuries. However, a challenge, then and now, is that a large class of theoretical models are complex and non-linear. Finding the general solution of these non-linear models is typically infeasible and analytically intractable. To find an analytic solution, typically a great number of simplifying assumptions are imposed. While these reduced solutions can provide insight or describe simple scenarios, they are not generalizable and in many cases they are erroneous. Analytical methods are thus insufficient to make reliable predictions specifically in non-linear or complex regimes. To overcome these difficulties, numerical techniques, now and then, have to be developed and employed.

The first evidence of the use of numerical techniques can be traced back at least to the Babylonian period. Before the advent of modern electronic computational machines, including computers, numerical methods often depended on manual interpolation of numbers, which could be read off of large tabulated data, or mechanical

machines. In the early twentieth century, electronic computational machines were invented. This invention immediately prompted a broad class of computational techniques that developed and performed aimed to solve analytically intractable problems. Consequently, computational discipline emerged as a new paradigm in scientific discovery. This new interdisciplinary field of research instigated a new class of techniques that have enhanced the rate of scientific discoveries. Central among these techniques are numerical simulations aimed at finding solutions to complex, nonlinear differential equations. In the mid-twentieth century, this new area of research instantly became popular among scientists of various disciplines and provided new insights and predictions which were not possible before. Ultimately, the numerical simulations were accepted as another pillar of the scientific discovery.

In the past decade, the scientific community has experienced another revolution in the way that science is carried out. Specifically, improvements in data acquisition techniques and technologies led to the exponential growth of empirical data. This growth has prompted the emergence of a new paradigm (Hey et al. 2009), big data analytics. This new paradigm soon produced a proliferation of new techniques and models. This fledgling discipline is expected to enhance the rate of scientific discovery by thoroughly exploiting massive datasets and extracting patterns from them. Due to the importance of this, still developing, interdisciplinary domain of research in enabling the scientific progress, it is considered as the most recent pillar of the scientific discovery.

Today, these four pillars of the scientific discovery, i.e. theory, observation, numerical modeling, and big data analytics, are empowering one another and accelerating the pace of uncovering the mysteries of Nature.

1.4 The Triumph of Modern Cosmology

The onset of Modern Cosmology. There are several landmark events that signal the maturation of cosmology from mere speculations to an active research discipline. The modern description of space-time began with a celebrated paper of Albert Einstein in 1905 who introduced the Special Theory of Relativity, positing that space and time are not separate continua. Ten years later in 1915, Einstein published the General Theory of Relativity, which reformulated Newton's theory of gravity. The General Theory of Relativity theory illustrates how matter and energy density warp space-time, and provided new sets of predictions. This theory predicts that the light bends in the presence of gravitational potential. For the first time, Eddington and his collaborators showed the deflection of starlight by the potential well of Sun during the total solar eclipse of May 29, 1919 (Dyson et al. 1920). This observation provided the first empirical evidence in favor of the General Theory of Relativity. However, this observation was not sufficient to say whether the Universe is dynamic or static.

In the meantime, scarce, but ground-breaking, observational studies played a central role in developing the hot big bang model. The pivotal discovery of the linear velocity-distance relation for galaxies by Edwin Hubble in 1929 ruled out the static Universe. Many scholars, including Willem de Sitter and Alexander Friedmann, developed the first pieces of the expanding Universe. They derived the expansion history of the Universe under the isotropic and homogeneous assumption. A key prediction of this expanding Universe is the existence of the Cosmic Microwave Background radiation (Gamow 1948), which was accidentally discovered by the Bell labs scientists in 1964 (Penzias & Wilson 1965). This discovery did set the stage for many discover-

ies to become possible with the cosmic microwave background radiation (CMB). The existing theories, however, could not explain the degree of flatness and homogeneity inferred from the observational data. To explain away these fine-tuning problems, the inflationary models, in which the Universe goes through a period of exponential expansion, were proposed during 1980's. These pieces slowly came together to build our current understanding of the Cosmos under the standard hot big bang model.

Rise of Surprises. Even though the standard model of particle physics and the General Theory of Relativity had successfully passed many empirical tests, a number of observational studies led to surprises. These unexpected observations required another sort of matter-energy which could not be explained by the standard model of particle physics. Central among them was a study of the Coma cluster by Fritz Zwicky who illustrated that the amount of matter associated with the light received from galaxies could not explain the gravitational well predicted by the velocity dispersion of the galaxies inside it (Zwicky 1937). In 1970, 37 years later, a measurement of the rotation curve of spiral galaxies at large radii by Vera Rubin and Kent Ford confirmed the need for a substantial amount of invisible matter (Rubin & Ford 1970), which is best known as dark matter. Another surprising, but important, discovery was the evidence for the acceleration of the expansion of the Universe by the Supernova Cosmology Project and the High-Z Supernova Project teams (Riess et al. 1998; Perlmutter et al. 1999).

The Advent of Large Scientific Collaborations. As new discoveries came along, the entire field of astronomy and cosmology started to grow significantly. In the past century, the field also experienced a growth in the number of theoretical models and theoretical predictions. To test these models random and scarce observational studies were insufficient; consequently, cosmology became a data-demanding

discipline. To address this need, the community evolved into teams supporting large-scale collaborative projects. Today, these collaborations are playing a key role in providing a large volume of data for the community and enabling new scientific discoveries. The primary aim of these collaborations is to design, build, and conduct large survey programs to provide data required for assessing the existing theories and making discoveries.

Among all the sub-fields in astronomy, projects which looked for the CMB pioneered building such a large and successful collaboration. For instance, the Cosmic Background Explorer (COBE), operating from 1989 to 1993, confirmed that the CMB has a blackbody spectrum, and more importantly, discovered the primordial anisotropy of CMB, which is on the order of $\frac{\delta T}{T} \approx 10^{-5}$. This experiment was the precursor to the NASA's Wilkinson Microwave Anisotropy Probe (WMAP) mission and ESA's Planck mission which verified these findings and constrained the cosmological parameters with unprecedented accuracy.

It was the Sloan Digital Sky Survey (SDSS), designed in the 1990's by James Gunn and many colleagues (Gunn et al. 1998), that brought astronomy into the big data era. This optical survey was designed to cover one-third of the available celestial sphere. The observations initially led to an acquisition of positions and brightness of about a billion stars, galaxies, and quasars. This dataset was then supplemented with spectra of about a million astronomical objects. To put these numbers into context, this survey produced about 200 GB of data every night, adding to a database that stands at tens of TB today. And the next generation of astronomical surveys, such as the Dark Energy Survey and the Large Synoptic Survey Telescope, would eventually produce several orders of magnitude larger datasets.

1.5 The Key Role of Scientific Computing in Modern Cosmology

The Onset of Cosmological Simulations. During the 1970's and 1980's, scientists tried to piece together the jigsaw puzzle of how the hot big bang model can give rise to non-linear structures such as galaxies, galaxy groups, and galaxy clusters. These systems are a result of gravitational instability in the primordial density fluctuations. Many analytical solutions for the evolution and collapse of the initial density field were proposed with the caveat of many simplified and unrealistic assumptions, such as spherical symmetry or simple equations of state. Various teams started examining the non-linear structure formation by employing N-body simulations, meanwhile significant progress in the numerical techniques and the invention of electronic computing machines enabled carrying out large-scale simulations. The advent of the cosmological simulations led to a better understanding of the formation and growth of the cosmic structure from the primordial density fluctuations. The first generation of computer simulations of the cosmic structure appeared in the 1980's; and the results supported the gravitational instability and the cold dark matter model (Aarseth et al. 1979; Turner et al. 1979; Bhavsar et al. 1981; Centrella & Melott 1983; Davis et al. 1985; Efstathiou et al. 1985). Since then the computational cosmology has grown as an influential branch of cosmology. Thus far, the computational cosmology has played the leading role in gaining insight regarding the non-linear evolution of the cosmic structure and constraining astrophysical phenomena occurring at small scales.

Thanks to the increase in the computational power and a significant investment in fast, parallelizable computational algorithms, running larger and more complex simulations become viable (see figure 1.3). The modern cosmological simulations are

evolving the large-scale structure and solving small-scale effects and astrophysical events. Among these phenomena, the most important ones are plasma cooling, galaxy formation and evolution, the formation of stars and star clusters and their fate, the energy injection via feedback of supermassive black and supernovae, and the effects of the cosmic rays on baryonic plasma (Borgani & Kravtsov 2011). Typically, astrophysical phenomena are tuned so that the resultant statistics reflect the current observational trends. Sometimes, the outcome leads to novel predictions which could be followed up with observational data. These predictions allow us to falsify or confirm the input model, thereby constraining the tuning parameters.

The modern application of numerical simulations falls into two primary categories. In the first and the original category, numerical simulations are concerned with modeling physical phenomena. The second class of emerging applications is centered on examining data analysis algorithms and inference models. Today, many research teams are working to provide technologies which can generate synthetic night skies or data catalogs for a specific observatory or a survey program. These synthetic data products are then employed by the community to develop new or evaluate existing statistical inference models or data analysis algorithms.

Where Data Meets with Theory. A major challenge is that the scale and complexity of these datasets are exceeding the capacity of traditional data analysis algorithms and models (see figure 1.3). Therefore, analyzing such large datasets poses formidable computational and modeling challenges. Another major challenge is how to deal with and extract unbiased insight from these empirical datasets. To extract unbiased information and conclusion from such unprecedentedly large volumes of data, the community will soon need to revisit outdated models and techniques. This demands scalable, fast, and realistic models that enable process automation,

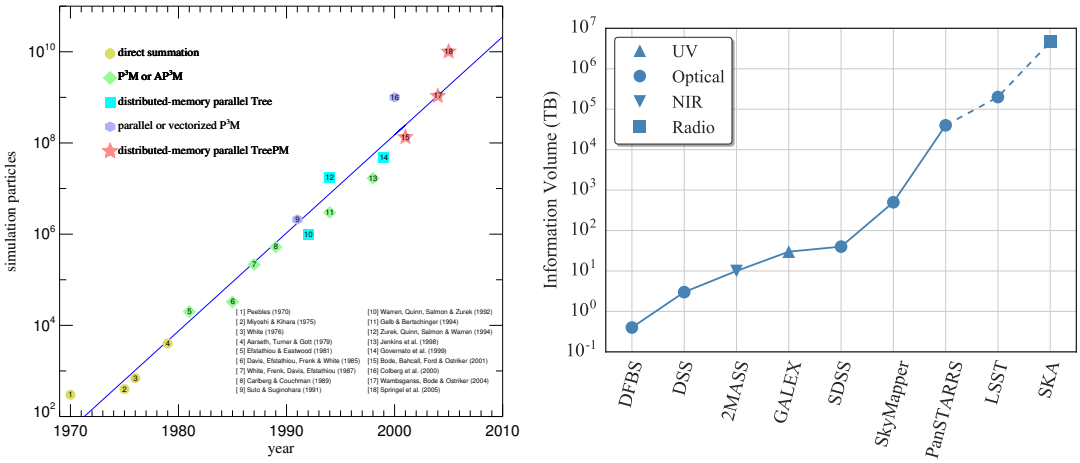

Figure 1.3: **Left.** The growth of the size of N-body simulations wit time. In the past 4 decades, the increase in the size of simulations is consistent with Moore's empirical law, i.e. the computing power doubles every 18 months. The figure is taken from Dolag et al. (2008). **Right.** The growth of data obtained in the past two decades. The plot shows volume of information stored in a number of past, present, and future astronomical surveys or databases and archives. The data points on this plot is taken from Mickaelian (2016).

enhanced insight, and precise inference. Thus, novel big data analytics is a key to unlocking the mysteries of the Universe.

1.6 Clusters of Galaxies, the Intersection of Cosmology and Astronomy

Clusters of Galaxies (Clusters) – the most massive virialized objects in the Universe formed at peaks of the initial density field – are acknowledged as one of the primary probes of cosmological parameters (Huterer & Shafer 2018). The population, spatial distribution, and internal structure of these massive systems are sensitive to dark energy models, the sum of neutrino masses, non-Gaussian primordial fluctuations, the nature of gravity, and astrophysics (see, Allen et al. 2011; Weinberg et al. 2013). Thus far, modern cluster samples have produced competitive cosmological constraints (e.g., Vikhlinin et al. 2009; Rozo et al. 2010; Mantz et al. 2014; de Haan et al. 2016); and now larger and deeper cluster samples are in the process of being assembled. Many observational campaigns are dedicated to identifying and measur-

ing the properties of these systems with high statistical significance. To put it into context, near-future multi-wavelength imaging survey programs (such as WFIRST, eRosita, Euclid, CMB-S4, and LSST) are expected to find hundreds of thousands of new systems. Soon we are going to reach cosmic completeness for clusters with masses above a few times $10^{14}M_{\odot}$ providing a unique discovery opportunity to study astrophysics and the growth of structure with unprecedented detail.

While the wealth of clusters samples derived from deep, wide-area surveys offer discovery opportunities in fundamental physics and astrophysics, modeling and analysis of these datasets introduce theoretical and computational challenges. The main theoretical challenge is how to formulate accurate and precise inference models which map observational space into theory space. These models are expected to be complex and non-linear, thereby computationally challenging.

To tackle the modeling challenge, this dissertation employs simulations. Observational systems, clusters, and theoretical systems, halos, are embedded in different spaces. Halos are typically three-dimensional structures predicted by theory while clusters are 2+N dimensional structures identified in the sky, and here N is redshift or flux or color space. The simulations of these systems provide a powerful means of a modeling formulation, parametrization, and calibration. Therefore, simulations are employed to develop a more accurate and precise inference model.

To tackle the computational challenge, a novel population model and a cluster mass calibration technique are proposed. The population model allows constructing a fast and accurate computation of a cluster cosmology likelihood model. Furthermore, this dissertation establishes a new, fast mass-calibration technique. The proposed mass calibration technique maps cluster observables into theoretical space by employing thousands of galaxy spectra.

The primary aim of this dissertation is to support the science goal of the Dark Energy Survey (DES), the Local Cluster Substructure Survey (LoCuSS), and the XMM-XXL Consortium programs. To achieve this goal, this dissertation establishes new data analysis tools and applies well-vetted and validated models to the data derived from these surveys. Another key contribution of this dissertation is establishing new computational expectations for cluster property covariance, which are tentatively confirmed with data derived from the LoCuSS program.

1.7 Astronomical Surveys to Which this Dissertation Contributes

The XMM-XXL Consortium. The XXL survey, an XMM-Newton Very Large Programme, is a large X-ray program conducted by XMM-Newton satellite to map two extragalactic regions of 25 square degrees, at a depth of $\sim 5 \times 10^{-15}$ erg/cm^2/s in the $[0.5-2]$ keV band. Similar to DES the main goal of this project is to constrain the dark energy equation of state and other key cosmological parameters. This survey is employing counts of clusters, identified via hot gas emitting X-ray, as the primary probe of the cosmological parameters. At intermediate redshift, the XXL coverage and depth enables detection of about 250 clusters for redshifts between $0.3 < z < 0.5$, which have masses around 5×10^{13} – 10^{14} M$_\odot$ (Pierre et al. 2016).

The Local Cluster Substructure Survey (LoCuSS). LoCuSS is a systematic multi-wavelength survey of clusters designed to measure the mass and structure of these massive systems as accurately as possible. LoCuSS provides a sample of about 50 X-ray luminous clusters in the redshift range $0.15 \leq z \leq 0.3$ selected from the ROSAT All-Sky Survey catalogs. This project measures a variety of observable properties, including total mass. The primary goal of this survey is to estimate the weak-lensing mass of each cluster and determine the scaling behavior of cluster

observables with the weak-lensing mass. These results will support the science goal of numerous big projects that aim to measure dark energy to high precision, including Euclid and LSST (Okabe & Smith 2016).

The Dark Energy Survey (DES). DES is an optical survey using a 570-megapixel camera installed on the Blanco telescope to map 5,000 square degrees *grizY* optical passbands to a limiting magnitude of $m_r = 24$, which is two magnitudes fainter than SDSS. The primary goal of DES is to determine the Dark Energy equation of state and other key cosmological parameters to high precision. To achieve this goal, DES will employ four complementary probes, including population statistics of galaxy clusters. DES is expected to map 5,000 square degrees in *grizY* and to provide a catalog of 300 million galaxies with photometric redshifts. The expectation is that about 200 million of these objects will have shape measurements for weak lensing analysis. The DES Science Verification, the first stage of DES which mapped about 250 square degrees to nearly the full DES depth, has a sky surface density of roughly 7 clusters per square degree with richness $\lambda > 20$, implying 35,000 clusters over the full DES survey area (Dark Energy Survey Collaboration et al. 2016).

1.8 The Structure and Contribution of this Dissertation

The work in this dissertation is motivated by the importance of galaxy clusters as a powerful probe of the cosmological parameters and astrophysics. While this dissertation falls within the realm of the standard model of cosmology, it makes an important contribution in enabling the search for the fundamental physics of the Universe via the population of galaxy clusters.

The simulations and models employed in this dissertation rely heavily on the standard model of cosmology, which will be discussed in Chapter II. In this Chapter,

I further discuss the notation and the halo-cluster language which is used throughout this dissertation. I also outline a simple halo population model which can enhance cosmological analysis with a set of galaxy clusters. The accuracy of this population model is evaluated in later chapters.

In Chapter III, I study the baryonic content of halos derived from hydrodynamical simulations. This study leads to a number of predictions for the gas and stellar content of halos which could be followed-up observationally. Moreover, the results of this chapter enable a more accurate parametrization of the cluster observables–mass relation. Finally, the halos derived from these cosmological simulations are employed to assess the precision of the halo population model discussed in Chapter II.

A key challenge in cluster cosmology is cluster mass calibration, i.e. estimating the expected conditional mass of a selected cluster sample. Chapter IV establishes a novel, independent cluster mass calibration technique. To estimate the expected conditional mass, my colleagues and I introduced a new ensemble mass calibration method based on member galaxy kinematics. In Chapter IV, this model is introduced and then evaluated with simulated data. Furthermore, all the potential systematics are identified, and their effects are studied. Subsequently, this new method is applied to an optically selected cluster catalog. The result provides a competitive and independent mass–optical richness relation which enhances the constraints on the cosmological parameters derived from an optical survey, such as SDSS and DES data.

The work presented in Chapter V is a result of a collaboration with the XMM-XXL Consortium project. To enable their ongoing cluster cosmology effort, a new cluster mass–X-ray temperature relation is estimated. This new cluster mass–X-ray temperature relation relies on the method developed in Chapter IV. This work em-

ploys spectra derived from many spectroscopic campaigns that have an overlap with the XXL survey. A major finding is that this new estimate is currently dominated by the systematic uncertainties; therefore, any progress in addressing the systematics would make a significant impact on the constraining power of clusters.

Unbiased statistical models are required to interpret the empirical data and infer the underlying population of halos. In Chapter VI, I develop an inference model which enables the multi-wavelength analysis of a cluster sample considering key systematic effects. These systematics include the effects of covariance between observables, the measurement error covariance, the sample selection, and the distribution of the covariate parameter. This model is, then, applied to a subset of the LoCuSS cluster sample to study the relation between the mass of dark matter halos and cluster observables. With this method and the LoCuSS sample, a number of theoretical predictions for the baryonic content of halos are examined. The results of this work have profound implications for how the Universe's largest structures form and evolve.

The results of Chapter III establishes new predictions for the massive halos. One prediction, which has not previously been studied with observational data, is that there is an anti-correlation between gas mass and stellar mass of massive halos. This finding prompts the postulate that these cosmic giants are closed boxes, meaning that they retaining their cosmic baryonic content. The LoCuSS sample and the model developed in Chapter VI enables testing this prediction, and provide the first observational constrain on this covariance parameter. There results are presented in Chapter VII. The findings of this chapter provide the first observational evidence that supports the closed box model.

Finally, I conclude this dissertation in Chapter VIII, and discuss future directions.

1.9 Beyond this Dissertation

During the course of this dissertation, I have been involved in several research projects and extra circular activities which led to novel research opportunities, which are not included in this dissertation. I have excluded them, because they are, for the most part, beyond the scope of this dissertation. However, I briefly describe these projects and their impact in the following.

Exploring the Halo Mass Function and the Merger History of Warm Dark Matter Universes. Astrophysicists and particle physicists are actively attempting to determine the nature of dark matter. To explain some of the small-scale observed phenomena, a new class of dark matter models, known as warm dark matter, have been studied in the literature. Adding to this stream of work, with Andrew Benson, a staff scientist at the Observatories of the Carnegie Institution for Science, I constructed an analytical model that determines the halo mass function and merger history of a warm dark matter Universe (Farahi & Benson 2013; Benson et al. 2013). In this research, we illustrated how to construct these functions without the need for running full N-body simulations. Our results show that the halo mass function and progenitor mass functions of a cold dark matter universe differ significantly compared to those of the warm dark matter universe.

Evaluating the XXL Cosmology Pipeline. Synthetic skies, which are "observed" by virtual telescopes patterned after real instruments, provide exceptional tools to study and estimate cross-correlation signals, survey specific characteristics, and examine various sources of confusion in the detection and measurement of clusters. To creat a realization of such skies, I developed a fast and scalable template technology[2]. This algorithm realizes synthetic X-ray emission maps from groups and

[2]Xtra: X-ray Template Realization Algorithm, `https://github.com/afarahi/XTRA`

clusters of galaxies. The resultant maps are used by the XMM-XXL program (Pierre et al. 2016) to refine estimates of the selection function and assess the accuracy and precision of their cosmological pipeline (Valotti et al. 2017).

Super-Massive Black Hole (SMBH) Science with eROSITA. It is typically believed that every large galaxy hosts an SMBH, and a fraction of these objects are revealed by an extremely bright Active Galactic Nucleus (AGN). In a collaboration with the XMM Cluster Survey (XCS) team, I performed a scaling relation analysis to study the scaling behavior of the mass of SMBHs, $M_{\rm SMBH}$, with their AGN observables, including the X-ray variability and the X-ray luminosity. The results of this study allow us to evaluate the reliability of $M_{\rm SMBH}$ estimates derived from short exposure L_X measurements. An important application of this work is that the mass inference algorithm, developed in this work, enhances SMBH science in the era of eROSITA where hundreds of thousands of AGN L_X values are expected to be measured (Mayers et al. 2018).

Gravitational Collapse in AdS. The interest in asymptotically Anti-de-Sitter (AdS) has seen a resurgence in the past two decades. This interest was motivated by applications from the AdS/CFT correspondence. Gravitational collapse is one of the dynamical processes in a gravitational theory, which could occur in AdS space as well. In a collaboration with Leo Pando Zayas, I developed a simulation technology which enables us to study the gravitational collapse and formation of a black hole in AdS space (Farahi & Pando Zayas 2014). The techniques involved in this study could help us to better understand the information paradox and quantum turbulence.

Dual Description of a Superconductor. The AdS/CFT correspondence also provides theoretical tools which allows us to study the quantum behavior of a strongly coupled system by solving a dual gravitational problem in a higher dimension. Re-

cent studies illustrated that gravitational theory can provide a holographically dual description of a superconductor. Extending this line of research, with Leo Pando Zayas and other collaborators, I performed numerical simulations to study the effects of disorder on a holographic superconductor using AdS/CFT techniques (Areán et al. 2015, 2014). As an important result, we noticed that moderate disorder, characterize by the amplitude of the noise, enhances superconductivity, which is consistent with experimental results.

In addition to Astronomy and Physics, I am deeply involved in a number of high-impact data science and outreach projects and initiatives. I will describe two of these projects which led to important publications.

The Michigan Data Science Team, A Data Science Education Program with Significant Social Impact. The Michigan Data Science Team[3] (MDST) is an outreach, student-led organization at the University of Michigan which teaches practical data science skills to students. For the past two years, I served as the Vice President of projects on the MDST leadership board. With MDST, I established a sustainable and high impact educational environment in which undergraduate and graduate students can serve the community by employing data science skills in a "Service-Learning" framework (Farahi et al. in preparation). The resultant projects often led to novel, data science research opportunities for students outside of conventional research labs. The most impactful projects emerged from a collaboration with the City of Flint (Abernethy et al. 2016; Chojnacki et al. 2017) and a collaboration with the City of Detroit (Gardner et al. 2017). As a noteworthy example, during the Flint Water Crisis, we collaborated with the City of Flint to develop predictive and classification algorithms to enhance the on-going water recovery efforts. During

[3] `http://midas.umich.edu/mdst/`

the course of this collaboration, we identified several risk factors associated with the elevated levels of lead in the city's drinking water (Abernethy et al. 2016). We also studied the effects of sample selection on the prediction outcomes and developed an unbiased predictive model to inform citizens who are at the risk of lead contamination in their drinking water (Chojnacki et al. 2017).

A Case Study of Education and Public Outreach (EPO). The need for improved communication between scientists and the general public is recognized worldwide. In the past decade, many large-scale astronomy programs have devoted resources to EPO programs and initiatives. To enhance future EPO programs, we conducted a survey of the DES collaboration members to provide data-driven recommendations (Farahi et al. in preparation). In this work, we studied scientists attitudes towards STEM EPO, their motives for participation, and any deterrents from engagement. This also included an analysis of collaboration members' opinions about DES EPO specifically and how to best manage EPO for large science collaborations. We finally explored the value of centralized EPO efforts and provided a list of recommendations for increasing scientists' engagement. As an important finding, we surprisingly noticed that there is a disparity between the types of EPO activities scientists deem valuable, such as on-air media and elementary or high school teacher development, and those in which they participate, such as public presentations.

CHAPTER II

The Standard Model of Cosmology

This chapter provides the basis for the rest of this dissertation. The primary goal of this chapter is to set the theoretical foundations of cluster cosmology and the notation used throughout. Throughout this dissertation, it is assumed that the Universe follows the ΛCDM model. Thus, I first overview this model and discuss its implications. I then present the essential ingredients of large-scale structure, most importantly the formation of bound structures known as halos. This chapter continues with a discussion of implications and ingredients of cosmological simulations. Finally, a halo population model is proposed which is employed in constructing a cosmological inference model. It is worth emphasizing that, in this chapter, I do not present, propose, or develop any new theory or alternative theory to explain the observable Universe.

2.1 The Geometry and the Evolution of the Universe

The standard model of cosmology is built upon the basis of the Cosmological Principle. Based on this principle there is no special place in the Universe, which is also known as the **Copernican principle**. From the cosmological perspective, this means that on large enough scale the Universe is the same everywhere. This postulates that the Universe ought to be homogeneous, the same in all positions, and

isotropic, the same in all directions. To date, empirical data have confirmed that the observable Universe is extremely homogeneous and isotropic. There has been no compelling observational evidence of deviations from the Cosmological Principle. Under this simplified, but accurate, assumption, modeling the geometry and the evolution of the Universe is a rather straightforward task. To describe the geometry and the evolution of the Universe, the General Theory of Relativity, proposed by Albert Einstein in 1915, is employed. Since 1915, the General Theory of Relativity has exhaustively tested with observational data (Will 2014). These studies all confirmed the predictions of the General Theory of Relativity at almost any relevant scale in cosmology and astronomy. Hence, there is no compelling evidence of a violation of the General Theory of Relativity.

Early observations conducted by Edwin Hubble demonstrated that the distant galaxies appear to be receding from us. This finding postulates that the Universe is not static and expands over time. Another implication of this finding is that the Universe cannot be homogeneous and isotropic in time. Therefore, the standard model of cosmology distinguishes between space and time, where spacelike slices are individually homogeneous and isotropic.

2.1.1 Friedmann Equation

The separation of space and time allows us to consider the spacetime as a product of two manifolds $T \times R$, where T represents the time direction, and R is a three-dimensional, spacelike, manifold. The isotropy implies that R is invariant under rotations, and homogeneity implies that R is invariance under translations. Therefore, the Universe's metric can be of the form

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a(t)\gamma_{ij}(u)\mathrm{d}x^i\mathrm{d}x^j \,, \tag{2.1}$$

where t is the timelike coordinate, and (x^1, x^2, x^3) are the spacelike coordinates. The function $a(t)$ is known as the **scale factor**, which describes the size of the Universe at time t and its evolution. The spacelike coordinate system, (x^1, x^2, x^3), is known as the **Comoving Coordinate**, in which cross terms, $\mathrm{d}t\,\mathrm{d}u^i$, are taken to be zero. An observer who stays at constant x^i, a "comoving" observer, observes the universe as isotropic[1]. The isotropy and homogeneity assumptions imply that the γ_{ij}, metric on x, should be maximally symmetric, thereby spherically symmetric (Carroll 1997). The metric can be put in the form

$$\gamma_{ij}(u)\mathrm{d}x^i\mathrm{d}x^j = \frac{\mathrm{d}r^2}{1-kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)\,. \tag{2.2}$$

This is the famous **Robertson-Walker metric**. Equation 2.2 is invariant under the transformation

$$k \longrightarrow \frac{k}{|k|} \quad , \quad r \longrightarrow \sqrt{|k|}r \quad , \quad a \longrightarrow \frac{a}{|k|}\,. \tag{2.3}$$

This invariance implies that the only relevant parameter is $\frac{k}{|k|}$. Hence, we only can imagine three not unique scenarios, $k = -1$ (open universe), $k = 0$ (flat universe), and $k = +1$ (closed universe).

2.1.2 Fluid Equation

Thus far, we have not made use of the General Theory of Relativity. The Universe may be described with many species of perfect fluids, matter and energy, with density ρ_i. The energy-momentum tensor for a perfect fluid can be written in form of

$$T_{\mu\nu} = (p_i + \rho_i)U_\mu U_\nu + p_i g_{\mu\nu} \tag{2.4}$$

where $g_{\mu\nu}$ is the metric tensor, and ρ_i and p_i are energy density and pressure, respectively, of fluid i as measured in the rest frame. U_μ is the four-dimensional velocity

[1]We note that the Earth is not a comoving observer as the earth moves respect to the CMB rest frame. This effect induces a dipole anisotropy in CMB which has been detected and measured by the CMB experiments.

of the fluid. We note that the fluid is at rest in the comoving coordinate. Thus, the velocity vector in natural units will be,

$$U^{\mu} = (1, 0, 0, 0)\,. \tag{2.5}$$

And finally, the energy-momentum tensor is

$$T_{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & & & \\ 0 & & g_{ij}p & \\ 0 & & & \end{pmatrix}\,. \tag{2.6}$$

2.1.3 Equation of State

We first consider the conservation of energy equation. This equation reads

$$\partial_0 \rho = -3(\rho + p)\frac{\dot{a}}{a}\,. \tag{2.7}$$

Equation (2.7) can also be derived from the First Law of Thermodynamics. This expression, however, does not tell anything about the relationship between density, ρ, and pressure, p. To make progress, it is necessary to choose an equation of state. The equation of state is typically defined as

$$p = p(\rho) = w\rho\,, \tag{2.8}$$

where w depends on the fluid being considered. Substituting this into Equation (2.7) and rearranging leads to

$$\frac{\dot{\rho}}{\rho} = -3(1 + w)\frac{\dot{a}}{a}\,. \tag{2.9}$$

Assuming w does not run with the scale factor, the above equation can be integrated to obtain

$$\rho = a^{-3(1+w)}\,. \tag{2.10}$$

Essentially all of the perfect fluids relevant to cosmology obey a constant equation of state. This relation applies to all components of the Universe, radiation, cold-matter and dark energy, but w takes on different values. There exists also baryonic matter which is considered as collisionless, non-relativistic matter at large-scales. But at small-scales the equation of state of the baryonic matter become important. This is further discussed in Section 2.4.

- Radiation may be used to describe either actual electromagnetic radiation or relativistic matter such as light neutrinos. The equation of state for relativistic particles follow $w = \frac{1}{3}$. Therefore, when the Universe is radiation-dominated the energy density in matter falls off as $\rho \propto a^{-4}$.

- Cold-matter is collisionless and non-relativistic matter, which obeys $w = 0$. Accordingly, the pressure is negligible in comparison with the energy density. To a high precision and for a large enough scale this is true for both cold-dark matter and baryonic matter, including stars and hot gas. When the Universe is matter-dominated, the energy density falls off as $\rho \propto a^{-3}$.

- The cosmological constant or Vacuum Energy or Dark Energy follows the equation of state in which $w = -1$.[2] For dark energy, the energy density is constant and independent of the scale factor.

2.1.4 Friedmann Equations

To describe the evolution of the Universe, we need to solve the Einstein's equations,

$$R_{\mu\nu} = 8\pi G(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T)\,. \tag{2.11}$$

[2]There are competing models in which the dark energy is not exactly $w = -1$ and evolve over time or take another value close to -1. For example, there are Quintessence models in which dark energy is a scalar field and does, in fact, vary slowly with time. Thus far, there is no statistically significant evidence for such model. We, therefore, do not explore the dark energy models beyond the cosmological constant, where $w = -1$.

There are two non-zero terms, terms with $\mu\nu = 00$ or $\mu\nu = ii$. These two terms are

$$-3\frac{\ddot{a}}{a} = 4\pi G(\rho + 3p)\,, \tag{2.12}$$

and

$$\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{k}{a^2} = 4\pi G(\rho - p)\,. \tag{2.13}$$

Rearranging the above two equations leads to

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p)\,, \tag{2.14}$$

and

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}\,. \tag{2.15}$$

Together these are known as the Friedmann equations. Metrics of the form Equation (2.2) which obey the Friedmann equations define the **Friedmann-Robertson-Walker (FRW)** universes.

2.1.5 Hubble Law

The rate of expansion of the Universe is characterized by the **Hubble parameter**,

$$H(a) = \frac{\dot{a}}{a}\,. \tag{2.16}$$

The value of the Hubble parameter at the present epoch is the Hubble constant, H_0. Accordingly, the evolution factor is defined as

$$E(a) \equiv \frac{H(a)}{H_0}\,, \tag{2.17}$$

where $E(a)$ is basically the normalized Hubble parameter. To solve the Friedmann equations first we define the density parameter. For each fluid specie, the corresponding density parameters Ω_i, is defined as following

$$\Omega_i = \frac{8\pi G}{3H^2}\rho_i = \frac{\rho_i}{\rho_{\rm crit}}\,, \tag{2.18}$$

where $\rho_{\rm crit} = \frac{3H^2}{8\pi G}$. One can easily show that the sum of densities are equal to one and the curvature term $\sum \Omega_i = 1 + \frac{k}{H^2 a^2}$. For convenience, the same quantity can be defined for the curvature as well,

$$\Omega_k = \frac{-k}{H^2}\,. \tag{2.19}$$

Therefore, the curvature term fall off as $\rho \propto a^{-2}$. Because the curvature decays slower than the matter and radiation, a small value for the curvature at the early Universe would become significant today [3].

Finally, dividing the Friedmann Equation by H_0^2 gives

$$\frac{H(a)^2}{H_0^2} = \frac{\Omega_{r,0}}{a^4} + \frac{\Omega_{M,0}}{a^3} + \frac{\Omega_{k,0}}{a^2} + \Omega_{\Lambda,0}\,, \tag{2.20}$$

where $\Omega_{i,0}$ is the density of component i at current time. The exact value of each component is subject of current studies. The cosmological probes approximately found the following values, $\Omega_{r,0} \approx 10^{-6}$, $\Omega_{M,0} \approx 0.3$, $\Omega_{k,0} \approx 0$, and $\Omega_{\Lambda,0} \approx 0.7$ (Planck Collaboration et al. 2014).

The ΛCDM model successfully explains the observational data that probe the geometry and the expansion history of our Universe since the end of inflation. In this dissertation, we study the Universe during the period of the structure formation; and this model is sufficient for our purpose during this period.

2.2 Cosmological Redshift

Typical observations in cosmology are dealing with photons. There are also experimental setups which are looking for the direct signature of dark matter particles, cosmic neutrinos, cosmic high energy particles, and more recently gravitational waves; however, photons still make up the dominant part of the observational experiments

[3] The inflationary models are proposed to solve the curvature fine-tuning problem and a few other fine-tuning problems.

and current astronomical data. Measuring the number of photons received from an event or an object and their energy, i.e. frequency, are two typical measurable quantities.

A comoving observer can measure the frequency of photons. The measured frequency is $\nu = -U_\mu V^\mu$, where $V^\mu = \mathrm{d}x^\mu/\mathrm{d}\lambda$ (Carroll 1997). Accordingly, the frequency of photons are changing as the Universe expands. This can be formulated with

$$\frac{\nu_i}{\nu_j} = \frac{a_j}{a_i}\,. \tag{2.21}$$

where ν is the frequency of photons. Instead of frequency or wavelength, cosmologists and astronomers prefer to map these quantities into space of **redshift** z. Redshift is defined between the two events by the fractional change in wavelength,

$$z = \frac{\lambda_0 - \lambda_1}{\lambda_1}\,. \tag{2.22}$$

Combining the above equation with Equation (2.21) gives the cosmological redshift

$$z = \frac{a_0}{a_1} - 1\,. \tag{2.23}$$

The scaling parameter at current time is usually set to be 1, $a_0 = 1$. Therefore, according to Equation (2.23), the scale factor of an astronomical object is

$$a = \frac{1}{1+z}\,, \tag{2.24}$$

where z is the redshift of that object. We can write all above equations in term of redshift, for example the Evolution Parameter become

$$E(z) = \frac{H(z)^2}{H_0^2} = \Omega_{r,0}(1+z)^4 + \Omega_{M,0}(1+z)^3 + \Omega_{k,0}(1+z)^2 + \Omega_{\Lambda,0}\,. \tag{2.25}$$

2.2.1 Observed Redshift

It is worth mentioning that the cosmological redshift, which is defined in Section 2.2, is not the only source of the photon redshift. This redshift differs from the conventional Doppler effect. The cosmological redshift is the direct effect of the expansion of space, while the Doppler redshift is induced by the relative velocities of the observer and the source. Astronomical objects are dynamical systems and move in comoving spacetime. Their relative velocity respect to the observer induces additional redshift, which is called the "Doppler shift".

There is one additional source of frequency change in collected photons. The General Theory of Relativity predicts that there is time dilation within a gravitational well. This effect is known as the "gravitational redshift". Assuming the Schwarzschild solution of the Einstein equations, the redshift associated with a photon traveling in the gravitational field of an uncharged, point mass would change according to

$$1 + z = \frac{1}{\sqrt{1 - \frac{2GM}{rc^2}}}\,, \tag{2.26}$$

where M is the mass creating the gravitational field, G is the gravitational constant, r is the distance of the source in the Schwarzschild coordinate, and c is the speed of light.

The gravitational redshift effect is significant near a black hole. As an object approaches the event horizon of a black hole, theoretically the redshift could be as large as infinity. However, this effect is typically minimal while it could still be measurable. For instance, this effect can be measured on Earth using the Mössbauer effect, first observed in the Pound-Rebka experiment (Pound & Rebka 1960). Another interesting example of such an effect is the Sachs-Wolfe effect (Sachs & Wolfe 1967).

The expansion of space changes gravitational potential wells of the large-scale structure. This change induces gravitational redshift which is imprinted on temperature fluctuations in the cosmic microwave background radiation at large scales.

These three competing effects, cosmological redshift, doppler redshift, and gravitational redshift, contribute to the measured redshift simultaneously. The magnitude of these effects depends on many factors, most importantly the distance of the source from the observer and the environment of the source. For instance, the measured redshift of near-by objects is dominated by the Doppler shift (e.g., the redshift of the Andromeda galaxy), while the measured redshift of very distant galaxies is dominated by the cosmological redshift. For the purpose of this dissertation though, we are not concerned about the gravitational redshift, as it is negligible.

The relative velocity of galaxies within the large-scale structure is of our particular interest, as this quantity can be employed to map the potential well of dark matter halos or estimate the mass of these objects. Zwicky (1937) is the first who measured a mass of such an object with measuring the relative velocity of galaxies inside a cluster of galaxies. This quantity is used in this work as well to construct a relation between the mass of dark matter halos and their observables (see Chapter IV for more detail).

2.3 Growth of Structure

During the inflationary epoch, the scale factor grew exponentially and causes quantum fluctuations to become classical. Because of the exponential expansion, these tiny, quantum fluctuations stretched to macroscopic scales, and eventually left the horizon. This process is called "**freeze in**". After the termination of the inflation, during the radiation-dominated and matter-domination epochs, these stretched

quantum fluctuations re-entered the horizon. These tiny fluctuations induce perturbations on the metric of the spacetime which couples them to matter and photon. This coupling led to fluctuations in matter and photon density which can be studied and characterized via its signatures on the CMB, the primordial temperature fluctuations.

The initial density modes which re-entered the horizon started to grow as a result of the gravitational force. To model this evolution, the relativistic Boltzmann equation is employed to evolve matter and photons. This model can describe the evolution of the density field and radiation fluctuation until the decoupling epoch when matter and photons are decoupled from one another. After this epoch, the matter, cold-dark matter and the baryonic matter, behaves like collisionless, non-relativistic particles, which could be described solely by the theory of gravity. Finally, the gravitational theory is employed to describe the growth of the cosmic structure. This growth can be modeled with the Newtonian physics on top of the expanding Universe background, which is discussed in Section 2.1. At the peaks of this density field, the dark matter halos, massive virialized objects in the Universe, begin forming and growing. The next section presents a model that describes how these halos were formed and evolved.

2.3.1 Growth of Matter Density Fluctuations and Halo Formation

The linear solution of the Boltzmann equation until the decoupling era would approximate the evolution of matter in the early Universe pretty accurately. This solution sets the initial condition within which growth of the cosmic structure started. We start with the initial density field to study the non-linear evolution and growth

of structure after the decoupling epoch. First, the density contrast is defined as

$$\delta(r,t) \equiv \frac{\rho(r,t) - \bar{\rho}(t)}{\bar{\rho(t)}}\,, \tag{2.27}$$

where $\bar{\rho}(t)$ is the average density at time t. To evolve this density field, the spherical collapse solution is followed. The spherical collapse solution was first proposed by Gunn & Gott (1972). Assuming that the seeds of gravitationally bound structures are spherical density, they illustrated that, in Einstein de Sitter Universe (EdS), the density field evolves as

$$\delta_{\rm EdS} \propto t^{\frac{2}{3}}\,. \tag{2.28}$$

What we observe today are discrete, gravitationally bound systems, which are formed at the peaks of the matter density fluctuations. In order to relate the initial density fluctuations in the early Universe to the structure we observe today, we need a model which describes the space density of these gravitationally bound structures, which are known as **dark matter halos**. Historically, the space density of dark matter halos is referred to as the **"halo mass function" (HMF)**. To determine HMF, theoretical models typically rely on the assumption that the smoothed over-density field $\delta(r,t)$ is a Gaussian random field with a scale-dependent variance $\sigma(r) = \langle\delta\rangle$. Press & Schechter (1974) were the first who attempted to model and determine HMF. They assumed that the probability that an overdense region has a value above a critical over-density δ_c is

$$P(>\delta_c) = \frac{1}{\sqrt{2}\sigma}\int_{\rho_c}^{\infty} \exp\left(-\frac{\nu}{2}\right)\mathrm{d}\delta = \frac{1}{2}\mathrm{erfc}\left(\frac{\nu}{\sqrt{2}}\right), \tag{2.29}$$

where $\nu = \delta_c/\sigma$ is the critical over-density in units of the variance. $P(>\delta_c)$, is equal to the fraction of mass collapsed into halos with mass greater than M, $F(>M)$. This model predicts that only half of the entire mass of the Universe collapse into halos,

$\lim_{\sigma\to\infty} P(>\delta_c) = 1/2$. To avoid this problem Press & Schechter (1974) multiplied the result by an ad hoc factor of two, leading to $F(>M) = 2P(>\delta_c)$ where F is the fraction of collapsed mass. This model predicts the cumulative number count. Observationally, however, it is more appealing to have an estimation of the number density of objects within a mass bin $[M, M+\mathrm{d}M]$, i.e. $\mathrm{d}n = n(M)\mathrm{d}M$ or HMF. This prediction can be directly compared with empirical data derived from observations. HMF is determined by manipulating $F(>M)$. Thus, HMF is

$$n(M)\mathrm{d}M = \frac{\rho}{M}F(M)\mathrm{d}M = \frac{\rho}{M^2}f(\nu)\left|\frac{\mathrm{d}\ln\sigma}{\mathrm{d}\ln M}\right|\mathrm{d}M\,, \tag{2.30}$$

where $\sigma(R)$ is the variance of the smoothed density field which is

$$\sigma^2(R) = \langle\delta^2(x,R)\rangle = \frac{1}{2\pi^2}\int \mathrm{d}k\ P(k)\ W_k^2(R)k^2\,. \tag{2.31}$$

$P(k)$ is the matter power-spectrum and $W_k(R)$ defines the smoothing function in the Fourier space. Note that the scale, R, and the corresponding mass, M, are uniquely related to each other through $M = \frac{4\pi}{3}\bar{\rho}R^3$. Thus, σ can be re-evaluated in term of M.

According to the Press & Schechter (1974) ansatz $f(\nu)$, the fraction of collapsed mass in a unit of $\ln\nu$, reads

$$f_{PS}(\nu) = \sqrt{\frac{2}{\pi}}\nu\exp\left(-\frac{\nu^2}{2}\right). \tag{2.32}$$

After this pioneering work, many other authors attempted to model and determine $f(\nu)$, thereby HMF, by employing peak theory or excursion set theory[4] (e.g., Bardeen et al. 1986; Bond et al. 1991; Paranjape et al. 2013). While these rather simple analytical approaches do a remarkable job of describing the observed trends, they

[4]I explored the excursion set theory and integrated this model into a semi-analytical simulation to construct HMF and merger history of a Warm Dark Matter universe (Farahi & Benson 2013; Benson et al. 2013). These studies are beyond the scope of this dissertation, thereby not discussed.

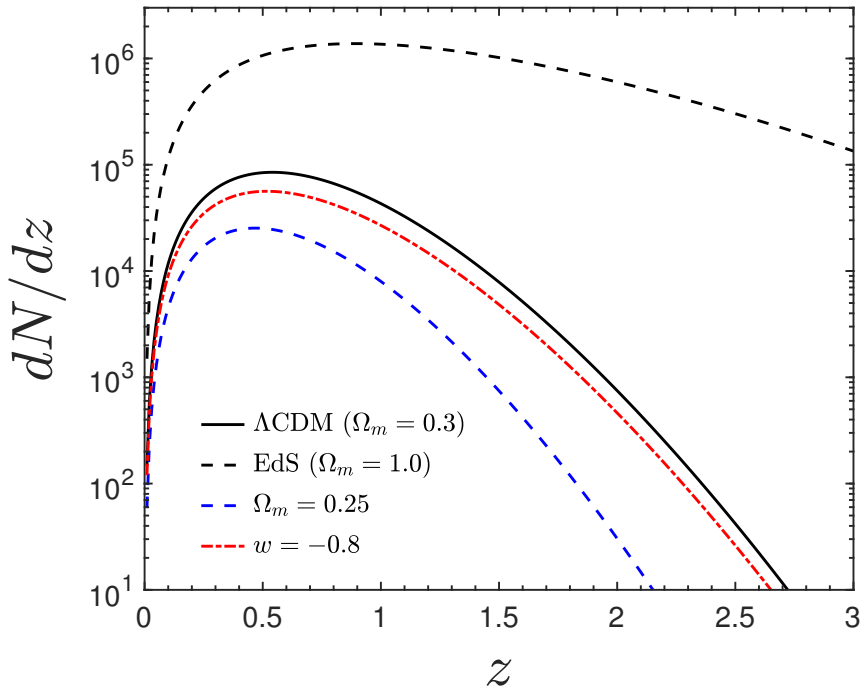

Figure 2.1: This figure illustrated that the number counts of halos as a function of redshift is sensitive to the cosmological parameters. This is a prediction for a survey covering 5,000 square degrees assuming a complete sample of halos more massive than $10^14\ M_\odot/\,h^{-1}$. The figure is taken from Huterer & Shafer (2018).

are inaccurate and imprecise, thereby insufficient, to be applicable in modeling the current observational data. The current attempts in constraining the cosmological parameters with 1% precision require percent level calibration of HMF as a function of the cosmological parameters. As an alternative, cosmological simulations have since been used to derive an accurate estimation of this quantity (e.g., Jenkins et al. 2001; Reed et al. 2007; Tinker et al. 2008).

A key lesson from the functional form of HMF is that the number density of halos is sensitive to the cosmological parameters and the initial condition, which itself is a function of the cosmological parameters. Figure 2.1 demonstrates the sensitivity of the number counts of halos as a function of redshift to the cosmological parameters. Vikhlinin et al. (2009) performed a cosmological analysis with a set of these systems derived from an observational study, successfully illustrated that the number density

of their sample rules out a dark matter only universe. Since this study clusters are considered as one of the primary probes of the cosmological parameters. To turn clusters into a competitive cosmological probe, however, there are many challenges. A major practical challenge is measuring the mass of these systems as it is not directly observable.

The mass of dark matter halos is typically estimated with properties derived from the luminous matter residing in the potential well of these systems. Dark matter halos provide gravitational potential wells within which baryonic hot plasma cools and forms stars and galaxies. Each dark matter halo contains hot plasma, which is in the form of ionized hot gas, and stellar population, typically in the form of galaxies. Considering the gravitation physics implies that the mass and redshift of halos determines the total amount of baryonic matter within these systems. However, how much of the initial gas turns into the stars depends on the efficiency of cooling and the feedback mechanisms that inject energy into the hot plasma. A major lesson in the past two decades is that the highest mass halos that host groups and clusters of galaxies are less efficient at converting baryonic gas into stars. Therefore, the majority of baryons end up in a hot intracluster medium (ICM, Briel et al. 1992). Measuring galaxy formation and their assembly across cosmic history is a key to understanding the astrophysical processes happening within halos and eventually inferring the mass of these systems.

The following section, I will present the self-similar model. This simple model describes how the observable quantities of dark matter halos scale with their total mass. Even though this model cannot capture all the essential ingredients of baryonic physics, it provides a basis for the rest of the work presented in this dissertation.

2.4 Galaxy Clusters

The primary goal of this dissertation is to model and study how the mass of halos scales with their observable quantities. It is useful to review the self-similar model, originally proposed by (Kaiser 1986), which attempts to quantify these scaling relations. This model assumes only gravitational physics and employs the spherical collapse model. Another key assumption of this model is that the total baryonic content of halos follows the dark matter content at all scales (Bertschinger 1985). In the following I will illustrate that the observable quantities from the luminous matter follow a power-law form with halo mass.

2.4.1 Self-similar Model of Galaxy Clusters

The gravitational force is a scale-invariant process. If gravity is the dominant force on the scale of massive dark matter halos, then it is expected that the derived quantities from the baryonic content of dark matter halos are scaled versions of each other. The observable properties of these massive systems are therefore determined only by their mass and redshift. Redshift determines the critical density

$$\rho_c(z) = \frac{3H^2(z)}{8\pi G} = \frac{3H_0^2 E^2(z)}{8\pi G} = E^2(z)\rho_{c,0}\,, \tag{2.33}$$

where $E(z) = H(z)/H_0$ defined in Section 2.2.

It is conventional to define halo mass as that, centered on a local potential minimum, contained within a sphere of radius r_Δ encompassing an over-density Δ relative to the critical density, thus

$$M_\Delta \;=\; \frac{4}{3}\pi r_\Delta^3 \Delta \rho_c(z) \;\propto\; E^2(z) r_\Delta^3\,. \tag{2.34}$$

The combination of mass and redshift sets the strength of the local Newtonian gravitational potential, Φ, and circular velocity of the halo, $v_{\rm circ}$, through the com-

bination $E(z)M_\Delta$:

$$v_{\rm circ}^3 \propto \Phi^{3/2} \propto E(z)M_\Delta\,. \tag{2.35}$$

The above relation motivates our use of the effective potential well depth, $E(z)M_\Delta$, as the independent degree of freedom in the scaling laws we consider in this dissertation. To reemphasis the role of the cosmological simulations, (Evrard et al. 2008) accurately and precisely calibrated by N-body simulations of cold dark matter particles (Evrard et al. 2008). They illustrated that the scaling power, derived from simulations, is very close to prediction of the self-similar model.

Intracluster medium (ICM) is the superheated plasma that falls into the potential well provided by the dark matter halos. The ICM particles are assumed to be in hydrostatic equilibrium inside these deep potential wells (Cavaliere & Fusco-Femiano 1978). This equilibrium satisfies the Virial Theorem between gravitational potential energy U and kinetic energy, K, of ICM particles, $\langle U\rangle = -2\langle K\rangle$. Thus, the total kinetic energy can be written in terms of the average kinetic energy of the ICM particles, i.e. the cluster X-ray temperature, T_X, leading to

$$T_X \propto [M_\Delta E(z)]^{2/3}\,. \tag{2.36}$$

The X-ray emission from the ICM is dominated by thermal bremsstrahlung emission, for which the resulting luminosity scales as $L_X \propto \rho_{\rm gas}^2 r^3 \Lambda(T_X)$, where there are two factors of the gas density $\rho_{\rm gas}$ because the radiation is produced by a two-body interaction, and $\Lambda(T_X)$ is the cooling function. In the soft-band range [$\sim 0.1-2.4$keV], the integral of the cooling function is nearly independent of T_X, while across the full energy range used for bolometric X-ray luminosity it scales with $T_X^{1/2}$ (Blanchard et al. 1992). This leads to

$$\frac{L_{X,\rm soft}}{E(z)} \propto M_\Delta E(z)\,, \qquad \frac{L_{X,\rm bol}}{E(z)} \propto [M_\Delta E(z)]^{4/3}\,. \tag{2.37}$$

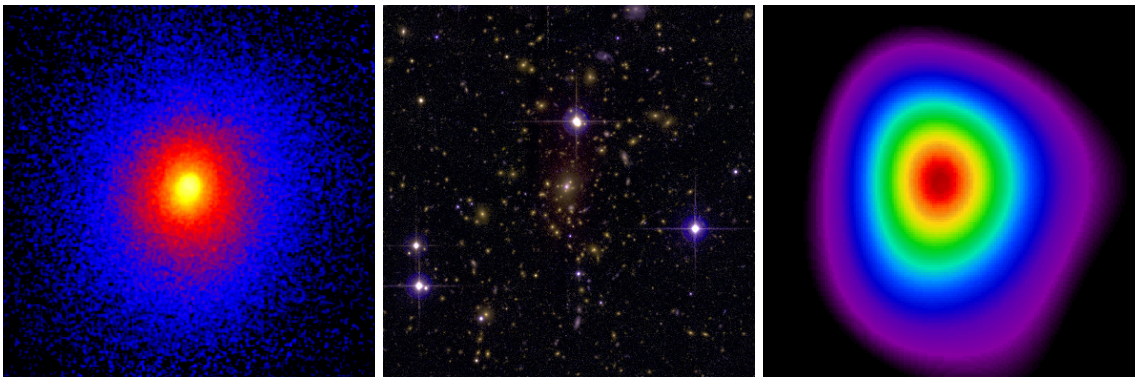

Figure 2.2: A multi-wavelength view of a massive, dunamically relaxed galaxy cluster. This figure illustrates the multi-wavelength realization of Abell-1835 (z = 0.25) at X-ray (left), optical (middle), and mm (right) wavelengths. The figure is taken from Allen et al. (2011).

As probes of the same thermal energy, Y_X and Y_{SZ} have the same self-similar scaling, as can be derived from the product of $M_{\rm gas}$ and T_X:

$$YE(z) \propto [M_\Delta E(z)]^{5/3} \,, \tag{2.38}$$

under the simple assumption of a constant gas fraction, $f_{\rm gas}$. The similar assumption of a constant stellar fraction, $f_\star$, leads to

$$M_{\rm gas} = f_{\rm gas} M_\Delta \propto M_\Delta \,, \qquad L_K = M_\star = f_\star M_\Delta \propto M_\Delta \,. \tag{2.39}$$

since K-band luminosity, L_K, is a good indicator of the total stellar mass. Finally, a relation between richness, the number counts of galaxies with the virial boundary of a halo, and mass would be in form of

$$\lambda = \frac{M_\star}{m_{\star,\rm gal}} \propto M_\Delta \,, \tag{2.40}$$

assuming each cluster has a galaxy population drawn from a universal luminosity function with some effective mean stellar mass, $m_{\star,\rm gal}$.

The term mass–property relation is employed to represent the functional form of conditional halo statistics, $p(\mathbf{S}|M,z)$, where $\mathbf{S}$ is a set of intrinsic properties of the population of halos of mass M at redshift z. To discuss the three-dimensional

spatial measurements of halo properties, here, instead of *observable*, the term *property* is used. While not directly observable, estimators for these quantities can be constructed from optical, infrared, X-ray or SZ observations. Figure 2.2 illustrates multi-wavelength observations of a massive galaxy cluster which can be used to estimate the properties and eventually the mass of this system. Even though the above idealistic model is based on a number of simplified assumptions, it provides a basis by which mass–property relation can be described. For instance, according to the self-similar model, mass–property relation can be modeled with a power-law form.

2.5 Cosmological Hydrodynamical Simulations

Cosmological hydrodynamical simulations, which evolve gravitationally-coupled baryons and dark matter, provide mass, redshift, and environment dependent predictions for the observables of massive halo. Historically, these simulations have played a leading role in studying the baryonic content of dark matter halos and guiding the modeling and parametrization of observational quantities (e.g., Evrard 1990; Evrard et al. 1996; Bryan & Norman 1998; Sembolini et al. 2013; Le Brun et al. 2017; Barnes et al. 2017; Pillepich et al. 2018). While significant progress has been made, multi-fluid hydrodynamic simulations remain challenged by the wide dynamic range and complex astrophysical elements involved in modeling the formation of stars, supernova feedback, and supermassive black hole effects. In a recent attempt to address these challenges, McCarthy et al. (2017) have taken a novel approach by tuning sub-grid control parameters to match the observed galaxy stellar mass function and the hot gas mass fractions of groups and clusters simultaneously. Their simulations include metal-dependent radiative cooling, star formation, and prescriptions for both supernova and active galactic nucleus (AGN) feedback.

A variety of numerical schemes for solving the coupled system of collisional baryonic matter and collisionless dark matter have been developed in the past decades. These methods generally fall into two categories: (i) particle-based methods, which discretize mass, and (ii) mesh-based methods, which discretize space (for a review see Dolag et al. 2008). In this section, we briefly overview some of the relevant pieces and ingredients of cosmological simulations.

2.5.1 N-body simulations

Large-scale structure formation started after the decoupling epoch when the Universe was dominated by matter. The leftover radiation is not playing a significant role in determining the evolution of the baryonic content of the Universe, thereby ignored in simulations employed in this work. In particle-base schemes, the cold-dark matter is typically described as a collisionless, nonrelativistic fluid of particles with mass m, position $\mathbf{x}$, and momentum $\mathbf{p}$, where $\mathbf{x}$ is taken to be the comoving position. Accordingly, the momentum is $\mathbf{p} = ma^2\dot{\mathbf{x}}$ and proper peculiar velocity is $\mathbf{v} = a\dot{\mathbf{x}}$. The phase-space distribution function $f(\mathbf{x}, \mathbf{p}, t)$ of the dark-matter fluid can be described by the collisionless Boltzmann equation,

$$\frac{\partial f}{\partial t} + \frac{\mathbf{p}}{ma^2}\nabla f - m\nabla\Phi\frac{\partial f}{\partial \mathbf{p}} = 0\,. \tag{2.41}$$

In this equation Φ is the gravitational potential which is described by the Poisson equation,

$$\nabla^2\Phi(\mathbf{x}, t) = 4\pi G a^2[\rho(\mathbf{x}, t) - \bar{\rho}(t)]\,. \tag{2.42}$$

In the above equation, ρ is the density field and $\bar{\rho}$ is the mean density of the Universe when the scale factor was a. The scale factor depends on the cosmological parameters and is solved independently. Having the phase-space distribution function, $f(\mathbf{x}, \mathbf{p}, t)$,

we can determine the density easily as following

$$\rho(\mathbf{x}, t) = \int f(\mathbf{x}, \mathbf{p}, t) \mathrm{d}\mathbf{p} \,. \tag{2.43}$$

The above equation is typically solved by sampling the phase-space density by a finite number of N representative particles. The solution is determined by integrating the equations of motion,

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -m\nabla\Phi \tag{2.44}$$

and

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \frac{\mathbf{p}}{ma^2} \,. \tag{2.45}$$

The gravitational evolution of dark matter particles describes the large-scale statistics quite well but fails to describe the statistical quantities and observables at small-scales. To describe these observables and the mass and redshift evolution of baryonic content of dark matter halos, it is essential to model and include the baryonic physics and the relevant astrophysical events. Hydrodynamical simulations have emerged to address these needs. New ingredients are added to capture relevant baryonic physics and the effects of feedback, basically any phenomenon other than gravity.

2.5.2 Baryonic physics

The baryonic plasma content of the Universe can typically be described as additional ideal fluid. Therefore, to follow the evolution of the fluid, one usually has to solve a set of hydrodynamic equations. These equations are

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\frac{\nabla P}{\rho} - \nabla\Phi \,, \tag{2.46}$$

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho\nabla\mathbf{v}\,, \tag{2.47}$$

and

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -\frac{P}{\rho}\nabla\cdot\mathbf{v} - \frac{\Lambda(u,\rho)}{\rho}\,. \tag{2.48}$$

The above equations are the *Euler equation*, *continuity equation*, and the *first law of thermodynamics*, respectively. One additional equation, which relating the pressure P to the internal energy u, is required to have a complete set of solvable equations. The latter equation is the *equation of state*. Assuming an ideal, non-relativistic, monatomic gas, this will be

$$P = (\gamma - 1)\rho u\,, \tag{2.49}$$

where $\gamma = 5/3$. Finally, $\Lambda(u,\rho)$ is the cooling function due to the radiative losses. The hydrodynamical equations for an expanding Universe are

$$\frac{\partial\mathbf{v}}{\partial t} + \frac{1}{a}(\mathbf{v}\cdot\nabla)\mathbf{v} + \frac{\dot{a}}{a}\mathbf{v} = -\frac{1}{a\rho}\nabla P - \frac{1}{a}\nabla\Phi\,, \tag{2.50}$$

$$\frac{\partial\rho}{\partial t} + \frac{3\dot{a}}{a}\rho + \frac{1}{a}\nabla\cdot(\rho\mathbf{v}) = 0\,, \tag{2.51}$$

and

$$\frac{\partial}{\partial t}(\rho u) + \frac{1}{a}\mathbf{v}\cdot\nabla(\rho u) = -(\rho u + P)\left(\frac{1}{a}\nabla\cdot\mathbf{v} + 3\frac{\dot{a}}{a}\right)\,. \tag{2.52}$$

There are four additional, required ingredients to model the baryonic content of these massive cosmic giants. Each one of these ingredients is briefly described the following.

(i) Cooling. The cooling function, $\Lambda(u,\rho)$, is added to the first law of thermodynamics to describe radiative cooling of the plasma that exists inside virialized

systems, dark matter halos. Therefore, this is a key ingredient in studying the baryonic content of halos. This function strongly depends on the temperature and the metallicity of the plasma, i.e. hot gas.

(ii) Star formation. Reservoirs of cold and dense gas can turn into collisionless stars. As the stellar population evolves, it feeds back mass and energy via stellar wind and supernova explosions, which would heat the cold gas. As for star formation, a relatively simple recipe is that originally introduced by Katz et al. (1996), which is often used in cosmological simulations. According to this prescription, for a gas particle to be eligible to form stars, it must have a convergent flow,

$$\nabla \mathbf{v}_{\rm gas} < 0\,, \tag{2.53}$$

and have density in excess of some threshold value. These criteria are complemented by requiring the gas to be Jeans unstable, which is described with

$$\frac{h_{\rm gas}}{c_{\rm gas}} > \frac{1}{\sqrt{4\pi {\rm G}\rho_{\rm gas}}}\,, \tag{2.54}$$

where h_i is the smoothing length and c_i is the local sound speed. Once a gas particle is eligible to form stars, it forms stars with the rate can be written as

$$\frac{{\rm d}\rho_{\rm star}}{{\rm d}t} = -\frac{{\rm d}\rho_{\rm gas}}{{\rm d}t} = \frac{c_{\rm star}\rho_{\rm gas}}{t_{\rm star}}\,, \tag{2.55}$$

where $c_{\rm star}$ is a dimensionless star formation rate parameter and $t_{\rm star}$ the characteristic timescale for star formation.

(iii) Supernovae Feedback. High mass stars eventually ran out of fuel, collapse, and then followed by an explosion as type-II supernovae (SN II). This violent event would release some amount of energy to the surrounding gas. Under the approximation that the typical lifetime of massive stars which explode as SN II does not exceed the typical time step of the simulation, this energy released instantaneously,

with the feedback energy deposited in the surrounding gas. A further improvement can be achieved by a more accurate model of the feedback energy, metal enrichment of the surrounding gas, and the effects of SN Ia (Borgani et al. 2008).

(iv) Active galactic Nucleus (AGN) Feedback. AGN feedback can heat up the surrounding gas via radiation. This process can also mix-up the hot and cold gas through the mechanical wave resulted from the mass accretion on a black hole (BH). The AGN mechanical feedback can be modeled with the following equations (Ostriker et al. 2010),

$$\frac{\mathrm{d}M_{\rm BH}}{\mathrm{d}t} = \frac{1}{1+\eta}\frac{\mathrm{d}M_{\rm in}}{\mathrm{d}t}\,, \tag{2.56}$$

$$\frac{\mathrm{d}M_{\rm out}}{\mathrm{d}t} = \eta\frac{\mathrm{d}M_{\rm BH}}{\mathrm{d}t}\,, \tag{2.57}$$

$$\frac{\mathrm{d}E_{\rm w}}{\mathrm{d}t} = \epsilon_{\rm w}c^2\frac{\mathrm{d}M_{\rm BH}}{\mathrm{d}t}\,, \tag{2.58}$$

$$\frac{\mathrm{d}p_{\rm w}}{\mathrm{d}t} = \mathbf{v}_{\rm w}\frac{\mathrm{d}M_{\rm out}}{\mathrm{d}t}\,, \tag{2.59}$$

where $\frac{\mathrm{d}M_{\rm BH}}{\mathrm{d}t}$ is the mass accretion rate on the BH, $\frac{\mathrm{d}M_{\rm out}}{\mathrm{d}t}$ is the mass outflow rate, $\frac{\mathrm{d}M_{\rm in}}{\mathrm{d}t}$ is the mass inflow rate, $\epsilon_{\rm w}$ is the efficiency of generating mechanical energy with an AGN wind, $\mathbf{v}_{\rm w}$ is the AGN wind velocity, and $\eta \equiv \dot{M}_{out}/\dot{M}_{in} = 2\epsilon_{\rm w}c^2/|\mathbf{v}_{\rm w}|^2$. Finally, $\frac{\mathrm{d}E_{\rm w}}{\mathrm{d}t}$ is proportional to the energy rate deposited into the surrounding gas, and a mechanism that describes a flux of hard X-rays photons.

2.6 Population Statistics of Massive Halos and Galaxy Clusters

As mentioned earlier, the number density of massive halos is an essential ingredient of a cluster cosmology analysis. Cluster surveys identify massive halos in the sky by measuring their bulk properties integrated within an angular aperture. Derived

observable properties populate a signal vector, $\mathbf{S}$, with elements that may include measures of optical richness, λ (number of red galaxies), galaxy velocity dispersion, $\sigma_{\rm gal}$, X-ray luminosity, $L_{\rm X}$, temperature, $T_{\rm X}$, derived hot gas mass, $M_{\rm gas}$, SZ decrement ,$Y_{\rm SZ}$, and surface mass density from weak-lensing shear, $\Sigma_{\rm lens}$. As the scaling relations are following a power-law relation it is more convenient to define the halo properties in natural log space. The natural log of a vector of properties $\mathbf{S}$ is defined with $\mathbf{s} \equiv \ln(\mathbf{S})$ and the natural log of the halo mass is defined with $\mu \equiv \ln(M/M_p)$, where M_p is a pivot mass. Note that a vector of properties is expressed with bold font, $\mathbf{s}$, and an element of this vector is expressed with italic font, s. The expected, comoving number density of halos expected within some specific property bin, i, at redshift bin, j, is given by the convolution,

$$\left\langle \frac{\mathrm{d}n_{i,j}(s,z)}{\mathrm{d}V} \right\rangle = \int_{z_j}^{z_{j+1}} \mathrm{d}z \frac{\mathrm{d}V}{\mathrm{d}z} \int_{s_i}^{s_{i+1}} \mathrm{d}s \int_{-\infty}^{\infty} \mathrm{d}\mu \; \frac{\mathrm{d}n(\mu,z)}{\mathrm{d}\mu} \; p(s|\mu,z)\,, \tag{2.60}$$

with $p(s|\mu,z)$ the conditional likelihood of the property used to select the halo sample, $\frac{\mathrm{d}n(\mu,z)}{\mathrm{d}\mu}$ is the halo mass function, and V is the volume element of universe at redshift z. The cosmological parameters define the halo mass function and the volume element in this equation. The dependency of these functions to the cosmological parameters for standard model of cosmology is pretty well understood.

The number abundance of clusters within a property bin and redshift bin can be easily measured from observational data. By comparing the theoretical prediction and observational data, one can constrain the cosmological parameters. To perform such an analysis, a knowledge of $p(s|\mu,z)$ is required, which is the subject of this dissertation. Thus, producing competitive cosmological constraints with a set of clusters requires an unbiased estimation of halo properties from cluster observables. Knowledge of these mapping functions is also critical for the understanding of the multi-phase baryon evolution. A considerable effort has gone into measuring and

calibrating the halo mass–cluster observables of the most massive halos in our Universe (see Giodini et al. 2013, for a recent review). However, these studies have been limited by samples of tens to low hundreds, systematic uncertainties in total mass estimates, and complex or ill-defined sample selection criteria. Recent efforts are improving on these fronts (e.g., Mantz et al. 2016a,b; Zou et al. 2016; Saro et al. 2017). The work in this dissertation extend these efforts to accurately and precisely model the baryonic content of galaxy clusters. Furthermore, I examine some of the above relations with the clusters derived from cosmological simulations and cluster identified in the observational data. In the following section, I discuss the notation which is employed throughout this work.

2.6.1 The Mass–Multi-Property Relation

The mass–multi-property relation, $p(\mathbf{s}|\mu, z)$, is the joint probability distribution for a vector of halo properties at fixed halo mass. It is typically assumed that this joint probability distribution has a log-normal form,

$$P(\mathbf{s} \,|\, \mu, z) \propto \exp\left\{-\frac{1}{2}(\mathbf{s} - \langle \mathbf{s}|\mu, z\rangle)^T \Sigma^{-1} (\mathbf{s} - \langle \mathbf{s}|\mu, z\rangle)\right\}\,, \tag{2.61}$$

where $\langle \mathbf{s}|\mu, z\rangle$ defines the log-mean behavior, and Σ defines the covariance of Gaussian deviations about the log-mean. Each diagonal element of the covariance matrix specifies the variance of properties while the off-diagonal elements are the property covariance, all at fixed halo mass. The assume log-normal form is explicitly evaluated in Chapter III of this dissertation. Assuming a power-law form, the expected value of the property vector, conditioned on halo mass and redshift, is

$$\langle \mathbf{s}|\mu, z\rangle = \boldsymbol{\pi}(z) + \boldsymbol{\alpha}(z)\mu\,, \tag{2.62}$$

where the vectors $\boldsymbol{\pi}$ and $\boldsymbol{\alpha}$ are the normalizations and slopes of the halo properties scaling law. The scatter of a property about the mean relation is expressed with

σ. In practice, the normalization and the slope can be redshift-dependent. The redshift-dependent of these quantities are studied for stellar mass and gas mass in Chapter III.

If $\mathbf{s}$ is an vector of multi-wavelength properties then to construct $p(\mathbf{s}|\mu, z)$ the full covariance between each pair of properties is required. The diagonal element of this covariance specifies the property variance about the mean relation, and the non-diagonal elements is proportional to the correlation between two property residuals about the mean relations. The latter is referred to as property covariance. This property covariance is a subject of our investigation in Chapter III and Chapter VII of this dissertation.

2.6.2 An Analytical Model of Conditional Statistics

Under the power-law and log-normal assumptions coupled with a simple parameterization of HMF, Evrard et al. (2014) derive closed-form expressions for multi-property population statistics. A few of these expressions are presented in the following. This analytic model exposes fundamental parameter degeneracies between the shape of HMF, which is driven mainly by the cosmological parameters, and mass–property relation parameters determined by astrophysical processes, discussed in Section 2.4. Another appealing advantage of this model is that the model supports fast computation of expectations for cosmological likelihood analysis, which can substantially speed up the evaluation of a cluster sample inference models.

The smoothness of the mass function allows a logarithmic polynomial expansion,

$$\frac{\mathrm{d}n(\mu, z)}{\mathrm{d}\mu} = \exp\left[\beta_0(z) - \sum_{j=1}^{3} \frac{\beta_j(z)}{j!}\, \mu^j\right], \tag{2.63}$$

consisting of an amplitude, $\exp\{\beta_0(z)\}$, and linear through cubic coefficients, $\beta_j(z)$, that control the shape. In theory, one can go beyond the cubic approximation; how-

ever, in practice the cubic approximation is sufficient for percent level accuracy. This claim is explicitly evaluated in Chapter III. (Evrard et al. 2014) derived a complete set of first order, second order, and third order expressions. These expressions are used in later chapters.

First Order Approximation of a Cluster Selected Sample. The non-uniform shape of HMF would have an effect of the log-mean total halo mass selected by a given observable, s_a. The expression for the log-mean total halo mass is

$$\langle \mu \,|\, s_a \rangle = \left[\left(\frac{s_a - \pi_a}{\alpha_a} \right) - \beta_1 \sigma^2_{\mu|a} \right] , \tag{2.64}$$

where $\sigma^2_{\mu|a} = \sigma^2_a/\alpha^2_a$ is the first-order estimate of the mass variance selected by property s_a. The other first order approximation of the key observable quantities are also shown in the following.

The probability density function of an observable s_b, for a selected sample of another observable, s_a, has a log-normal form with mean and variance

$$\langle s_b | s_a \rangle = \pi_b + \alpha_b [\, \langle \mu | s_a \rangle + \beta_1 \, r_{ab} \, \sigma_{\mu|a} \, \sigma_{\mu|b} \,] \,, \tag{2.65}$$

$$\sigma^2_{b|a} = \alpha^2_b \, [\, \sigma^2_{\mu|a} + \sigma^2_{\mu|b} - 2 r_{ab} \, \sigma_{\mu|a} \sigma_{\mu|b} \,] \,, \tag{2.66}$$

where r_{ab} is the intrinsic correlation coefficient between properties s_a and s_b at fixed mass. This is a key quantity which is studied in this work both using halos derived from simulations and observational data. This intrinsic correlation induces a bias in the mean of s_b. This effect can be understood by the fact that the dominant lower mass halos that scatter upward into the chosen s_a bin will also have a positive deviation from the mean s_b if r_{ab} is positive. And if r_{ab} is negative, the effect is reversed. It is understood that the above functions can be a function of redshift. However, a redshift correction is not required as all these derived quantities are conditioned on redshift.

This work pays careful attention to the intrinsic property correlation matrix, as this quantity is not yet carefully studied. Improved knowledge of this matrix should have a major impact on inference analysis using multi-wavelength cluster data (Cunha 2009).

Second Order Approximation. As it is illustrated in the first order approximation, the convolution brings HMF coefficients into the expression for the log-mean total halo mass selected by a given observable, s_a,

$$\langle \mu \,|\, s_a, z \rangle = x_s \left[\left(\frac{s_a - \pi_a}{\alpha_a} \right) - \beta_1 \sigma^2_{\mu|a} \right] , \tag{2.67}$$

where $\sigma^2_{\mu|a} = \sigma^2_a / \alpha^2_a$ is the first-order estimate of the mass variance selected by property s_a, and

$$x_s \equiv (1 + \beta_2 \, \sigma^2_{\mu|a})^{-1} \simeq (1 - \beta_2 \, \sigma^2_{\mu|a}) \, , \tag{2.68}$$

is a compression factor less than unity that is sensitive to the curvature of HMF. This compression factor appears in other statistical measures. It also worth noting that the magnitudes of the corrections due to the shape of HMF is directly related to the property scatter. If this scatter is small (large) then the corrections would be small (large) accordingly.

CHAPTER III

Simulated Halo Population Properties: scalings, log-normality, and covariance

Philosophy and Contribution

In this chapter, I study the scaling behavior of the baryonic - stellar and hot gas - content of dark matter halos derived from hydrodynamical simulations. The work in this chapter provides new insights regarding these scaling relations. Additionally, this work establishes new theoretical, model-dependent predictions which could be falsified or confirmed via future observational data. Finally, I explicitly evaluate the accuracy of population model which is discussed in Chapter II. This chapter is taken from Farahi et al. (2017a), "Localized massive halo properties in BAHAMAS and MACSIS simulations: scalings, log-normality, and covariance". The analysis and plots presented in this Chapter are my own work. The simulations have been performed and the halo catalogs are generated by the co-authors.

3.1 Chapter Introduction

The total baryonic content of dark matter halos appears in inform of hot gas or cold, stellar material. In the previous chapter, we discussed how this baryonic plasma, trapped inside the potential well of a dark matter halo, evolve, cools and form stars. According to the self-similar model of Chapter II, the mass of these luminous matters

should follow the mass of the host halo with a power-law relation with a power index of one. Deviation from this prediction is expected as the self-similar model does not capture all astrophysical phenomena occurring in these systems. In the past two decades, measuring this potential deviation motivated a number of observational studies (Giodini et al. 2013). These studies attempted to measure the statistical relation between the halo mass and its gas and stellar content (Mantz et al. 2016a,b; Saro et al. 2017; Schellenberger & Reiprich 2017). However, no consensus has been reached yet regarding the exact value of the index of these power-laws.

The primary goal of this Chapter is to study the functional form of multi-wavelength properties of a halos population conditioned on halo mass and redshift, $p(\mathbf{S}|M, z)$, which defines the mass–property relation (MPR). In addition to the mean relation, multi-wavelength population statistics requires a good understanding of the covariance between pairs of intrinsic properties or observable quantities. This covariance is an essential element in modeling multi-wavelength cluster samples, as pointed out by Nord et al. (2008) for the case of inferring luminosity evolution from X-ray flux-limited samples. The diagonal elements of the covariance matrix linking mass to observable properties are becoming better measured, but currently off-diagonal elements are poorly known (Mantz et al. 2016a). Cosmological hydrodynamics simulations, however, are a great tool for gaining insight into the detailed form of the MPR, including property covariance.

The likelihood of little or no loss of baryons from the deepest potential wells motivates an expectation of anti-correlation in the gas and stellar mass fractions in the highest massive halos. If all clusters of fixed halo mass are closed baryon boxes with baryons partitioned into stars and gas, then a particular system with slightly more (less) gas than average must contain a lower (higher) stellar mass than average,

meaning a strong anti-correlation between gas mass and stellar mass. Such an anti-correlation is apparent in the Rhapsody-G simulations of Wu et al. (2015), where a correlation coefficient $r = -0.7$ is found for gas and stellar mass deviations about the mean in a sample of ten $10^{15}\,\mathrm{M}_{\odot}$ halos and their progenitors. In lower-mass halos hosting groups and poor clusters of galaxies, feedback can effectively drive baryons outside of the virial radius (e.g., Lau et al. 2010; Sembolini et al. 2013; Le Brun et al. 2017; Truong et al. 2018), reducing or eliminating the degree of anti-correlation.

Another key assumption in modeling MPRs is the form of the conditional distribution of properties at fixed halo mass, usually assumed to take a log-normal form. Under a log-normal assumption coupled with a simple parameterized approximation to the halo space density, or mass function, Evrard et al. (2014, hereafter E14) derive closed-form expressions for multi-property population statistics. The analytic model exposes fundamental parameter degeneracies between the shape of the mass function, which is driven by cosmology, and MPR parameters determined by astrophysical processes (see Chapter II). Practically, the model supports fast computation of expectations for cosmological likelihood analysis.

The goals of this Chapter are: i) to measure the mass and redshift dependencies of MPRs for stellar mass and hot gas mass; ii) evaluate the statistical form of the MPR likelihood, and; iii) test the accuracy of the E14 model in a simulation setting where the intrinsic properties are measured directly. Unlike previous "zoom-in" simulations (e.g., Wu et al. 2015), the BAHAMAS simulation models baryon behavior in a large cosmic volume, enabling study of a wide range of halos hosting groups and clusters. The large samples from BAHAMAS allow us to apply a localized regression approach to estimate mass-dependent MPR parameters. However, the $400\,h^{-1}\,\mathrm{Mpc}$ simulation size limits the number of the most massive halos; BAHAMAS statistical coverage drops

off above $3 \times 10^{14}\,\mathrm{M}_{\odot}$. We therefore also include the MACSIS simulation ensemble which, like Wu et al. (2015), uses the zoom-in technique to extend the mass range of the BAHAMAS sample while employing the same astrophysical model, resolution, and cosmology (Barnes et al. 2017).

This Chapter organized as follows. In Section 3.2 we present the simulation samples used in this work while Section 3.3 describes our non-parametric local linear regression (LLR) model. The LLR results, including covariance of hot gas and stellar mass at fixed halo mass, are presented in Section 3.4. In Section 3.5 we test the performance of the E14 analytic model, followed by discussion in Section 3.6 and a summary in Section 3.7.

Chapter's Notation. Throughout this Chapter, we use radial and mass scales defined by a spherical density contrast with respect to the critical density of the universe, $\rho_{\rm crit}(z)$; M_{Δ} indicates the mass within which the average total mass density is $\Delta\rho_{\rm crit}(z)$. Halo masses are expressed in units of $\mathrm{M}_{\odot}$, not $h^{-1}\mathrm{M}_{\odot}$).

3.2 Simulations

We use the BAHAMAS cosmological hydrodynamical simulation (McCarthy et al. 2017) run using the Gadget-3 SPH code with subgrid prescriptions for metal-dependent radiative cooling, star formation, and stellar and AGN feedback developed as part of the OverWhelmingly Large Simulations project (Schaye et al. 2010). The periodic $400\,h^{-1}\,\mathrm{Mpc}$ cube we use here adopts a flat ΛCDM cosmology with *Planck 2013* cosmological parameters (Planck Collaboration et al. 2014), namely $\Omega_m, \Omega_b, \Omega_\Lambda, \sigma_8, n_s, h = 0.3175, 0.049, 0.6825, 0.834, 0.9624, 0.6711$ where Ω_m, Ω_b and Ω_Λ are the normalized densities in matter, baryons and vacuum energy, σ_8 sets the power spectrum normalization, n_s is the primordial spectral index, and $h \equiv H_0/(100\ \mathrm{km\ s^{-1}\ Mpc^{-1}})$ is the

dimensionless Hubble constant.

The wind velocity associated with stellar feedback and the heating temperature associated with the AGN feedback in BAHAMAS are adjusted so as to reproduce the observed local galaxy stellar mass function and the amplitude of the relation between hot gas mass and halo mass of local X-ray-selected galaxy groups and clusters. Non-tuned features match an unprecedentedly wide range of observed properties, including galaxy and hot gas radial profiles as well as the behavior of stacked SZ and X-ray luminosity as a function of galaxy stellar mass (McCarthy et al. 2017).

Cosmological simulations featuring volume-complete hydrodynamics with full sub-grid physics at high spatial and mass resolution are very computationally expensive. The $400\,h^{-1}$ Mpc BAHAMAS simulation has spatial resolution of $4\,h^{-1}$ kpc and resolves a $10^{14}\,\mathrm{M}_{\odot}$ halo with $\sim 30,000$ particles. Because of the limited number of very high mass halos in the realized volume, the MACSIS project (Barnes et al. 2017) was developed to extend the sample to higher mass halos. The MACSIS ensemble consists of 390 "zoom-in" simulations (Tormen et al. 1997) of individual halo regions drawn from a parent 3.2 Gpc N-body simulation. The hydrodynamic resimulations employ the same resolution and sub-grid prescriptions as BAHAMAS in a Planck cosmology with nearly identical parameters as BAHAMAS (parameter values typically differ in the third significant digit, see Barnes et al. 2017).

As described in McCarthy et al. (2017), halos are identified using a "friends-of-friends" percolation method. The spherically integrated quantities used here are measured using the minimum of the local gravitational potential as the halo center, and any sub-halos that lie outside the characteristic radii, R_{Δ} are ignored.

The samples we use, listed in Table 3.1, include all halos with $M_{500} > 10^{13}\mathrm{M}_{\odot}$ at redshifts $z = 0$, 0.5 and 1.0. Note that there the redshift slice for MACSIS sample

Table 3.1: Halo sample sizes with $M_{500} > 10^{13}\,\mathrm{M}_{\odot}$.

Redshift	BAHAMAS	MACSIS
1	11387	377
0.5/0.46[a]	17668	377
0	21987	385

[a] 0.5=BAHAMAS , 0.46=MACSIS

is 0.46. The combined BAHAMAS and MACSIS simulations offer tens of thousands of halo realizations covering a wide dynamic range in total mass.

The halo properties we study are the aggregate stellar mass, $\mathrm{M_{star}}$, and the hot phase gas mass, $\mathrm{M_{gas}}$, measured within spheres enclosing densities of $\Delta = 500$ and 200 times the critical density, $\rho_{\mathrm{crit}}(z)$. Note that the hot gas mass includes particles with temperatures greater than 10^5 K while the stellar mass uses all star particles within R_{Δ}.

For this study, we combine BAHAMAS and MACSIS samples into a super-sample. Since the BAHAMAS and MACSIS are not using exactly the same cosmology, we re-normalize the baryonic contents of the MACSIS sample to align the global baryon fraction, Ω_b/Ω_m, to that assumed in the BAHAMAS cosmology; however, the magnitude of this correction is small, $< 2\%$. We also note that there is small difference in the redshift of BAHAMAS and MACSIS samples, 0.5 versus 0.46. Since we show below that the redshift evolution of the properties we examine is relatively weak, we do not apply any correction for this redshift.

The complex interactions of mergers, turbulence, cooling, chemical enrichment, and feedback from supernovae and AGN play out within the evolving cosmic web network of large-scale structure to determine the overall statistical nature of the baryon component masses within the halo population. While matching observed mean stellar and gas fraction behavior, within the limits of current observational uncertainties, has been done in the BAHAMAS and MACSIS simulations by tuning a small number of

sub-grid parameters, higher-order features of the property statistics should be considered *model-dependent predictions* of the underlying astrophysical theory. Within the context of these simulations' numerical and astrophysical treatments, we focus this Chapter on the model's expectations for running of the slope and scatter of the MPR with mass and redshift. Future work can examine the robustness of these features using multiple simulations by independent groups.

3.3 Mass-localized Regression

In this section, we describe a localized linear regression model to characterize the conditional joint property likelihood, $p(\mathrm{M_{star}},\ \mathrm{M_{gas}}|\ \mathrm{M_{halo}}, z)$, of the simulated halo ensemble. In practice, the power-law nature exhibited by most properties with respect to mass motivates the use of logarithmic variables.

The method produces *mass localized* estimates of the intercepts, slopes and covariance of this pair of properties as a function of halo mass at fixed redshift. The assumption of a log-normal form for the conditional likelihood underlies this model, and we demonstrate the validity of this assumption in Section 3.4.2.

Following E14, our underlying population model considers a vector of properties, $\mathbf{S}$, associated with halos of total mass, M_Δ, at redshift, z. Using natural logarithms of the properties, $\mathbf{s} = \ln \mathbf{S}$, and mass, $\mu = \ln M_\Delta$, the log-mean scaling of property a at a fixed redshift is locally linear

$$\langle s_a \,|\, \mu, z \rangle \ = \ \pi_a(\mu, z) + \alpha_a(\mu, z)\mu \ , \tag{3.1}$$

with redshift- and scale-dependent parameters that we measure by differentially weighting halos in the simulation ensemble around a chosen mass scale. In this model the normalization of the property element, S_a, is $e^{\pi_a(\mu,z)}$.

At a fixed redshift, we determine local fit parameters — the slope $\alpha_a(\mu)$, intercept,

$\pi_a(\mu)$, and intrinsic sample variance, $\sigma_a^2(\mu)$ — for property s_a by minimizing the weighted square error,

$$\epsilon_a^2(\mu) = \sum_{i=1}^{n} \; w_i^2 \; (s_{a,i} - \alpha_a(\mu)\mu_i - \pi_a(\mu))^2 \;\; , \tag{3.2}$$

where the sum i is over halos, $\mu_i \equiv \ln(M_{\rm halo,i}/M)$, and w_i is the local weight centered on the mass scale, $M \equiv e^{\mu}$. We sweep through values of M covering the mass scale of poor groups to rich clusters, $M_{500} \in \{10^{13}, 10^{15}\}\,{\rm M}_\odot$, in the joint BAHAMAS and MACSIS halo samples.

We use a Gaussian weight in log-mass,

$$w_i \;\; = \;\; \frac{1}{\sqrt{2\pi}\sigma_{\rm LLR}} \exp\left\{-\frac{\mu_i^2}{2\sigma_{\rm LLR}^2}\right\} \;\; , \tag{3.3}$$

with $\sigma_{\rm LLR} = 0.46$, equivalent to 0.2 dex in halo mass. As the central halo filter scale, μ, is varied, we record the local slope and intercept fit parameters. Ideally we want the weighting scheme has the smallest possible width; however practically this is not achievable as we ran out of clusters, as a result the fit parameters become noisy. If the width is too large, then it smooths out the effect of running. We test whether decreasing the width of the weighting scheme changes our results, and we find that the estimates become noisier, but the shapes and the estimated values are effectively the same. We therefore conclude that the size of the width is sufficiently small for the purpose of this work.

With a local slope and intercept for each property, j, we can compute the local property covariance using the same weighting scheme. We use an unbiased weighted estimator of the property covariance matrix, C (Gough 2009),

$$C_{a,b} = A \sum_{i=1}^{n} w_i \; \delta s_{a,i} \; \delta s_{b,i} \;\; , \tag{3.4}$$

where $\delta s_{a,i} \equiv s_{a,i} - \alpha_a \mu_i - \pi_a$ is the residual deviation from the local best-fit, (a, b) are labels representing either stellar mass or hot gas mass, and the pre-factor is

$$A = \frac{\sum_{i=1}^{n} w_i}{\left(\sum_{i=1}^{n} w_i\right)^2 - \sum_{i=1}^{n} w_i^2} \quad . \tag{3.5}$$

The covariance matrix for our pair of halo properties has one correlation coefficient,

$$r_{\rm gas,star} = \frac{C_{\rm gas,star}}{\sqrt{C_{\rm gas,gas}\ C_{\rm star,star}}} \quad . \tag{3.6}$$

We note that fitting a global power-law to MPRs that run with scale could induce covariance as an artifact of the poor, i.e. underfit, regression model. The locally estimated covariance is unbiased, easily computable, and asymptotically approaches the population true value in the limit of $\sigma_{\rm LLR} \to 0$ and $N_{\rm halo} \to \infty$.

3.4 Results

In this section, we begin by presenting the LLR scaling behavior of log-mean stellar mass and hot gas mass as a function of halo mass and redshift. We then examine the form of the conditional likelihood PDF, finding excellent agreement with a log-normal form, the assumption behind the weighted Pearson covariance, equation (3.4). Finally, we investigate the redshift and mass dependence of the star-gas covariance.

Unless otherwise stated, error bars and shaded regions in the figures below are one standard deviation based on bootstrap estimates of 1000 re-sampled halo datasets.

3.4.1 LLR fits to scaling relations

Figure 3.1 shows how the hot gas mass (top) and stellar mass (bottom) of the BAHAMAS and MACSIS halo population scale with total mass at three redshifts and for two critical overdensity scales, $\Delta = 500$ and 200. LLR fit lines are also shown.

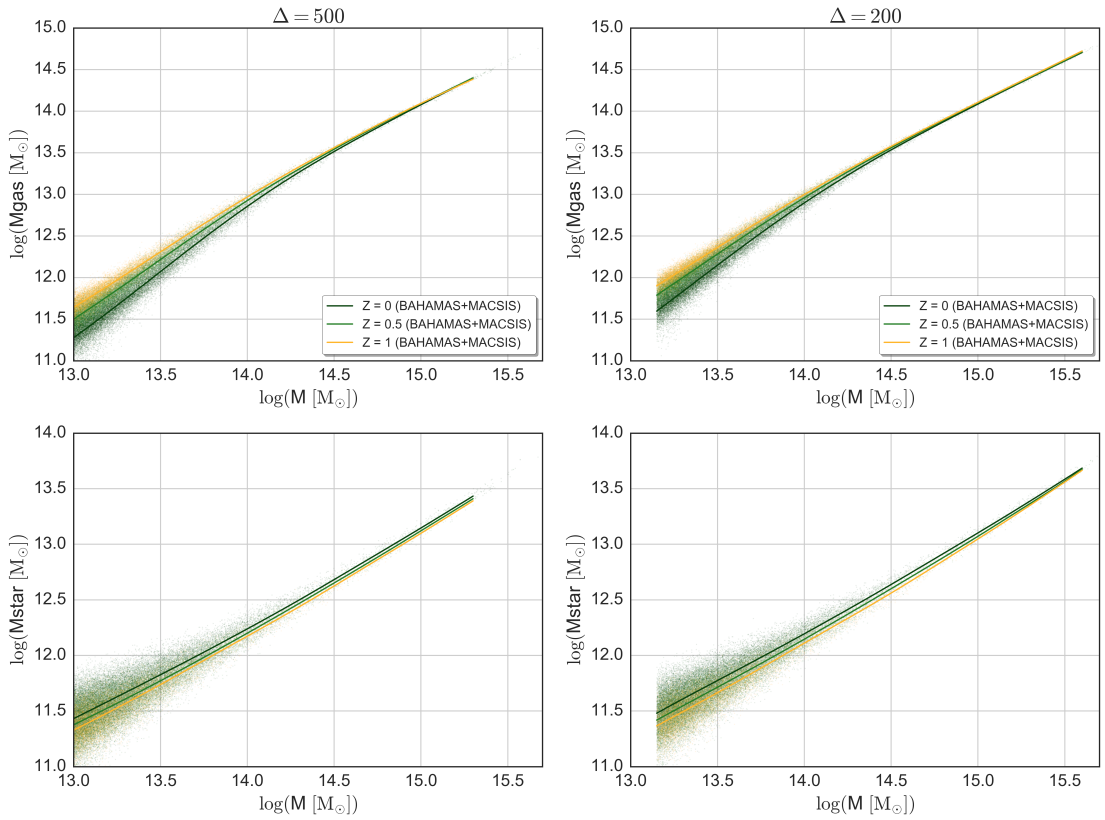

Figure 3.1: Halo baryon contents (points) measured within over-densities, $\Delta = 500$ (left) and 200 (right), for $M_{\rm gas}$ (top) and $M_{\rm star}$ (bottom) as a function of total halo mass at three redshifts indicated in the legend. Lines show the LLR fits. Parameters for the $\Delta = 500$ case are shown in Figures 3.2 and 3.3.

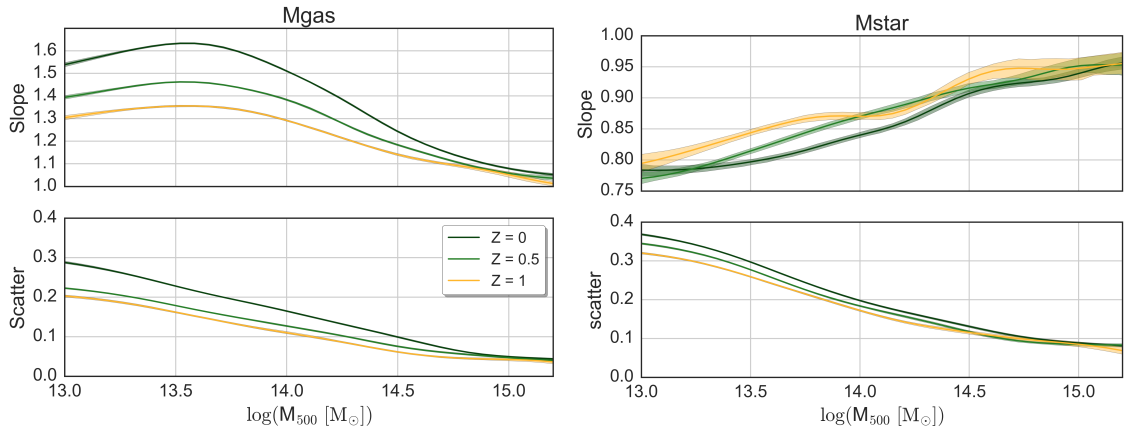

Figure 3.2: Dependence of the slope and scatter of hot gas mass (left) and stellar mass (right) MPRs on total halo mass for $\Delta = 500$. Lines show the LLR estimates and shaded regions give 1σ confidence bootstrap errors in the parameters. The scatter is the root-mean square of the natural log of the measured property.

Overall, the conditional statistics display similar forms at different overdensities and redshifts, but the fit parameter values depend on scale, redshift and halo mass.

Figure 3.2 shows the mass and redshift dependence of the gas/star LLR slope and rms scatter at $\Delta = 500$. There is strong scale dependence in the slopes of the MPR scalings in both $\mathrm{M_{gas}}$ and $\mathrm{M_{star}}$, with milder redshift dependence. For $\mathrm{M_{gas}}$ both the slope and scatter at fixed halo mass increase at lower redshifts, and the running behavior of the slope is non-monotonic with halo mass, exhibiting a peak value near a group-scale mass, $M_{500} \sim 3 \times 10^{13}\,\mathrm{M_\odot}$. For $\mathrm{M_{star}}$ the redshift sensitivity of the MPR parameters at fixed halo mass is more modest, and the slope at tends to slightly decrease toward lower redshifts. The running of the $\mathrm{M_{star}}$ slope is approximately linear in the log of halo mass.

In the BAHAMAS simulation study of Le Brun et al. (2017), a broken (piece-wise constant) power-law is used to fit the scaling of hot gas mass with halo mass. The broken power-law approach introduces a particular mass scale — the transition, or break, mass — that is not anticipated by the relatively smooth astrophysical processes operating within halos. The LLR approach enables the detection of continu-

ously varying, scale-dependent features without introducing an arbitrary halo mass scale. Indeed, the smooth behaviors of the local slopes in Figure 3.2 do not support a broken power-law approximation for either hot gas mass or stellar mass.

For cluster-scale systems above $\sim 5 \times 10^{13}\,\mathrm{M}_{\odot}$, the slopes in both gas mass and stellar mass run nearly linearly with log-mass, approaching the naive self-similar expectation of one in the highest mass systems from above and below, respectively. This is in agreement with Barnes et al. (2017) who find a slope ~ 1 when only the most massive systems are considered, but find a steeper slope using the superset of BAHAMAS and MACSIS halos more massive than 10^{14} $\mathrm{M}_{\odot}$.

As hierarchical clustering progresses and halos grow larger and develop deeper potential wells, feedback driven by the central galaxy becomes more confined to the core region, allowing gravity to become dominant and self-similar scalings to recover. The simulations show this type of progression, with slopes at $z = 0$ in $\mathrm{M}_{\mathrm{gas}}$ and $\mathrm{M}_{\mathrm{star}}$ lying within 1.00 ± 0.05 at masses, $M_{500} > 10^{15}$ $\mathrm{M}_{\odot}$. Furthermore, for the highest-mass systems, the MPR parameters do not vary significantly with redshift, but there are statistically significant changes in the slope and normalization for group-scale systems. The above trends persist at both overdensity scales presented in this work. We confirm, but do not present here, similar behavior at $\Delta = 2500$.

Figure 3.3 shows the scale and redshift behavior of the $\Delta = 500$ LLR normalizations for stellar and hot gas masses. The normalizations are presented as halo mass fractions normalized by mean cosmic baryonic fraction. Recall that we have aligned the MACSIS cosmic baryon fraction to that of the BAHAMAS simulation.

Above a halo mass of $\sim 3 \times 10^{14}\,\mathrm{M}_{\odot}$, the total gas mass and stellar mass fractions become nearly constant; however, there is strong mass and redshift evolution for lower mass systems. The nearly fixed high mass behavior provides strong evidence

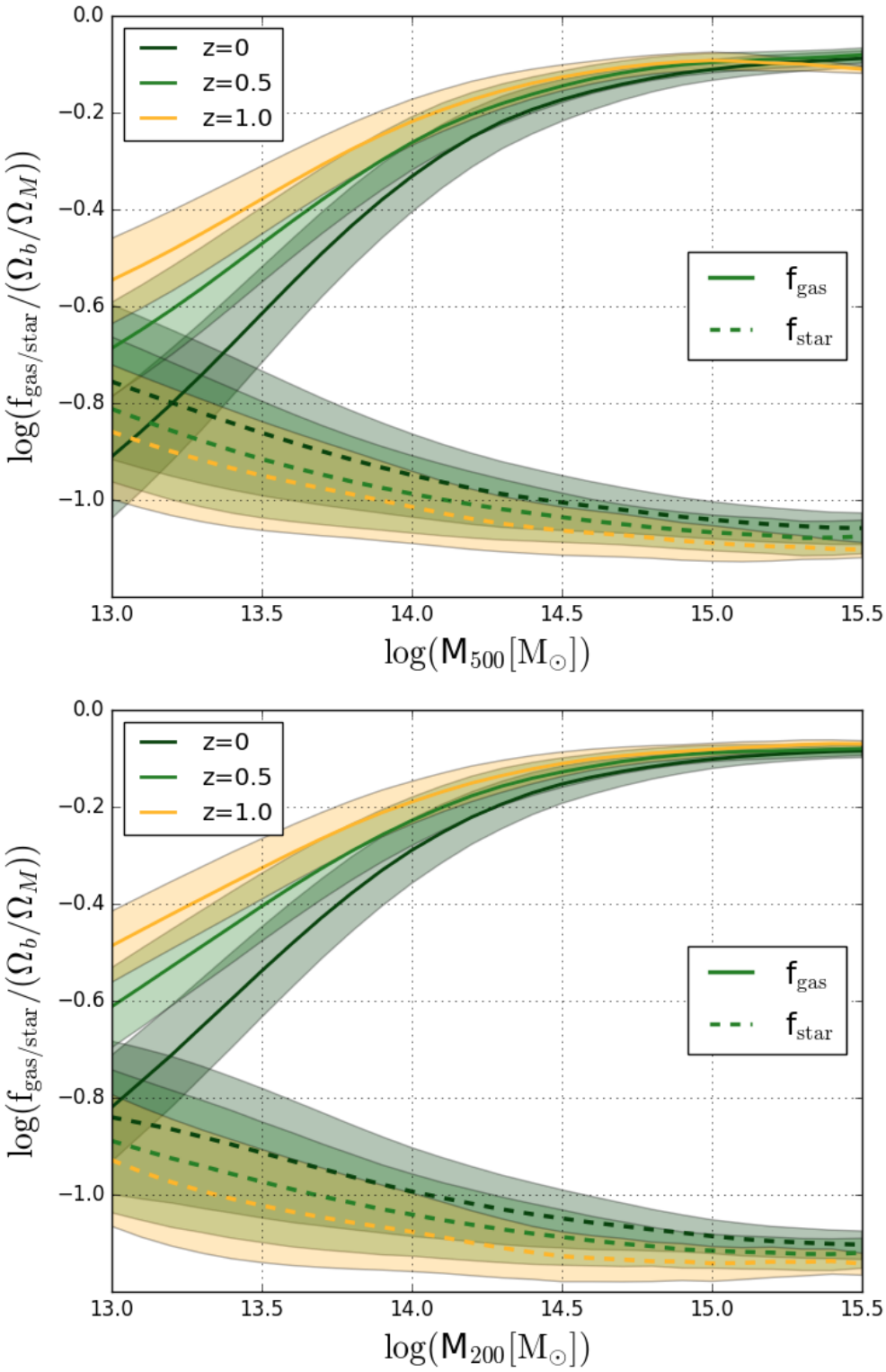

Figure 3.3: LLR normalizations of hot gas mass (solid) and stellar mass (dashed), expressed as mass fractions, $f_a = e^{\pi_a(\mu,z)}/M$, where $\pi_a(\mu, z)$ is the scale- and redshift-dependent log-mean, equation (3.1), normalized by the cosmic mean baryon fraction of the BAHAMAS universe. Shaded regions show the intrinsic scatter within the population rather than uncertainty in the mean behavior. The top and bottom panels are for $\Delta = 500$ and 200 respectively.

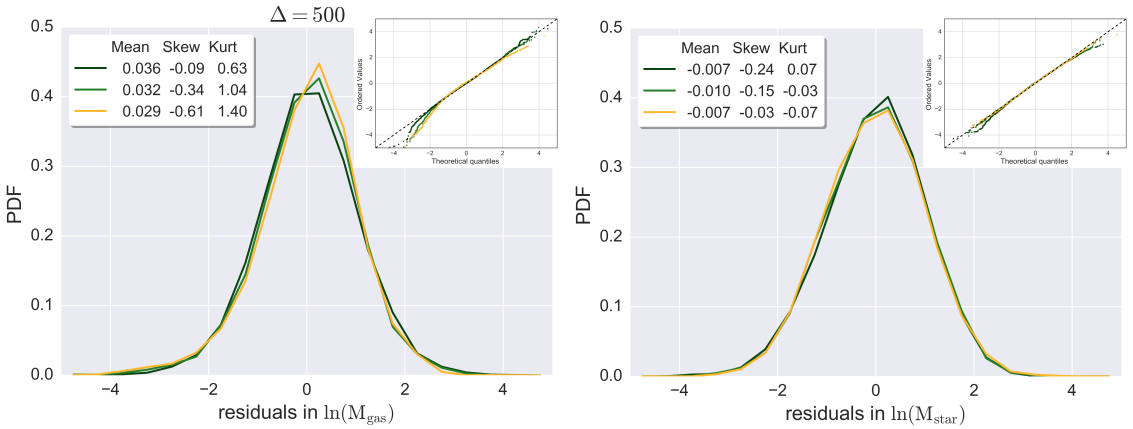

Figure 3.4: Conditional likelihood distribution derived from scaling relation residuals, equation (3.7) in hot gas mass (left) and stellar mass (right). Colors indicate redshift as in Figure 3.1. The mean bias is typically less than 1%, skewness is less than 1, and kurtosis is less than 5 which are strong indicators of log-normality. Rank (Q-Q) comparison, shown in the inset of each panel, indicate only mild deviations in log-normality in the wings of each distribution.

that baryon venting is negligible, while considerable venting occurs at the mass scale of groups. The weak redshift dependence at high mass is in good agreement with trends observed from a joint analysis of South Pole Telescope (SPT) and Dark Energy Survey (DES) data in a sample of 93 massive SPT clusters (Chiu et al. 2017).

The interplay between cooling and feedback controls the relative mean proportions of the integrated gaseous and stellar masses in a way that introduces considerable variance at the group mass scale, but the variance decreases for richer clusters with deeper potential wells. Associated with this, the covariance of gas and stars determines the scatter in overall baryon content. We find evidence for a "closing box" scenario at the high-mass end, with increasing anti-correlation of stellar mass and gas mass at later times. We present this result in Section 3.4.3.

3.4.2 Log-normality of conditional statistics

The log-normal shape of conditional statistics, an implicit assumption in previous analyses, is a core ingredient of the E14 population model. In the context of modeling star formation, a log-normal shape for final stellar masses is expected when random

multiplicative factors govern the evolution of the system (e.g. Larson 1973; Adams & Fatuzzo 1996). Observational studies of galaxy clusters broadly support this form, although with currently modest sample sizes (e.g., Pratt et al. 2009; Mantz et al. 2010; Czakon et al. 2015; Mantz et al. 2016a).

Non-Gaussian terms in MPR statistics can introduce bias in cosmological analysis based on cluster counts (Erickson et al. 2011; Weinberg et al. 2013). Such terms cannot be characterized through measurement of the scatter alone. We use the large BAHAMAS halo samples to study the PDF shape in detail, and assess the degree to which conditional property statistics of the simulated halo sample follow a log-normal frequency distribution.

Previous simulation studies have addressed this issue with generally smaller samples. Using an ensemble of N-body and non-radiative hydrodynamics simulations, Evrard et al. (2008) show that the PDF of dark matter velocity dispersion at fixed halo mass is very close to log-normal, with some samples showing a modest skew caused by a minority population of post-merger, transient systems. The construction of the BAHAMAS and MACSIS halo samples effectively filters out the small fraction of such secondary objects. Stanek et al. (2010) demonstrate log-normal PDFs for multiple properties within a sample of ~ 4000 halos drawn from the Millennium Gas Simulations, as do other hydrodynamic simulations with smaller samples (Fabjan et al. 2011; Biffi et al. 2014; Le Brun et al. 2017; Truong et al. 2018).

Given the LLR fit for property s_a (with a a label indicating either $\ln \mathrm{M_{star}}$ or $\ln \mathrm{M_{gas}}$), we calculate the normalized deviation of halo i from the mean relation,

$$\tilde{\delta}_{a,i} \equiv \delta s_{a,i}/\sigma_a(\mu_i) = \frac{s_{a,i} - \alpha_a(\mu_i)\mu_i - \pi_a}{\sigma_a(\mu_i)}, \tag{3.7}$$

where $\alpha_a(\mu_i)$ and $\sigma_a(\mu_i)$ are the local slope and scatter of the MPR evaluated at the total mass of the i^{th} halo (see, Figure 3.2).

Figure 3.4 presents the PDF of the normalized residuals of gas mass (top panels) and stellar mass (bottom panels) for $\Delta = 500$ at $z = 0$, 0.5 and 1. These results are consistent for all overdensities. The inset of each panel provides a Q-Q plot[1] to illustrate deviations from the normal form. The residuals in the log of stellar mass are extremely Gaussian, while the gas mass displays slight negative skewness and non-zero kurtosis. We note that only a small fraction halos, $< 1\%$, are outliers with low gas mass. Understanding the physical causes of this minor deviation from normality lies beyond the scope of this work. The Gaussian form persists for both $\mathrm{M_{gas}}$ and $\mathrm{M_{star}}$ and over all over-density scales considered in this work.

These results provide strong evidence that the log-normal form is adequate to model the *intrinsic* quantities of halos. In Section 3.5 we demonstrate that employing a local form of the E14 model achieves sub-percent accuracy in estimating the population mean mass selected on baryon mass.

Within the scope of cluster cosmology, non-Gaussian MPR shapes were formulated by Shaw et al. (2010) in terms of an Edgeworth series expansion,

$$P(M_{\mathrm{proxy}}|M_{\mathrm{true}}) \approx G(x) - \frac{\gamma}{6}\frac{\mathrm{d}^3 G}{\mathrm{d}x^3} + \frac{\kappa}{24}\frac{\mathrm{d}^4 G}{\mathrm{d}x^4} + \frac{\gamma^2}{72}\frac{\mathrm{d}^6 G}{\mathrm{d}x^6}, \tag{3.8}$$

where the skewness, γ, is defined as,

$$\gamma = \frac{\langle (M_{proxy} - M_{true})^3 \rangle}{\sigma^2}, \tag{3.9}$$

and the kurtosis, κ, is defined as,

$$\kappa = \frac{\langle (M_{proxy} - M_{true})^4 \rangle}{\sigma^4} - 3, \tag{3.10}$$

and $G(x)$ is a Gaussian distribution. We note that achieving sub-percent level systematic uncertainty in cluster number counts under a log-normal approximation with

[1]The quantile-quantile (Q-Q) plot is a visualization technique for determining if a population sample comes from an assumed distribution. Axes compare rank quantiles of the model to quantiles of the sample.

a mass proxy having 20% scatter requires roughly $\gamma < 7$ and $\kappa < 90$ (see, equation (156) of Weinberg et al. 2013). The skewness and kurtosis values for our halo samples are at least an order of magnitude smaller than what is needed to achieve sub-percent uncertainty in number count statistics, but more work is needed to confirm this result for realistic cluster samples.

In principle, if the form of an observable conditional statistics at fixed halo mass is known, it can be easily incorporated into a cosmological analysis without introducing additional source of systematic error due to the uncertainty in the form of distribution. When modeling observational data, the form of the conditional statistics of measured quantities may differ from a log-normal form, for example due to projection effects (e.g., Cohn et al. 2007; Erickson et al. 2011). Analysis of such data using a log-normal assumption in the likelihood leads to systematic biases in halo mass that in turn can bias cosmological parameter constraints. These additional uncertainties are strongly dependent on survey characteristics and data reduction pipeline and so must be modeled explicitly (e.g., Juin et al. 2007; Farahi et al. 2016; Pacaud et al. 2016; de Haan et al. 2016).

3.4.3 Stellar–hot gas covariance

A complete multi-wavelength MPR likelihood model will include property covariance. For cosmology, knowledge of property covariance improves dark energy constraints when performing analysis of joint, multi-wavelength cluster samples (Cunha 2009). For astrophysical studies, Nord et al. (2008) demonstrate how covariance between temperature and luminosity can confuse studies of luminosity-temperature redshift evolution. Covariance of observed hot gas properties has recently been measured in X-ray selected samples (Mantz et al. 2010, 2016a; Andreon et al. 2017).

In simulations, a covariance matrix of dark matter and hot gas properties was first

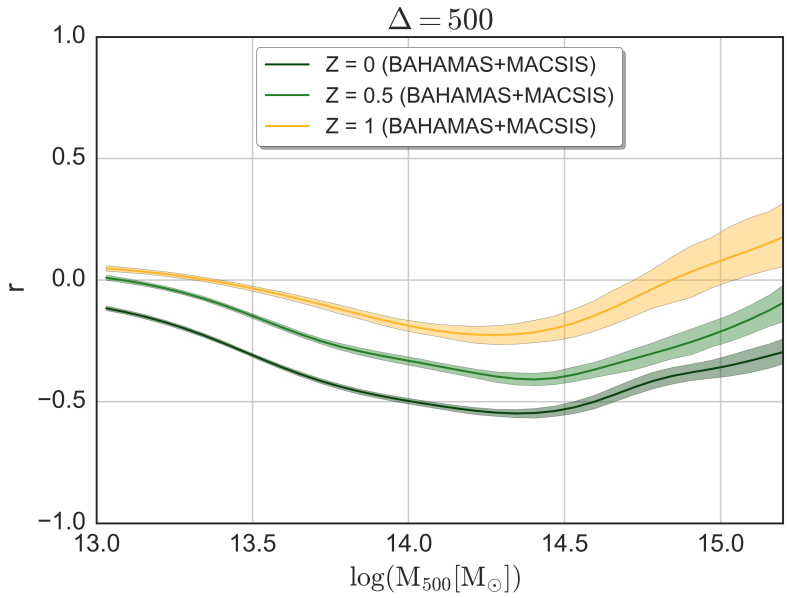

Figure 3.5: The LLR correlation coefficient between stellar mass and gas mass at fixed halo mass, equation 3.4 at the redshifts indicated. Anti-correlation is favored at low redshifts and masses above $10^{14}\,M_{\odot}$.

presented by Stanek et al. (2010) for halo samples in the Millennium Gas simulation. Based on a small sample of high mass halos and their progenitors run with RAMSES hydrodynamics including AGN feedback, Wu et al. (2015) published the first non-zero correlation of hot gas and stellar mass fractions. We perform a similar measurement here on a much larger sample of halos evolved with an independent numerical method.

The correlation coefficient of gas and stellar mass at fixed total mass, equation 3.4, is plotted as a function of halo mass in Figure 3.5. The color scheme is consistent with that used in Figure 3.1. The correlation coefficient begins near zero at $10^{13}\,M_{\odot}$ and becomes increasingly negative at higher halo mass. The values plateau around $3 \times 10^{14} M_{\odot}$ and decline in amplitude for the highest mass halos. The results at $\Delta = 200$ follows a similar pattern as $\Delta = 500$.

The lack of correlation for group size halos can be explained through an "open box" scenario in which the total baryonic content of a halo is not conserved. Feedback effects at low masses are efficient at venting material out of the relatively shallow

potential well. As shown by McCarthy et al. (2011), the gas ejection takes place at high-redshifts, $2 \lesssim z \lesssim 4$, in the progenitors of present-day groups. The ejection is sufficiently energetic that the gas is not re-accreted later on. For higher mass halos, however, the gas is re-accreted. The anti-correlation above $10^{14}\,\mathrm{M}_{\odot}$ is indicative of a more "closed box" nature in which the overall baryon fraction of halos more closely resembles the global value, Ω_b/Ω_m. The redshift behavior in Figure 3.5 indicates that the box is closing more tightly over time, with the extremal value of r decreasing from -0.25 at $z = 1$ to -0.5 at $z = 0$.

Wu et al. (2015) find a correlation coefficient of -0.68 at $\Delta = 500$, stronger than what is found here. The different behaviors appears are likely due to the smaller variance in stellar mass in the BAHAMAS and MACSIS samples for the most massive systems, $\gtrsim 10^{15}\mathrm{M}_{\odot}$. We return to this issue in more detail in Section 3.6.

Another application of the covariance is in estimating the expected value of two properties, for example $\langle \log \mathrm{M_{star}} | \log \mathrm{M_{gas}} \rangle$. According to the E14 analytic model, which will be discussed in detail in Section 3.5, the expected value stellar mass about fixed gas mass requires an estimation of the correlation coefficient between the two property about fixed halo mass. Figure 3.6 illustrates the effect of ignoring this covariance. Figure 3.6 shows the estimated $\langle \log \mathrm{M_{star}} | \log \mathrm{M_{gas}} \rangle$ excluding (dashed line) and including (solid line) the correlation coefficient. We note that excluding this factor leads to a few percent bias, which is proportional to the full covariance, i.e. the gas mass scatter, stellar mass scatter, and the correlation coefficient.

3.5 Validating the analytic population model

Cluster population statistics are linked to the constituents of the universe through the growth of cosmic structure, and many ongoing and future cluster surveys are

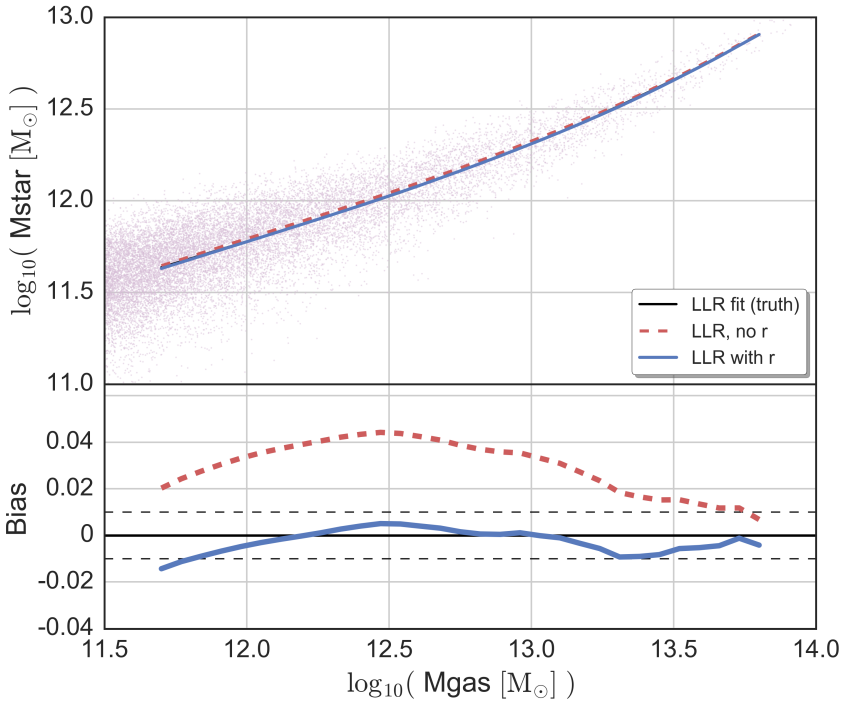

Figure 3.6: The effect of correlation coefficient on estimating $\langle \log \mathrm{M_{star}} | \log \mathrm{M_{gas}} \rangle$ for halos in the BAHAMAS simulation at redshift zero. The red (dashed) line is the predicted $\langle \log \mathrm{M_{star}} | \log \mathrm{M_{gas}} \rangle$ assuming the correlation coefficient of zero; and the red (dashed) line is the predicted $\langle \log \mathrm{M_{star}} | \log \mathrm{M_{gas}} \rangle$ assuming the correlation coefficient estimated in 3.5.

focused on using cluster population statistics to constrain models of dark energy and cosmic acceleration (e.g., Mantz et al. 2015; de Haan et al. 2016; Mantz et al. 2016a; Dark Energy Survey Collaboration et al. 2016; Pierre et al. 2016). The multi-property space density and conditional statistics of the population of massive halos are essential ingredients of such efforts. The evidence presented above indicates that the BAHAMAS and MACSIS halo populations obey the log-normal statistics assumed by the E14 analytic model. In this section we explicitly test the accuracy of that model by examining the expected log-mass of halos, $\langle \ln M | s_a \rangle$, selected by an intrinsic property, s_a. This model is introduced in Chapter II. However, for the sake of completeness, we remind some of the relevant elements and equations. For the complete set of mathematical expressions, we refer the reader to Chapter II.

The smoothness of the mass function allows a logarithmic polynomial expansion,

$$\frac{\mathrm{d}n(\mu, z)}{\mathrm{d}\mu} = \exp\left[\beta_0(z) - \sum_{j=1}^{3} \frac{\beta_j(z)}{j!}\, \mu^j\right] \quad , \tag{3.11}$$

consisting of an amplitude, $e^{\beta_0(z)}$ and linear through cubic coefficients, $\beta_j(z)$, that control the shape. These coefficients vary smoothly with redshift. Figure 3.7 shows the differential number counts as a function of halo mass for redshift $z = 0$ slices as points, and the corresponding mass function fits as lines. We analyze the $z = 0$ sample and fit the number counts of halos to the above third-order polynomial. To estimate the β's we fit a third order polynomial to the $\ln \frac{\mathrm{d}n(M,z)}{\mathrm{d}\ln M} - \ln M$. We find values of $\beta_0 = 8.42$, $\beta_1 = 2.93$, $\beta_2 = 0.86$, and $\beta_3 = 0.42$ [2]. The β_1 term represents Eddington bias from convolution of a pure power-law mass function. Generally, the slope of the mass function lies in the range $\beta_1 \in [2, 4]$, the curvature term $\beta_2 \simeq 1$, and the variance ranges from $(0.05)^2$ to $(0.3)^2$ (see Fig. 3.2).

[2]Note that the β_1 and β_2 terms in E14 are the *local* first and second derivatives of HMF evaluated at a pivot mass, while the β_1 and β_2 in this work are derived from fitting the halo mass function over the mass range shown in Figure 3.7.

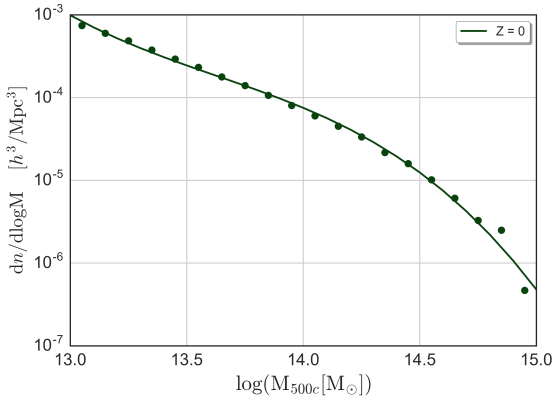

Figure 3.7: The halo mass function derived from the BAHAMAS simulation. The line is a third-order polynomial fit to the data points, equation (3.11), for redshift $z = 0$.

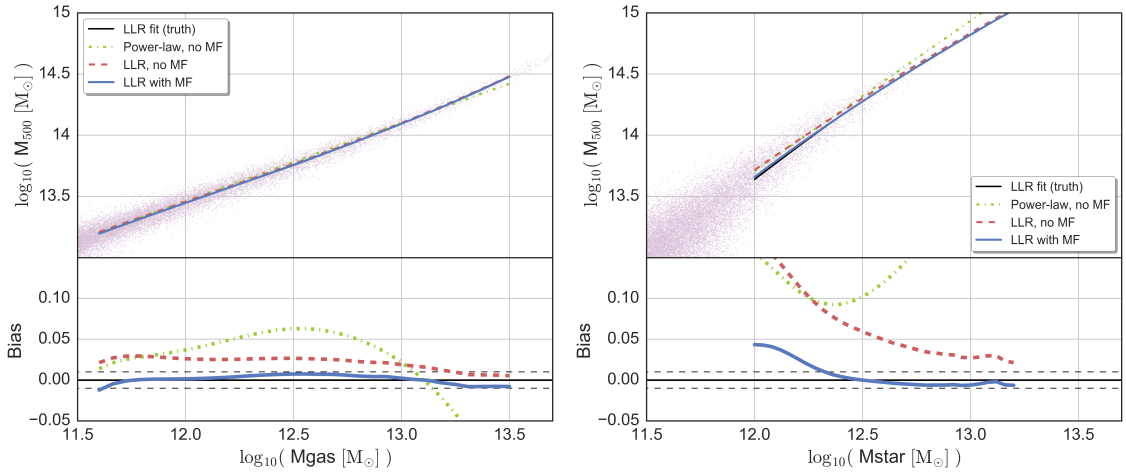

Figure 3.8: Tests of the E14 model for halos selected by hot gas mass (left) and stellar mass (right). In each panel the upper sub-panels show the total halo mass of individual halos as a function of the selection mass, with black curves showing the LLR estimates of the underlying true $\langle \ln M_{500} | s_a \rangle$ relation, where $s_a = \ln \mathrm{M_{gas}}$ or $\ln \mathrm{M_{star}}$. The red dashed (green solid) lines are predictions from inverting the global (local) MPRs, ignoring Eddington bias, while the blue lines show E14 model expectations that include the mass function convolution at second order. The lower sub-panels show the bias in the estimated halo mass, with dashed black lines showing $\pm 1\%$ accuracy with respect to the LLR true estimate.

The model estimate can be compared to the true log-mean halo mass in the simulations. To determine the underlying "true" values of $\langle \mu | s_a, z \rangle$, we perform the inverse LLR fit to that used above, meaning we fit for the mean total halo mass, M_{500}, as a function of either stellar mass or gas mass. We perform this regression above $\mathrm{M_{star}} = 10^{12}\,\mathrm{M_\odot}$ and $\mathrm{M_{gas}} = 4 \times 10^{11}\,\mathrm{M_\odot}$. The results are shown as black lines in the upper panels of Figure 3.8.

The lower panels of Figure 3.8 show the accuracy of various estimates compared to the direct LLR fits. Green lines show the naive estimator, $\langle \mu | s_a, z \rangle = (s_a - \pi_a)/\alpha_a$, using best fit with constant slopes over halos with total masses $> 10^{13}\mathrm{M_\odot}$. This naive estimator, which ignores both the mass dependence of the slope and the Eddington bias, struggles to achieve mass accuracy at the level of 10%.

Red dashed lines improve on this naive estimate by using the local slope from the LLR model, Figure 3.2, while still ignoring the Eddington correction. This model is an improvement but it does not reach percent-level mass accuracy, given by the horizontal dotted lines in the lower panels of Figure 3.8.

Applying the full expression of equation (2.67), with the bias term and local estimates of the slope and scatter, leads to the blue line in Figure 3.8. This estimate recovers the true mean mass within 1% for selection by $\mathrm{M_{gas}}$ over the entire mass range shown.

Equation (2.67) is similarly accurate for selection by $\mathrm{M_{star}}$ above a stellar mass of $10^{12.3}\mathrm{M_\odot}$. Below this the error grows, approaching a 5% bias at the lowest stellar masses. In halos near $10^{13}\mathrm{M_\odot}$ that host poor groups of galaxies, the scatter in cumulative stellar mass within halos is large, $\sigma \simeq 0.3$. The equivalent mass scatter at fixed $\mathrm{M_{star}}$, given by $\sigma_\mu = \sigma/\alpha$ is larger, $\sigma_\mu \simeq 0.4$, since the LLR slope is sub-linear, $\alpha \sim 0.8$. The magnitude of the bias correction, proportional to the MPR

variance, is largest for the low-mass halos selected by M_{star}. In addition, there may be some non-Guassianity beginning to appear in $p(M_{star}\ M_{halo})$ at these low masses, as close inspection of Figure 3.8 indicates.

What we have shown is that simple properties of simulated halos, namely M_{gas} and M_{star}, follow the E14 model form at a level sufficient to achieve sub-percent accuracy in estimated log-mean total halo mass. The test here, involving intrinsic *halo* properties, $\mathbf{S}^{int}$, measured directly within the simulations, is a prelude to more realistic tests using mock observables. Projection and telescope/instrument effects introduce an extra convolution, $p(\mathbf{S}^{obs}|\mathbf{S}^{int}, z)$, that may introduce non-Gaussianity into the form of the measured observables, $\mathbf{S}^{obs}$. We defer such survey and instrument-specific studies to future work.

Future work will extend this analysis to include additional observable properties such as X-ray temperature or luminosity. Support for cosmological analysis also requires mapping intrinsic to observed properties in a survey-specific manner, a process that could induce non-Gaussian features into the conditional statistics.

3.6 Discussion

Here we discuss our findings in the context of previous simulation work. We offer some initial thoughts on observations, but leave detailed study of modeling observed MPRs to future work.

3.6.1 Mean MPR behavior

The cosmo-OWLS simulations, precursor to those used here, display hot gas scaling trends similar to those of BAHAMAS and MACSIS simulations. Le Brun et al. (2017) fit the median behavior in mass bins for halos above $10^{13}\,M_{\odot}$ and $0 < z < 1.5$ to both single and broken power law forms. For $\Delta = 500$ they find a single power-law

slope in M_{gas} of 1.32 ± 0.02, intermediate to the values shown in Figure 3.2. Using a break point of $M_{500} = 10^{14}\,M_\odot$, they find a high-mass slope of 1.18 ± 0.02, similar to our LLR values at $3 \times 10^{14}\,M_\odot$. For low masses between the break and sample limit, they find redshift-dependent behavior with a slope of 1.74 at $z = 0$ declining to 1.32 at $z = 1$. The BAHAMAS and MACSIS samples behave similarly; the local LLR slope of the M_{gas} MPR is most sensitive to redshift below $10^{14}\,M_\odot$.

Using an independent smoothed particle hydrodynamics code, Truong et al. (2018) simulate 24 massive halos with astrophysical treatment that includes AGN feedback. While their methods are not directly calibrated to match the observed gas content of clusters, their estimate of the M_{gas} MPR slope is ~ 1.07, near the value found for halo masses $3 \times 10^{14} M_\odot$ in the BAHAMAS and MACSIS simulations.

The IllustrisTNG project (Springel et al. 2018) produces full-physics simulations of 100 and 300 Mpc volumes with a moving-mesh code and an updated feedback model. Pillepich et al. (2018) study the stellar contents of a subset of halos at redshift $z < 1$ derived from the TNG100 and TNG300 simulations. Fitting a single power-law to the total stellar mass MPR around a mass scale of $M_{500} = 10^{14}\,M_\odot$, they find a slope of 0.84, in very good agreement with our findings.

The trend toward a self-similar slope of one in the M_{gas} MPR is supported by the observational sample of relaxed, high mass clusters by Mantz et al. (2016a). Using weak lensing masses, they find a slope of 1.04 ± 0.05 in the $M_{gas} - M_{WL}$ relation for 40 clusters with $kT_X > 5$ keV. Studies of lower mass clusters typically find super-linear scaling of gas mass with halo mass, such as the slope of 1.22 ± 0.04 found by Lovisari et al. (2015) for a sample of 82 clusters. Nevertheless, a fair comparison between simulation results and observational study should include various systematic and observational effects ignored in this analysis, such as selection effect of clusters and

projection effects.

3.6.2 Diagonal elements of the property covariance

The intrinsic scatter in the MPR for a certain property sets its quality as a proxy for total halos mass. Among observable X-ray properties, it has previously been noted that M_{gas} has low scatter in both observations (Okabe et al. 2010; Mantz et al. 2016a) and hydrodynamic simulations (Stanek et al. 2010; Le Brun et al. 2017; Barnes et al. 2017; Truong et al. 2018).

For cosmo-OWLs, Le Brun et al. (2017) find a scatter of 0.11 in M_{gas} at fixed halo mass of $10^{14}\,M_\odot$ at $z = 0$, which agrees well with our results. They find redshift and mass trends similar to those found here. Wu et al. (2015) find M_{gas} scatter of 0.08 in the Rhapsody-G simulations of ten massive halos, including their progenitors. Truong et al. (2018) find a somewhat smaller scatter of 0.06 in their sample of 24 halos.

We note that the scatter derived in this work is an intrinsic halo property whereas the observational data are measured in a projected space. Given the incoherent nature of projections, the scatter derived from observational data should be larger that the intrinsic values derived in this work. For instance, Mantz et al. (2016a) find 0.09 ± 0.02 for M_{gas} for halos above $3 \times 10^{14} M_\odot$ which is marginally larger than what is found in this work.

On the scatter in overall stellar mass at fixed halo mass, relatively little work has been published from either simulations or observations. Pillepich et al. (2018) find scatter of 0.16 in M_{star} the TNG100 and TNG300 simulations for halos $\sim 10^{14}\,M_\odot$, in good agreement with the BAHAMAS and MACSIS results. A more detailed comparison is needed to compare trends with mass and redshift more precisely. In the Rhapsody-G sample, Wu et al. (2015) find M_{star} a larger scatter of 0.34 in a combined sample

comprised of ten massive halos at $z = 0$ and their progenitors at $z = 0.5$ and 1.

The previous observational constrains on the scatter of stellar mass at fixed halo mass have been consistent with our findings (e.g., Andreon 2010, 2012). In a different work, Zu & Mandelbaum (2015) combine the galaxy stellar mass function with galaxy-galaxy lensing and galaxy clustering from a sample of Sloan Digital Sky Survey (SDSS) clusters and find a scatter in the natural log of central galaxy stellar mass of 0.4 for clusters with masses near $10^{14}\,M_{\odot}$. They also find statistically significant evidence in favor of the scatter in M_{star} decreasing with increasing halo mass, but this refers only to the central galaxy, not the total stellar content.

3.6.3 The off-diagonal element of the property covariance

In contrast to the diagonal elements which determine the mass proxy quality of individual properties, the off-diagonal covariance elements of the joint property matrix have received far less attention.

The results presented in Section 3.4.3 are from hydrodynamics simulations that have been carefully calibrated to reproduce the observed mean relations between gas mass and halo mass and stellar mass and halo mass. While model-dependent, these theoretical predictions are testable empirically with current and future multi-wavelength survey data.

The Rhapsody-G simulation by Wu et al. (2015) established the first estimate of anti-correlation between stellar and gaseous content of halos. In this work, we extend their analysis by using a much larger halo sample that extends to galaxy group scales.

In agreement with Wu et al. (2015), we find that the most massive systems are approximately "closed boxes", but our correlation coefficient peaks at a smaller magnitude than the value of -0.68 found in that work. For the group size halos, the

link between the stellar mass and hot gas mass is strongly reduced (see Figure 3.5). This trend is due to more efficient feedback in low mass halos that ejects a significant fraction of the gas from the progenitors of the groups to radii outside R_{500}, which is evident from the change in the normalization of the total baryonic content.

Furthermore, we see redshift evolution in the correlation coefficient toward larger anti-correlation at later times. This evolution might suggest that halos of fixed mass vent their baryonic content more efficiently at high redshift. This interpretation indicate that baryon fractions *increase* with increasing redshift at fixed halo mass. However, this scenario is not supported by the LLR normalizations (Figure 3.3). Instead, we observe increasing scatter at lower redshift for both gas mass and stellar mass at fixed halo mass, which allows more a longer lever arm to support correlation. This increase in the scatter could be the primary factor which explains the observed redshift evolution. Accretion events might be the key in understanding this trend. Massive halos gain mass through merging and accretion, and the rate of accretion declines with redshift (Fakhouri et al. 2010). Due to the stochastic nature of these events, these events add additional "irreducible scatter" which could weaken the strength of anti-correlation.

A key difference between the Rhapsody-G simulation results of Wu et al. (2015) and ours is the scatter in M_{star} at fixed halo mass, which for high mass halos is much larger in Rhapsody-G ($> 30\%$) than BAHAMAS and MACSIS simulation ($< 10\%$). We note that the Rhapsody-G sample combines all halos progenitors into a single sample. The different sample definitions, along with different numerical and modeling treatments for star formation and feedback, are likely both conspiring to create the difference in property correlation behavior.

The return toward zero of the correlation coefficient for high mass systems most

likely has a simple origin: the very small effect of scatter in M_{star}. Comparing Figures 3.2 and 3.3, we see that a typical $10^{15}\,M_\odot$ halo at $z = 0$ will have converted 10% of its baryons into stars, with 75% remaining in hot gas within R_{500}. The *fractional* deviations in these components are 0.1 and 0.05, respectively, meaning the contributions to the baryon fraction scatter are roughly 0.01 for stars and 0.04 for hot gas. These small values leave little room for coupling deviations in gas mass with those in stellar mass. By comparison, the contributions to the baryon fraction scatter at $10^{14}\,M_\odot$ are larger by roughly a factor of two, 0.02 for stars and 0.07 for hot gas.

Put another way, we expect *irreducible scatter* in the baryon content of halos when masses are defined using a simple spherical threshold. Deviations are sourced by the basic nature of the dynamics — collisionless for dark matter and stars but collisional for gas — as well as edge effects introduced by the spherical filter, including choice of center. A measure of this irreducible scatter can be found from the gravity-only models of Stanek et al. (2010), which show a fractional scatter in gas/baryon mass (there are no stars) at fixed halo mass of 0.036 ± 0.001. This value is very close to the level seen in the hot gas phase of BAHAMAS and MACSIS halos above $10^{15}\,M_\odot$.

We remind the reader that these are results from a model-dependent simulation. These predictions await testing by future empirical studies, which will ultimately be capable of constraining the baryon content covariance of clusters with high accuracy.

3.6.4 Observational prospects for stellar-hot gas mass covariance

The historical absence of well-defined, uniform, multi-wavelength cluster samples explains the sparsity of observational attempts to constrain the off-diagonal elements of the property covariance matrix. The few extant studies focus on covariance between X-ray observables (e.g., Mantz et al. 2010; Maughan 2014; Mantz et al. 2016a;

Andreon et al. 2017). To the best of our knowledge, no constraint on the correlation between an optical and X-ray property pair has been reported. Finally, modeling the mapping between cluster observables and intrinsic halo properties is an important task.

A minimum requirement is to obtain both stellar mass and gas mass estimates for a large cluster sample with a well-defined selection function. Uniformity of the sample is a key factor; combining several heterogeneous datasets is not an option due to complexity in modeling the full selection function.

The Local Cluster Substructure Survey (LoCuSS) survey[3] is taking the lead to make such a measurement possible by combining multi-wavelength observables for a well-defined cluster sample of moderate size. The results from the LoCuSS sample and the constraints are presented in Chapter VII of this dissertation. The results in Chapter VII are the first empirical test of the findings of this Chapter.

3.6.5 Sensitivity to Cosmological Parameters

To test whether our findings are sensitive to the underlying cosmology, we analyzed the *WMAP9* cosmology suite of the BAHAMAS simulation at $z = 0$, 0.5, and 1.0. We obtain results in good agreement with results from the *Planck* cosmology. Specifically, we find evidence for a log-normal PDF and see trends in LLR scaling parameters, including off diagonal elements, similar to those we report here. This reaffirms that the log-normal assumption is a sufficient statistical model independent of cosmological parameters. We also note that the actual values for the slope and scatter is not appreciably different from what have been reported in this Chapter.

[3] http://www.sr.bham.ac.uk/locuss/

3.7 Chapter Conclusion

We present population statistics for volume-limited samples of massive halos selected from the BAHAMAS simulation and its high-mass extension, MACSIS. The combination of these two sets of simulations provides large sample sizes across a wide dynamic range in halo mass realized with consistent, sub-grid physics treatments for star formation and feedback from supernovae and active galactic nuclei. We introduce local linear regression to measure conditional statistical properties of stellar mass and hot gas mass given total halo mass, including their covariance. We assess the validity of the log-normal assumption in MPR models, and investigate the accuracy of the multi-property analytical model of E14.

Our main findings are as follows.

- The scalings of $\langle \ln \mathrm{M_{gas}} | \mathrm{M_{halo}}, z \rangle$ and $\langle \ln \mathrm{M_{star}} | \mathrm{M_{halo}}, z \rangle$ with halo mass are well approximated by power laws with running exponents. For clusters with masses above $10^{14}\,\mathrm{M_\odot}$, the local slope and scatter behave monotonically with mass. The local slope and scatter in stellar mass are nearly redshift independent, while the hot gas slope and scatter tend to increase with increasing redshift. Above $5 \times 10^{14}\,\mathrm{M_\odot}$, the behavior approaches simple self-similarity, with slopes approaching one and very small fractional scatter in baryon component masses: 0.04 in hot gas and 0.08 in stellar mass. The component fractional scatter in galaxy groups near $\sim 3 \times 10^{13}\,\mathrm{M_\odot}$ is significantly larger: 0.2 in hot gas and 0.3 in stellar mass.

- The PDF of residuals in gas and stellar mass about the local regression fit is very close to log-normal. The deviations from normality in the intrinsic halo population are too small to bias cosmological constraints from cluster counts,

but further modeling of sample selection effects and of how intrinsic properties map to those observed remains to be done.

- Studying the hot gas and stellar property covariance, we find that massive halos display anti-correlation indicative of a "Closed Box" nature, with the box closing increasingly tighter at later times. The correlation coefficient is suppressed in lower mass halos, which are capable of venting a significant fraction of their baryons outside their virial regions, as well as in the highest mass halos, where small deviations about a small mean contribution in stellar mass has little effect on the overall baryon content of these systems.

- We verify that the model proposed by E14 can predict the expected log total mass of property-selected halo samples with sub-percent accuracy when local MPR scaling parameters are used.

These theoretical predictions need to be confirmed or falsified through empirical evidence from analysis of observational data. Chapter VII presents comparison with an observational study. Future campaigns of multi-wavelength observational studies, such as XXL (Pierre et al. 2016) and DES (Dark Energy Survey Collaboration et al. 2016), have the opportunity to test these predictions and enrich our knowledge of baryon component physics.

CHAPTER IV

A Novel Galaxy Cluster Mass Estimator from Stacked Spectroscopy

Philosophy and Contribution

In this chapter, I develop a new stacked cluster mass calibration technique. I then evaluate accuracy and precision of this technique with a realistic synthetic data catalog. Next, I identify and investigate all potential sources of systematics. Finally I apply this model to a subset of optically-selected cluster sample to get an estimation of the mean conditional halo masses. This chapter is taken from Farahi et al. (2016): "Galaxy Cluster Mass Estimation from Stacked Spectroscopic Analysis". The analysis and plots presented in this Chapter are my own work. The simulations have been performed and the halo and cluster catalogs are generated by the co-authors, and the observational data is taken from the literature.

4.1 Chapter Introduction

As spelled out in chapter II, predicting cluster counts for a given cosmology requires convolving the halo mass function (spatial number density as a function of mass and redshift) with a likelihood function linking observable cluster properties to total halo mass. As a result, the true halo mass of clusters is a crucial element in the methodology of cluster count cosmology.

Ongoing and near-future cosmological surveys are dedicated to identifying clusters for the purpose of studying cosmology and fundamental physics through spatiotemporal counts and other statistical properties of the cluster population. The largest cluster samples are identified using photometric data, through color-based (Gladders & Yee 2005; Koester et al. 2007; Dong et al. 2008; Murphy et al. 2012; Oguri 2014; Stanford et al. 2014; Bleem et al. 2015; Licitra et al. 2016) or photometric redshift-based (Milkeraitis et al. 2010; Durret et al. 2011; Soares-Santos et al. 2011) algorithms.

Because photometric data provides only coarse resolution in redshift, projection of galaxies along the line of sight to a massive halo limits the ability of cluster-finding algorithms to uniquely identify the galaxies that are members of a particular massive halo. Spectroscopic data provides improved distance and mass estimators for group and cluster selection (e.g., Robotham et al. 2011), but projection and miscentering still pose challenges for these methods (see e.g., Duarte & Mamon 2015, and references therein).

These sources of confusion are fundamentally rooted in the fact that clusters and halos are identified in different spaces: sky-redshift or sky-color space for clusters and 3D real space or 6D phase space for halos. Peculiar velocities can blend distinct halos in real space into a single structure in redshift space (e.g., van den Bosch et al. 2004; Biviano et al. 2006; Wojtak et al. 2007; Saro et al. 2013; Duarte & Mamon 2015). In addition, the fact that high mass halos in cold dark matter cosmologies are dynamically evolving at late times means that substructure and mergers can create complex, transient phase-space structure. In simulations, this complexity can confuse assignment of subhalos hosting galaxies to their parent halos (Knebe et al. 2011).

In practice, assigning galaxies as members of either clusters or halos is a matter of convention, defined by application of specific, algorithm-dependent rules to galaxy samples. Regardless of the particulars, the joint likelihood, $P_{\alpha,i}(k)$ that a galaxy, k, is a member of both cluster α and halo i offers a means to map from one space to the other (Gerke et al. 2005).

The total galaxy content, or *richness*, of a cluster can then be considered as a sum of partial contributions from halos closely aligned along a common sightline. In this Chapter, we apply such a membership-matching approach in simulations to build a network linking clusters to halos, with network edges weighted by fractional cluster membership.

We investigate the membership properties of the redMaPPer cluster finding algorithm (Rykoff et al. 2014). The method, which identifies clusters through their red sequence galaxy population, outputs background-corrected membership probabilities (Rozo et al. 2009; Rykoff et al. 2012) to each galaxy in a cluster as well as central galaxy probabilities for up to four cluster members. The method is designed to make optimal use of data from large, multi-color photometric surveys such as the Sloan Digital Sky Survey (SDSS, York et al. 2000) and the Dark Energy Survey (DES, Dark Energy Survey Collaboration et al. 2016). The SDSS redMaPPer cluster catalog (Rykoff et al. 2014) has been extensively studied with multiwavelength data, including comparisons to existing X-ray and Planck satellite Sunyaev-Zel'dovich catalogs (Sadibekova et al. 2014; Rozo & Rykoff 2014; Rozo et al. 2015; Planck Collaboration et al. 2016b).

The latest study in the redMaPPer series uses stacked spectroscopic analysis of cluster member pairwise velocities to investigate photometrically assigned membership probabilities (Rozo et al. 2015, hereafter RMIV). In that work, very good

agreement was found between spectroscopic and photometric definitions of cluster membership after a small number of modest corrections for blue cluster members, correlated line-of-sight structure, and photometric noise.

Using only SDSS data, the RMIV study could not study membership from the perspective of the underlying halo population. Instead, spectroscopic members are defined in velocity space using an assumed Gaussian form for the pairwise velocity probability density function (PDF) of central and satellite cluster members. In this Chapter, we use simulations to link spectroscopic cluster members to the underlying halo population, leading to an estimate of the log-mean matched halo mass.

In Section 4.2, we apply the redMaPPer algorithm to a 10,000 $\deg^2$ synthetic photometric galaxy catalog derived from lightcone outputs of N-body simulations. We then employ a membership-based matching algorithm, described in Section 4.3, to build bipartite graphs[1] in which each cluster links to a set of halos ranked by their fractional member contribution to that cluster, a measure we term membership *strength*. This method is used to deconstruct the stacked pairwise velocity distribution of central-satellite galaxies in Section 4.4.

In Section 4.5, we apply the N-body simulation-based virial scaling of Evrard et al. (2008) to estimate the total mass at fixed cluster richness from the velocity dispersion model of Section 4.4. We show that this dynamical mass recovers the log-mean mass of halos matched by cluster membership to better than one percent. We also test the robustness of our results to the details of the synthetic galaxy population by implementing our analysis on an independent, higher-resolution simulation, populated with a different galaxy prescription. Confounding effects of mis-centering and velocity bias are then discussed. Using current estimates for the magnitudes of

[1] A bipartite graph links two disjoint sets of nodes, U and V, with edges, each of which connects a node in U with one in V. In our case U is the set of clusters and V the set of halos.

these sources of systematic error, in Section 4.6 we estimate the halo mass scale of the RMIV sample using their stacked velocity dispersion measurements. Our results are summarized in Section 4.7.

Chapter's Notation. Unless otherwise noted, our convention for the mass of a halo is M_{200c}, the mass contained within a spherical region encompassing a mean density equal to 200 times the critical density of the universe, $\rho_c(z)$.

4.2 Simulation samples and synthetic cluster catalog

We employ N-body simulations produced with a lightweight version of the Gadget code developed for the Millennium Simulation (Springel et al. 2005). Three simulations, of 1.05, 2.6 and $4.0\,h^{-1}$ Gpc volumes, are used to produce a sky survey realization covering 10,000 deg^2 that resolves all halos above $10^{13}\,\mathrm{M}_\odot$ within $z \leq 2$. We refer to this suite of runs as the Aardvark simulation.

The resultant sky catalog is built by concatenating continuous lightcone output segments from the three different N-body volumes using the method described in Evrard et al. (2002). The smallest volume maps $z < 0.35$, the intermediate maps $0.35 \leq z < 1.1$ and the largest volume covers $1.1 \leq z < 2$. The simulations employ 2048^3 particles, except for the $1.0\,h^{-1}$ Gpc volume which uses 1400^3, and corresponding particle masses are 0.27, 1.3 and $4.8\times10^{11}\,h^{-1}\,\mathrm{M}_\odot$. The Aardvark suite assumes a ΛCDM cosmology with cosmological parameters: $\Omega_m = 0.23$, $\Omega_\Lambda = 0.77$, $\Omega_b = 0.047$, $\sigma_8 = 0.83$, $h = 0.73$, and $n_s = 1.0$. The Rockstar algorithm is used for halo finding (Behroozi et al. 2013a).

4.2.1 Galaxy population and halo membership

Galaxy properties are assigned to particles using the ADDGALS algorithm (Busha et al. 2013; Chang et al. 2015). The algorithm is empirical, using the observed r-band

luminosity function and trend of galaxy color with local environment as input. The method assigns central galaxies to resolved halos, but satellites as well as centrals in unresolved halos are assigned to dark matter particles in a probabilistic manner weighted by a local dark matter density estimate. This density assignment scheme is tuned to match the clustering properties of a sub-halo assignment matching (SHAM) approach applied to a $400\,h^{-1}\,\mathrm{Mpc}$ simulation using 2048^3 particles.

Central galaxies are placed at the center of resolved halos and assigned a velocity at rest relative to the halo's mean dark matter velocity within $R_{200\mathrm{c}}$. We explore the issue of non-zero central galaxy velocities in the analysis below. All other galaxies are assigned the positions and velocities of the corresponding particles to which they are assigned. Note that no particle can host more than one galaxy. The velocity assignment implies that the velocity dispersion of central–satellite pairs is expected to follow the same scaling with halo mass as that identified in the simulation ensemble of Evrard et al. (2008).

Regarding halo membership, our convention is that a galaxy, n, is assigned to one and only one halo. Thus, if galaxy n is assigned to halo j, then the probability that galaxy n belongs to halo i is $P_{\mathrm{halo},i}(n) = \delta_{ij}$. A spherical region of radius $R_{200\mathrm{c}}$ is used when defining halo membership. This region approximately defines the hydrostatic region of massive halos but it does not extend to the outer caustic, or backsplash, edge which contains a mix of infalling and outgoing material (Busha et al. 2005; Cuesta et al. 2008; More et al. 2015). We note that $R_{200\mathrm{c}}$ is similar in scale to the search radius used by the redMaPPer cluster finding algorithm. In regions where two or more halos spatially overlap, the galaxy is assigned to the nearest halo. In the ADDGALS algorithm, galaxies can reside outside of a resolved N-body halo; 13% of $m_i < 19$ galaxies reside beyond $R_{200\mathrm{c}}$ of a resolved halo.

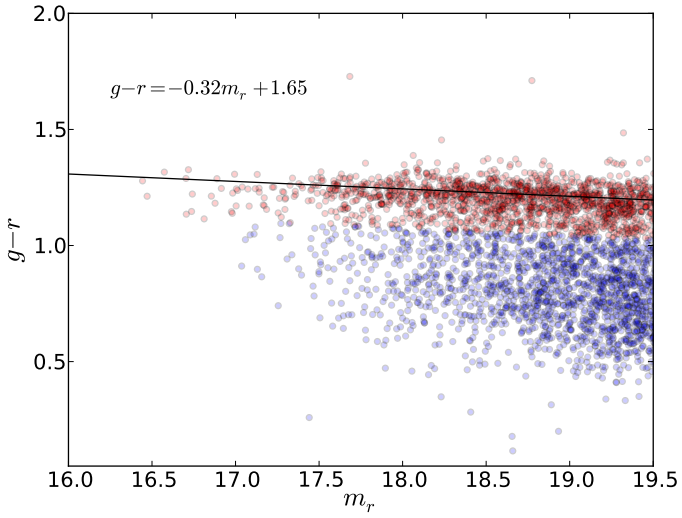

Figure 4.1: Color–magnitude diagram for Aardvark simulation galaxies occupying halos of mass $M_{200c} > 10^{14}\,h^{-1}\,\mathrm{M}_\odot$ in the redshift interval $0.19 < z < 0.21$. The line indicates the red sequence ridge-line, $g - r = 1.65 - 0.32\,m_r$; 78% of galaxies brighter than $m_i = 19$ lie within 0.2 mag of this ridge-line.

While not strictly a halo occupation distribution (HOD) method, ADDGALS produces an effective HOD for which intrinsic richness scales as a power law with halo mass. At low redshift, $\lambda_{\rm int}$, defined as the number of galaxies with $M_r - 5\log h \leq -19$ within R_{200c}, scales with halo mass in a sub-linear fashion, $\lambda_{\rm int} \propto M^\alpha$ with $\alpha \sim 0.8$. To test the robustness of our conclusions to the intrinsic HOD structure of massive halos, we repeat the analysis on the galaxy catalogs of Hearin & Watson (2013) extracted from the Bolshoi simulation, which have a slightly steeper slope, $\alpha \sim 1.0$, and smaller intrinsic scatter in $\lambda_{\rm int}$ compared to the Aardvark galaxy catalog. We find similar results using the Bolshoi simulation.

The redMaPPer algorithm assumes that red galaxies are the prominent population occupying high mass halos. In Figure 4.1, we show the distribution of $g - r$ color as a function of r-band magnitude, m_r, for Aardvark galaxies in halos of mass $M_{200c} > 10^{14}\,h^{-1}\,\mathrm{M}_\odot$, and in the narrow redshift interval, $0.19 < z < 0.21$. A red sequence is evident, containing 78% of galaxies brighter than $19^{\rm th}$ magnitude. The line shows

the ridge-line approximate red sequence population. The slope and intercept are consistent with those found in SDSS analysis of Hao et al. (2009, see their Fig. 11) for the same redshift range.

While the ADDGALS method uses a local dark matter density to assign galaxy luminosity to particles, the smoothing scale employed to calculate the local density leaves the inner $\sim 100\,\mathrm{kpc}$ of high mass halos relatively devoid of galaxies other than the central. As a result of this and possibly other factors, the frequency of mis-centered clusters is larger in the Aardvark redMaPPer cluster catalog than in the observed SDSS sample. We therefore work with two different cluster samples, consisting of the correctly centered subset (denoted CEN) as well as the full set of identified redMaPPer clusters (ALL). The exact definition of these two samples is given in Section 4.2.3.

4.2.2 Cluster finding with redMaPPer

Cluster finding methods that use only optical photometry fall into two main categories based on whether the method uses colors directly or photometric redshifts derived from those colors. The redMaPPer algorithm is in the former category; it uses colors, along with training spectroscopy, to track the multi-band location of the red sequence as a function of redshift (Rykoff et al. 2014). We note that redMaPPer is continuously updated, so there is no unique redMaPPer catalog. Here, we rely on the redMaPPer v5.10 SDSS catalog, as this constitutes the most recently publicly available version.

The redMaPPer cluster finder is a matched filter algorithm with components that characterize the luminosity function, red-sequence color, and projected number density of cluster galaxies. Writing the projected galaxy distribution in sky-magnitude space as a sum of cluster members and a locally-uniform background component, the

algorithm works iteratively to eventually tag each galaxy in the vicinity of a cluster, α, with a probability, $P_{\mathrm{mem},\alpha}$ of being a member of that cluster. The richness, λ, is defined as the sum of the membership probabilities over the set, G_α, of all member galaxies

$$\lambda_\alpha = \sum_{n \in G_\alpha} P_{\mathrm{mem},\alpha}(n)\,. \tag{4.1}$$

The redMaPPer algorithm applied to the Aardvark galaxy sample yields 3927 clusters with $\lambda > 20$ and redshift of $[0.1 - 0.3]$ over 10,400 square degrees. By comparison, there are 4522 clusters in the redMaPPer v5.10 DR8 cluster sample. Figure 4.2 shows differential sky number counts, $\mathrm{d}n/\mathrm{d}z$, in units of number per $10,000$ square degrees, for clusters with $\lambda > 20$ (upper lines) and 80 (lower lines) in the Aardvark and SDSS DR8 samples.

The number of clusters with $\lambda > 20$ in our simulation is lower by $\sim 24\%$ relative to the SDSS DR8 catalogs. While this suppression may partly reflect the underpopulation of the inner $\sim 150\,\mathrm{kpc}$ regions of the most massive simulated halos, which suppresses the membership probability PDF at high P_{mem} values for cluster members, this effect is not the only potential cause. The lower central galaxy density of massive Aardvark halos also makes it more difficult for redMaPPer to center clusters correctly. We note that the simulation matches well the observed trend of increasing counts with redshift. Finally, the difference may reflect differences in the underlying cosmological parameters. The Aardvark simulation has a smaller dark matter density ($\Omega_m = 0.23$) than most current observational constraints, which implies a lower space density at fixed halo mass. The small difference in overall counts does not influence the spectroscopic analysis below.

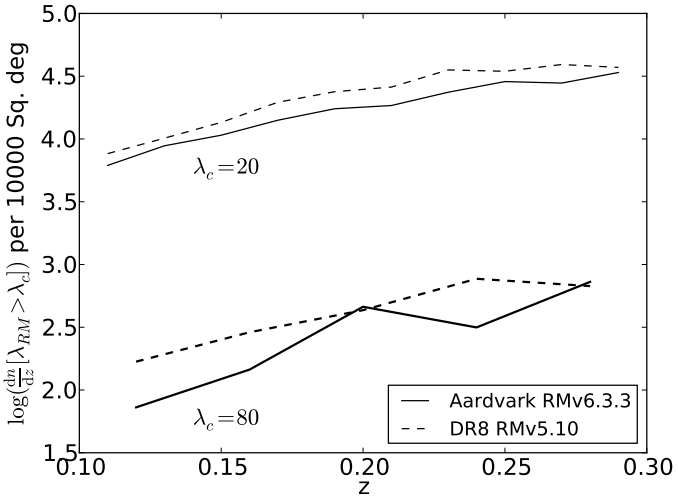

Figure 4.2: Differential sky number counts per $10,000$ square degree of clusters with richness, $\lambda > 20$ (thin lines) and 80 (bold lines) are shown for the Aardvark simulated galaxy catalog run with RMv6.3.3 (solid) and SDSS DR8 run with RMv5.10 (dashed, Rozo et al. 2015) samples.

Table 4.1: Aardvark cluster samples, including the number of redMaPPer clusters, $N_{\rm cl}$, the number of galaxies in the spectroscopic samples, $N_{\rm spec}$, and the number of spectroscopic, central-satellite pairs, $N_{\rm pair}$.

Name	$N_{\rm cl}$	$N_{\rm spec}$	$N_{\rm pair}$	Sample description
ALL	3927	134464	130537	full sample with $\lambda > 20$
CEN	2294	78794	76500	correctly centered subsample of ALL

4.2.3 Cluster and Spectroscopic samples

The full redMaPPer cluster catalogs for both observations and simulated galaxy catalogs consist of clusters with $\lambda > 20$ in the redshift range $z = 0.1$ to 0.3.

To evaluate the sensitivity of our analysis to mis-centering, we identify a correctly centered sub-sample of simulated clusters, those for which the central cluster galaxy is also the central galaxy of the top-ranked, matched halo. Throughout this Chapter, we refer to this correctly centered sub-sample as CEN, and denote the full simulated cluster sample as ALL.

Our spectroscopic membership study is limited to cluster member galaxies with $m_i < 19$. The limit of $m_i < 19$ is a compromise value lying between the SDSS and GAMA limits used by RMIV. Because satellite galaxies in halos trace the dark matter kinematics by construction, our results are not strongly sensitive to the choice of magnitude limit.

Table 4.1 summarizes the number of clusters, number of galaxies, and number of central–satellite galaxy pairs in the simulation samples used below.

4.3 Cluster–Halo membership matching

To match redMaPPer clusters to halos, we build a bipartite network between clusters and halos with edges weighted by joint cluster–halo membership. The network is built using all photometric redMaPPer members of the cluster. Edges are weighted by the membership **strength** between cluster α and halo i, defined as

$$S_{\alpha,i} = \frac{1}{\lambda_\alpha} \sum_{n \in G_\alpha} P_{mem,\alpha}(n) P_{halo,i}(n) \tag{4.2}$$

where $G_\alpha \equiv \{ID\}_\alpha$ is the list of galaxy ID's associated with cluster α, $P_{halo,i}(n)$ is a boolean set to 1 if galaxy n is a member of halo i, as described in Section

4.2. The strength, normalized to lie between 0 and 1, gives the fraction of the total membership of cluster α contributed by halo i.

Recall that λ_α is the cluster richness defined in Equation (4.1). In essence, the measured optical richness of a cluster can be expressed as a series of decreasing halo contributions

$$\lambda_\alpha = \sum_{r=1}^{N} S_{\alpha,i(r)} \,, \tag{4.3}$$

where the halo list, $i(r)$, is rank ordered such that $S_{\alpha,i(1)} \geq S_{\alpha,i(2)} \geq ...S_{\alpha,i(N)}$. The matched halo of a cluster is defined as the halo with the highest strength; we use the terms "matched halo" and "top-ranked halo" interchangeably throughout this Chapter. The mapping is not exclusive; two clusters can be mapped to one halo. In practice this happens infrequently. Out of 3927 redMaPPer clusters of redshift 0.1 to 0.3 only 38 clusters shared top rank halo. These 38 clusters mapped to 19 halos.

Our approach is similar to that of Gerke et al. (2005), who introduced the concept of the largest joint member fraction to map clusters to halos. However, that work uses a boolean measure of cluster membership. The probabilistic approach of redMaPPer makes the strength definition equivalent to the largest group fraction used in Gerke et al. (2005). Note that Rozo & Rykoff (2014) use a similar approach to match pairs of clusters derived from different search algorithms applied to the same SDSS data.

4.4 Pairwise Velocity PDF: Halo Contributions to Spectroscopic Membership

The study of RMIV assessed the validity of redMaPPer photometric membership probabilities by using spectroscopic redshifts. That work models the line-of-sight velocity distribution of central–satellite pairs as a Gaussian distribution with zero mean and a dispersion that scales with cluster richness and, implicitly, with halo

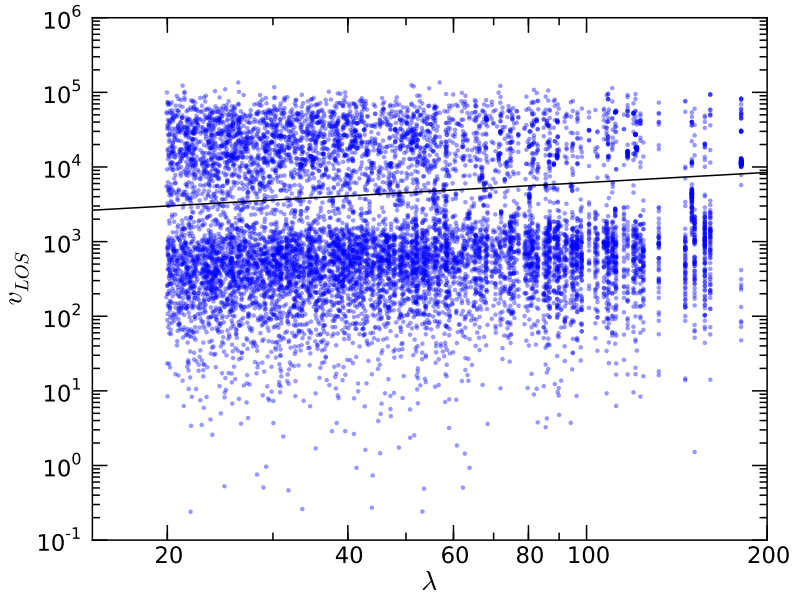

Figure 4.3: The line-of-sight magnitude of central–satellite pairwise velocities for all spectroscopic cluster members in the Aardvark simulation. The line shows the cut applied applied to the SDSS sample by RMIV to separate cluster members (below) from projected contaminants (above). We apply this cut to the Aardvark sample, eliminating $\sim 23(25)\%$ of galaxy pairs from CEN (ALL) samples.

mass. After removing projected pairs having larger than escape velocities, the PDF of the remaining normalized pairwise velocities is modeled as a Gaussian plus a uniform background.

We begin by demonstrating that the simulated galaxy sample displays similar characteristics to the observations. Unlike the observations, our knowledge of the halo membership of each galaxy allows us to deconstruct the spectroscopic likelihood into distinct halo contributions.

4.4.1 Constructing the velocity PDF of cluster central–satellite pairs

Using redshifts of cluster members in the spectroscopic samples described in Section 4.2.3, we determine pairwise velocities of each cluster's satellite galaxies relative to its central galaxy

$$v = c \left(\frac{z_{\rm gal} - z_{\rm cen}}{1 + z_{\rm cen}} \right) , \tag{4.4}$$

where c is the speed of light, and $z_{\rm gal}$ and $z_{\rm cen}$ are redshifts of satellite and central galaxies, respectively. The galaxy redshifts in the simulation are used with zero measurement error. Recall that central galaxies of resolved halos are at rest with respect to their host halo.

In Equation (4.4), the central galaxy is defined by the redMaPPer cluster-finding algorithm. In the CEN sub-sample this is also the central galaxy of the matched halo. For clusters in the CEN sample with high strength, we expect the root mean square velocity to be an unbiased estimate of the dark matter velocity dispersion of the matched halo.

Figure 4.3 shows the distribution of pairwise velocity magnitudes against cluster richness for the ALL sample. The structure is very similar to that found by RMIV for the SDSS+GAMA spectroscopic data (see their Fig. 2), with a main component at low velocities, referred to as the signal, and a cloud of projected pairs lying at high velocities.

We apply the RMIV velocity cut, shown by the line in Figure 4.3, to remove the projected contamination, eliminating $\sim 23(25)\%$ of galaxy pairs in CEN (ALL) sample.

As per RMIV, we model the velocity of a central-satellite pair as a random draw from a Gaussian distribution with a richness and redshift dependent velocity dispersion, σ_v, modeled via

$$\sigma_v(\lambda, z_{cen}) = \sigma_p \left(\frac{1+z_{cen}}{1+z_p}\right)^{\beta} \left(\frac{\lambda}{\lambda_p}\right)^{\alpha} \tag{4.5}$$

where σ_p is the characteristic dispersion at the pivot point, $\lambda_p = 30$ and $z_p = 0.2$ [2], corresponding to the approximate median cluster richness and redshift of our sample, respectively.

[2]In this Chapter, the RMIV normalization is calculated using pivot richness, $\lambda_p = 30$, and redshift, $z_p = 0.2$, slightly different from the published RMIV pivot values.

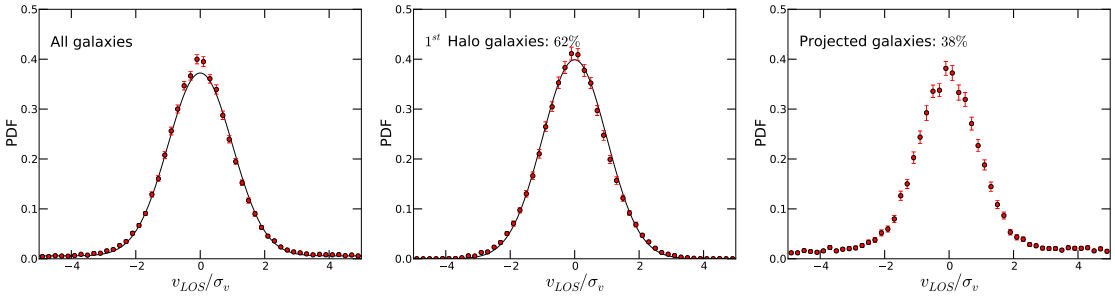

Figure 4.4: **Left:** The PDF of LOS pairwise velocities, normalized according to Equation 4.5, for the correctly centered (CEN) sample of redMaPPer clusters in the Aardvark simulation. The black line shows the best fit likelihood model, Equation (4.6), with parameters given in Table 4.2. **Middle:** Same as left but using only galaxy pairs in the matched (top-ranked) halo associated with each cluster. The black line shows the likelihood model, Equation 4.6, but with $p = 1$. Error bars are 2σ based on bootstrap resampling. **Right:** Velocity PDF of galaxy pairs not belonging to the matched halo.

To incorporate non-physically associated pairs, a flat velocity component is added to the distribution. The likelihood of the stacked velocity distribution is given by the following sum over pairwise velocities, v_i,

$$\mathcal{L} = \prod_{i=1}^{N_{\rm pair}} \left[p\, G(v_i, \sigma_v(\lambda, z)) + (1-p)\frac{1}{2v_{max}} \right] , \tag{4.6}$$

where $G(v_i, \sigma_v(\lambda, z))$ is a Gaussian of zero mean and width $\sigma_v(\lambda, z)$, and p, α, β, and σ_p are free parameters to be determined by maximizing the likelihood. Each v_i is the line-of-sight (LOS) satellite–central pair velocity, Equation (4.4), and the product is over all pairs in the sample.

As we shall see in Section 4.4.3, the fraction of pairs contained in the central Gaussian, given by the parameter p, is not the same as the fraction of cluster members contributed by the top-ranked halo.

4.4.2 Velocity PDF analysis

We maximize our likelihood to recover the scaling relation parameters between cluster richness and velocity dispersion. We assume flat priors on all parameters to calculate the posterior probability, and find the best-fit values given in Table 4.2.

Table 4.2: Best fit parameters of the velocity dispersion model, Equation (4.5), using the likelihood, Equation (4.6) for the simulations (ALL, CEN, and Bolshoi), and the observational data of RMIV. Note that RMIV normalization is calculated at the pivot point, $\lambda_p = 30$ and $z_p = 0.2$, used in this Chapter. The Bolshoi simulation used only the $z = 0$ simulation snapshot so cannot constrain β. The quantity $\langle f_{\rm h1} \rangle$ is the mean fraction of spectroscopic cluster members contributed by the top-rank, matched halo. The notations are defined in equation 4.5 and equation 4.6.

sample	σ_p [km s^{-1}]	p	α	β	$\langle f_{\rm h1} \rangle$
ALL	585 ± 2	0.885 ± 0.002	0.387 ± 0.007	0.83 ± 0.07	0.58
CEN	547 ± 2	0.919 ± 0.002	0.405 ± 0.008	0.87 ± 0.08	0.62
Bolshoi	535 ± 4	0.884 ± 0.003	0.295 ± 0.010	-	0.70
RMIV	598 ± 6	0.916 ± 0.004	0.435 ± 0.020	0.54 ± 0.19	-

The left panel of Figure 4.4 shows the PDF of the pair velocities normalized by the expected velocity dispersion for the CEN cluster sample. The structure of the full sample is similar. We bootstrap the cluster sample to compute means and standard deviations of the PDF in 50 bins between -5 and 5 in v/σ_v, shown as the points with error bars. The black line is a Gaussian of zero mean and unit variance plus the constant distribution, with amplitude given by the best fit model. The model is not a good fit to the data ($\chi^2/dof = 82/16$ over the signal region, $v/\sigma_v \in [2.5, 2.5]$).

We find parameters that are similar to the RMIV fit to the SDSS redMaPPer sample. The CEN sample's Gaussian magnitude, $p = 0.919 \pm 0.002$, and velocity–richness slope, $\alpha = 0.405 \pm 0.008$, are very similar to the SDSS values of 0.916 ± 0.004 and 0.44 ± 0.02, respectively. The ALL sample has reduced magnitude, $p = 0.885 \pm 0.002$ and a slightly shallower slope, $\alpha = 0.387 \pm 0.007$, differences that we discuss further in Section 4.5.2 below.

The velocity normalization, σ_p, is generally $\sim 10\%$ lower than the RMIV value. As we discuss in Section 4.5, non-zero central galaxy velocities, satellite galaxy velocity bias, cosmology, and mis-centering frequency all play a role in setting the normalization.

As found by RMIV, the best-fit model does not have an acceptable χ^2, as re-

flected by the deviations seen in the left panel of Figure 4.4 ($\chi^2/dof = 82/16$). We show below that the deviations from the simple flat-plus-Gaussian model arise from galaxies lying along the line of sight in halos outside the matched halo.

As an independent check that explores the sensitivity of these parameters to the galaxy assignment scheme, we repeat the analysis using measurements at known halo locations of the Bolshoi simulation catalogs of Hearin & Watson (2013). That work uses age distribution matching, a method for predicting how galaxies of magnitude r and color $g-r$ occupy haloes, to populate halos with galaxies at redshift $z = 0$. When using the galaxy catalog from the Bolshoi simulation, we rely on a $z = 0$ snapshot rather than a properly constructed lightcone. We note the Hearin & Watson (2013) catalog has only g and r data available, rather than the full 5-band photometry available in the SDSS and Aardvark.

Results of this exercise produce a velocity PDF of similar shape to the Aardvark CEN sub-sample. The best-fit parameters show a similar Gaussian magnitude, $p = 0.89$, but a shallower slope, $\alpha = 0.30$, that reflects the steeper HOD slope in the Bolshoi galaxy catalog compared to the Aardvark galaxy catalog. However, the results and conclusions remain the same.

4.4.3 Halo-ranked contributions to the velocity PDF

The cluster–halo membership network allows us to determine what fraction of pairs in the main Gaussian PDF component arise from the matched halo. For the CEN sample, we find that, on average, 62% of galaxy pairs arise from within the matched halo. For the full sample, the mean value decreases somewhat, to 58%. For the Bolshoi catalog, in which all clusters are correctly centered by construction, the mean is somewhat larger, 70%.

The middle panel of Figure 4.4 shows only the matched halo's contribution to the

pairwise velocity PDF of the CEN sample. As before, error bars are produced via bootstrap resampling of the cluster sample using 50 bins between -5 and 5 in v/σ_v. The black line shows a Gaussian with dispersion given by the best fit to the entire spectroscopic sample (left panel), listed in Table 4.2. The principal difference with the left panel is that we force $p = 1$, meaning no background component. While there exists moderate kurtosis in this distribution, the high velocity wings of the PDF are not well populated. Relative to the full CEN sample, the goodness of fit is improved by nearly a factor of two ($\chi^2/dof = 47/16$ over $v/\sigma_v \in [2.5, 2.5]$).

The good match seen in the middle panel is important in that it indicates that the best-fit velocity derived from the spectroscopic data set accurately recovers the velocity dispersion of the top-ranked halo. This finding offers leverage for a mean dynamical mass estimate as a function of cluster richness that we explore in the next section.

The right hand panel of Figure 4.4 shows the contribution from satellite galaxies outside of the matched halo. Clearly, a constant background does not adequately capture this component, which is a sum over second and higher-ranked halos. For the CEN sample, an average of 38% of spectroscopic pairs are not contributed by the top-ranked halo. Of this total, an average of 10% and 5% come from the second- and third-ranked halo, respectively. The remaining 23% is contributed by fourth and higher ranked halos, with 12% in unresolved halos below our mass resolution limit. Developing an accurate model for the projected galaxy contribution, while potentially feasible within the context of the halo model, would involve choosing a number of currently uncertain model elements that describe the halo occupation and kinematic biases as a function of galaxy magnitude. We leave detailed modeling of the projected component to future work.

For the full sample (ALL), the overall non-matched halo fraction is slightly higher, 42%, with 12% and 6% arising from the second and third halo terms.

Similar results have been found in prior simulation studies. Using a spectroscopic group finder based on a Voronoi-Delaunay tesselation, Gerke et al. (2005) and Gerke et al. (2012) find that 70% of cluster galaxies truly belong to the matched host halo, on average. Though they use a completely different group finder algorithm, their conclusion regarding the level of interloper galaxies is consistent with the results of our spectroscopic analysis. In a different study, Mamon et al. (2010) finds the density of interloping dark matter particles in redshift space around massive halos takes the form of a constant component plus a quasi-Gaussian component, similar to the structure seen in the right panel of Figure 4.4.

4.5 Mass Estimation

In this section we derive a scaling relation between total mass and optical richness by applying the virial velocity scaling of massive halos to the pairwise velocity dispersion model described above. We compare this stacked dynamical mass to that derived from membership matching to halos, and find excellent agreement with the log-mean matched mass at fixed richness.

We begin by using the CEN sample to avoid uncertainties caused by mis-centering, then investigate mis-centering in Section 4.5.2.

4.5.1 Cluster mass from dark matter virial scaling

The classical virial theorem balances the kinetic energy with (modulo surface terms) half the gravitational potential energy of a halo, thereby offering a scaling law between velocity dispersion and mass within an enclosed radius. In a study of multiple, independent N-body and adiabatic hydrodynamic simulations, Evrard

et al. (2008, hereafter, E08) calibrated the dark matter virial relation.

In that work, the one-dimensional velocity dispersion of a halo, σ_h, is defined in an orientation-averaged fashion using particles within R_{200c}. The dispersion is measured with respect to the mean dark matter velocity within that radius.

E08 showed that the halo velocity dispersion of the population follows a power-law form with approximately log-normal scatter, meaning the conditional probability, $P(\ln(\sigma_h)|M,z) = \mathcal{N}(\ln(\sigma_{\rm DM}(M,z)), 0.046)$, where $\mathcal{N}$ denotes a normal distribution, $\sigma_{\rm DM}(M,z)$ is the log-mean velocity dispersion at fixed mass and redshift, and 0.046 is the scatter in $\ln(\sigma_h)$ at fixed mass.

The log-mean velocity dispersion follows the scaling

$$\ln(\sigma_{\rm DM}(M_{200c},z)) = \pi_\sigma + \alpha_\sigma \ln(h(z)M_{200c}/10^{15}\ {\rm M}_\odot)\,, \tag{4.7}$$

with amplitude $\pi_\sigma = \ln(1082.9 \pm 4.0)$ and slope $\alpha_\sigma = 0.3361 \pm 0.0026$. Here, $h(z) = H(z)/100\,{\rm km\ s^{-1}\,Mpc}$ is the dimensionless Hubble parameter. The ellipsoidal collapse model of Okoli & Afshordi (2016) offers a first-principles explanation of the form and parameter values of this calibration.

At fixed mass, the distribution of velocity dispersion seen in the E08 simulation ensemble is very close to log-normal, with a modest tail to higher values driven by actively merging systems. Saro et al. (2013) show that the 1D LOS velocity dispersion has higher scatter compared to angle-averaged velocity dispersion. The normalization and slope of their scaling relation, found using sub-halos as galaxy tracers, are within $\lesssim 3\%$ of the E08 values.

For a halo ensemble uniformly sampled in mass, the inverse of the above scaling relation provides an unbiased estimate of the log-mean halo mass at fixed velocity dispersion, $P(\ln(M)|\ln(\sigma_h),z)$. For samples drawn from the expected cosmic mass function, the log-mean mass will be biased low by approximately 5%, as detailed in

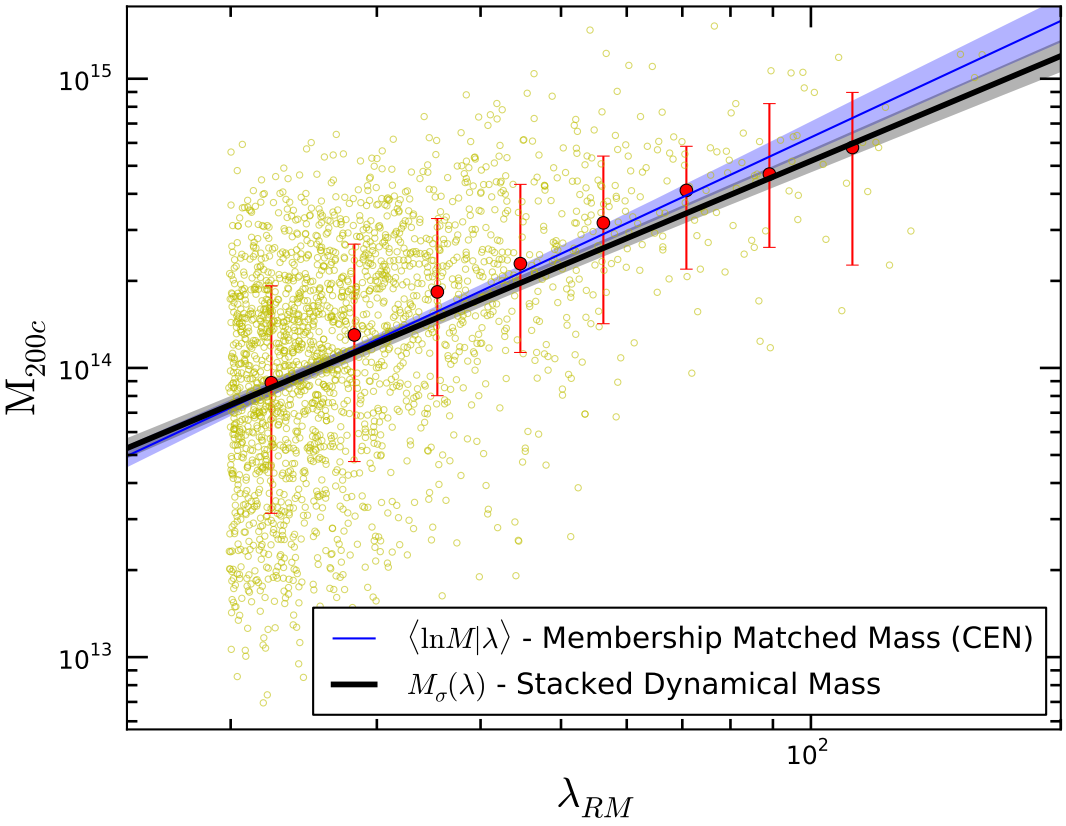

Figure 4.5: The mass–richness scaling relationship derived from application of the virial relation to stacked central satellite velocities, Equation (4.8), (solid black line) at redshift 0.2 is compared to halo masses of correctly-centered redMaPPer clusters derived from galaxy membership matching in the redshift range $[0.1, 0.3]$ (yellow circles). The red dots with error bars show the median and 68% inclusion region of matched halo mass in different richness bins. The blue line is the best fit to the membership-matched masses in this redshift range, with shaded region showing 95% confidence uncertainties in this mean relation at redshift 0.2.

Evrard et al. (2014). The magnitude of this correction is sub-dominant to systematic errors discussed below, so we choose to ignore it in this work.

To estimate halo mass as a function of richness in the redMaPPer cluster population, we apply the inverse to the log-mean halo virial scaling relation found in E08,

$$\ln(h(z)M_\sigma(\lambda,z)/10^{15}\,\mathrm{M}_\odot) \;=\; 3\;\ln\left(\frac{\sigma_v(\lambda,z)}{1083\,\mathrm{km\ s^{-1}}}\right)\,, \tag{4.8}$$

where $\sigma_v(\lambda,z)$ is the velocity dispersion scaling of central–satellite pairs analyzed in Section 4.4 and the simple cubic power is consistent with the slope found in the E08 simulation ensemble.

If intrinsic galaxy richness, λ, were a nearly perfect tracer of halo mass, and if cluster finders cleanly identified halo members, then the log-normal form of the PDF relating velocity to mass (or vice-versa) implies that the virial-scaled mass, $\ln(M_\sigma(\lambda,z))$, should accurately measure the log-mean mass, $\langle\ln(M)|\lambda,z\rangle$, at fixed richness and redshift. Introducing (log-normal) scatter in richness at fixed mass can produce shifts that depend on the local slope and curvature of the mass function as well as the covariance of λ and σ_h at fixed M (Evrard et al. 2014). We defer a more detailed examination of these issues to future work.

Galaxy joint member matching provides an independent mass estimate for each cluster — the matched halo mass — that can used to assess the meaning of the stacked dynamical mass estimate, Equation (4.8).

Figure 4.5, a key result of this Chapter, compares the mass scale inferred from the scaled velocity dispersion with membership matched masses for the CEN sample. The thick black line shows the mass–richness scaling at redshift 0.2 inferred from virial scaling, Equation 4.8, while the points show individual M_{200c} values of matched halos for individual correctly-centered clusters of redMaPPer richness, λ

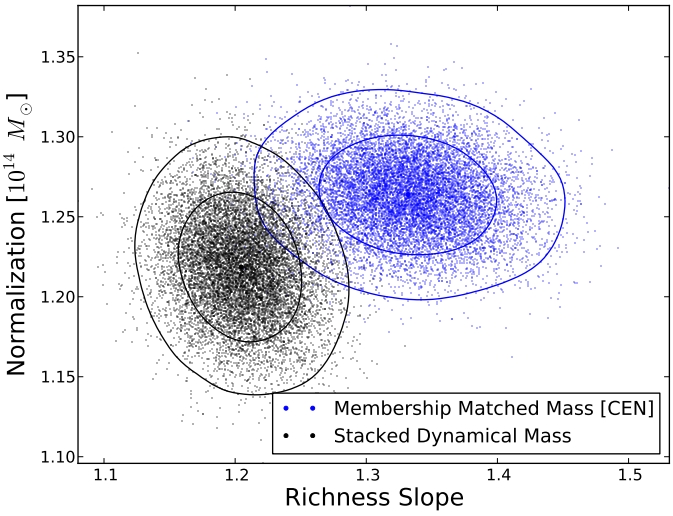

Figure 4.6: The normalization and slope of mass–richness scaling at redshift 0.2 inferred from stacked dynamical masses (black contours) and membership matching in the redshift range $[0.1, 0.3]$ (blue) for correctly-centered redMaPPer clusters. Contours show 68% and 95% statistical uncertainties.

within redshift range of $[0.1, 0.3]$. The red dots with error bars show the median and 68% inclusion region of matched halo mass in different richness bins.

The blue line and shaded blue region are the mean and 95% uncertainty of a least-squares fit to the form, $\langle \ln M | \lambda, z \rangle = \pi_h + \alpha_h \log(\lambda/\lambda_p) + \beta_h \log((1+z)/(1+z_p))$. We find parameters $\pi_h = \log(1.26 \pm 0.02\ [10^{14}\ M_\odot])$, $\alpha_h = 1.33 \pm 0.05$ and $\beta_h = -0.48 \pm 0.43$. The line is the $z = 0.2$ relation while the shaded area shows combined uncertainties in the intercept and richness slope. We find that the slope with redshift is consistent with zero with large uncertainties.

This virial scaling of stacked pairwise velocities is remarkably accurate in capturing the scaling with richness of the log-mean membership matched halo mass. Differences are less than 1% at the pivot point and within $\sim 5\%$ over a broad range in richness. Note that the E08 dark matter virial scaling is measured independently of the Aardvark simulation, so the level of agreement between the M_{200c} and membership matched masses is a non-trivial result.

Constraints on the slope and normalization of the mass–richness scalings for the CEN sample are compared in Figure 4.6. Black contours for the stacked dynamical mass include only statistical uncertainties in the constrained velocity parameters, not systematic errors discussed below. The blue contours are based on bootstrap resampling of membership matched halos within the redshift range $[0.18, 0.22]$. The normalizations at the $\lambda = 30$ pivot are consistent, while the slopes are in tension at the level of 0.13, or roughly 2.2σ, in their central value. An ensemble of simulated samples would be useful to reduce the statistical error on the membership matched slope.

We turn next to discuss sources of systematic uncertainty before applying this method to derive a constraint the matched halo masses of RMIV clusters.

4.5.2 Sources of Systematic Uncertainty

The good agreement between stacked dynamical mass and membership-matched masses offers strong incentive to combine large photometric and spectroscopic galaxy samples to relate cluster richness to halo mass.

Applying this method to survey data introduces several sources of systematic error that must be modeled. The Aardvark synthetic sky realization is idealized in several respects; central galaxies are at rest with respect to their underlying halo and satellite galaxies trace the kinematics of the dark matter. Also, the differences in stacked pairwise velocity model parameters for the CEN and ALL samples indicate that mis-centering plays an additional role. In addition, variance in the velocity dispersion of clusters of fixed richness, reflective of the variance in matched halo mass, can introduce bias.

The following sections address these issues in turn, finding that the first two are more important than the third. How satellite galaxies trace dark matter kinematics

is the key source of systematic error.

Central galaxy velocities and satellite galaxy velocity bias

The degree to which galaxy velocities trace the kinematics of dark matter particles in halos is a central issue for virial mass calibration. By construction, the central galaxy is at rest with respect to its host halo in our simulations. In reality, central galaxies are measured to have a non-zero velocity dispersion with respect to their host clusters.

In cases of actively merging systems the rest frame of a cluster is often difficult to define. In the post-merger phase, the central galaxy will settle to the center of cluster due to dynamical friction on a timescale on the order of 1 Gyr (White 1976; Bird 1994), during which time the central galaxy will have a net velocity with respect to the full halo. Based on a sample of nearly 500 Abell clusters with 10 or more redshifts, Coziol et al. (2009) find that brightest cluster galaxies have velocities with root mean square magnitude $\sim 0.3\sigma_{\rm cl}$, with $\sigma_{\rm cl}$ the line-of-sight velocity dispersion of the host cluster. A similar ratio of 0.25 is found by Lauer et al. (2014) using 178 clusters with 50 or more member spectra. Martel et al. (2014) find a similar thermal motion for central galaxies in a sample of 18 massive halos extracted from a large cosmological, hydrodynamic simulation.

Redshift-space distortion studies also support non-zero values for central galaxy velocities (Skibba et al. 2011; Guo et al. 2015b,a). If the central galaxy population has velocity dispersion scaling as some fraction, $\alpha_{\rm c}$, of the host halo dispersion, $\sigma_{\rm cen} = \alpha_{\rm c}\sigma_{\rm halo}$, then the central–satellite pairwise velocity normalization, σ_p, will be enhanced by a factor $(1 + \alpha_{\rm c}^2)^{1/2} \simeq 1 + \alpha_{\rm c}^2/2$, the latter if $\alpha_{\rm c}$ is small compared to unity. Mass estimates derived from virial scaling will be boosted by a factor $(1 + \alpha_{\rm c}^2)^{3/2} \simeq 1 + 3\alpha_{\rm c}^2/2$ relative to the case of cold centrals ($\alpha_c = 0$). These factors

assume that the satellite galaxy velocities are unbiased with respect to the dark matter.

The velocity dispersion of satellite galaxies relative to the halo rest frame may also biased (Carlberg 1994), so that $\sigma_{\rm sat} \; = \; \alpha_{\rm s} \; \sigma_{\rm DM}$, where $\alpha_{\rm s}$ is the satellite galaxy velocity bias. The simulation study of Wu et al. (2013) that combines N-body and hydrodynamic models indicates that $\alpha_{\rm s}$ lies near unity, with brighter galaxies tending to have values less than one and fainter galaxies slightly above unity, asymptotically reaching a value of 1.05. This pattern is not seen in the redshift-space distortion work of Guo et al. (2015b), discussed below.

Let $\sigma_{p,0}$ be the normalization of the central–satellite pairwise velocity dispersion determined through the simulation analysis presented in Section 4.4.2. Recall that the simulations are constructed to have $\alpha_{\rm c} = 0$ and $\alpha_{\rm s} = 1$. Introducing uncorrelated central and satellite galaxy velocity biases modifies the pairwise velocity PDF normalization to

$$\sigma_p \; = \; (\alpha_{\rm s}^2 + \alpha_{\rm c}^2)^{1/2} \sigma_{p,0}. \tag{4.9}$$

If these effects alone are responsible for the normalization difference between the SDSS and Aardvark CEN samples (see Table 4.2), then we would require $(\alpha_{\rm s}^2 + \alpha_{\rm c}^2)^{1/2} = 1.13$.

Cluster mis-centering

While the analysis of Section 4.4.2 focused on the well-centered subsample of clusters, the pairwise velocity PDF of the full sample has a similar form. However, the fit parameters in Table 4.2 indicate that the normalization of the full sample is enhanced, 585 (ALL) versus 547 km s^{-1} (CEN), and the slope α is slightly decreased. Because of the simulation limitations discussed in Section 4.2, the mis-centered frac-

tion of simulated redMaPPer clusters in the ALL sample is larger than that of the SDSS sample. Comparing to X-ray observations of a joint sample of more than 100 clusters, Rozo & Rykoff (2014) find that $86 \pm 4\%$ of high mass clusters are correctly centered on the X-ray counterpart. This statistic is weighted toward higher richness values, $\lambda \sim 100$, but preliminary results of ongoing redMaPPer sample analysis indicate that the full sample of $\lambda > 20$ redMaPPer clusters has a similar fraction of well-centered clusters.

We exploit the differences in the CEN and ALL samples to estimate, using a weighted sampling approach, how velocity PDF parameters shift as the fraction of mis-centered clusters is varied.

The ALL cluster sample contains both mis-centered and correctly centered clusters. Let $f_{\rm cen}$ be the fraction of ALL galaxy pairs lying in the latter (CEN) sample. Our approach is to simply create simulated central-satellite pairs drawn in proportion from the CEN and (ALL-CEN) cluster samples in order to achieve a desired $f_{\rm cen}$ value.

Specifically, for a given $f_{\rm cen}$ value, we randomly draw without replacement a total of 10,000 galaxy pairs from these two cluster sub-populations in a way that satisfies the $f_{\rm cen}$ fraction. We run the MCMC chains for these samples to find the best fit velocity PDF parameters for a total of 2000 realizations uniformly spanning $0.5 \leq f_{\rm cen} \leq 1$.

Figure 4.7 shows how the velocity PDF parameters change with correctly centered fraction, $f_{\rm cen}$. The black lines are the best linear fits as a function of $f_{\rm cen}$, with fit parameters and their root mean square deviations, σ, listed in the legend of each panel.

As the fraction of mis-centered clusters increases (lower $f_{\rm cen}$ values), the velocity

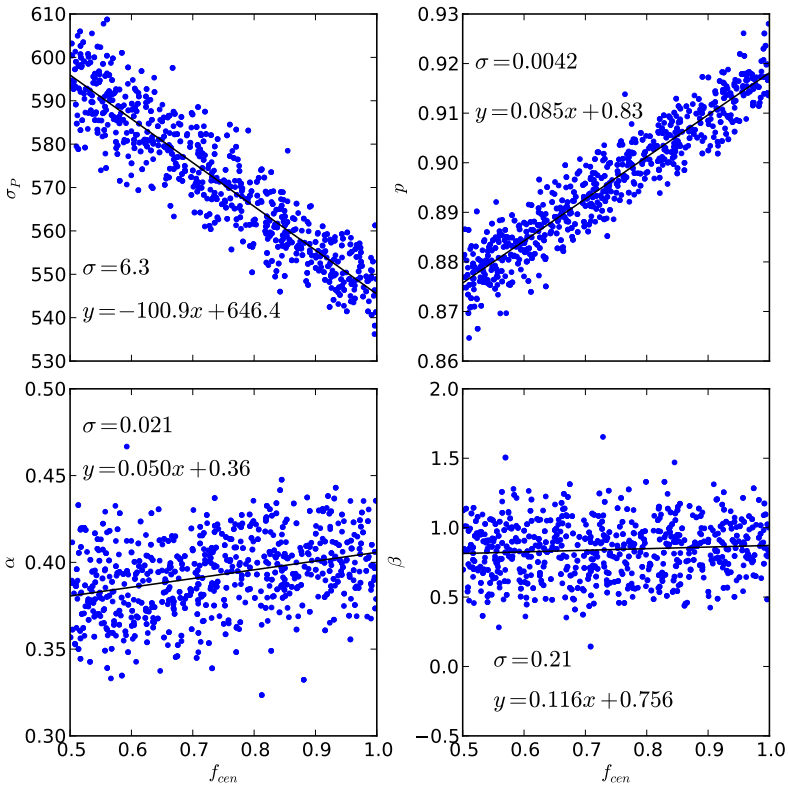

Figure 4.7: Sensitivity of the pairwise velocity PDF model to the fraction of correctly centered clusters, $f_{\rm cen}$. Each point is derived from 10,000 galaxy pairs drawn randomly from the Aardvark CEN and (ALL-CEN) catalogs weighted to achieve the desired $f_{\rm cen}$. The black lines show the best linear fit for each parameter, with the fit and standard deviation, σ, quoted in each panel.

dispersion normalization, σ_p, increases while the slope, α, and Gaussian amplitude, p, both decrease. As expected, the limit of $f_{cen} = 1$ recovers parameters of the CEN catalog (see Table 4.2).

We use this behavior to correct for the effect of mis-centering on the RMIV pairwise velocity normalization. Assuming the fraction of correctly centered SDSS redMaPPer clusters with $\lambda > 20$ to be $f_{\rm cen} = 0.85 \pm 0.05$[3] leads to a $\sim 3\%$ normalization correction for correctly centered systems,

$$\sigma_{p,\rm RMIV,CEN} = 582 \pm 8 \ \ {\rm km\ s^{-1}}\,. \tag{4.10}$$

We use this value to evaluate the mass scale of SDSS redMaPPer clusters in Section 4.6 below. The mis-centering correction to the slope, α, is smaller than 0.01 and is not applied below.

Velocity dispersion variance at fixed richness

The satellite–central velocity likelihood model employs a single Gaussian of width $\sigma_p(\lambda, z)$ at fixed richness, λ, but there is non-zero variance in velocity dispersion values of a fixed-λ population that reflects the variance in matched halo mass. Scatter in halo mass at fixed lambda is already incorporated into the simulations; the scatter in matched halo masses shown in Fig. 4.5 is 0.85 in $\ln M$. We perform here an explicit test, independent of the simulated samples, to confirm that this scatter does not strongly affect the recovered velocity PDF parameters.

We create ensembles of 10,000 galaxy pairs drawn from Gaussian distributions with dispersion values log-normally distributed about a scaling mean relation, $\sigma_p(\lambda, z)$, Equation 4.5 with variance $\sigma^2_{\ln\sigma}$. Sampling in λ and redshift uniformly covers the observed ranges of $[20, 200]$ and $[0.1, 0.3]$, respectively. We then perform the stacked

[3]The value of 0.85 ± 0.05 is slightly more conservative than that published for higher richness clusters in Rozo et al. (2015).

velocity PDF analysis on each simulated pair ensemble.

We find that the model parameters remain unbiased until $\sigma_{\ln\sigma} > 0.2$, after which the tails of the velocity distribution begin to affect the normalization p at the one percent or greater level. The recovered values of σ_p and α, the key parameters involved in mass estimation, are unaffected up to values of $\sigma_{\ln\sigma} = 0.5$, or 1.5 scatter in $\ln M$. This degree of mass scatter is larger than either the simulated or observed (Rozo & Rykoff 2014) values. Variance in host halo velocity dispersion at fixed richness is therefore a negligible source of systematic error in the velocity PDF modeling and resultant mass estimates.

Orientation and Shape selection bias

Because dark matter halos are aspherical, optical cluster selection and richness estimation on the sky are sensitive to halo orientation, with preferential selection of structures elongated along the line of sight (e.g. Dietrich et al. 2014). The Hubble Volume simulation analysis of Kasun & Evrard (2005) finds alignment of position and velocity ellipsoids in massive halos, with median alignment angle of 22°. Orientation biases in an optically-selected cluster sample such as redMaPPer could produce shifts in the mean stacked velocities. Along these lines,Skielboe et al. (2012) show that the LOS velocity dispersion of galaxies lying along the major axis of SDSS clusters is slightly larger than that of galaxies lying along the minor axis. Simet et al. (2017) use analytic arguments to estimate a $4\% \pm 2\%$ orientation bias (overestimate) in stacked weak lensing masses for redMaPPer selected clusters.

A potentially counteracting effect, found by Ragone-Figueroa et al. (2010) in the The MareNostrum Universe simulation, is that, at fixed mass, more elongated halos have smaller 3D-averaged velocity dispersion than less elongated systems. They link this effect to formation epoch, hence it is a form of assembly bias.

We perform two tests on the Aardvark simulation to estimate orientation biases on stacked dynamical.

First we ask whether the redMaPPer finder preferentially selects elongated halos. To measure halo shape, we assume an ellipsoidal model and determine the three eigenvalues, λ_i, of the shape tensor in position space for galaxy members. The largest eigenvector gives the orientation. We define the elongation as c/a, where c is is the minor axis and a the major axis of the shape tensor (see Section 2.4 of Kasun & Evrard (2005) and Section 2 of Zemp et al. (2011) for more detail).

We find that the distribution of shapes for matched halos selected by redMaPPer matches well the that of the overall halo population. Using bins of width 0.2 dex in mass, the median and quartile values of c/a for the two populations match to within ~ 0.02 for halos more massive than $10^{13.5}\,\mathrm{M}_\odot$. Shape selection bias is not a large effect for this sample.

The second test concerns possible orientation bias of redMaPPer selection. The unbiased velocity dispersion is the 3D-averaged velocity dispersion of galaxies within the halo. We measure the LOS and 3D velocity dispersion for all galaxies inside matched halos. Regressing both velocity dispersion values against mass, we find that the normalization of the LOS velocity dispersion is larger than the 3D value by $\sim 1.1\%$. This implies a 3.3% overestimation of the stacked dynamical mass at fixed richness.

Because this level of bias is smaller that the other sources of uncertainty described in Section 4.6), we do not explicitly apply a correction. We note that the specific correction will depend on the algorithm employed for optical cluster selection.

4.6 Stacked Dynamical Mass Scaling of SDSS redMaPPer Clusters

The above analysis indicates that the mass determined through virial scaling of the pairwise velocity PDF normalization offers an unbiased estimate of the log-mean mass of halos matched via joint galaxy membership.

We now turn to estimate the characteristic M_{200c} mass scale of correctly centered redMaPPer clusters as a function of richness λ at the pivot redshift $z_p = 0.2$. Recall from Section 4.5.2 and 4.5.2 that the pairwise velocity normalization depends on the mis-centering frequency and the velocity bias of central and satellite galaxies. We need to estimate the magnitudes of these effects, and their uncertainties, into our mass estimate.

The normalization correction for mis-centering, assuming $f_{\rm cen} = 0.85 \pm 0.05$ for the SDSS redMaPPer sample, is already incorporated into the correctly-centered estimate given in Equation (4.10).

To estimate the velocity dispersion of the underlying dark matter from the pairwise satellite–central galaxy measurements, we need to divide the latter by the quadrature sum of the respective velocity bias factors,

$$\sigma_{p,\rm RMIV,DM} = \frac{\sigma_{p,\rm RMIV,CEN}}{(\alpha_{\rm s}^2 + \alpha_{\rm c}^2)^{1/2}} \,. \tag{4.11}$$

The velocity bias of galaxies has been recently investigated by Guo et al. (2015a,b) using SDSS galaxy clustering measured both in projected separation and in redshift space. We employ the Guo et al. (2015b) estimates for the velocity bias factors of bright ($M_r \sim -21.5$, as appropriate for the bulk of the spectroscopic galaxies in this study) galaxies (see their Fig. 8) of $\alpha_{\rm c} = 0.30 \pm 0.05$ and $\alpha_{\rm s} = 1.05 \pm 0.08$. Their central galaxy dispersion is in line with previous estimates based on explicit spectroscopy of cluster members (Coziol et al. 2009; Lauer et al. 2014) as well as

with recent simulation expectations (Martel et al. 2014). There is more contention on the velocity bias of satellite galaxies. In recent simulations, values less than one have been measured for bright galaxies in massive halos (Munari et al. 2013; Old et al. 2013; Wu et al. 2013). We note that the 2σ range of $\alpha_s \in [0.89, 1.21]$ admits values less than unity.

These velocity bias estimates imply a correction factor, $(\alpha_{\rm s}^2+\alpha_{\rm c}^2)^{-1/2} = 0.92\pm0.07$, which leads to the dark matter velocity dispersion at the pivot richness and redshift of

$$\sigma_{p,\rm RMIV,DM} = 535 \pm 41 \ \ {\rm km\ s^{-1}}\,. \tag{4.12}$$

Note that the uncertainty in this velocity is dominated by systematic error in the velocity bias estimate.

Finally, using this value in Equation (4.8), we obtain an estimate of the log-mean mass of redMaPPer clusters at the pivot richness and redshift of

$$M_\sigma(\lambda_p = 30, z_p = 0.2) = (1.56 \pm 0.35) \times 10^{14}\,{\rm M}_\odot\,, \tag{4.13}$$

where to infer above mass scale we assume a ΛCDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $h(z = 0) = 0.7$.

The scaling of the pairwise velocity normalization, $\sigma_p(\lambda, z)$, determines how the mean dynamical mass, $M_\sigma(\lambda, z)$, scales with richness and redshift. Because of the relatively weak constraint on the redshift scaling behavior of the SDSS cluster sample velocities, we defer analysis of redshift evolution to a later study and concentrate here on the behavior with richness at the pivot redshift of 0.2. The simulations indicate the the mean dynamical mass, $M_\sigma(\lambda, z)$, matches the log-mean membership matched mass at the pivot richness, but as shown in Figure 4.6, the best-fit slope of log-mean

mass with richness differs by 0.10 from the slope of $M_\sigma(\lambda)$. We therefore include this difference as a systematic error term when quoting the slope.

The result is an estimate for the log-mean membership matched mass of the SDSS redMaPPer sample at redshift 0.2 of

$$\langle \ln(M_{200c}/10^{14}\,\mathrm{M}_\odot)|\lambda, z_p = 0.2\rangle = \pi + \alpha_m\,\ln(\lambda/30) \tag{4.14}$$

with normalization $\pi = 0.44 \pm 0.22$ and slope $\alpha_m = 1.31 \pm 0.06_{stat} \pm 0.13_{sys}$.

Of the 22% error in the derived mass normalization, 21.5% arises from systematic uncertainty in the velocity bias terms, particularly that of satellite galaxies. Miscentering contributes 2.6%, and statistical uncertainties from the stacked pairwise velocity and virial calibration parameters are 3.2%. The error in $\ln(M)$ is essentially triple the uncertainty in $\ln(\alpha_s)$. As a result, achieving ten percent error in mean mass would require knowing α_s to a fractional accuracy of ~ 0.03. It remains to be seen whether future spectroscopic campaigns, coupled with improved hydrodynamic simulations of galaxy formation in massive halos to pin down systematic errors, can achieve this level of precision.

4.7 Chapter Conclusion

Using galaxy catalogs derived from large N-body simulations, we study the mapping of galaxy clusters identified in sky-photometry space to the underlying real-space population of halos through membership matching. We measure membership strength, defined as the fraction of a cluster's richness contributed by a given halo, and build bipartite graphs linking clusters to halos with strength-weighted edges. The matched halo of a cluster maximizes this strength.

We then study pairwise velocities, and derived masses, from stacked spectroscopic analysis of clusters patterned after the spectroscopic analysis of SDSS redMaPPer

clusters developed by RMIV. The structure in the simulated data is similar to that of the observations, with galaxy pairwise velocities having a main Gaussian provisionally identified as cluster members. We employ a sub-sample of correctly centered clusters — those for which the central cluster galaxy is also the central galaxy of the matched halo — as well as studying the full simulated cluster sample.

We then use our findings to estimate the log-mean, membership-matched mass of SDSS redMaPPer clusters at $z = 0.2$. Our detailed results are as follows.

- Although the pairwise velocity PDF model is not a good fit to data, the richness and redshift dependent width of the PDF adequately reflects the log-mean velocity dispersion of matched halos. Decomposing this main component into halo contributions, we find that the top-ranked, matched halo contributes an average of 62% (58%) of pairs in the correctly centered (full) cluster samples. The second-ranked halo contributes $\sim 10\%$, the third $\sim 5\%$, and the remainder contribute $\sim 20\%$, in the mean. The projected component, consisting of all galaxy pairs not contributed by the top-ranked matched halo, has a pairwise velocity PDF described roughly by a Gaussian plus constant form.

- Converting the velocity dispersion–richness relation to a mass–richness relation using the dark matter virial relation calibrated by independent simulations (Evrard et al. 2008), we find this stacked dynamical mass recovers, to within a few percent, the log-mean mass determined from membership matching between clusters and halos.

- We model effects of cluster mis-centering and galaxy velocity bias in order to correct the measured redMaPPer cluster velocity dispersion to reflect that of correctly centered, dark matter halos. Using central and satellite velocity bias

parameters $\alpha_c = 0.30 \pm 0.05$ and $\alpha_c = 1.05 \pm 0.08$, respectively (Guo et al. 2015b), we infer a log-mean matched halo mass of $M_{200,p} = (1.56 \pm 0.35) \times 10^{14}\,\mathrm{M}_\odot$ at the pivot richness, $\lambda_p = 30$, and redshift $z_p = 0.2$, and a slope with richness of $1.31 \pm 0.06_{stat} \pm 0.13_{sys}$ for SDSS redMaPPer clusters.

Kinematic biases of central and, especially, satellite galaxies, are the dominant source of systematic error. Further work is needed, both empirically and through hydrodynamic simulations, to better constrain the relationship between galaxy velocities and dark matter. One possible approach is to invert the analysis presented here; comparing the stacked dynamical masses with stacked weak lensing masses of the same sample with the aim of constraining velocity bias.

CHAPTER V

The Mass Scale of XXL Clusters from Ensemble Spectroscopy

Philosophy and Contribution

In this chapter, I apply the ensemble mass calibration technique developed in the previous chapter to estimate the mass–temperature relation for clusters selected from the XXL survey program. The work in this chapter provides a new, independent mass calibration which supports the science goal of the XXL cluster cosmology analysis. This chapter is taken from Farahi et al. (2017b): "The XXL Survey. XXIII. The Mass Scale of XXL Clusters from Ensemble Spectroscopy". The statistical analysis, tests and validations, and plots presented in this chapter are my own work. The observations obtained with XMM-Newton via XXL Survey program. The observational data is reduced and analyzed by the co-authors who are members of the XXL collaboration. The co-authors provided the observational data catalogs, including the cluster catalog and the galaxy catalog.

5.1 Chapter Introduction

Chapter II presents how the cosmic web of dark matter drives the gravitational potential wells. The self-similar model presented in that chapter specifies how halo properties scales with the redshift and mass of halos. This simple model predicts

a power-law relation for the mass-property relations (MPRs). The idea that both galaxies and hot gas are in virial equilibrium within a common gravitational potential, originally proposed by Cavaliere & Fusco-Femiano (1976), leads to the expectation that galaxy velocity dispersion scales as the square root of X-ray temperature, $\sigma_{\rm gal} \propto T_X^{0.5}$. This behavior reflects MPR scalings with total mass $M \propto T_X^{3/2}$ and $M \propto \sigma_{\rm gal}^3$ at fixed redshift (for the derivation see Chapter II). However, astrophysical processes within halos, such as star formation and associated supernova and AGN feedback, can drive deviations from self-similarity.

For the most massive clusters in the sky, multiple surveys and follow-up observations are enabling individual halo masses to be estimated from gravitational lensing, hydrostatic, and dynamical methods (see Allen et al. 2011; Kravtsov & Borgani 2012, for reviews). These methods are subject to different sources of systematic uncertainty (e.g., Meneghetti et al. 2014), and the samples to which they are applied may have additional systematic shifts, relative to a sample complete in halo mass, due to sample selection. The resulting biases pose limits on the accuracy of empirically derived MPRs.

Multiple, independent mass proxies allow for consistency tests that can expose and help mitigate systematic errors. We present here a Virial analysis of 132 spectroscopically confirmed clusters identified in the XMM-XXL Survey (Pierre et al. 2016, hereafter XXL paper I). The method extends the stacked spectroscopic technique developed in Chapter IV, originally applied to optically selected clusters in SDSS (Rykoff et al. 2014).

We focus first on the virial scaling of galaxy velocity dispersion with hot gas temperature, then infer how mean total mass scales with temperature using an additional degree of freedom that relates galaxy velocity dispersion to the underlying

dark matter. This galaxy velocity bias is the largest source of uncertainty in our mass estimate.

Early N-body simulations established virial scaling for purely dark matter halos (Evrard 1989) and ensemble analysis of billion-particle and larger simulations provides a highly accurate calibration, with sub-percent error in the intercept of dark matter velocity dispersion at fixed halo mass (Evrard et al. 2008).

Inferring a virial, or dynamical, mass of an individual cluster requires a large number of spectroscopic members and a reliable interloper rejection algorithm (e.g., Biviano et al. 2006) such as that provided by the caustic technique (Rines et al. 2007; Rines & Diaferio 2010; Gifford et al. 2013). For large cluster samples emerging from surveys, a complementary approach to infer mean MPR scaling behavior is to employ ensemble population analysis, effectively stacking the local velocities of galaxies in multiple clusters to extract a mean velocity dispersion signal.

Here we have employed a large collection of galaxy spectroscopic redshifts assembled from multiple sources for groups and clusters identified in the north field of the XMM-XXL survey. The 132 systems span X-ray temperatures $\mathrm{k}T_{300\mathrm{kpc}} \in [0.48 - 6.03]\,\mathrm{keV}$, and redshift $z \in [0.03 - 0.6]$, and the spectroscopic sources include GAMA, SDSS-DR10, VIPERS, and VVDS Deep and Ultra Deep surveys.

The mass-temperature scaling has been studied extensively (e.g., Xue & Wu 2000; Ortiz-Gil et al. 2004; Arnaud et al. 2005; Vikhlinin et al. 2006; Kettula et al. 2015; Mantz et al. 2016b; Lieu et al. 2016). Observational relations generally steepen from close to the self-similar for hot systems to a slope of $\sim 1.6 - 1.7$ once cooler systems ($\mathrm{k}T_{300\mathrm{kpc}} \lesssim 3\,\mathrm{keV}$) are included (Arnaud et al. 2005; Lieu et al. 2016). More than half of the clusters in the work presented in this chapter will be systems with $\mathrm{k}T_{300\mathrm{kpc}} \lesssim 3\,\mathrm{keV}$, which allows us to test deviation from the self-similar model, with

yet another mass calibration technique.

As part of the first series of XXL papers, (Lieu et al. 2016, hereafter XXL paper IV) estimates the mass–temperature scaling relation of X-ray bright systems using weak-lensing mass measurements from the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS) shear catalogue (Heymans et al. 2012; Erben et al. 2013). The work presented here is complementary to that study where it provides a mean dynamical mass as a function of X-ray temperature. The X-ray sample differs from that used by XXL Paper IV, but the pipeline for deriving X-ray properties from the XMM data is identical.

We describe the sample, data, and selection criteria in Section 5.2. The likelihood model used to constrain the galaxy velocity dispersion scaling with temperature is described in Section 5.3. In Section 5.4, we present results for this relation, followed by a discussion of a range of systematic uncertainties and sensitivity analysis in Section 5.5. A key result of this work, the dynamical mass-temperature relation, is presented in Section 5.6. Finally we conclude this chapter in Section 5.7

Chapter's Notation. Throughout this Chapter, we have assumed WMAP9 consistent cosmology with $\Omega_m = 0.28$, $\Omega_{DE} = 0.72$, and local Hubble constant $h = H_0/100$ km s^{-1} Mpc$^{-1} = 0.7$. Unless otherwise noted, our convention for the mass of a halo is M_{200}, the mass contained within a spherical region encompassing a mean density equal to 200 times the critical density of the Universe, $\rho_c(z)$. Similarly, r_Δ is defined as the radius of the sphere inside which the mean density is a factor Δ times the critical density of the universe at that redshift, and M_Δ is the total mass within that radius.

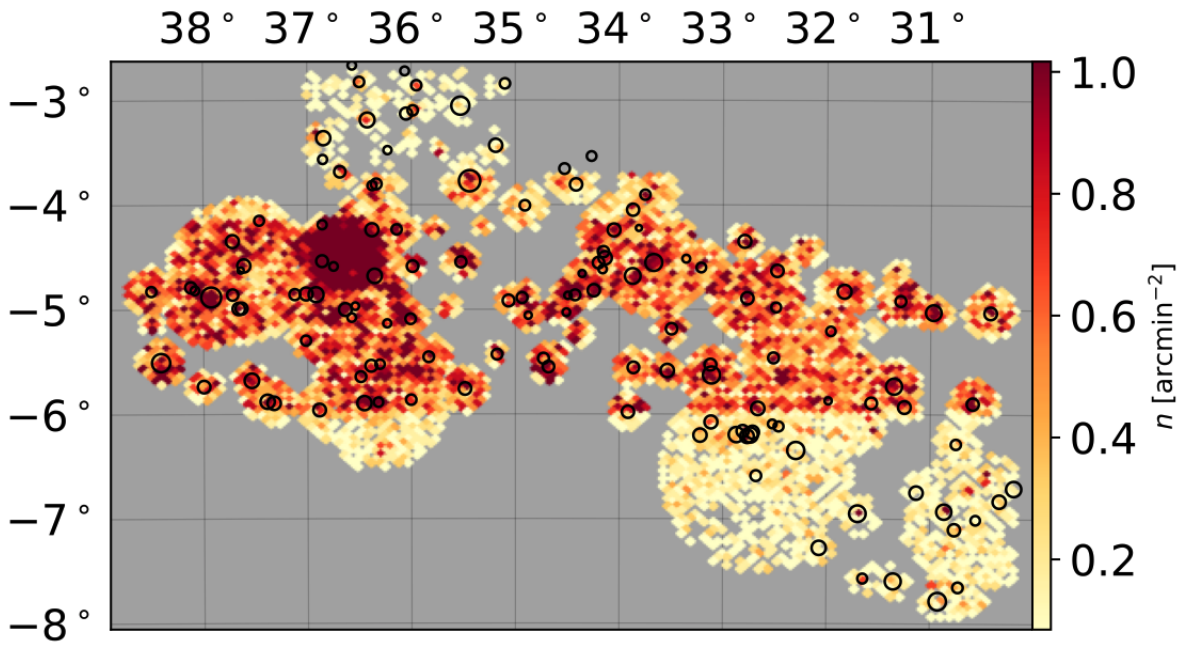

Figure 5.1: Spatial distribution of galaxies and clusters in the XXL north field used in this chapter. Black circles show cluster centres with $z \leq 0.6$ with area proportional to temperature. The heat map shows the sky surface density of spectroscopic galaxies lying within a projected aperture of $3r_{500}$ around cluster centres.

5.2 Cluster and spectroscopic sample

The XXL survey consists of tiled 10 ks (or longer) exposures across two fields of roughly 25 deg^2 each. The observing strategy and science goals of the survey are described in XXL Paper I while source selection and a resultant brightest 100 cluster sample are published in Pacaud et al. (2016, hereafter XXL Paper II). The X-ray images were processed with the XAMIN v3.3.2 pipeline (Pacaud et al. 2006), which produces lists of detections of varying quality. The overall catalogue with point sources will be available in computer readable form via the XXL Master Catalogue browser `http://cosmosdb.iasf-milano.inaf.it/XXL` and at the Centre de Données astronomiques de Strasbourg (CDS)[1] (Chiappetti et al. 2017, hereafter XXL Paper XXVIII), while cluster candidates are grouped by detection classes (C1, C2,

[1] `http://cdsweb.u-strasbg.fr`

C3) and hosted in the same places as catalogue XXL-365-GC (Adami et al. 2017, hereafter XXL Paper XX). The 2016 series of XXL papers, including (XXL Paper II), pertained to the brightest 100 clusters and 1,000 AGN, while for the second series, including the work presented in this chapter, we are publishing much deeper samples: 365 clusters and 20,000 AGN, with slightly revised cluster properties and scaling relations.

Of the XXL cluster sample 46% are classified as high-quality (C1) detections, 43% are intermediate quality (C2) and the remaining 11% are marginal quality (C3) sources. We discard C3 sources in this work as they do not have reliable luminosity and temperature measurements. The subject of this work is a subset in the XXL-N area, with spectroscopically confirmed redshifts and with redshifts $z < 0.6$, generating a sample of 132 systems. A detailed discussion of the sample selection is provided by XXL Paper XX and Guglielmo et al. (2017, hereafter XXL Paper XXII). [2]

The sky distribution of the systems used in this work is shown in Fig. 5.1. X-ray extended sources are shown as black circles and the color map shows the sky surface density of spectroscopic galaxies lying in an aperture of radius $r \leq 3r_{500}$ with respect to their centres. The r_{500} estimates are determined from weak lensing mass estimates presented in XXL Paper IV. We next provide additional details of the group/cluster and galaxy spectroscopic samples.

5.2.1 X-ray Temperatures

Of the 132 spectroscopically confirmed C1 and C2 clusters with $z < 0.6$, X-ray temperatures are available for 106, 81 C1 and 25 C2 clusters. All are C1 clusters and most but not all are included in the XXL 100 brightest sample of XXL Paper II. The temperature determination, described in detail by Giles et al. (2016, hereafter XXL

[2]The cluster optical and X-ray images can be found in the XXL cluster database: `http://xmm-lss.in2p3.fr:8080/xxldb`.

Paper III), outputs the temperature measured within a physical 300 kpc aperture for sufficiently high signal-to-noise-ratio systems.

After detection by XAMIN v3.3.2 - a detection pipeline piloted by the XMM-LSS project (Pacaud et al. 2006) - as an extended X-ray source, a background subtracted radial profile is extracted in the $[0.5-2]$ keV band. The detection radius is defined as that at which the source is detected at 5σ above the background. A spectrum is then fit from a circular aperture of radius of 300 kpc centred on the X-ray centroid, using a minimum of five counts per energy bin, resulting in a temperature measurement we refer to as $T_{300\mathrm{kpc}}$. Cluster spectral fits were performed in the $0.4-7.0$ keV band with an absorbed APEC model with the absorbing column fixed at the Galactic value, and a fixed metal abundance of $Z = 0.3Z_{\odot}$. For more detail on the data processing, we refer the reader to Pacaud et al. (2016). We note that the measured X-ray temperatures are non-core excised owing to the limited angular resolution of XMM-Newton and the modest signal-to-noise-ratio of most detections. These temperatures are taken from XXL Paper XX.

For the systems that lack direct temperature estimates, we estimate temperatures from X-ray luminosities using published XXL scaling relations as follows. First, background-corrected XMM count-rates within 300 kpc from the cluster centre in the $[0.5-2]$ keV band are extracted. This forms the basis of a first luminosity estimate, the starting point for an iterative scheme that uses the $L-T$ scaling relation from XXL Paper XX and the $T-M_{500}$ relation from XXL Paper IV. The process assumes isothermal β-model emission with parameters $(r_c, \beta) = (0.15r_{500}, 2/3)$, and iterations continue until convergence. This method outputs temperature, mass, and r_{500} estimates. Details of the steps above are described and reported in XXL Paper XX.

To check the internal consistency of the derived X-ray temperature, XXL Pa-

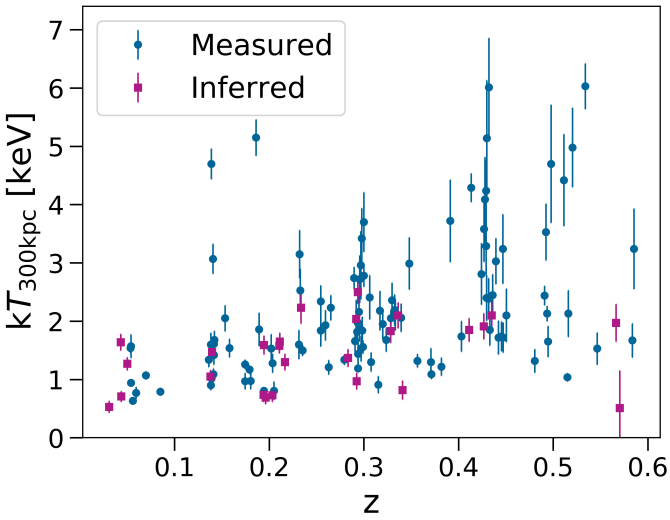

Figure 5.2: Temperature vs. redshift of the full 132 XXL-N cluster sample. Blue circles are clusters with measured temperature and magenta squares show clusters with inferred temperature.

per XX performs a comparison of $T_{300\mathrm{kpc}}$ derived using the above approach with direct temperature measurements for a subset of systems, finding good agreement. Below, we show that the velocity dispersion scaling parameters using the subset of systems with directly measured temperatures are consistent with those of the full cluster sample.

Figure 5.2 shows redshifts and temperatures of the XXL-N clusters. At a given redshift, higher mass systems that are both brighter and hotter tend to have direct temperature measurements. As explained in Section 5.2.5, the sample size shrinks, by roughly 3% (four clusters), after we apply velocity and aperture cuts discussed below.

5.2.2 Spectroscopic sample

Concerning the spectroscopic database of galaxies, reduced spectra from several public surveys are combined with XXL dedicated observing runs to create a large, heterogeneous collection of redshifts. The surveys and observing programmes, listed in Table 2 in XXL Paper XXII, include GAMA (45%, Hopkins et al. 2013; Liske

et al. 2015), SDSS-DR10 (5% Ahn et al. 2014), VIPERS (32% Guzzo et al. 2014), VVDS Deep and Ultra Deep (9% Le Fèvre et al. 2005, 2015). The remaining 9% are obtained mainly by ESO Large Programme + WHT XXL dedicated observational campaigns which are individually contributing less than 2%. The typical error in redshift for galaxies is $\sim 0.00041(1+z)$, equivalent to $120(1+z)$ km/s. The full list of spectroscopic catalogues are listed in XXL Paper XXII. We note that the spectroscopic sample adopted in this work is a subset of the spectroscopic sample of XXL Paper XXII.

Given that the catalogue sources overlap in the sky, a non-negligible number of objects are observed by more than one project. The cleaning of catalogue duplicates follows the selection criteria designed to identify the best spectrum in the final catalogue, as described by XXL Paper XXII. The selection procedure is based on two sets of priorities, the first regarding source origin and then the second regarding the reliability flag attributed to the redshift estimate.

The full sample contains 120506 galaxies in the north XXL region, 63681 of which are at $z \leq 0.6$. For our default analysis, we employ a sub-sample comprised of those galaxies lying within a projected distance of r_{500} from the centres of the clusters, shown in Fig. 5.1, yielding 7751 galaxies. [3]

5.2.3 Spectroscopic redshifts of XXL-selected clusters

All C1 and C2 candidate clusters identified within the XXL survey are followed up for spectroscopic redshifts using an iterative semi-automatic process similar to that used for the XMM-LSS survey (Adami et al. 2011).

First, spectroscopic redshifts from public and private sources lying within the

[3]The spectroscopic information for these galaxies, as well as for spectroscopically confirmed groups/clusters, is hosted in the CeSAM (Centre de donnéeS Astrophysiques de Marseille) database in Marseille (CeSAM-DR2), publicly available at `http://www.lam.fr/cesam/`.

X-ray contours are selected. These are sorted to identify significant (more than 3 galaxies) concentrations, including a preliminary "cluster population" based on projected separation from the X-ray centroid. For the large majority of cases, a single concentration appears, allowing for relatively unambiguous redshift determination.

A preliminary measure of the cluster redshift is the mean value of the redshift of the preliminary cluster population. From this redshift, a physical region of 500 kpc radius is defined, and all galaxies within this radius were selected as cluster members. This procedure is iterated with all available redshifts within a 500 kpc physical radius to get the final mean cluster redshift. However, for ambiguous cases where there are not more than three galaxies with spectroscopic redshifts, the redshift is measured by looking for the putative brightest cluster galaxy (BCG) in the i-band located close to the X-ray centroid (see XXL Paper XX for a detailed discussion).

The cluster centre is defined by the peak in the detected X-ray emission. Because X-ray emission is continuous and the gas traces the gravitational potential, we expect fewer mis-centered clusters (mis-centered with respect to the dark matter potential minimum) compared to photometrically-defined samples (Rykoff et al. 2012). We defer a detailed treatment of cluster mis-centering to future work.

5.2.4 Galaxy-cluster velocities

Given the redshift, z_c, of each XXL-N group or cluster, we measure the rest-frame relative velocity of each galaxy within the target field of that cluster,

$$v_{\rm gal} \;=\; c\,\left(\frac{z_g - z_c}{1+z_c}\right), \tag{5.1}$$

where c is the speed of light and z_g is the redshift of the galaxy.

In this chapter the original spectroscopic galaxy selection for each cluster is defined only by sky location, not cluster redshift. Therefore, each cluster field contains a

mix of galaxies residing within and outside the cluster environment. We describe below the probabilistic method originally applied to SDSS redMaPPer systems by Rozo et al. (2015), which involves a two-stage approach to handling foreground and background galaxies.

5.2.5 Signal component and final cluster sample

The model framework, wherein observable properties scale with halo mass as power laws with some intrinsic covariance, motivates the modeling process. For systems with a given temperature, $T_{300\mathrm{kpc}}$, and redshift, we expect a log-normal distribution of halo mass with some intrinsic $(10-20\%)$ scatter (Le Brun et al. 2017). The galaxy velocities internal to these halos are assumed to follow a Gaussian distribution with a dispersion that increases with halo mass. Because the intrinsic scatter of these relations is not very large, the expected distribution of galaxy velocities, v_{gal}, at fixed $T_{300\mathrm{kpc}}$ and z will also be close to Gaussian (see Becker et al. 2007, for a specific model applied to galaxy richness instead of temperature). This collective component is the fundamental signal we seek to model and extract from the data.

The first stage of the process removes projected interlopers with large v_{gal} offsets, much larger than those expected from the underlying Gaussian model. The threshold value, $v_{\mathrm{max}}(T_{300\mathrm{kpc}})$, is set empirically by examination of the absolute magnitude of the line-of-sight galaxy velocities as a function of cluster temperature, given in Fig. 5.3. Similar to the analysis of Chapter IV, where redMaPPer optical richness plays the role of $T_{300\mathrm{kpc}}$, two populations emerge: a signal component at low velocities and a projected population offset to higher velocities.

Based on the structure of Fig. 5.3, we define a maximum, rest-frame galaxy ve-

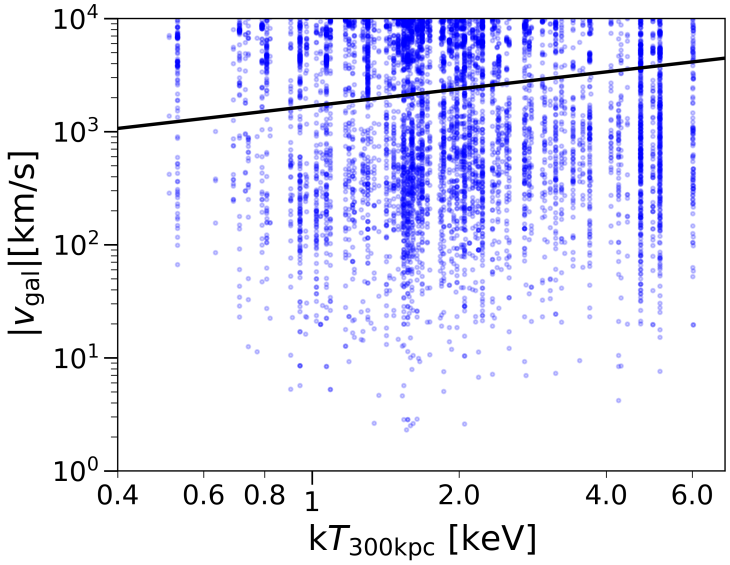

Figure 5.3: Magnitude of the rest-frame velocity of cluster galaxies, Equation 5.1, as a function of cluster temperature. Each dot is one galaxy, and some galaxies appear in the fields of multiple clusters. The black line shows the cut, Equation 5.2, that separates the lower signal population from a projected background. Points above the black line are disregarded in our analysis.

locity for the signal region of

$$v_{\rm max}(T_{\rm 300kpc}) \;=\; 2500 \left(\frac{{\rm k}T_{\rm 300kpc}}{2.2\,{\rm keV}}\right)^{0.5} {\rm km\ s^{-1}}\,. \tag{5.2}$$

Applying this cut along with the radial cut, $r \leq r_{500}$, eliminates four clusters from the sample because no galaxies satisfy these cuts. The final cluster sample involves 1592 galaxies across 128 clusters, 103 of which have directly measured temperatures.

Figure 5.4 shows the distribution of spectroscopic galaxy counts within r_{500} in the cluster sample after applying the velocity threshold, Equation 5.2. The modal, median, and mean values of the distribution are 3, 9, and 12.4 respectively. After applying the velocity and aperture cuts, the main contribution of spectroscopic sample came from GAMA (45%), VIPERS (30%), VVDS Deep and Ultra Deep (11%), SDSS-DR10 (5%). The remaining catalogues individually contribute less than 2%.

In Section 5.5, we investigate the sensitivity of our results to $v_{\rm max}$ and r_{500} selection thresholds, not finding statistically significant change.

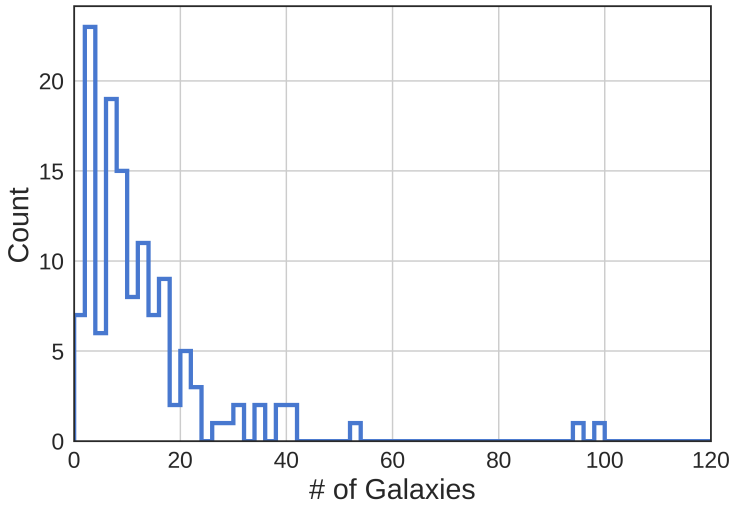

Figure 5.4: Frequency distribution of the number of spectroscopic members per cluster within r_{500} after removing the high-velocity background component using the velocity cut, Equation 5.2.

5.3 Cluster ensemble velocity model

The study of Rozo et al. (2015) introduced an ensemble likelihood model for stacked cluster spectroscopy with the goal of assessing the quality of photometric membership likelihoods computed by the redMaPPer cluster finding algorithm (Rykoff et al. 2012). This model was designed to take advantage of sparse, wide-area spectroscopic samples, for which each cluster may have only a few member redshifts. Subsequently, the approach was extended in the previous chapter to infer the scaling of mass with optical richness, $\lambda_{\rm RM}$. In the present chapter we follow a similar approach, with X-ray temperature replacing $\lambda_{\rm RM}$.

5.3.1 Ensemble galaxy velocity likelihood

Power-law scaling relations, originally motivated by the self-similar model (Kaiser 1986, see also Chapter II for more discussion), are confirmed in modern hydrodynamic simulations, which model baryonic processes in halos (e.g., McCarthy et al. 2017; Truong et al. 2018). Consequently, we assume a power-law scaling relation between characteristic galaxy velocity dispersion, $\sigma_{\rm gal}$, and X-ray temperature of the

form,

$$\sigma_{\rm gal}(T_{\rm 300kpc}, z) \;=\; \sigma_p \left(\frac{{\rm k}T_{\rm 300kpc}}{{\rm k}T_p}\right)^{\alpha} \left(\frac{E(z)}{E(z_p)}\right)^{\beta}, \tag{5.3}$$

where ${\rm k}T_p = 2.2$ keV and $z_p = 0.25$ are the pivot temperature and redshift, and $E(z) = H(z)/H_0$ is the normalized Hubble parameter.

The probability distribution function (PDF) of galaxy velocity at a given cluster temperature is taken to be Gaussian with the above dispersion. The ensemble likelihood for the signal component allows for a residual, constant background atop this cluster member signal. The likelihood for the ensemble cluster-galaxy rest-frame velocity sample is thus

$$\mathcal{L} \;=\; \prod_{i=1}^{n} \left[\; p \; G(v_{{\rm gal},i} | 0, \sigma_{\rm gal}(T_i, z_i)) + \frac{1-p}{2v_{\rm max}(T_i)} \right], \tag{5.4}$$

where G is the Gaussian distribution with zero mean and standard deviation, σ_{gal}, $v_{\rm gal}$ is the line-of-sight (LOS) velocity, Equation 5.1, and the sum i is over all galaxy-cluster pairs in the spectroscopic sample lying below the maximum cutoff, Equation 5.2. The parameter p is the fraction of galaxies that contribute to the Gaussian component, while $1-p$ is residual fraction of projected systems that are approximated by a uniform distribution in the signal portion of velocity space.

We maximise this likelihood with respect to the four model parameters, σ_p, α, β, and p. Below we find that the redshift evolution parameter, β, is both relatively poorly constrained and consistent with zero. We therefore also perform a restricted analysis in which we assume self-similar evolution (SSE), with $\beta = 0$.

5.3.2 Ensemble velocity model in simulations

This model has been tested against simulation in Chapter IV, using cluster richness instead of X-ray temperature, with several key findings. First, the spectroscopic

Table 5.1: Expectation values and standard deviations of the marginalized posterior distributions of free parameters of the model defined in Eqs. 5.3 and 5.4. Parameters listed below are for the fiducial model; the self-similar evolution model, with β set to zero, returns identical central values and errors for the other parameters and so are not listed.

σ_p [km/s]	α	β	p
539 ± 16	0.63 ± 0.05	-0.49 ± 0.38	0.88 ± 0.015

mass estimate is a nearly unbiased estimator of $\langle \ln M_{\rm mem} | \lambda_{\rm RM} \rangle$, where $M_{\rm mem}$ is the mass of the underlying halo that contributes the maximum fraction of the cluster's photometric member galaxies assigned by redMaPPer. Second, galaxies lying in the signal region consist of a majority coming from the top-ranked, member-matched halo ($\sim 60\%$) as well as locally projected galaxies ($\sim 40\%$) lying outside the matched halo. Finally, the main source of systematic uncertainty in the SDSS cluster mass estimate of Chapter IV is uncertainty in the magnitude of the galaxy velocity bias.

5.4 Velocity scaling results

In this section, we present the inferred $\sigma_{gal} - {\rm k}T_{\rm 300kpc}$ scaling relation for the full cluster sample. The fiducial analysis uses the signal velocity threshold of Equation 5.2, an angular limit of r_{500}, and solves for the four degrees of model freedom using the entire sample. Sensitivity tests of the angular and velocity thresholds used in our fiducial treatment are presented in the next section.

We run the MCMC analysis module `PyMC` (Patil et al. 2010) to maximise the likelihood and recover the scaling relation parameters between velocity dispersion of galaxy members and temperature of hot cluster gas. We assume a uniform priors on all parameters, with the following domain limits: $p \in [0, 1]$, $\sigma_p \in [50, 1000]\,{\rm km\ s^{-1}}$, $\alpha \in [-10, 10]$, and $\beta \in [-10, 10]$. The best-fit parameter values for the fiducial model and the restricted SSE model are given in Table 5.1.

For the fiducial treatment, the posterior constraint on the slope of galaxy velocity

dispersion scaling with temperature is $\alpha = 0.63 \pm 0.05$, is in tension with the self-similar expectation of 0.5. A slope steeper than self-similar could potentially arise from AGN feedback effects on the ICM. Recent simulations including AGN feedback exhibit shifts in the global ICM temperature of halos that are mass-dependent, with larger increases seen at lower masses (Le Brun et al. 2017; Truong et al. 2018). Since the galaxy velocity dispersion is not directly coupled to AGN activity, the impact on the ICM would lead to $\alpha > 0.5$.

We find no significant change in the scaling amplitude with redshift but our constraint is weak, $\beta = -0.49 \pm 0.38$. Since the fiducial analysis yields no evidence of redshift evolution, it is no surprise that the posterior SSE parameter values are identical to those of the fiducial analysis.

The Gaussian component amplitude, p, is close to, but significantly different from unity. While the value of 0.88 ± 0.02 is consistent with the 0.916 ± 0.004 value found by Rozo et al. (2015) in their study of SDSS redMaPPer clusters, differences in selection and measurement preclude a direct comparison. Besides sample selection differences, the SDSS galaxy velocities are pairwise with respect to the central galaxy's velocity, whereas ours are determined by the mean cluster redshift, z_c. Some of the difference could reflect mis-centering, as a larger fraction of mis-centered clusters both reduces p and increases σ_p (see the previous chapter). We defer detailed modeling of such selection effects to future work.

Normalized velocity residuals about the mean scaling behavior in the fiducial analysis are shown in Fig. 5.5. We bootstrap the galaxy sample to compute means and standard deviations of the PDF in 64 bins between -4 and 4 in v/σ_{gal}, and these are shown as points with error bars in the figure. The line is the model, a Gaussian of zero mean, unit variance and amplitude given by the fiducial best fit plus a constant

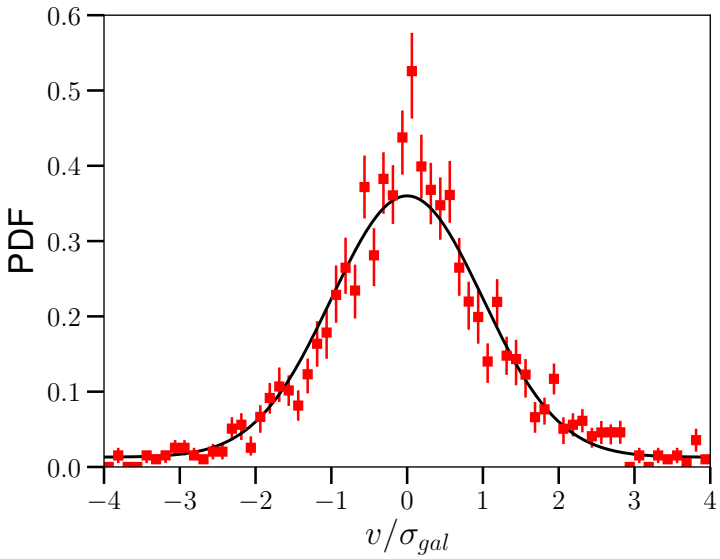

Figure 5.5: Normalized residuals of galaxy velocity about the mean scaling relation in the fiducial analysis. Red points show the data and the black line is the model, Equation 5.4, a mixture of a Gaussian and a uniform distribution. Error bars are calculated by bootstrapping the velocities of the spectroscopic sample, using 64 bins between -4 and 4 in $v_{\rm gal}/\sigma_{gal}$. See text for discussion of the goodness of fit.

background.

From Fig. 5.5, it is evident that our fit is not a good fit to data in the standard chi-squared sense. The normalized velocity PDF structure is very similar to that seen by Rozo et al. (2015) and in Chapter IV for redMaPPer clusters and simulations, respectively. We find $\chi^2/{\rm dof} = 74/44$ for $v_{\rm gal}/\sigma_{gal} \in [-3, 3]$. The estimated $\chi^2/{\rm dof}$ is less than that for the best-fit value found by Rozo et al. (2015) for SDSS redMaPPer clusters, $\chi^2_{\rm SDSS}/{\rm dof} = 96/26$.

While the centrally peaked nature of the normalized velocity PDF remains to be carefully modeled, two potential sources of systematics are likely to be important. One is projected large-scale structure; the simulations of Chapter IV show that only $\sim 60\%$ of the galaxies in the signal component of velocity space actually lie within r_{200} of the halo matched to each member of the cluster ensemble. Another is intrinsic scatter in $\sigma_{\rm gal} - T_X$, which will distort the Gaussian shape. The fact that the $\chi^2/{\rm dof}$

Table 5.2: Summary of published $\sigma_{gal} - kT_X$ scaling relation parameters, using the notation[1] of Equation 5.3

Source	σ_p (km s^{-1})	α	β	fitting method	N	redshift
This work	539 ± 16	0.63 ± 0.05	-0.49 ± 0.38	Ensemble ML	132	$z < 0.6$
Wilson et al. (2016)	497 ± 85	0.86 ± 0.14	-0.37 ± 0.33	ODR [2]	38	$z < 1$
Nastasi et al. (2014)	508 ± 147	0.64 ± 0.34	-	BCES bisector	15	$0.64 \leq z \leq 1.46$
Xue & Wu (2000)	523 ± 13	0.61 ± 0.01	-	ODR [2]	145	$z < 0.2$

[1] We note that sample definitions, analysis methods and notation vary across sources. Published intercepts are renormalized to the fixed pivot temperature and redshift used in Equation 5.3.
[2] Orthogonal Distance Regression

is smaller for the XXL sample compared to SDSS redMaPPer may reflect the fact that the intrinsic scatter in galaxy velocity dispersion is smaller at fixed temperature than at fixed richness, but differences in selection may also play a role.

Although the best fit is not a good fit to a Gaussian, the simulations of Chapter IV show that the derived galaxy velocity dispersion scaling is unbiased with respect to the log-mean value obtained by matching each cluster to the halo that contributes the majority of its galaxy members. Because the galaxy velocities in that simulation are unbiased relative to the dark matter by construction, the virial mass scaling derived from the galaxy velocity dispersion, $M(\lambda_{\rm RM}, z) \propto \sigma_p^3(\lambda_{\rm RM}, z)$, presents an unbiased estimate of the log-mean, membership-matched halo mass of the cluster ensemble. The reader interested primarily in mass scaling estimates can move directly to Section 5.6.

We turn next to comparing our scaling of galaxy velocity dispersion with gas temperature to previous work, and then explore the robustness of our parameter values in Section 5.5.

Comparison with previous studies

Soon after early observations of extended X-ray emission from clusters indicated a thermal gas atmosphere, a dimensionless parameter of interest emerged: the ratio of specific energies in galaxies and hot gas, $\beta_{\rm spec} = \sigma^2_{\rm gal}/(kT_X/\mu m_p)$, where μ is the

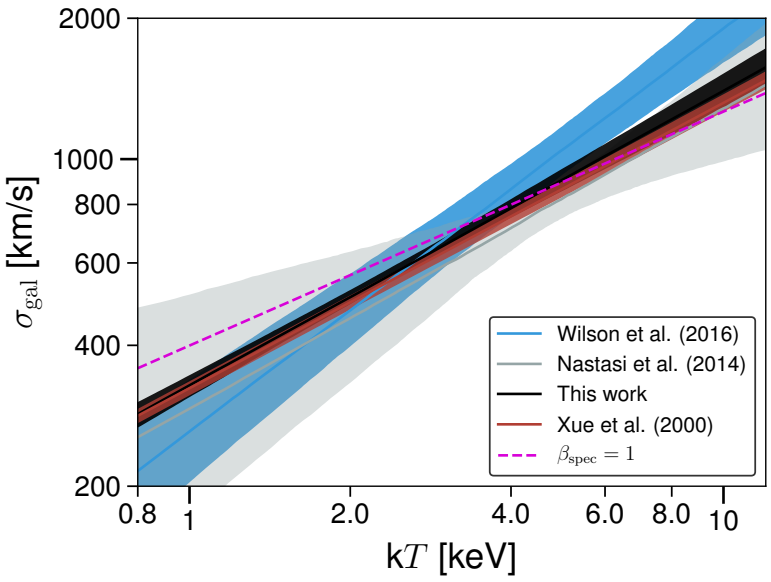

Figure 5.6: Comparison of the $\sigma_{\rm gal} - {\rm k}T_{\rm 300kpc}$ scaling relation of this work with prior literature, as labeled. Shaded regions are 1σ uncertainty on the expected velocity dispersion at given temperature. The magenta line is the locus of constant specific energy ratio, $\beta_{\rm spec} = \sigma^2_{\rm gal}/({\rm k}T_X/\mu m_p) = 1$ with $\mu = 0.6$. The slope of Wilson et al. (2016) suffers from a potential bias discussed in the text.

mean molecular weight of the plasma and m_p is the proton mass (note this beta is fundamentally different from the symbol used in Section 5.3).

Early estimates of this ratio in small observational samples (Mushotzky et al. 1978) and gas dynamic simulations (Evrard 1990; Navarro et al. 1995) yielded $\beta_{\rm spec} \approx 1$, consistent with a scenario in which both components are in virial equilibrium within a common gravitational potential. More recently, this ratio has been explored at high redshift; Nastasi et al. (2014) find $\beta_{\rm spec} = 0.85 \pm 0.28$ for 15 clusters with $z > 0.6$.

Figure 5.6 compares the fiducial scaling relation of this work to previous determinations in the literature. In addition, the dashed (magenta) line shows $\beta_{\rm spec} = 1$ assuming mean molecular weight $\mu = 0.6$, appropriate for a metal abundance of $0.3 Z_\odot$. Shaded regions show 1σ uncertainty on the expected velocity dispersion at a given temperature.

Table 5.2 summarizes the comparison with previous studies. The published scaling relations are re-evaluated at the pivot point of this work to be directly comparable. When appropriate, errors in the published slope are propagated to the normalization error.

The measured slope in temperature is consistent between our work and previous works. Wilson et al. (2016) find a slope 0.86 ± 0.14 for a sample of 38 clusters from the XMM Cluster Survey. Using simulations, however, they show that the orthogonal fitting method on their sample produces a substantial overestimate in slope, by ~ 0.3, in the test shown in their Table 7 and Fig. 9. They caution that their fit overestimates the velocity dispersion of clusters above 5 keV. Similarly Ortiz-Gil et al. (2004) uses the orthogonal fitting method and find a steep slope $\sim 1.00 \pm 0.16$ for a sample of 54 clusters.

If a bias correction is applied, the slope of Wilson et al. (2016) reduces to ~ 0.55, consistent with our findings. We note that a smaller shift of ~ 0.2 would bring the Ortiz-Gil et al. (2004) result into consistency with self-similarity at the 2σ level. For a heterogeneous sample constructed from the literature, Xue & Wu (2000) report a slope of 0.61 ± 0.01, consistent with our result.

The velocity dispersion normalizations given in Table 5.2 at the pivot temperature and redshift are all in good agreement within their stated errors. The 3% fractional uncertainty in our quoted normalization is among the tightest published constraints, comparable to the statistical error of the more heterogeneous sample of Xue & Wu (2000).

Table 5.3: Sensitivity analysis of $\sigma_{\rm gal} - {\rm k}T_{\rm 300kpc}$ inferred parameters. See text for further discussion.

Model	σ_p [km/s]	α	β	p	# Clusters	# Galaxies
Fiducial	539 ± 16	0.63 ± 0.05	-0.43 ± 0.38	0.88 ± 0.02	128	1592
Measured k$T_{\rm 300kpc}$ only	547 ± 17	0.60 ± 0.05	-0.39 ± 0.39	0.87 ± 0.02	103	1421
$r < 0.5 r_{500}$	509 ± 20	0.67 ± 0.07	-1.29 ± 0.50	0.90 ± 0.02	127	891
$r < 2.0 r_{500}$	557 ± 13	0.56 ± 0.04	0.42 ± 0.32	0.82 ± 0.02	131	2810
$v_{max} = 2000\,{\rm km\ s^{-1}}$ [1]	526 ± 18	0.62 ± 0.05	-0.50 ± 0.40	0.88 ± 0.02	128	1557
$v_{max} = 3000\,{\rm km\ s^{-1}}$ [1]	549 ± 15	0.63 ± 0.05	-0.45 ± 0.37	0.88 ± 0.02	128	1617
$\alpha_{V_{max}} = 0.3$ [2]	539 ± 16	0.61 ± 0.05	-0.46 ± 0.39	0.88 ± 0.02	128	1591
$\alpha_{V_{max}} = 0.7$ [2]	543 ± 16	0.65 ± 0.05	-0.48 ± 0.38	0.88 ± 0.02	128	1589
z_c [3] > 0.25	550 ± 32	0.58 ± 0.09	-0.82 ± 0.79	0.87 ± 0.02	84	814
z_c [3] ≤ 0.25	576 ± 48	0.63 ± 0.06	0.63 ± 1.42	0.88 ± 0.02	44	778

[1] Normalization of the maximum velocity threshold in Equation 5.2
[2] Slope in temperature of the maximum velocity threshold in Equation 5.2
[3] Cluster redshift.

5.5 Systematic errors and sensitivity analysis

In this section, we investigate sources of uncertainty in the scaling presented in the previous section, including survey selection and the sensitivity of the posterior parameters to the details of the spectroscopic sample used to define the signal region.

Table 5.3 summarizes the results of the tests presented below. A cursory look at the table indicates that most parameters shift by modest amounts, typically within one or two standard deviations of the fiducial result, with the exception of the Gaussian amplitude, p, discussed further below.

5.5.1 Temperature estimates

As presented in Section 5.2.1, the XXL temperatures are directly determined for 103 of the 128 clusters in our sample. A natural question to ask is whether our results are sensitive to the temperature estimation method applied to the remaining 25 clusters.

We first note that the 103 systems with measured $T_{\rm 300kpc}$ tend to be more massive at a given redshift, with higher galaxy richness. The higher richness translates into more galaxies with spectroscopy, and it turns out that this subset holds most of the

statistical weight of the spectroscopic sample. Within the fiducial r_{500} aperture, there are 1421 galaxies in the 103 clusters with direct temperatures, compared with 171 galaxies in the 25 clusters with inferred temperatures. So $\sim 90\%$ of the statistical weight comes from clusters with measured temperatures.

As a consistency check, we refit the scaling relation after removing all clusters with inferred temperature from the sample. The parameter constraints remain consistent with our fiducial analysis.

5.5.2 Angular aperture

The velocity dispersion of dark matter particles in simulations varies weakly as a function of distance from the halo centre (Old et al. 2013), and this effect has been confirmed observationally (Biviano & Girardi 2003). We test the sensitivity of our fit parameters by varying the angular aperture of inclusion by factors of $2^{\pm 1}$ from the fiducial value of r_{500}. We note that the size of the sample varies slightly as the aperture is changed. The main change is that a larger aperture induces a larger projection effect, evident from the Gaussian normalization, $p = 0.82 \pm 0.02$ for $2r_{500}$ versus $p = 0.90 \pm 0.02$ for $0.5r_{500}$. There are modest trends in the other parameters, including a slightly steeper slope $\alpha = 0.67 \pm 0.07$ at $0.5r_{500}$, and β is not consistent with 0 at the $\sim 2\sigma$ level at $0.5r_{500}$, but the statistical power of the sample is insufficient to determine these trends with high precision.

5.5.3 Signal component maximum velocity

Recall that the likelihood model is applied to a subset of all spectroscopic galaxies that lie in the signal region, with rest-frame velocities below a maximum value, $v_{\rm max}(T_{\rm 300kpc})$, given by Equation 5.2. We test the effect of this maximum by independently varying the amplitude by $\pm 500\,{\rm km\ s^{-1}}$ (or $\pm 20\%$) and the power-law index

by ± 0.2. The number of signal galaxies does not vary much with these changes, indicating that our fiducial cut is roughly identifying the caustic edge that separates bound and unbound galaxies in clusters (Miller et al. 2016). All parameters remain within 1σ of their fiducial values as these changes are made.

5.5.4 Redshift range

We take the pivot redshift in this chapter, $z_p = 0.25$, and split the full sample into high and low redshift subsets. For these, we do not find statistically significant deviations from the fiducial model parameters. The changes in the normalization, slope, redshift evolution, and parameter p are all less than 1σ. Although, as to be expected, there remains no effective constraints on the redshift evolution factor.

5.5.5 X-ray selection and Malmquist bias

The aim of our analysis is to produce unbiased estimates of the scaling relations inherent to the population of dark matter halos. Selection by X-ray flux and angular size (Pacaud et al. 2006) can introduce bias in the inferred $\sigma_{\rm gal} - {\rm k}T_{\rm 300kpc}$ scaling relation if there is non-zero covariance between X-ray selection properties and galaxy velocity dispersion (see Section 5.1 in Kelly 2007). Such data sets are said to be "truncated", and the truncation effects need to be explicitly modeled in the likelihood.

There have not yet been observational estimates of the correlation between galaxy velocity dispersion and X-ray properties at fixed halo mass. Halos in the Millennium Gas simulations of Stanek et al. (2010) show intrinsic correlation coefficients of ~ 0.3 for L_X and $\sigma_{\rm DM}$, where $\sigma_{\rm DM}$ is the velocity dispersion of dark matter particles in the halos. However, translating this estimate into correlations involving $\sigma_{\rm gal}$ projected along the line-of-sight is non-trivial and lies beyond the scope of this work. Redshift-

space projection presumably dilutes any intrinsic halo correlation, unless the source of the projected velocity component also carries associated X-ray emission.

The magnitude of potential selection biases can be addressed by simulating the entire process of survey selection and subsequent spectroscopic analysis, along the lines of that done by the previous chapter for redMaPPer optical selection. We defer that work to future analysis. From the perspective of halo mass estimation, corrections to the velocity dispersion scaling from sample selection are likely to be smaller than the systematic uncertainty associated with galaxy velocity bias, as discussed below.

5.6 Ensemble dynamical mass scaling of XXL clusters

In Chapter IV we use sky realizations derived from lightcone outputs of cosmological simulations to show that the mass determined through virial scaling of the ensemble, or stacked, pairwise velocity dispersion offers an unbiased estimate of the log-mean mass of halos matched via joint galaxy membership. Here, we apply this approach to the fiducial velocity dispersion scaling in order to estimate the characteristic mass scale, $\langle \ln M_{200} | T_X \rangle$ of XXL clusters as a function of temperature at the pivot redshift, $z_p = 0.25$.

The simulation of the previous chapter assumed galaxies to be accurate tracers of the dark matter velocity field, but real galaxies may be biased tracers. To estimate the velocity dispersion of the underlying dark matter from the galaxy redshift measurements, we introduce a velocity bias factor, b_v, defined as the mean ratio of galaxy to dark matter velocity dispersion within the target projected r_{200} region used in our analysis. The normalization of the dark matter velocity scaling with temperature is

then

$$\sigma_{p,DM} = \frac{\sigma_p}{b_v}\,, \tag{5.5}$$

where σ_p is the galaxy normalization with temperature, Equation 5.3.

Following the analysis presented in Chapter IV, we proceed by: i) imposing an external b_v estimate to derive the normalization of the *dark matter* virial velocity scaling with X-ray temperature, then ii) applying the dark matter virial relation calibrated by Evrard et al. (2008) to determine the scaling of total system mass with temperature.

We use $b_v = 1.05 \pm 0.08$ which is an empirical estimate derived from redshift-space clustering of bright galaxies by Guo et al. (2015a). A similar value of 1.06 ± 0.03 is found in the simulation study of Wu et al. (2013), although that study found galaxy bias slightly below 1 for the brightest galaxies.

According to Guo et al. (2015a) the velocity bias runs with the absolute magnitude of selected galaxies. Figure 5.7 show the distribution of absolute r-band magnitude of selected galaxies in this work. We note that the peak of distribution of absolute r-band magnitude of selected galaxies in this work is $M_r = 21.5$, which is consistent with the brightest galaxy sample of (Guo et al. 2015a). This result justifies the choice of our prior distribution, $b_v = 1.05 \pm 0.08$ found by Guo et al. (2015a) for this magnitude threshold.

Using a velocity bias of 1.05 ± 0.08 leads to an estimate of the dark matter velocity dispersion at the pivot temperature and redshift,

$$\sigma_{p,DM} = 516 \pm 43 \text{ km/s}\,. \tag{5.6}$$

We note that $\sigma_{p,DM}$ uncertainty has contribution from the b_v prior and σ_p posterior.

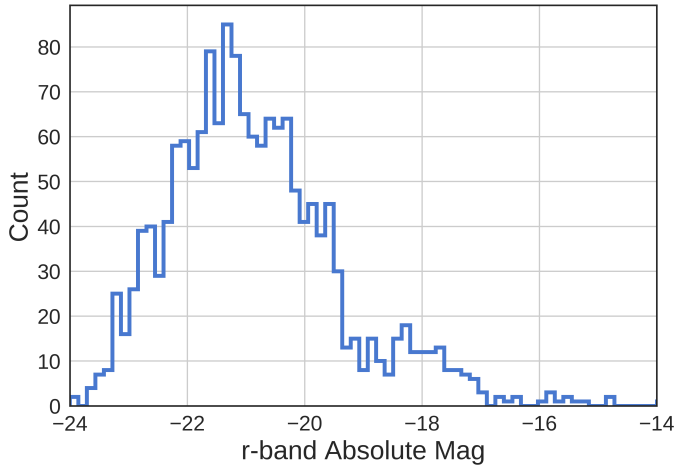

Figure 5.7: The distribution of r-band absolute magnitude for selected galaxies after applying the fiducial aperture and velocity cuts.

The virial scaling of halos in simulations displays a linear relationship between the cube of the dark matter velocity dispersion, $\sigma^3_{p,DM}$, and a mass measure, $E(z)M_\Delta$, where $E(z) = H(z)/H_0$ is the normalized Hubble parameter. Using Equation 6 and Table 3 of Evrard et al. (2008) along with $h = 0.7$, the total mass within r_{200} at the pivot temperature and redshift is

$$\langle \ln(M_{200}/10^{14}\ \mathrm{M}_\odot)\rangle = 0.33 \pm 0.24\,, \tag{5.7}$$

corresponding to $M_{200} = (1.39^{+0.37}_{-0.30}) \times 10^{14}\ \mathrm{M}_\odot$.

The full velocity scaling implies a log-mean mass for the XXL selected cluster sample of

$$\left\langle \ln\left(\frac{E(z)M_{200}}{10^{14}\mathrm{M}_\odot}\right) | T, z \right\rangle = \pi_T + \alpha_T \ln\left(\frac{T}{T_p}\right) + \beta_T \ln\left(\frac{E(z)}{E(z_p)}\right)\,, \tag{5.8}$$

with intercept $\pi_T = 0.45 \pm 0.24$, temperature slope $\alpha_T = 3\alpha = 1.89 \pm 0.15$, redshift slope $\beta_T = 3\beta = -1.29 \pm 1.14$. Recall that this result is based on 300 kpc temperature estimates, $T \equiv T_{300\mathrm{kpc}}$.

Biviano et al. (2006) have examined the robustness of virial mass estimates in a cosmological hydrodynamic simulation. They find that dynamical mass estimates are

reliable for densely sampled clusters (over 60 cluster members). Due to the ensemble technique adapted here, this work does not suffer from sparse sampling of cluster members. Generally speaking, stacking techniques reduce the noise associated with sparse samples, at the price of not constraining the intrinsic scatter.

While we explicitly remove extreme projected outliers in velocity space (see Fig. 5.3) and account for a residual, constant contribution in the velocity likelihood, it is worth noting that the central Gaussian component has contributions from galaxies that do not lie in the main source halo. While this component retains some degree of projected galaxies, we previously showed that the dynamically-derived mass is a robust estimate of log-mean mass at a given observable, in that case $\langle \ln M_{200} | \lambda_{\rm RM}, z \rangle$. While the optical and X-ray samples are selected differently, not enough is known about hot gas and galaxy property covariance to model selection effects precisely. We discussed in Section 5.5.5 why selection effects are unlikely to imprint significant bias into the inferred scaling relation.

5.6.1 Comparison with previous studies

Figure 5.8 compares the mass-temperature scaling relation, a dynamical mass estimates, derived in this chapter with previous studies that use weak lensing (XXL Paper IV) and hydrostatic (Arnaud et al. 2005) mass estimates. Overall, there is a good agreement within the uncertainties.

The data points with error bars are weak lensing estimates of M_{200} for a subsample of the 100 brightest clusters in XXL (XXL Paper IV). In order to directly compare our MPR with XXL Paper IV and other works, we evaluate all results at $z = 0$ using $h = 1$. When shifting the normalization, we assume SSE, $\beta_T = 0$, yielding $\pi_T = 0.09 \pm 0.25$.

Assuming self-similar redshift evolution, XXL Paper IV estimated the mass - tem-

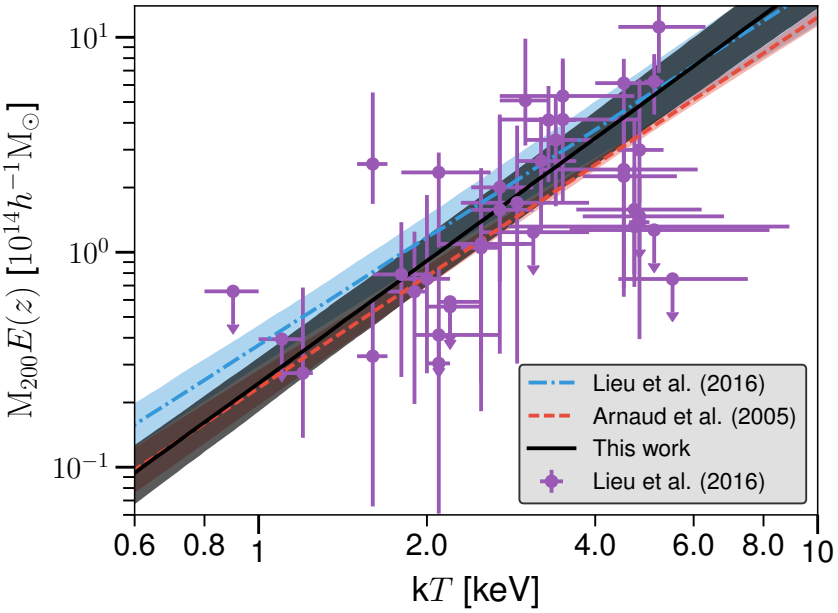

Figure 5.8: The $M_{200} - kT$ scaling relation from this work (black line and dark shaded region) is compared with published relations given in the legend and Table 5.4. Shaded regions are the 1σ uncertainty in the expected mass at a given temperature. See the text for more discussion.

perature scaling relation using a subsample of 38 out of 100 brightest XXL clusters. To improve their constraint, their sample is complemented with weak lensing mass measurements from clusters in the COSMOS (Kettula et al. 2013) and CCCP (Hoekstra et al. 2015) cluster samples. While the data points plotted in Fig. 5.8 are taken directly from XXL Paper IV, their published MPR is framed in terms of M_{500}. We therefore convert the normalization to M_{200} using an NFW profile with concentration $c = 3.1$, the median value of the XXL Paper IV sample, for which $M_{200}/M_{500} = 1.4$. The slope of the weak lensing relation lies within $\sim 1\sigma$ of the self-similar expectation of 1.5.

The assumption of hydrostatic equilibrium is commonly used to derive masses from X-ray spectral images, and Arnaud et al. (2005) apply this method to a sample of ten nearby, $z < 0.15$, relaxed clusters in the X-ray temperature range $[2 - 9]$ keV. The masses are derived from NFW fits to the mass profiles, obtained under the

Table 5.4: Comparison of the mass normalization, $\ln A = \langle \ln(M_{200}/10^{14}\,h^{-1}\mathrm{M}_\odot) \,|\, \mathrm{k}T_X = 2.2\,\mathrm{keV}, z = 0 \rangle$, and slope of the mass–temperature determined by the works listed.

Paper	$\ln A$	Slope	Mass Proxy	Number of Clusters	redshift
This work	0.09 ± 0.25	1.89 ± 0.15	Dynamical Mass	132	$z < 0.6$
XXL Paper IV [1]	0.31 ± 0.23	1.67 ± 0.14	Weak-lensing Mass	96	$0.1 < z < 0.6$
Kettula et al. (2015) [2]	0.43 ± 0.17	1.73 ± 0.19	Weak-lensing Mass	70	$0.1 \leq z \leq 0.5$
Arnaud et al. (2005) [3]	-0.09 ± 0.09	1.72 ± 0.10	Hydrostatic Mass	10	$z < 0.16$

[1] The normalization is converted from M_{500} to M_{200} as described in the text.
[2] CFHTLenS + CCCP + COSMOS cluster sample.
[3] Spectroscopic temperature within the $0.1r_{200} \leq r \leq 0.5r_{200}$ region. All clusters.

hydrostatic assumption using measurements from the XMM-Newton satellite. We note that they use a core-excised spectroscopic temperature from a $0.1r_{200} \leq r \leq 0.5r_{200}$ region. Our result is consistent with that of Arnaud et al. (2005) within their respective errors.

Kettula et al. (2015) combine 12 low mass clusters from the CFHTLenS and XMM-CFHTLS surveys with 48 high-mass clusters from CCCP (Hoekstra et al. 2015) and 10 low-mass clusters from COSMOS (Kettula et al. 2013). From this sample of 70 systems, they measure a mass - temperature scaling relation with slope 1.73 ± 0.19 for M_{200}. When M_{500} is used, they find a slope of 1.68 ± 0.17 which they argue may be biased by selection. Applying corrections to this (Eddington) bias, they find a slope of 1.52 ± 0.17, consistent with self-similarity.

Table 5.4 summarizes these comparisons, showing the slopes and normalizations scaled to $z = 0$ for a pivot X-ray temperature of 2.2 keV. The expected log mass is the largest for weak-lensing proxies, and smallest under the hydrodynamic assumption, but they are statistically consistent within their stated $10 - 20\%$ errors. The slope derived in this work is statistically consistent with the scalings derived from weak-lensing and hydrostatic techniques. In agreement with prior work, we find a significantly ($> 2.5\sigma$) steeper slope than the expected self-similar value of 1.5. A more precise comparison would need to take into account different approaches to measuring

X-ray temperature, as well as potential instrument biases (Zhao et al. 2015; Schellenberger et al. 2015). For example, Arnaud et al. (2005) and Kettula et al. (2015) measure core-excised temperatures within r_{200} while the temperatures used in this work are measured within fixed physical radius. Comparing the non-core excised temperatures of XXL clusters with the core excised temperatures used by Kettula et al. (2013), XXL Paper IV found a mean ratio of $\langle T_{300\ \mathrm{kpc}}/T_{0.1-0.5r_{500,WL}}\rangle = 0.91 \pm 0.05$.

Several independent hydrodynamic simulations that incorporate AGN feedback, including models from variants of the Gadget code (cosmo-OWLS; Le Brun et al. 2017; Truong et al. 2018) as well as RAMSES Rhapsody-G (Hahn et al. 2017a), find slopes near 1.7 for the scaling of mean mass with spectroscopic temperature. These results are in agreement with our finding. We note that the cluster sample used in this work is dominated by systems with $kT_X < 3$ keV, while Lieu et al. (2016)'s cluster sample is dominated by clusters with temperature above 3 keV. Slopes steeper than the self similar prediction for low temperature systems have been noted in preceding observational works as well (e.g., Arnaud et al. 2005; Sun et al. 2009; Eckmiller et al. 2011).

5.6.2 Velocity bias

Similar to the previous chapter the dominant source of systematic uncertainty in ensemble dynamical mass estimates comes from the uncertainty in the velocity bias correction.

Dynamical friction is a potential physical cause for the velocity bias that would generally drive galaxy velocities to be lower than that of dark matter particles within a halo (e.g., Richstone 1975; Cen & Ostriker 2000; Yoshikawa et al. 2003). On the other hand, clusters that are undergoing mergers tend to have galaxy members with a larger velocity dispersion relative to the dark matter particles (Faltenbacher &

Diemand 2006), and merging of the slowest galaxies onto the central galaxy could also tend to drive b_v to be greater than one. These competing effects are subject to observational selection in magnitude, color, galaxy type, star formation activity and aperture which need to be addressed with larger sample size. There is growing observational evidence that velocity bias is a function of the aforementioned selection variables (e.g., Guo et al. 2015a; Barsanti et al. 2016; Bayliss et al. 2017).

The space density of clusters as a function of velocity dispersion also constrains the velocity bias in an assumed cosmology, and (Rines et al. 2007) find $b_v = 0.94 \pm 0.05$ and 1.28 ± 0.06 for WMAP1 and WMAP3 cosmologies, respectively. The quoted errors are statistical and based on a sample of 72 clusters in the SDSS DR4 spectroscopic footprint. The study of Maughan et al. (2016) compares caustic masses derived from galaxy kinematics (e.g., Diaferio 1999; Miller et al. 2016) with X-ray hydrostatic masses. Such a comparison yields a measure of relative biases in hydrostatic and caustic methods, and their finding of $1.20^{+0.13}_{-0.11}$ for the ratio of hydrostatic to caustic M_{500} estimates is consistent with unity at the $< 2\sigma$ level. If incomplete thermalization of the intracluster plasma leads hydrostatic masses to underestimate true masses by 20% (e.g., Rasia et al. 2006, and references therein), then the central value of Maughan et al. (2016) indicates that caustic masses would further underestimate true masses. Because of the relatively strong scaling $M \propto b_v^{-3}$, a value $b_v \simeq 0.9$ would suffice for consistency.

Redshift space distortions provide another means to test velocity bias (Tinker et al. 2007). The current constraints from Guo et al. (2015b,a) indicate a magnitude-dependent bias, with b_v^{-1} changing from slightly above one for bright systems — the value $b_v = 1.05 \pm 0.08$ we employ in Section 5.6 to infer total mass — to slightly below one for fainter galaxies. Oddly, this trend is opposite to that inferred for

galaxies from both hydrodynamic and N-body simulations, where bright galaxies are kinematically cooler than dimmer ones (Old et al. 2013; Wu et al. 2013). The recent observational study of (Bayliss et al. 2017) finds a similar trend.

In summary, studies are in the very early stages of investigating velocity bias in the non-linear regime, both via simulations and in observational data. The statistical precision of future spectroscopic surveys, such as DESI (DESI Collaboration 2016), will empower future analyses that may produce more concrete estimates of b_v as a function of galaxy luminosity and host halo environment.

Given the current level of systematic error in mass calibration, our ensemble velocity result is consistent with the weak-lensing mass calibration results of XXL Paper IV. Similarly, the weak lensing results of Simet et al. (2017) and Melchior et al. (2017) for redMaPPer clusters agree with the estimates of Chapter IV. Better understanding of the relative biases of weak lensing, hydrostatic and other mass estimators will shed light on the magnitude of velocity bias in the galaxy population.

5.7 Chapter Conclusion

We model ensemble kinetic motions of galaxies as a function of X-ray temperature to constrain a power-law scaling of mean galaxy velocity dispersion magnitude, $\langle \ln \sigma_{\rm gal} | T_{\rm 300kpc}, z \rangle$ for a sample of 132 spectroscopically confirmed C1 and C2 clusters in the XXL survey. Spectroscopic galaxy catalogues derived from GAMA, SDSS DR10, VIPERS, VVDS and targeted follow-up surveys provide the input for the spectroscopic analysis. From the kinetic energy, we derive total system mass using a precise dark matter virial calibration from N-body simulations coupled with a velocity bias degree of freedom for galaxies relative to dark matter.

Following Chapter IV, we employ a likelihood model for galaxy–cluster relative

velocities, after removal of high-velocity outliers, and extract underlying parameters by maximizing the likelihood using an MCMC technique. The analysis constrains the behavior of a primary Gaussian component, containing $\sim 90\%$ of the non-outlier galaxies, the width of which scales as a power law with temperature, as anticipated by assuming self-similarity (Kaiser 1986).

Based on 1908 galaxy-cluster pairs, we find a scaling steeper than self-similarity,

$$\left\langle \ln\left(\frac{\sigma_{\rm gal}}{\rm km\ s^{-1}}\right) \mid T_{\rm 300kpc}, z = z_p \right\rangle = \ln(\sigma_p) + \alpha \ln\left(\frac{T_{\rm 300kpc}}{2.2\,\rm keV}\right) , \tag{5.9}$$

with $\sigma_p = 539 \pm 16$ and $\alpha = 0.63 \pm 0.05$ at a pivot redshift of $z_p = 0.25$. While redshift evolution is included in the likelihood, the data are not sufficiently dense at high redshift to establish a meaningful constraint on evolution.

We identify and characterise several sources of systematic error and study the sensitivity of inferred parameters to the galaxy selection model and assumptions of the stacked model. The method is largely robust (Table 5.3). It is worth noting that these systematic error sources are generally different from those of other mass calibration methods, such as weak-lensing and hydrostatic X-ray methods, which allows the XXL survey to have an independent estimate of the cluster mass scale.

Employing the precise N-body virial mass relation calibrated in Evrard et al. (2008) coupled with an external constraint on galaxy velocity bias, $\sigma_{\rm gal}/\sigma_{\rm DM} = 1.05 \pm 0.08$, we derive a halo mass scaling

$$\begin{aligned} &\left\langle \ln\left(\frac{E(z) M_{200}}{10^{14}\,\rm M_\odot}\right) \mid T_{\rm 300kpc}, z = z_p \right\rangle = \\ &\pi_T + \alpha_T \ln\left(\frac{T_{\rm 300kpc}}{2.2\,\rm keV}\right) + \beta_T \ln\left(\frac{E(z)}{E(0.25)}\right) , \end{aligned} \tag{5.10}$$

with normalization, $\pi_T = 0.45 \pm 0.24$, and slopes, $\alpha_T = 1.89 \pm 0.15$ and $\beta_T = -1.29 \pm 1.14$.

Within the uncertainties, our result is consistent with mass scalings derived from

both weak-lensing measurements of the XXL sample (XXL Paper IV) and provides an independent X-ray analysis using the hydrostatic assumption to obtain mass. But uncertainties in the scaling normalization remain at the level of $10 - 25\%$ (see Table 5.2), and fractional errors in slope are also of order ten percent.

We note that the dominant source of uncertainty in our mass estimator is not statistical, but systematic uncertainty due to the galaxy velocity bias. Deeper and denser spectroscopic surveys, partnered with sophisticated sky simulations, will enable richer analyses than that performed here. As the accuracy of weak lensing and hydrostatic mass estimates improve, the ensemble method we employ here could be inverted to constrain the magnitude of velocity bias at small scales from future surveys such as DESI (DESI Collaboration 2016). Such an approach has recently been applied to a small sample of Planck clusters by Amodeo et al. (2017).

Larger numbers of spectroscopic galaxies at $z > 0.5$ are needed to constrain the redshift evolution. In recent hydrodynamic simulations that incorporate AGN feedback, Truong et al. (2018) present evidence for weak redshift evolution in the slope of the mass-temperature scaling relation at $z < 1$, with stronger evolution at $z > 1$. Next generation X-ray missions, such as eROSITA (Merloni et al. 2012) and Lynx (Gaskin et al. 2015), will offer the improved sensitivity needed to identify and characterise this population. In the meantime, deeper XMM exposures over at least a subset of the XXL area can be used to improve upon the modest constraints on evolution we obtain using the current 10 ksec exposures.

The best practice in comparing the forthcoming, more sensitive observational data with theoretical models will require generating synthetic light-cone surveys from simulations and applying the same data reduction techniques to the models and observations.

An extension that we leave to future work is to properly include temperature errors into the ensemble spectroscopic likelihood model. Richer data will allow investigation of potential modifications to the simple scaling model assumed here, including testing for deviations from self-similarity (in the redshift evolution of the normalization or a redshift-dependent slope, for example) and potential sensitivity to the assembly history or large-scale environment of clusters.

CHAPTER VI

A Powerful Hierarchical Bayesian Model for Analyzing Multi-wavelength Observables of Galaxy Clusters

Philosophy and Contribution

In this chapter, I develop and implement a new likelihood model which accounts for the effects of sample selection, error covariance, and unobserved true halo masses. I then apply this model to a sub set of LoCuSS cluster sample to estimate the scaling relation parameters. The results provide a full calibration of conditional mass–observables probability distribution for a wide range of observables. These relations are a key element that enables a cosmological analysis with cluster samples. This chapter is a modified version of Mulroy et al. (in preparation): "LoCuSS: Galaxy Cluster Scaling Relations" to which I have contributed substantially. The analysis and plots, except Figure 1, presented in this chapter are my own work. The data is obtained and reduced by the collaborators.

6.1 Chapter Introduction

In Chapter II, it is illustrated that the abundance of galaxy clusters is a strong function of the halo mass hosting the system. A cosmological analysis with these systems requires an accurate calibration of the probability distribution of observables conditioned on halo mass. This requirement has motivated a significant effort to

find and calibrate observable quantities which correlate with halo mass, i.e. scaling relations (e.g., Giodini et al. 2013). As discussed in Chapter II, the baryon content of these systems is observable, either in stellar material or in hot intracluster gas (e.g., Gonzalez et al. 2013; Chiu et al. 2016). The derived properties of the stellar and gas content of clusters are sensitive to a wide range of physical effects, including cooling, star formation, feedback and accretion, which are not quantified with the self-similar model of Chapter II. The observable properties of gas and stellar material and their scaling with respect to the total cluster mass, which can give direct insight into the physics of these processes, are the subject of this Chapter. In Chapter IV and Chapter V, we develop an algorithm which estimates the mean-log mass of clusters. That algorithm, however, does not determine the full conditional probably distribution that is required for a cosmological analysis. To determine this probability distribution, a mass estimation for each individual cluster is required. In this Chapter our primary goal is to estimate the full conditional mass–observables relation, $p(\mathbf{s}|\mu, z)$. This includes mean relations, the scatter about each mean relation, and the off-diagonal elements of the property covariance. The first two relations are discussed in this Chapter and the latter is discussed in Chapter VII.

Ideally we would like to constrain the scaling relation of an observable with the "true" mass of the cluster; however, in practice the true halo mass is not directly measurable. A popular method of mass measurement uses X-ray properties together with the simplifying assumption of hydrostatic equilibrium (e.g., Mathews 1978; Sarazin 1988; Vikhlinin et al. 2006). More recently, significant progress has been made in using the weak-lensing signal to probe the mass of galaxy clusters. When carefully accounting for systematic effects, these masses are thought to be on average unbiased with respect to the "true" mass (e.g., Oguri & Hamana 2011; Becker &

Kravtsov 2011; Bahé et al. 2012). Crucially, these measurements do not rely on the assumption of hydrostatic equilibrium; therefore are more reliable.

Perhaps the most often overlooked requirement for calibrating robust scaling relations is a clear understanding of the cluster sample selection and correction for the resulting selection biases. As each observable has a non-zero scatter in its relation with mass, selection based on anything but 'true' mass can cause biases in the derived relations relative to those of the underlying halo population. The latter are often characterized by cosmological simulations (e.g., Le Brun et al. 2017). Cluster samples are commonly selected from optical, X-ray or Sunyaev-Zel'dovich (SZ) surveys (Rozo et al. 2009; Böhringer et al. 2004; Bleem et al. 2015), and constraints are ultimately limited by understanding of the selection function and the sample size.

The 41 clusters in this chapter and the next chapter are particularly well studied over a wide range of wavelengths (e.g., Marrone et al. 2012; Martino et al. 2014; Mulroy et al. 2014; Haines et al. 2015; Okabe & Smith 2016). Combined with a well described selection function, they provide the first cluster sample with which to simultaneously constrain scaling relations for X-ray, SZ and optical observables.

In Section 6.2, we describe our cluster sample, its selection and the wide range of multi-wavelength data which we use in this chapter and next chapter. In Section 6.3, we describe a new analysis tool, Hierarchical Bayesian method, to fit the scaling relations. We present our results in Section 6.4. We, then, discuss these results and compare them to the literature in Section 6.5. Finally, we conclude in Section 6.6.

Chapter's Notation. We assume $\Omega_{\rm M} = 0.3$, $\Omega_{\Lambda} = 0.7$ and $H_0 = 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$. In this cosmology, at the average cluster redshift, $\langle z \rangle = 0.22$, 1 arcsec corresponds to a projected physical scale of 3.55 kpc. We employ a spherical mass and radius convention, M_{500} and R_{500}, based on a mean, enclosed density of 500 times the critical

density evaluated in the above cosmology.

6.2 Data

We study a sample of 41 X-ray luminous clusters from the "High-L_X" sample of the LoCuSS, which was selected from the ROSAT All Sky Survey catalogues (RASS, Ebeling et al. 1998, 2000; Böhringer et al. 2004). These are all the clusters satisfying a clearly defined selection criteria: $n_H < 7 \times 10^{20}\mathrm{cm}^{-2}$; $-25^\circ < \delta < +65^\circ$; and an X-ray luminosity threshold of $L_{X,RASS}E(z)^{-1} > 4.4 \times 10^{44}$erg/s for clusters between $0.15 < z < 0.24$, and $L_{X,RASS}E(z)^{-1} > 7.0 \times 10^{44}$erg/s for clusters between $0.24 < z < 0.30$ (Figure 6.1), where $E(z) \equiv H(z)/H_0 = \sqrt{\Omega_\mathrm{M}(1+z)^3 + \Omega_\Lambda}$ is the evolution of the Hubble parameter. Therefore the only physical selection variable for this sample of galaxy clusters is the RASS X-ray luminosity, $L_{X,\mathrm{RASS}}$.

The $L_{X,\mathrm{RASS}}$ measurements cover the soft X-ray band from 0.1 to 2.4 keV, and are taken from the ROSAT Brightest Cluster Sample and its low flux extension (BCS, Ebeling et al. 1998; eBCS, Ebeling et al. 2000) for objects in the northern hemisphere, and the ROSAT-ESO Flux Limited X-ray galaxy cluster survey (REFLEX, Böhringer et al. 2004) for objects mostly in the southern hemisphere ($\delta < 2.5^\circ$). For the clusters in the overlap between surveys (Abell0267: BCS, REFLEX and Abell2631: eBCS, REFLEX) we average the luminosities and errors. RASS luminosities are not core-excised due to the angular resolution of the instrument, and so are sensitive to the presence, or absence, of a cool-core.

We observed this sample of clusters at X-ray, optical, near-infrared, and millimeter wavelengths over the period 2005-2014, building up a unique and comprehensive dataset. The main facilities that we used are *Chandra*, *XMM-Newton*, Suprime-CAM on the Subaru telescope, Hectospec on the Multiple Mirror Telescope (MMT),

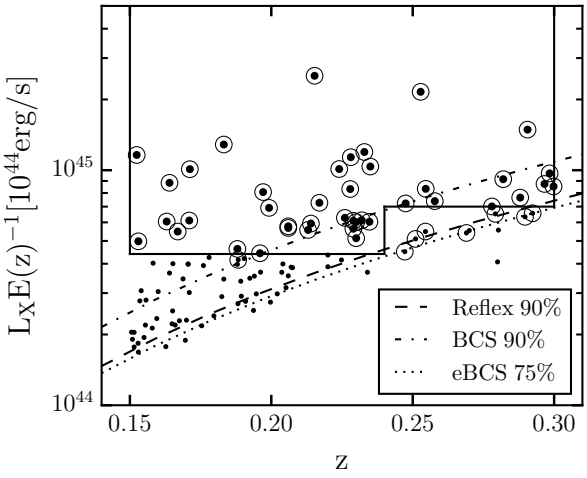

Figure 6.1: The $L_{X,\mathrm{RASS}}E(z)^{-1}-$ redshift distribution of the LoCuSS clusters; The large points show the 41 clusters passing the selection criteria and therefore used in this chapter and the next chapter, while the circles show the LoCuSS "High-L_X" clusters. The straight lines show the selection criteria, the curves show the completeness limits for (e)BCS (Ebeling et al. 1998, 2000) and REFLEX (Böhringer et al. 2004).

WFCAM on the United Kingdom Infrared Telescope (UKIRT), and the Sunyaev-Zeldovich array. The total investment of telescope time amounts to several million seconds. The details of the respective observations are provided in the wavelength specific articles that are cited in the following sections, that describe the measurements of galaxy cluster mass and observables that are used in this article.

6.2.1 Gravitational Weak-Lensing Masses

We use weak-lensing masses from Okabe & Smith (2016) (as tabulated in their Table 2), who calculate masses by fitting an NFW (Navarro, Frenk & White 1997) mass profile to the shear profile obtained from Subaru/Suprime-Cam observations. We use M_{500} values, defined as the mass within radius r_{500}, the radius within which the average density is $500 \times \rho_{\mathrm{crit}}(z)$, where $\rho_{\mathrm{crit}}(z) = 3H(z)^2/8\pi G$, the critical density of the Universe. We adopt these weak-lensing determined radii, $r_{500,\mathrm{WL}}$, as the radii within which we measure the other observables in this work (except Y_X and λ). The systematic biases in the ensemble calibration of the weak-lensing mass calculations

are controlled at $\sim$ 4 per cent level, based on careful selection of red background galaxies, extensive tests of both faint galaxy shape measurement methods and mass profile fitting methods (Okabe & Smith 2016). The measurement errors on M_{500} include contributions from shape noise, photometric redshift uncertainties and uncorrelated large-scale structure.

6.2.2 X-Ray Observables

We use X-ray measurements of the ICM described in Martino et al. (2014), where most clusters were observed with the *XMM-Newton* EPIC or *Chandra* ACIS-I detectors, except for Abell0611 and ZwCl0949.6+5207 which were only observed with the *Chandra* ACIS-S detectors. As shown in Martino et al. (2014), The emission measure profiles were robust to X-ray telescope cross-calibration issues for the selected energy band.

We consider bolometric luminosity $L_{X,\mathrm{ce}}$ and the average gas temperature $T_{X,\mathrm{ce}}$ within an annulus of $[0.15 - 1]r_{500,\mathrm{WL}}$ to avoid the measurements being contaminated by emission from the core. However the gas mass, M_{gas} , is measured within $r_{500,\mathrm{WL}}$. We also measure the integrated pressure proxy, Y_X , for all but the two clusters with ACIS-S observations. Defined as the product of gas mass and average temperature (Kravtsov et al. 2006), it is the X-ray equivalent of the SZ parameter described in Section 6.2.3.

Both the luminosity and the Y_X parameter derive from spherically symmetric templates of the X-ray emission measure per volume unit, $[\mathrm{n_p n_e}](\mathrm{r})$, that were projected along the line of sight, radially averaged and fitted to radial profiles of the soft [0.5 - 2] keV X-ray surface brightness. The bolometric estimate of $L_{X,\mathrm{ce}}$ derives from an extrapolation of the soft surface brightness assuming the spectral energy distribution of the ICM to correspond to a redshifted isothermal plasma with average

temperature $\langle k_B T \rangle$.

For each cluster we estimated the Y_X parameter by iterating about an existing Y_X - M_{500} scaling relation, yielding a characteristic radius r_{500}, different from the weak-lensing $r_{500,WL}$ radius within which the other X-ray observables are measured. For clusters observed with *XMM-Newton* we use the relation of Arnaud et al. (2010), and for those observed with *Chandra* we use the relation of Vikhlinin et al. (2009). Both relations are calibrated using hydrostatic mass estimates in a nearby cluster sample. The gas masses were computed from spherical integrals of the gas density profiles, $\mathrm{n_p(r)}$, and the gas temperatures correspond to spectroscopic measurements within projected [0.15 - 0.75] r_{500} and [0.15 - 1] r_{500}, following the prescription of the relevant scaling relation study.

6.2.3 Millimetre Observables – Sunyaev-Zel'dovich Effect

The SZ effect is caused by the inverse compton scattering of CMB photons by hot electrons, in this case in the ICM. These interactions boost the photon energy by $\sim \mathrm{k_B T/m_e c^2}$, leading to a characteristic distortion of the CMB spectrum in the direction of galaxy clusters. The CMB intensity is decreased below ∼220 GHz and increased above, in proportion to the 'comptonization' parameter, Y, which is an integral of the product of the electron density and temperature through the cluster. This integral of thermal pressure in the ICM, which is roughly in hydrostatic equilibrium with the gravitational potential well, should therefore be closely related to cluster mass.

Sunyaev-Zel'dovich Array – One of the SZ measurement data sets employed in this chapter is based on observations with the Sunyaev-Zel'dovich Array (SZA), an interferometer comprising eight 3.5-meter antennas observing at 27-35 GHz. During the period of these observations, from 2006 to 2014, the SZA initially observed from the floor of the Owens Valley, near Big Pine, CA, and later was relocated to the

nearby Cedar Flat site of the Combined Array for Research in Millimeter-wave Astronomy (CARMA). For all observations presented here the SZA antennas observed as an 8-element array, rather than in concert with other CARMA antennas as in, e.g., Plagge et al. (2013). The SZA was configured with six antennas in a compact configuration to maximize sensitivity to the large-scale cluster signal, with the remaining two antennas placed as 'outriggers' to discriminate the emission from point-like radio sources from the SZ signature of clusters. The resolution of the compact array was approximately 2 arcmin, while baselines to the outrigger antennas yield a resolution closer to 20 arcsec.

The SZ signal for each cluster was modeled as a generalized-NFW pressure profile (Nagai et al. 2007) using the parameters determined by Planck Collaboration et al. (2014) from a joint fit to SZ and X-ray profiles of 62 massive clusters. These parameters include a concentration parameter, c_{500}, the ratio of r_{500} to the scale radius ($r_{\rm s}$) of the pressure profile. The WL-derived values of r_{500} and their uncertainties were used to define a gaussian prior for the value of the scale radius, $r_{\rm s} = r_{500}/c_{500}$.

We are able to measure $Y_{\rm SZA}$ for 30 of the 41 clusters, finding that the fields for nine are contaminated and that two clusters (RXCJ2102.1-2431 and ZwCl0857.9+2107) are non-detections. The two non-detections are near the low end of the sample lensing mass limit. The contaminated clusters contain 30 GHz sources that are not point-like at the 20 arcsec resolution of the SZA long baselines. In such cases, the interferometric measurement cannot cleanly distinguish between emission from spatially extended radio sources and the spatially extended SZ effect signal, which appears as 'negative' emission. The degeneracy between extended radio source emission and cluster SZ signal makes the SZ measurements unreliable.

Planck – We also calculate the Y parameter from the six Planck High Frequency

maps (Planck Collaboration et al. 2016a) using a template fitting program similar to the method described in Section 2.3 of Bourdin et al. (2017). The maps are high-pass filtered to remove large scale signals from the cosmic infrared background, SZ background, and instrumental offsets. On cluster scales, we subtract a spatially and spectrally variable model of the CMB and galactic thermal dust anisotropies.

An Arnaud et al. (2010) pressure profile template is fit to the residual flux within $5r_{500,\mathrm{WL}}$ using χ^2 minimisation, from which we calculate the cylindrical signal within $r_{500,\mathrm{WL}}$. While we use the known cluster positions, the Planck team identify clusters as peaks in the signal map with a signal to noise above 4, and as such identify 38 of the 41 clusters in our sample. For this overlap, our flux measurements agree within 10 per cent with those measured by the Matched Multi-Filter 3 (MMF3) algorithm (Planck Collaboration et al. 2016b). We attribute this difference to the possible offsets of 1-2 arcmin in the cluster positions in the MMF3 analysis.

6.2.4 Optical and Infrared Observables

We also use optical and near-infrared observations of the member galaxies, calculating the K-band luminosity of the BCG, the total cluster K-band luminosity, and the optical richness.

Near-Infrared Luminosity – To investigate the stellar content of the clusters, we use near-infrared (NIR) data (Haines et al. 2009), where 38 clusters were observed with WFCAM on UKIRT, and two (Abell0963 and ZwCl0857) with NEWFIRM on the Mayall 4-m telescope at Kitt Peak National Observatory. We lack NIR data for Abell2697. From these data we calculate both the K-band luminosity of the BCG, $L_{K,\mathrm{BCG}}$, and the total K-band luminosity of the cluster members, $L_{K,\mathrm{tot}}$. The data is analyzed similar to Mulroy et al. (2014).

We convert from apparent K-band magnitude to rest-frame luminosity, using a

k-correction consistent with Mannucci et al. (2001), and the absolute K-band Vega magnitude of the sun, $M_{K,\odot} = 3.39$. For the total luminosity, we select cluster members as galaxies lying along a ridge line in $(J-K)/K$ space. We select those within $r_{500,WL}$ of the cluster centre down to a magnitude of $K \leq K^*(\mathrm{z}) + 2.5$, basing $K^*(\mathrm{z})$ on Lin et al. (2006) and choosing this limit because $2 < K - K^* < 2.5$ is the faintest 0.5mag width bin for which the average K-band magnitude error is < 0.1 for all clusters. To account for the background we perform this same calculation on a control field (The UKIDSS-DXS Lockman Hole and XMM-LSS fields, Lawrence et al. 2007) within 40 apertures of radius $r_{500,\mathrm{WL}}$, subtracting the average from $L_{K,\mathrm{tot}}$ and adding the standard deviation the measurement error. The other component of the measurement error is calculated by propagating the error on the weak-lensing radius. Note that the uncertainties in Mulroy et al. (2014) included a term calculated using bootstrap resampling of the members that we do not include here, because we are interested in the individual cluster measurement error and not the statistical properties of an ensemble of similar clusters.

We note that the consistency found in Mulroy et al. (2014) between color-magnitude selected luminosity and spectroscopically confirmed luminosity indicates the accuracy of color-magnitude member selection in $(J-K)/K$ space, due to the sensitivity of near-infrared data to old stars and its relative insensitivity to recent star formation.

Richness – We calculate the richness, λ, defined in Rozo et al. 2009 and improved in Rykoff et al. 2012, for the 33 cluster overlap between our sample and the SDSS sample (Gunn et al. 1998; Doi et al. 2010; Alam et al. 2015). This matched filter richness estimator is defined as the sum of the membership probabilities of all the galaxies, and was constructed as a low scatter optical mass proxy through extensive tests on the maxBCG cluster catalog (Koester et al. 2007).

For all potential cluster members, their membership probability is calculated considering their clustercentric radius, g-r color and i-band magnitude. The richness estimator is the sum of these probabilities integrated down to $M^* + 1.75$, while the measurement error is derived from the variance of the sum. The corresponding radius is not equivalent to an overdensity radius such as r_{500}, but rather scales deterministically as $\lambda^{0.2}$. The mean value for our sample is 1.4Mpc. While the scale misalignment with respect to the other measures may add some additional variance, we retain the algorithm's choice so as to preserve consistency with other redMaPPer applications (Rykoff et al. 2012, 2016). From a purely statistical point of view, λ is simply another label tagged to each cluster. We leave it to future work to identify physically meaningful, minimum variance estimators of these labels.

6.3 Linear Regression

Scaling relations between observable properties and mass are characteristically power-law in form. We linearize the problem by using the natural log of the values and perform a Bayesian analysis to infer scaling parameters. To do so correctly we have to take into account measurement errors, the halo mass function and the selection criteria. Most commonly used regression methods (e.g., BCES, Akritas & Bershady 1996, and FITEXY, Press et al. 1992; Tremaine et al. 2002) can handle measurement errors, while methods from Kelly (2007) and Mantz (2016) also take into account the independent variable distribution by modeling it as a Gaussian mixture model inferred from the data.

However the selection function can still introduce significant biases, either directly when the selection variable is considered directly in the regression, or indirectly due to covariance between this selection variable and the observables of interest. It is

Table 6.1: Elements of galaxy cluster observable vector

Element, S_i	Unit	Description
$L_{X,\mathrm{RASS}}\ E(z)^{-1}$	$10^{44}\,\mathrm{erg/s}$	Selection variable: RASS, soft-band X-ray luminosity
$L_{X,\mathrm{ce}}\ E(z)^{-1}$	$10^{44}\,\mathrm{erg/s}$	Core-excised, bolometric X-ray luminosity
$T_{X,\mathrm{ce}}$	keV	Core-excised ICM temperature
$M_{\mathrm{gas}}\ E(z)$	$10^{14}\,M_\odot$	ICM gas mass within WL R_{500}
$Y_X\ E(z)$	$10^{14}\,M_\odot$keV	ICM (spherical) X-ray thermal energy within WL R_{500}
$Y_{\mathrm{SZA}}\ E(z)$	$10^{-5}\,\mathrm{Mpc}^2$	ICM (spherical) SZ thermal energy
$Y_{\mathrm{Pl}}\ E(z)$	$10^{-5}\,\mathrm{Mpc}^2$	ICM (cylindrical) SZ thermal energy
$L_{K,\mathrm{BCG}}\ E(z)$	$10^{11}\,L_\odot$	BCG K-band luminosity
$L_{K,\mathrm{tot}}\ E(z)$	$10^{12}\,L_\odot$	Total K-band luminosity within WL R_{500}
$\lambda\ E(z)$	none	redMaPPer richness (count of galaxies)
$M_{WL}\ E(z)$	$10^{14}\,M_\odot$	weak-lensing mass

possible, in principle, to use the methods of Kelly (2007) and Mantz (2016) to correct for selection effects when the selection variable is on the dependent axis, by using upper limits and generating 'censored' or missing data below the selection limit in an iterative process (Gelman et al. 2014). We noticed that this feature of the Kelly (2007) linear regression code by applying it to our $L_{X,\mathrm{RASS}}$ scaling relation. However it is not so straight forward to correct for the bias caused by covariance with the selection variable, i.e. when considering a dependent variable which is not the selection variable, and this approach can be computationally challenging for a larger dataset.

We therefore develop a hierarchical Bayesian model similar to the methods of Kelly (2007) and Mantz (2016), which simultaneously considers the selection variable alongside all other observables in order to explicitly model property covariance, i.e. the intrinsics covariance between two observables at fixed halo mass, and correctly propagate selection effects.

6.3.1 Hierarchical Bayesian Model

We define log-space variables, $\mu \equiv \ln(M)$, where M is the total halo mass, and $\mathbf{s} \equiv \ln(\mathbf{S})$, where $\mathbf{S}$ is the vector of observables given in Table 6.1. In practice we

normalize mass using the median weak-lensing mass of the LoCuSS sample. At a fixed redshift, the joint probability that there exists a cluster with given observables and mass can be written as the product

$$P(\mathbf{s}, \mu \,|\, \boldsymbol{\theta}, \psi) = P(\mathbf{s} \,|\, \mu, \boldsymbol{\theta}) P(\mu \,|\, \boldsymbol{\psi}) \,, \tag{6.1}$$

where $\boldsymbol{\theta}$ is the set of parameters that characterize the scaling relation of observable properties with mass, and $\boldsymbol{\psi}$ characterizes the distribution of the independent variable, in this case the cosmological mass function of halos. For the analysis presented here, we simplify the latter term by assuming a fixed cosmology and use the second-order mass function model of Evrard et al. (2014) at redshift 0.22. Since the mass function shape has only a modest effect on the posterior scaling parameter constraints, we do not attempt to marginalize over cosmology and so drop $\boldsymbol{\psi}$ from the equations below.

We note that the mass discussed above is the true unobserved halo mass which we marginalize over. The small sample size and limited set of observables force us to make the simplifying assumption that weak-lensing mass is an unbiased measure of true halo mass, albeit with non-zero scatter of $\sim 20\%$ (e.g., Oguri & Hamana 2011; Becker & Kravtsov 2011; Bahé et al. 2012). We retain weak-lensing mass, $M_{\rm WL}$, in the vector of observables $\mathbf{s}$ and treat it in a special way to avoid parameter severe degeneracies of the type discussed in Penna-Lima et al. (2017).

We model $P(\mathbf{s} \,|\, \mu, \boldsymbol{\theta})$, the first term in the joint probability distribution in Equation (6.1), as a log-normal distribution,

$$P(\mathbf{s} \,|\, \mu, \boldsymbol{\theta}) \propto \exp\left\{ -\frac{1}{2} (\mathbf{s} - \boldsymbol{\alpha}\mu - \boldsymbol{\pi})^T \Sigma^{-1} (\mathbf{s} - \boldsymbol{\alpha}\mu - \boldsymbol{\pi}) \right\} \,, \tag{6.2}$$

where the model parameters, $\boldsymbol{\theta} = \{\boldsymbol{\pi}, \boldsymbol{\alpha}, \Sigma\}$, include the intercepts, $\boldsymbol{\pi}$, and slopes, $\boldsymbol{\alpha}$, of the log-mean behavior, as well as the property covariance matrix, Σ, of Gaussian

deviations about the log-mean. Each diagonal element of the covariance matrix specifies the variance of a property while the off-diagonal elements are the property covariance, all at fixed true halo mass. Except for the parameters connected to weak-lensing mass, which are fixed as explained below, the remainder are unknown parameters to be constrained. Parameter priors are uninformative, as specified in Table 6.2.

We impose a delta function prior on the scaling of $M_{\rm WL}$ that assumes unit slope and intercept with true mass, and a fixed log-normal scatter of 0.2. We tested values for the scatter of 0.1 and 0.3, finding that our results and inferred parameters are insensitive to this choice. We assume zero *intrinsic* correlation between weak-lensing mass and all other observable properties, a, at fixed true halo mass: $r_{M_{\rm WL},a} = 0$. We include the correlation of its measurement uncertainty with the other observables defined within the weak-lensing radius (so-called 'aperture bias').

In practice we do not measure the true values of $\mathbf{s}$; our measurements, $\mathbf{s}_o$, include observational uncertainties. We again assume a log-normal form for the measurement errors,

$$P(\mathbf{s}_o|\mathbf{s}) \propto \exp\left\{-\frac{1}{2}(\mathbf{s}_o - \mathbf{s})^T \Sigma_{\rm err}^{-1} (\mathbf{s}_o - \mathbf{s})\right\} , \tag{6.3}$$

where $\Sigma_{\rm err}$ is the measurement error covariance. This matrix includes both diagonal elements, given by the square of the fractional errors in each cluster's measured properties, and off-diagonal "aperture bias" terms for $M_{\rm gas}$, $L_{K,\rm tot}$ and $Y_{\rm SZA}$ properties measured within the characteristic radius inferred from weak-lensing mass. While most other observables are measured within the weak-lensing determined radius, the measurements are largely unaffected by small radial changes and so do not require these off-diagonal terms.

The probability of measuring the observable properties, $\mathbf{s}_{o,i}$, of a specific cluster,

i, is found by marginalizing over the true quantities, $\mathbf{s}$, resulting in

$$P(\mathbf{s}_{o,i}|\mu_i, \boldsymbol{\theta}) \propto \exp\left\{-\frac{1}{2}(\mathbf{s}_{o,i} - \langle \mathbf{s}_o \rangle_i)^T \Sigma^{-1}_{\mathrm{tot},i}(\mathbf{s}_{o,i} - \langle \mathbf{s}_o \rangle_i)\right\}, \tag{6.4}$$

where $\langle \mathbf{s}_o \rangle_i = \boldsymbol{\alpha}\mu_i + \boldsymbol{\pi}$, with μ_i the lensing mass estimate of the i^{th} cluster, and $\Sigma_{\mathrm{tot},i} = \Sigma + \Sigma_{\mathrm{err},i}$. We make a similar log-normal assumption about the weak-lensing mass measurements, $\mu_{o,i}$, and include the measurement error and its aperture-driven covariance with other measured property uncertainties in the regression analysis.

Our method is able to handle missing data, meaning systems for which not all elements of the data vector are available. We marginalize over these missing quantities by setting the missing values at the median of that observable quantity and assuming a large error, 999 in the natural-log, on the missing value.

Finally, and most significantly, we are able to account for the effect of selection, as the vector of observables includes the selection property (Kelly 2007; Gelman et al. 2014). Our selection function is simply a redshift dependent $L_{X,\mathrm{RASS}}$ threshold (see Figure 6.1), which is taken into account using a redshift dependent step function. The likelihood of the model parameters is based on the selection-normalized properties,

$$\mathcal{L}(\mathbf{s}_{o,i} \,|\, \boldsymbol{\theta}) = \prod_{i \in \mathcal{C}} \int \mathrm{d}\mu_i \; \Phi_i^{-1}(\mu_i, \boldsymbol{\theta}) \; P(\mathbf{s}_{o,i}, \mu_i \,|\, \boldsymbol{\theta}), \tag{6.5}$$

where $\mathcal{C}$ is the cluster sample and $\Phi_i(\mu_i, \boldsymbol{\theta})$ is a normalization factor due to the selection function for cluster i given the set of model parameters, $\boldsymbol{\theta}$, and its estimated weak-lensing mass. We interchange the order of integration and multiplication operator to get,

$$\mathcal{L}(\mathbf{s}_{o,i} \,|\, \boldsymbol{\theta}) \propto \int \mathrm{d}\mu_i \; \prod_{i \in \mathcal{C}} \Phi_i^{-1}(\mu_i, \boldsymbol{\theta}) \; P(\mathbf{s}_{o,i}, \mu_i \,|\, \boldsymbol{\theta}), \tag{6.6}$$

Letting $y \equiv \ln L_{X,\mathrm{RASS}}$, full distribution for $L_{X,\mathrm{RASS}}$, and denoting the z-dependent threshold luminosity as $y_t(z)$, the normalization factor in Equation (6.5) and Equa-

tion (6.6) becomes

$$\Phi_i(\mu_i, \boldsymbol{\theta}) = \int \mathrm{d}y \; \Theta(y - y_t(z_i)) \; P(y, \mu_i \,|\, \boldsymbol{\theta}) \,, \tag{6.7}$$

where $\Theta(z)$ is the Heaviside function (see Figure 6.1). For each cluster and each iteration of the Markov Chain Monte Carlo (MCMC) analysis, the likelihood is renormalized according to Equation (6.7).

For practical reasons, we do not perform the integrations over mass explicitly in Equation (6.6). Instead, we consider the set of 41 lensing masses as additional model degrees of freedom and perform the MCMC algorithm to derive the broader posterior distribution,

$$P(\boldsymbol{\theta}, \mu_i \,|\, \mathbf{s}_{o,i}) \propto \left[\prod_{i \in \mathcal{C}} \; \Phi_i^{-1}(\mu_i, \boldsymbol{\theta}) \; P(\mathbf{s}_{o,i}, \mu_i \,|\, \boldsymbol{\theta}) \right] P(\boldsymbol{\theta}) \tag{6.8}$$

where $P(\boldsymbol{\theta})$ is the prior distribution specified in Table 6.2. We then determine the model parameter constraints, $P(\boldsymbol{\theta} \,|\, \mathbf{s}_{o,i})$, by marginalizing over the posterior distributions of the 41 cluster masses.

The MCMC algorithm is based on the `PyMC` library (Patil et al. 2010) and proceeds as follows. For each iteration, a mass is assigned to each cluster drawn randomly from the halo mass function, i.e. the prior distribution. Then a new set of model parameters, $\boldsymbol{\theta}$, are drawn randomly from the prior distribution specified in Table 6.2. With the assigned cluster masses and chosen set of parameters, the selection function is evaluated and the likelihood evaluated. The initial seeds are adapted in a way to minimize the number of steps needed to reach the equilibrium. We choose the central value of the weak-lensing masses as the initial seed for each cluster mass, μ_i, and the scaling parameters are initialized with the Kelly (2007) regression estimate. This choice of initial seeds allows us to reach equilibrium faster and does not have an

Table 6.2: Prior distributions of scaling parameters for any property, a, other than weak lensing mass. The same priors are used for all properties and pairwise combinations, a,b.

Variable	Description	Prior
π_a	Intercept	$\mathcal{N}(0, 100)$
α_a	Slope	$\mathcal{N}(0, 100)$
$\sigma_{a\,\vert\,\mu}$	Scatter (natural log)	$\mathcal{U}(0, 5)$
$r_{a,b\,\vert\,\mu}$	Correlation coefficient	$\mathcal{U}(-1, 1)$

effect on the posterior distribution. The performance of this method is demonstrated and compared with other methods in Section 6.3.2.

6.3.2 Performance of the Hierarchical Bayesian Method

We test the performance of the Hierarchical Bayesian method on 1,000 mock datasets. To generate mass values for the independent axis we assume a Tinker mass function as a function of redshift (Tinker et al. 2008), and use the `hmf` code (Murray et al. 2013). The process of generating this vector is as follows:

1. Generate catalogue of $X \equiv M_{\rm True}$ and z using `hmf` code (Murray et al. 2013).
2. Generate $M_{\rm WL}$ assuming $\sigma_{\ln M_{\rm WL}|M_{\rm True}} = 0.1$.
3. Generate Y, an observable selection variable, values assuming a Y-X scaling relation.
4. Generate Z, an observable, values assuming a Z-X scaling relation and an intrinsic correlation coefficient -0.7.
5. Apply a correlated measurement errors of 0.1 with correlation coefficient 0.7 to Y and Z values about fixed halo mass.
6. Select systems those above a Y limit.

After applying the Y selection, each dataset contains ~ 50 objects, similar to our LoCuSS sample. We calculate the best fit parameters for each dataset, and show the

distribution of these parameters in Figure 6.2. We find that the estimated posteriors for these parameters are consistent with the input values. We compare the best fit parameters calculated using different methods:

- H-Bayes: the Hierarchical Bayesian Model presented in Section 6.3.1
- H-Bayes (no r): the same model, without modeling the error covariance r
- Kelly: the method of Kelly (2007), without modeling the selection
- OLS: Ordinary Least Squares

As expected, the methods that do not consider the selection function (Kelly (2007) and OLS) perform poorly and estimate a shallower slope (and higher intercept) for the selection variable Y and a steeper slope (and lower intercept) for Z due to its negative covariance with Y. This leads the method of Kelly (2007) to underestimate the intrinsic scatter in both relations. We note that while both H-Bayes methods are accurate in the Y relation where modeling error covariance is unimportant, the H-Bayes method that does not model error covariance is less accurate in the Z relation. This also emphasis on the effect of the error covariance on the scatter parameter of a non-selection variable.

6.4 Results

In this section we apply the Hierarchical Bayesian method described in Section 6.3.1 to the LoCuSS data described in Section 6.2. We discuss the resulting scaling relation parameters below, focusing on the individual properties in turn. Constraints on property correlations are presented in Chapter VII.

In order to characterize the scaling relations between cluster observables and mass, we use a fixed pivot mass defined by the sample average, $M_p = 7.41 \times 10^{14} M_\odot$, and

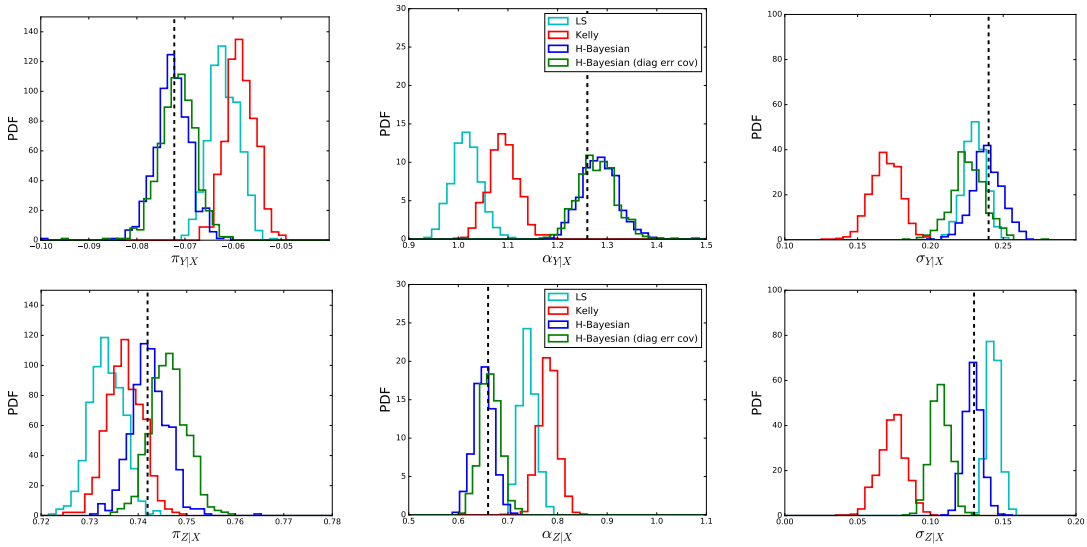

Figure 6.2: Distribution of best fit parameters for 1,000 mock datasets, constrained by four different methods: H-Bayes - the Hierarchical Bayesian Model presented in Section 6.3.1 (blue); H-Bayes (no r) - the same model, without modeling the error covariance (green); Kelly - the method of Kelly (2007), without selection function (red); OLS - Ordinary Least Squares (cyan).

fit the log-mean behavior of property a to the form,

$$\langle s_a \rangle = \alpha_a(\mu + e(z)) + \pi_a \tag{6.9}$$

where $\mu = \ln(M/M_p)$, $e(z) = \ln E(z)$ and the normalization is the natural log using units given in Table 6.1. The mass, M, is the weak lensing mass which is assumed to be an unbiased estimator of true mass, as discussed above. Since our method constrains the covariance between observables at a fixed mass, we use the same independent variable, $\mu + e(z)$, for all properties. Where this is not the natural independent variable derived in Chapter II (i.e. for $M_{\rm gas}$, L_K and λ) we include an additional factor of $e(z)$ to the dependent axis, as listed in Table 6.1.

As a check, we also perform the fits with μ as the independent variable and appropriately modified ϵ factors on the dependent axes. As expected within such a narrow redshift range, the results are consistent.

Table 6.3: Scaling relation parameters

Observable	Intercept $\exp(\pi_a)$	Slope α_a	Scatter $\sigma_{a\|\mu}$	Self-Similar Slope
$L_{X,\mathrm{RASS}}$	$4.17^{+1.63}_{-1.63}$	$1.23^{+0.41}_{-0.41}$	$0.57^{+0.15}_{-0.15}$	1.00
$L_{X,\mathrm{ce}}$	$7.89^{+0.94}_{-0.94}$	$0.98^{+0.20}_{-0.20}$	$0.37^{+0.05}_{-0.05}$	1.33
$T_{X,\mathrm{ce}}$	$6.97^{+0.54}_{-0.54}$	$0.49^{+0.12}_{-0.12}$	$0.20^{+0.04}_{-0.03}$	0.66
M_{gas}	$0.97^{+0.05}_{-0.05}$	$0.78^{+0.10}_{-0.10}$	$0.15^{+0.04}_{-0.04}$	1.00
Y_X	$6.08^{+0.80}_{-0.80}$	$1.27^{+0.20}_{-0.20}$	$0.33^{+0.06}_{-0.06}$	1.66
Y_{SZA}	$7.75^{+1.20}_{-1.20}$	$1.57^{+0.23}_{-0.23}$	$0.30^{+0.09}_{-0.09}$	1.66
Y_{Pl}	$11.04^{+1.11}_{-1.11}$	$1.15^{+0.16}_{-0.16}$	$0.28^{+0.04}_{-0.04}$	1.66
$L_{K,\mathrm{BCG}}$	$0.97^{+0.10}_{-0.10}$	$0.21^{+0.16}_{-0.16}$	$0.34^{+0.04}_{-0.04}$	–
$L_{K,\mathrm{tot}}$	$16.82^{+0.88}_{-0.88}$	$0.75^{+0.11}_{-0.11}$	$< 0.19^{*}$	1.00
λ	$125.27^{+10.44}_{-10.44}$	$0.73^{+0.14}_{-0.14}$	$0.24^{+0.05}_{-0.05}$	1.00

* The $L_{K,\mathrm{tot}}$ scatter is not bounded from below (see Figure 6.5), so the value quoted is a 95^{th} percentile upper limit.

6.4.1 Scaling Relations Parameters

The resulting posterior estimates of the scaling relation parameters are summarized in Table 6.3, shown in Figure 6.3, and discussed below. In ensuing subsections, we begin by presenting results for the selection variable, $L_{X,\mathrm{RASS}}$, then proceed to examine hot gas and stellar scaling behaviors. Subsequent sections discuss intrinsic property variance and the physical origins of deviations about the mean relations.

Selection variable – The posterior parameter constraints on the scaling of $L_{X,\mathrm{RASS}}$ with mass, listed in the first row of Table 6.3, entail large uncertainties that are driven by significant sample incompleteness as a function of mass. The upper left panel of Figure 6.3 shows that all but two of the 41 clusters lie above the best-fit underlying scaling relation; the selection skims off only the brightest systems as a function of mass. This behavior is a textbook example of Malmquist bias (Allen et al. 2011; Mantz et al. 2016a; Giles et al. 2017).

While the inferred slope of 1.23 ± 0.41 agrees with the self-similar expectation, the 30% uncertainty in slope dilutes the impact of this statement. The intrinsic scatter

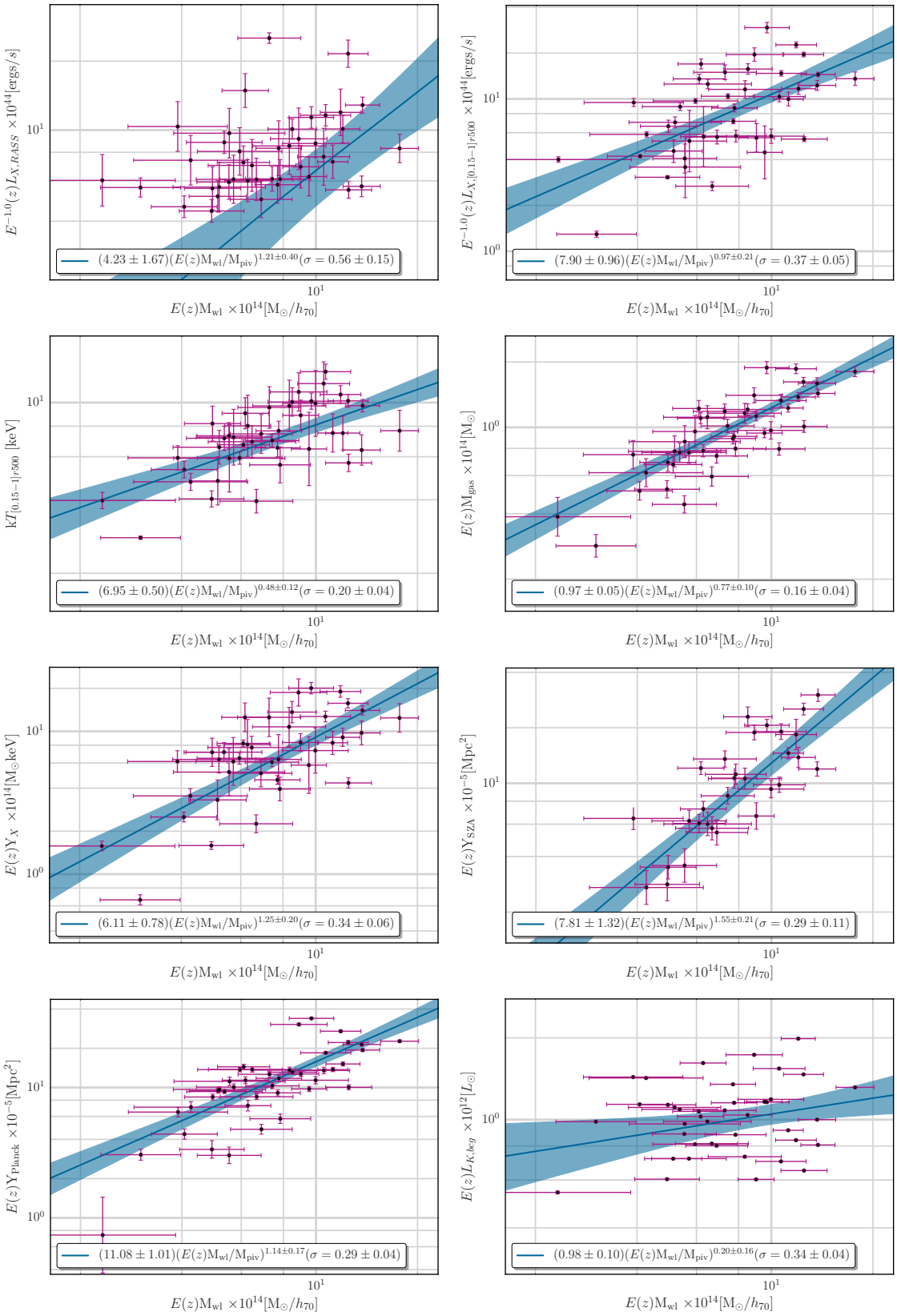

Figure 6.3: Scaling relations between cluster observable properties and potential well depth, $E(z)M_{WL}$. Individual cluster points with error bars are shown while the Hierarchical Bayesian fits and 68 per cent confidence regions of the mean behaviors are given by solid lines and grey-scales, respectively.

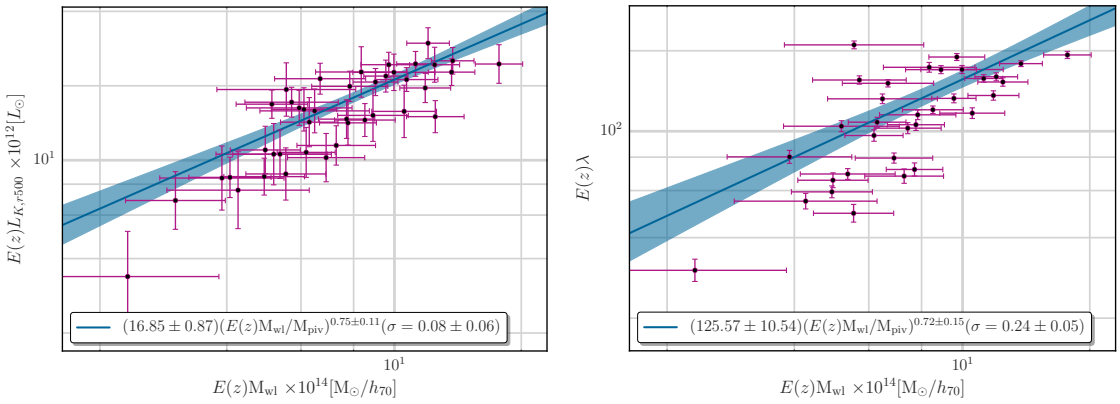

Figure 6.4: Same as Figure 6.3.

(in natural log) of $0.57^{+0.15}_{-0.15}$ is much higher than the $0.37^{+0.05}_{-0.05}$ seen for the core-excised counterpart, $L_{X,\mathrm{ce}}$, which we interpret as the consequence of including the core. We have also performed analysis using *Chandra*/*XMM-Newton* luminosities that include the core, finding an intrinsic scatter of $0.51^{+0.08}_{-0.08}$, consistent with the $L_{X,\mathrm{RASS}}$ value.

The relatively large uncertainty in $L_{X,\mathrm{RASS}}$ scaling parameters allows only weak estimates of the correlation coefficients between $L_{X,\mathrm{RASS}}$ luminosity and other cluster properties. The largest coefficients, with values between 0.4 and 0.5 and uncertainties of roughly 0.2, are with follow-up X-ray measures and Y_{SZA} . The full set of coefficients includes hint of an anti-correlation between hot gas mass and stellar mass discussed further Chapter VII.

X-ray Observables – For the X-ray properties (rows 2 through 5 of Table 6.3), posterior constraints on the slopes of the scaling relations are consistently shallower than self-similar model expectations at the $\sim$ 1 - 2 σ level, with uncertainties ranging from 0.1 (M_{gas} and $T_{X,\mathrm{ce}}$) to 0.2 ($L_{X,\mathrm{ce}}$ and Y_X). The shallow behavior for M_{gas} is unexpected, as previous studies covering a wider dynamic range in cluster mass have found that mean gas mass increases with halo mass in a super-linear fashion, M_{gas} $\propto M^{1.2}$ (e.g., Pratt et al. 2009). However, as discussed below, the slope we find is only

in 2σ tension with the Weighing the Giants study of Mantz et al. (2016b), who find a slope of 1.004 ± 0.014 for a high-mass sample of clusters. A trend toward self-similar behavior in the highest halo masses is seen in Chapter III, a set of hydrodynamical simulations that include AGN heating.

We highlight that there is a degeneracy between the posterior slope of a property and the covariance between that property and the selection variable, $L_{X,\mathrm{RASS}}$. Physically, we expect a positive correlation between M_{gas} and $L_{X,\mathrm{RASS}}$ residuals, but we find the correlation coefficient to be only $0.19^{+0.25}_{-0.24}$. If this value were higher, the slope of the M_{gas} relation would also increase.

SZ Observables – We find that the slopes of the two SZ-Y relations are consistent with each other, with Y_{SZA} being steeper than Y_{Pl} at the level of 1.5σ. They are also within 2σ of the self-similar slope of 5/3, and the two SZ values bracket the Y_X slope of 1.27 ± 0.20.

Regarding normalization, the cylindrical measurement of Y_{Pl} can be converted to a spherical estimate by dividing by factor, $Y_{\mathrm{cyl}}/Y_{\mathrm{sph}} = 1.2$ (Arnaud et al. 2010). The resulting value of 9.3 ± 0.8 compares well with the SZA spherical intercept of 7.9 ± 1.0.

To compare to the X-ray normalization, we follow Arnaud et al. (2010) and normalize Y_X by

$$C_{\mathrm{XSZ}} = \frac{\sigma_T}{m_e c^2} \frac{1}{\mu_e m_p} = 1.416 \times 10^{-19} \frac{\mathrm{Mpc}^2}{\mathrm{M}_\odot \mathrm{keV}}, \tag{6.10}$$

giving a Y_X intercept of 8.8 ± 0.8. To summarize, we find good agreement between the normalizations of all three relations that measure the electron thermal energy content.

We note that while the Y_{SZA} slope is in agreement with the self-similar relation, the Y_{Pl} is shallower. Assuming self-similarity and the high resolution measurement

$Y_{\rm SZA}$ reflect the true relation, we interpret this as a consequence of the low resolution of the Planck maps. The angular resolution of the Planck HFI channels is $\sim$ 5 - 10 arcmin, comparable to the angular size of intermediate and low mass clusters at the redshift studied here. The range of weak-lensing r_{500} values for our sample is 3.7 - 9.1 arcmin.

The $Y_{\rm Pl}$ measurement errors for the low mass clusters are large, so they do not have a strong influence on the fit. The fit parameters are largely constrained by the intermediate and high mass clusters, and an increase in the $Y_{\rm Pl}$ measurement of intermediate mass clusters would act to shallow the fitted slope. Indeed we find the highest ratios of $Y_{\rm Pl}$ to $Y_{\rm SZA}$ in low and intermediate mass clusters. As the $Y_{\rm Pl}$ values are calculated by fitting an Arnaud et al. (2010) profile to the signal at a cluster radius limited by the HFI resolution, this is consistent with the unresolved clusters being sensitive to any elevated signal in the outskirts, for example from infalling substructure.

Stellar Observables – The measures of galactic stellar content, $L_{K,\rm BCG}$, $L_{K,\rm tot}$ and λ, provide complementary insights into the star formation history of high mass halos. Both $L_{K,\rm tot}$ and λ attempt to measure the total stellar content of a cluster, but they differ in detail. The K-band total luminosity, $L_{K,\rm tot}$, is a background-corrected estimate that uses all member galaxies within the weak-lensing estimate of r_{500}, whereas λ is a red-sequence weighted estimate determined within an aperture scaling as $\lambda^{0.2}$. The former is luminosity weighted while the latter is number-weighted. We highlight that any interpretation of the stellar content derived from these galaxy observable scaling relations relies on the assumption that they are reliable tracers of the stellar mass. This is likely sensitive to the details of the measurement, and determining the best stellar mass estimate would require further study.

Despite their differences, the slopes of the $L_{K,\mathrm{tot}}$ and λ scaling relations are consistent, and in both cases shallower than the self-similar prediction. As both measures scale with total stellar mass, they point to a stellar fraction that decreases with increasing halo mass, implying that star-forming efficiency is a decreasing function of halo mass (Gonzalez et al. 2007; Laganá et al. 2011). This result is supported by abundance matching arguments (Behroozi et al. 2013b; Kravtsov 2013) and AGN-based feedback scenarios in cosmological hydrodynamics models are tuned to produce this feature (Croton et al. 2006; De Lucia & Blaizot 2007; Planelles et al. 2013; Pillepich et al. 2018, also the work presented in Chapter III) Both weak lensing (Simet et al. 2017) and ensemble spectroscopic (Chapter IV and Chapter V of this dissertation) mass estimate methods find mean mass scaling behavior, $M \propto \lambda^{1.3}$, consistent with the our findings.

The close agreement in the $L_{K,\mathrm{tot}}$ and λ slopes values may be somewhat fortuitous. The radius within which λ is measured scales more slowly ($\lambda^{0.20}$) than the halo radius implied from the scaling of weak-lensing mass ($\lambda^{0.46}$), within which $L_{K,\mathrm{tot}}$ is measured. While this could potentially lead to proportionally larger increases in λ compared to $L_{K,\mathrm{tot}}$ as halo mass decreases, a secondary factor such as a declining red galaxy fraction in lower mass halos may compensate for the scale mismatch effect.

The $L_{K,\mathrm{BCG}}$ scaling relation is very shallow, almost consistent with zero, demonstrating that the luminosity of the BCG is not a strong function of mass for clusters in this mass range. As halo mass increases, so does the galaxy velocity dispersion, and accretion onto the BCG slows relative to the total mass growth of the cluster. As these two processes are largely uncoupled it leads to large scatter in the relation, consistent with our finding that the $L_{K,\mathrm{BCG}}$ relation has a larger intrinsic scatter than the $L_{K,\mathrm{tot}}$ relation.

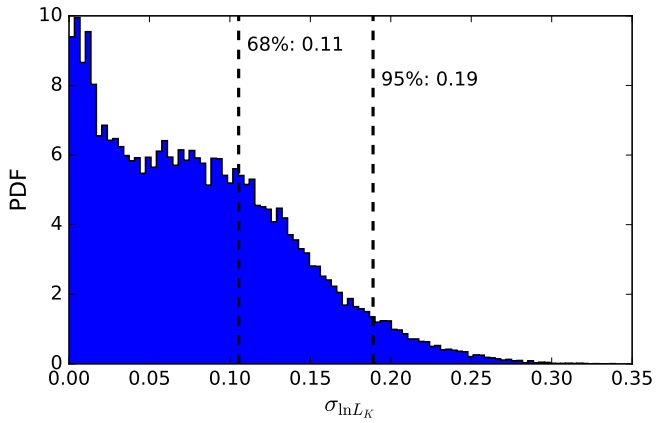

Figure 6.5: Posterior PDF of the scatter in total K-band luminosity, $\sigma_{\ln L_K}$, with the 68$^{\rm th}$ and 95$^{\rm th}$ percentile upper limits indicated.

The normalizations of the BCG and total $L_{K,\rm tot}$ relations provide a simple estimate of the fraction of stellar mass associated with the BCG. We find a value of $5.8 \pm 0.5\%$, with the uncertainty dominated by the BCG normalization error.

6.4.2 Intrinsic Variance

Knowledge of the intrinsic variance in cluster properties is important for precise cosmological studies with the population, but empirical estimates of the full covariance matrix, including both on-diagonal scatter and off-diagonal pair correlations have only recently begun to emerge (Maughan 2014; Mantz et al. 2016b).

Caution is required when estimating the covariance of sample properties, as the statistical (measurement) errors must be accurately determined and the selection model must be correctly described. Considerable interest lies in the intrinsic scatter of an individual property, σ_a, and its related scatter in halo mass, σ_a/α_a, where α_a is the slope with mass of that property.

Regarding selection, the effect of including selection has a significant affect on the posterior intrinsic scatter estimates. The "naive" regression model (see Section 6.3.2) produces scatter estimates that differ significantly from Table 6.3 for several

X-ray properties, including the $L_{X,\mathrm{RASS}}$ selection variable. Note, however that the intrinsic scatter constraints on M_{gas} and $T_{X,\mathrm{ce}}$, as well as all of the SZ and optical properties are consistent between the two treatments.

Since the model that includes selection effects should be closer to unbiased, we employ the values in Table 6.3 as our primary results, with a note of caution that posterior scatter constraints for $L_{X,\mathrm{ce}}$ and Y_X appear to be most sensitive to the selection model.

Reviewing the intrinsic scatter values, we note that M_{gas} and $L_{K,\mathrm{tot}}$ have the lowest values, while the $L_{X,\mathrm{RASS}}$ selection variable is highest. The posterior in $L_{K,\mathrm{tot}}$ scatter has no finite lower bound. As shown in Figure 6.5, the PDF of the intrinsic scatter in the $L_{K,\mathrm{tot}}$ relation is not well fit by a Gaussian, so we quote 68^{th} and 95^{th} percentile upper limit of 0.11 and 0.19 respectively. The 95^{th} percentile upper limit is slightly below with the intrinsic scatter in the λ relation. We note that the definition of membership for the two observables is different and therefore recalculate $L_{K,\mathrm{tot}}$ using membership as determined in the λ calculation, finding the result unchanged. We interpret this as an indication that $L_{K,\mathrm{tot}}$, as a tracer of the stellar mass, is a slightly better proxy for cluster mass than the richness.

From Table 6.3 we can estimate the mass proxy power using the inferred scatter in mass $\sigma_{\mu|a} = \sigma_{a|\mu}/\alpha_a$. BCG K-band luminosity is by far the least effective, with a wide scatter of 1.6 in logarithmic mass. Total K-band light, on the other hand, is much more tightly correlated, with an upper limit of $\sim 30\%$. Gas mass provides 0.20 ± 0.05 fractional accuracy in mass, as do all measures of Y. We stress that these estimates are with respect to the lensing mass values, and the inference with respect to true mass is dependent on our simplifying assumptions discussed in Section 6.3. Larger homogeneous samples of the type used here are needed to provide more

accurate estimates of the intrinsic property covariance.

6.4.3 Posterior Distribution on True Halo Mass

Our model fits for the cluster halo mass, and so generates a posterior distribution for the true mass of each cluster. We display them next to our weak-lensing mass estimates in Figure 6.6. Any differences are due to a combination of two effects – the mass function favoring low mass systems, and the scaling relations favoring systems that lie near the expectation value. The latter effect can be understood with looking at the residuals. Clusters with negative residuals from the scaling relations tend to have posterior masses smaller than their weak-lensing masses (e.g., Abell0907 and Abell0291), while those with positive residuals have the opposite (e.g., Abell2219 and Abell0781).

6.5 Discussion

6.5.1 Scaling Relations in the Literature

As we have discussed, to obtain robust scaling relations requires an unbiased measurement of the true mass, an understanding of and correction for the selection of the sample, and a method which allows for the covariance between the selection function and the observable property. Mainly due to the paucity of high signal-to-noise, uniform, multi-wavelength data for well-defined cluster samples, the number of studies in the literature which meet all of these criteria is small. We will largely restrict ourselves to these studies for comparison.

The most similar study to our own is that of Mantz et al. (2016b), who use weak-lensing measurements and gas mass as estimators of the 'true' mass, and who attempt to model the selection of their clusters. For the ICM properties, they also allow for the covariance of those properties with the selection variable. Their sample

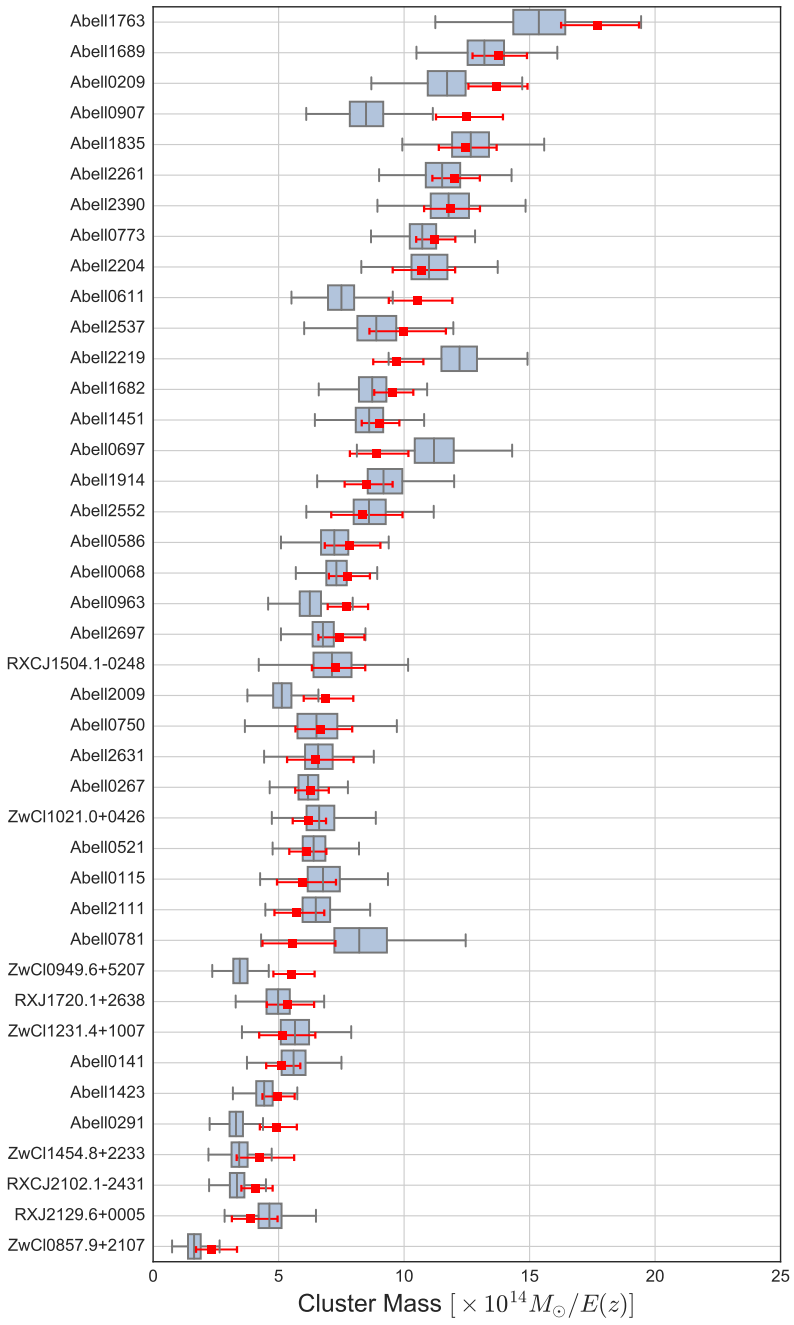

Figure 6.6: The posterior masses on true halo mass from the Hierarchical Bayesian fit in grey, alongside the measured weak-lensing cluster masses in red. The grey box plots show the 25 and 75 percentiles of the posterior masses, with the whiskers showing 99.7 and 0.3 percentiles, while the errors on the yellow points show the 25 and 75 percentiles according to the measurement errors on the weak-lensing measurements.

includes 27 clusters with weak-lensing masses and a larger sample with gas mass measurements, and span a slightly wider redshift range than ours. In mild conflict with our results, Mantz et al. (2016b) report that the core-excised gas temperature and the gas mass agree with the self-similar predictions. They find a $T_{X,\rm ce}$ relation slope of 0.62 ± 0.04, consistent with the self-similar expectation of 0.67 but only 2σ discrepant with our estimate of 0.49 ± 0.12. Their estimate of the $M_{\rm gas}$ relation slope is 1.007 ± 0.012, in agreement with unity and again marginally consistent with our estimate of 0.78 ± 0.10. It is unclear what causes the differences in our results, however given our method, selection and data analysis are all different from Mantz et al. (2016b), a difference of this magnitude is not unexpected.

Similar to our results, Mantz et al. (2016b) also find that the soft-band X-ray luminosity is steeper than the self-similar expectation and suggest that this is due to non-gravitational heating and cooling processes in cluster cores.

Our study is the first to look at the simultaneous scaling of X-ray, SZ and optical properties, and so there are few results to compare to the SZ and optical properties. Mantz et al. (2016b) provide an empirical scaling (without modelling the covariance and correcting for sample selection) and find a shallower $Y_{\rm SZ}$ slope than self-similarity would predict (1.31 ± 0.03). Note that this measurement is using $M_{\rm gas}$ as the mass parameter, but Mantz et al. (2016b) find a one-to-one relation between $M_{\rm gas}$ and $M_{\rm WL}$. This is similar to our $Y_{\rm Pl}$ slope, but quite different from our $Y_{\rm SZA}$ value.

Although not corrected for selection effects, studies have placed constraints on the optical scaling relations of $L_{K,\rm tot}$ (e.g., Lin et al. 2003, 2004; Mulroy et al. 2014, 2017) and λ (e.g., Rykoff et al. 2012; Mantz et al. 2016b; Simet et al. 2017; Melchior et al. 2017), finding the slopes to be shallower than the self-similar predictions, consistent with our results.

Results from recent numerical simulations indicate that AGN heating produces departures from self-similar scaling relations. Several independent groups find that galactic physics with AGN feedback steepens the ICM scaling relations (Planelles et al. 2013; Le Brun et al. 2017; Hahn et al. 2017b; Pillepich et al. 2018), in moderate tension with our X-ray findings. The overall star formation efficiency declines with increasing halo mass in these simulations, producing stellar mass scalings that are sub-linear with M, in agreement with the LoCuSS behavior. One concern when making sample comparisons is the likelihood that the scaling relation slopes may run with halo mass and, to a lesser extent, redshift (see Chapter III of this dissertation). However, the statistical power of this sample is not good enough for constraining these running parameters.

6.6 Chapter Conclusion

The task of constraining scaling relations is complicated by the effects of the selection function and the error covariance. In this chapter we have presented a new multivariate approach to correct for these effects, and applied it to a multi-wavelength observational dataset for which the selection function is well defined. For the first time, we have provided well-constrained scaling relation parameters with mass for a range of galaxy cluster observables, and our main results are as follows:

- We find that the ICM scaling relations are shallower than the self-similar expectations at the 1-2 σ level, which could be affected by inexact selection model.
- The results of integrated optical observables, $L_{K,\mathrm{tot}}$ and λ , are in good agreement, with slopes of ~ 0.7 suggesting that star-forming efficiency is a decreasing function of cluster halo mass.
- We find no distinction between clusters of different dynamical state in the core-

excised X-ray and high resolution SZ relations.

Our results indicate no difference between the scaling relations of core-excised X-ray or high resolution SZ measurements for clusters of different dynamical states, suggesting that selection on these variables would lead to dynamically diverse samples of clusters. Further investigation with samples including lower mass clusters is needed to fully understand any dependence of the cluster stellar fraction on its dynamical state. While our results in this chapter are limited by the modest number of observed clusters, our method will be applicable to future surveys and will lead to excellent constraints on the physics of clusters and the cosmological parameters. This chapter presents diagonal elements of the property covariance, the scatter parameter; and in the next chapter, we discuss the off-diagonal elements of the property covariance.

CHAPTER VII

An Empirical Study of Intrinsic Halo Property Covariance

Philosophy and Contribution

In this chapter, I study the intrinsic property covariance at fixed halo mass and compare the results with the predictions provided in Chapter III of this dissertation. In this chapter, I will close the chain, which starts with performing simulations, continues with making model-dependent predictions, acquiring an empirical dataset to test this prediction, modeling the data, and finally ends with comparing with the predictions. This chapter is a modified version of Farahi et al. (in preparation): "The First Empirical Evidence that Galaxy Clusters are Closed Baryon Systems" which is intended for submission to Nature Astronomy. This chapter heavily relies on the data obtained and the analysis method developed in Chapter VI. The analysis and plots presented in this chapter are my own work. The data is obtained and reduced by the collaborators.

7.1 Specifying the Model and the Notation

The analysis model assumes a multi-variate log-normal probability distribution. In Chapter VI, we develop and employ a hierarchical Bayesian inference model which accounts for the effect of the sample selection alongside the property covariance and measurement error covariance. The analysis model and the performance test are

discussed in Chapter VI of this dissertation. With a slope and intercept for each pair of observables, a and b, we can estimate the covariance matrix. This covariance is essentially defined as,

$$C_{a,b} = \frac{N}{N-1} \sum_{i=1}^{n} \delta s_{a,i} \; \delta s_{b,i}, \tag{7.1}$$

where $\delta s_{a,i} \equiv s_{a,i} - \alpha_a \mu_i - \pi_a$ is the residual deviation from the mean scaling relation and N is the total number of clusters. Finally the property correlation is

$$r_{a,b} = \frac{C_{a,b}}{\sqrt{C_{a,a} \; C_{b,b}}}. \tag{7.2}$$

This correlation coefficient is the quantity of interest that we are after in this letter. We note that our likelihood model constrains the correlation coefficients and the scaling parameters simultaneously.

7.2 Closing the Loop – From Theory to An Observation

Historically, there has been a strong tendency to describe galaxy clusters as closed, virialized objects, meaning that their baryonic content should be conserved and an unbiased estimator of the cosmic mean baryon fraction. The first application of this model dethroned the "standard CDM" model with $\Omega_m = 1$ (White et al. 1993), and suggested $\Omega_m = 0.3$ (Evrard 1997). Subsequently, five years later the discovery of dark energy dominated Universe, i.e. the ΛCDM model, came along with Type Ia supernova (Riess et al. 1998). After two decades of obtaining data, that there now exists precise, multi-wavelength observational data with a well-defined sample selection, capable of testing the predictions of this long-standing model.

The original spherical collapse model postulates that the baryons and dark matter trace one another within the virial boundary of a dark matter halo. In modeling the formation of halos, assuming only the gravitational force and collisional shocks leads

to self-similar solutions (Bertschinger 1985). In the self-similar solution, ideal fluids, both collisionless and collisional fluids, take similar radial profiles when expressed in units of the turnaround radius (Gunn & Gott 1972). Another implication of this model is that this multi-phase, ideal fluid shows no radial separation, and the content of radial shells, which are the building blocks of a cluster, reflects the comic mean baryon fraction.

Within this formulation, massive halos are closed systems, retaining all their gaseous and stellar matter. An implication of this closed box model is that there would be zero scatter in total baryon mass conditioned on halo mass and redshift. Most importantly, this model entails that a particular system with slightly more (less) hot gas than average must contain a lower (higher) stellar mass than average, leading to a strong anti-correlation between hot gas mass and stellar mass. This anti-correlation is the primary observable we are after.

The simple self-similar model, thereby closed box postulations, does not capture the effects of random perturbations – e.g. mergers and chaotic accretion from the halo vicinity – and systematic astrophysical phenomena – e.g. the supernova feedback and the active galactic nuclei (AGN) feedback. The mergers and chaotic accretion potentially induce uncorrelated scatter, and leading to a weaker anti-correlation. The feedback events can pull out the stellar or gaseous particles outside of the virial radius of halos, which messes up the anti-correlation. In low-mass halos, such as those hosting a single massive galaxy like the Milky Way, feedback from supernovae and black holes is energetic enough to vent baryons in the gas phase out of these relatively shallow gravitational potentials. The effect of these events on the reservoir of the gas and stellar particles depends on the rate of energy deposition, and radial scale by which halos are defined.

The most massive halos, the hosts of rich galaxy clusters, could still preserve their closed box nature. Feedback phenomena are not energetic enough to drive baryons out of the host halo. Cosmological, hydro-dynamical simulations are indispensable to predict the influence of these complex phenomena on total baryonic scatter and the degree of anti-correlation between hot gas and stellar mass. The predictions for the magnitude of this anti-correlation have recently emerged from two independent hydrodynamical simulations, AMR-based Rhapsody-G (Wu et al. 2015) and SPH-based BAHAMAS simulations (Chapter III of this dissertation).

Numerical simulations have established in great detail and with high statistical significance how cold dark matter halo formation and evolution is driven by gravitational amplification and collapse of the primordial dark matter density field (Springel et al. 2005; Tinker et al. 2008). On the baryonic side, these simulations have also successfully achieved high enough resolution and sub-grid complexity in order to explore, with high fidelity, how the baryonic components, such as stellar and gaseous content, evolve within the potential wells of dark matter halos (Vogelsberger et al. 2014; McCarthy et al. 2017). The state-of-the-art simulations take the approach of tuning sub-grid control parameters, that includes metal-dependent radiative cooling, star formation, and prescriptions for both supernova and AGN feedback, to match the observed galaxy stellar mass function and the hot gas mass fractions of groups and clusters simultaneously (McCarthy et al. 2017).

Thus far, observational studies have thoroughly explored the scaling behavior of baryonic observables (Giodini et al. 2013), including the scatter, but the constrains for anti-correlation between the two baryonic phases of halos, the gaseous and stellar phase, is unknown. This latter, unconstrained observable provides a novel and complementary means of studying the closed box nature of halos. While both baryon

scatter and the expected anti-correlation can examine the closed box scenario, the former approach is sensitive to how well the measurement uncertainties and the sample selection are understood and modeled. The latter approach is insensitive or weakly sensitive to these systematics. To complement the scaling study of Chapter VI, this Chapter establishes the first-ever empirical evidence of anti-correlation that supports the closed box predictions and the predictions arose from the simulations.

Despite the few extant studies investigated the covariance of hot gas observables (Mantz et al. 2010; Maughan 2014; Mantz et al. 2016a; Andreon et al. 2017), no empirical study has attempted to constrain the intrinsic covariance for any pair of a gas mass observable and a stellar mass observable. Typically the hot gas content of clusters are measured via X-ray or millimeter bands, and the stellar content of these systems are studied via optical or infrared bands. The historical absence of a well-defined, multi-wavelength cluster sample explains the sparsity of observational attempts to constrain the off-diagonal elements of the property covariance matrix. The minimum requirement for this analysis is to obtain both stellar mass and gas mass estimates for a cluster sample with a well-defined selection criterion and robust halo mass measurements. These requirements are fulfilled with the advent of the Local Cluster Substructure Survey (LoCuSS). In this Chapter, we report the first empirical evidence of the anti-correlation for pairs of hot gas mass and stellar mass observables, by employing a sub-sample of the LoCuSS original cluster sample.

The multi-wavelength nature of LoCuSS cluster sample offers a unique opportunity to study both the stellar and hot gas content of the cosmic giants. Observations of the LoCuSS cluster sample that utilized in this study spanned a decade, 2005-2014, and includes data from the Subary 8.2-m telescope and the 3.8-m United Kingdom Infrared Telescope on Mauna Kea, and the *Chandra* and *XMM-Newton*

X-ray satellites. These data, the combination of X-ray and optical/infrared, allow us to investigate the stellar and gaseous components with a set of well-defined halos with robust mass measurements (Okabe & Smith 2016) (see Table 6.1 for the complete list of observables employed in this Chapter). Another superior feature of this sample is the existence of a well-defined selection criterion. Clusters are selected from the ROSAT All-sky Survey (RASS) catalogs by applying a redshift-dependent luminosity cut, where the sample is deemed to be complete; and to correct for the effect of the sample selection, this selection criterion is incorporated into our analysis model. We take a hierarchical Bayesian approach which accounts for the effects of the sample selection, the measurement errors covariance, and the halo mass density function. Our model simultaneously constrains the scaling parameters, i.e. the intercept and the slope, and the full property covariance. The property covariance is the covariance between a pair of observables about the mean mass–observable relation, Equation 7.1. In this Chapter, we report the "property correlation", which is merely the property covariance divided by the intrinsic scatter of each observable, Equation 7.2.

Table 7.1 summarizes the estimated correlation coefficient for each pair of observables (off-diagonal, lower triangle elements) as well as the estimated scatter for each observable (diagonal elements). These results provide the tightest and the most comprehensive constraints for the property correlation among X-ray observables and the first constraints for the property correlation among pairs of X-ray and optical/infrared observables.

Highlighted in Figure 7.1, there is a strong positive property correlation between hot gas mass and X-ray luminosity as well as hot gas mass and integrated electron thermal energy at fixed halo mass. These findings are in excellent agreement with

previous observational findings (Mantz et al. 2016a; Andreon et al. 2017) and a set of model-dependent predictions from hydrodynamical simulations (Stanek et al. 2010; Truong et al. 2018). Despite the agreement between the empirical estimates and simulations predictions of the property correlations, the scatters derived from simulations are significantly smaller than the observational estimates. This systematic difference could be an indication of systematic bias in the quoted measurement errors for the observational sample. The effect of this systematic uncertainty on the estimation of the property correlation is sub-dominant compare to statistical uncertainties as discussed in the next section.

The estimated positive correlation between gas mass and other X-ray observables reflect the fact that the X-ray observables may be used as a noisy proxy for the gas mass of halos. On the optical side, the optical-richness, λ, and near-infrared luminosity, L_K, may be used as a noisy proxy of the stellar content of halos, thereby a positive correlation between these two observables are expected. We, indeed, find $\sim 2\sigma$ evidence that the correlation between L_K and λ is positive.

Figure 7.2 depicts the posterior estimate for the property correlation of each pair of hot gas mass proxy and stellar mass proxy. These trends are in agreement with the closed box model expectation as well as the predictions from the hydrodynamical simulations (Wu et al. 2015, and Chapter III of this dissertation). The uncertainties on these quantities are broad; we, therefore, cannot rule out a particular prediction from a set of simulations. These results are not currently dominated by the systematic uncertainties, nevertheless systematic effects would induce positive correlation on the estimate of two anti-correlated observables.

There are two primary systematic effects which can wash out or weaken the estimated anti-correlation signal. The first systematic effect would be due to a bias

in the estimated scaling relation. A bias in the normalization and slope of the scaling relation induces a positive correlation, but the magnitude of this bias out to be substantial to have a significant effect on our results. The details are provided in Section 7.3. To reduce the potential bias on the inferred scaling relation, we carefully incorporate the sample selection, which is the primary source of such a bias, into our model. The second systematic effect is linked to the fact that these observables are a noisy estimator of stellar mass or hot gas mass of the underlying halo population. This additional scatter would dilute the strength of an anti-correlation. This is discussed in detail in the next section. We note that this effect is implicitly revealed in our results. The integrated light is a better indicator of the total stellar mass, i.e. has less scatter, than the number of galaxies. The significant and systematic negative shift in the inferred anti-correlation for L_K and X-ray observables with respect to the inferred anti-correlation for λ and X-ray observables indicates that the degree of anti-correlation for the total stellar mass should be even stronger. Because this additional scatter is not quantified, we cannot correct for this effect. Nevertheless, we do not expect this correction to be large, as the intrinsic scatter on the $M_{\rm gas}$ and L_K is itself small. Both of the above systematics are inducing positive correlation. Thus, these results are a lower bound on the value of anti-correlation between the stellar mass and the gas mass of the underlying halo population, and our main conclusion remains unchanged.

The above results extend the evidence for the closed box model in which the most massive halos maintain their baryonic content. If the massive halos are entirely closed boxes and stellar content is a result of the cooling process, a property correlation of negative one is expected. Feedback from supernovae and AGNs helps the mixing hot gas and cold gas, but are not energetic enough to extract baryons from the

very deep gravitational potential of the most massive halos, thereby only affects the inner part of these cosmic giants. But, mergers and chaotic accretion from the halos vicinity may induce uncorrelated perturbations in the baryonic content of halos. This additional noise would also weaken the anti-correlation. Both of these effects have been shown to operate observationally. These two competing effects, the cooling star formation and the accretion, has an opposite effect which could be better understood by the proposed property covariance, yet the exact value to be determined.

On a final note, it worth to mention that this property covariance is an essential ingredient of a multi-wavelength cluster cosmology likelihood analysis (Evrard et al. 2014, and Chapter III of this dissertation). Hence, not only is this quantity an interest of astrophysics studies, but a multi-wavelength cluster inference model would also be incomplete, thereby biased, without this covariance element in it.

Now larger multi-wavelength cluster samples are in the process of being assembled (Dark Energy Survey Collaboration et al. 2016). We, thus, expect the accuracy of these measurements gets better with these upcoming sample. To support these emerging cluster samples, a future research direction could provide realistic error covariance for the observed quantities, which would be the primary source of systematics. In that regard, large simulated multi-wavelength cluster samples can play a key role in better understanding of these systematics.

7.3 Systematic Effects

Here, we study the effect of systematic uncertainties in the estimated property correlation, which support the claims made in the beginning of this chapter. We are primarily interested in constraining the correlation coefficient between hot gas mass and stellar mass of the underlaying halos population, which is expected to be

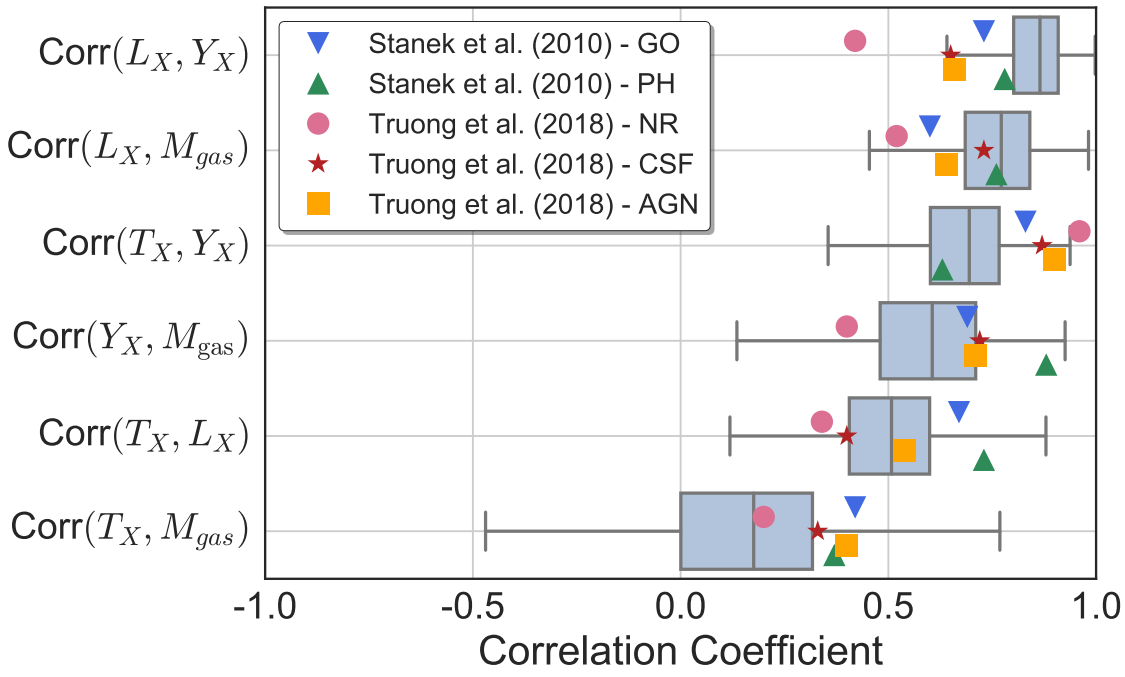

Figure 7.1: The X-ray observables correlation coefficients at fixed halo mass. This compares predictions emerged from SPH-based cosmological hydrodynamical simulations (markers) and the empirical constraints from the observational data (the box plots). The box plots are showing the posterior constraints from the LoCuSS cluster sample (this work). The markers are model-dependent simulations predictions, each of which assumes different hydrodynamical treatment or provided by different team. The statistical error bars on simulation predictions are all negligible ($< 10^2$); therefore they are not shown. Box plot is a standardized way of displaying a probability distribution. The middle line shows the median of the posterior distribution; the box edges show the first and third quartiles, which are equivalent at 25 and 75 percentiles respectively, and the whiskers extend to show the rest of the distribution, 0.35 and 99.65 percentiles.

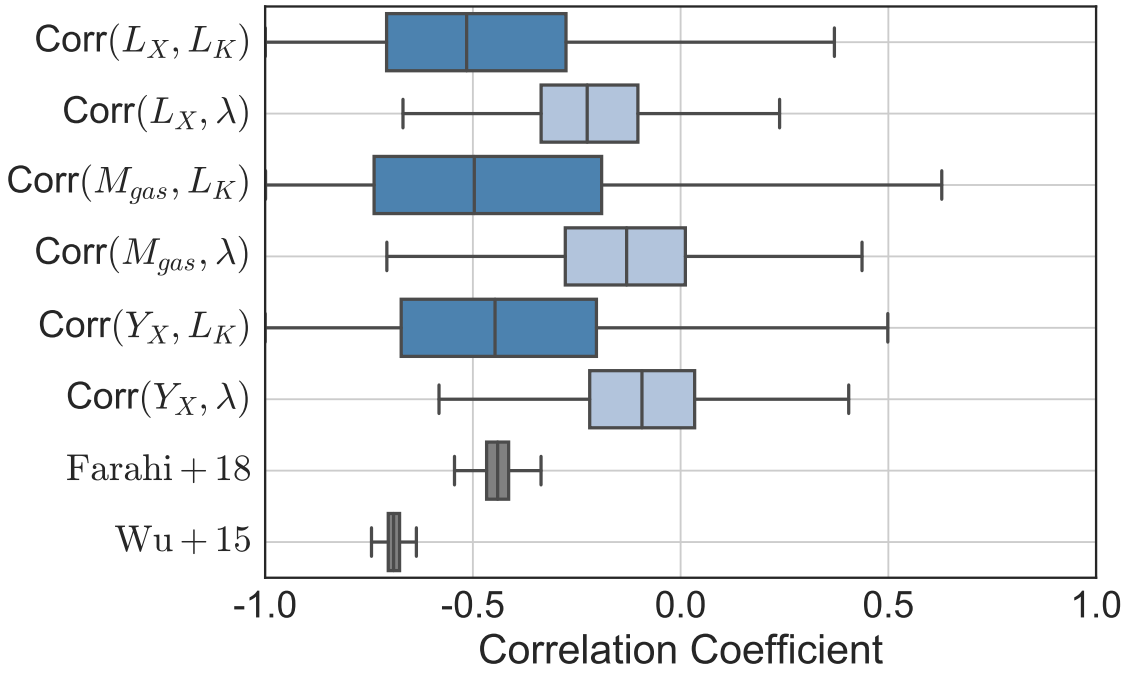

Figure 7.2: The X-ray and optical observables correlation coefficients at fixed halo mass. This compares predictions emerged from two independent hydrodynamical simulations, AMR-based Rhapsody-G (Wu et al. 2015) and SPH-based BAHAMAS simulations (Chapter III of this dissertation), and the empirical results from the LoCuSS cluster sample (the box plots). The gray boxes are model-dependent, simulations predictions, each of which assumes different hydrodynamical treatment. Box plot is a standardized way of displaying a probability distribution. The middle line shows the median of the posterior distribution; the box edges show the first and third quartiles, which are equivalent at 25 and 75 percentiles respectively, and the whiskers extend to show the rest of the distribution, 0.35 and 99.65 percentiles.

Table 7.1: **Lower Triangle**: The median and 68 percentile of the posterior distribution for the correlation coefficient. **Upper Triangle**: The statistical significance of the estimated property correlation assuming no-correlation, $r = 0$, as the null hypothesis. This has been calculated by measuring the probability of having positive (negative) correlation according to the posterior distribution if the median is negative (positive). **Diagonal**: The intrinsic scatter of an observable about the fixed weak-lensing mass.

	L_X	$\mathrm{k}_B T_x$	$M_{\rm gas}$	Y_X	Y_{pl}	Y_{SZA}	L_K	λ
L_X	$\mathbf{0.37^{+0.05}_{-0.05}}$	0.002	0.002	$< 10^{-4}$	0.003	0.04	0.12	0.10
$\mathrm{k}_B T_X$	$0.51^{+0.13}_{-0.16}$	$\mathbf{0.20^{+0.04}_{-0.03}}$	0.24	$< 10^{-4}$	0.19	0.08	0.15	0.15
$M_{\rm gas}$	$0.77^{+0.09}_{-0.13}$	$0.17^{+0.21}_{-0.24}$	$\mathbf{0.16^{+0.03}_{-0.03}}$	0.009	0.006	0.32	0.14	0.27
Y_X	$0.86^{+0.06}_{-0.10}$	$0.69^{+0.10}_{-0.14}$	$0.61^{+0.14}_{-0.19}$	$\mathbf{0.34^{+0.06}_{-0.05}}$	0.001	0.06	0.13	0.31
Y_{pl}	$0.57^{+0.12}_{-0.16}$	$0.18^{+0.19}_{-0.20}$	$0.58^{+0.13}_{-0.17}$	$0.62^{+0.12}_{-0.16}$	$\mathbf{0.28^{+0.04}_{-0.04}}$	0.48	0.41	0.18
Y_{SZA}	$0.44^{+0.17}_{-0.22}$	$0.39^{+0.20}_{-0.25}$	$0.12^{+0.22}_{-0.30}$	$0.43^{+0.18}_{-0.24}$	$0.01^{+0.22}_{-0.24}$	$\mathbf{0.30^{+0.08}_{-0.09}}$	0.31	0.27
L_K	$-0.52^{+0.41}_{-0.27}$	$-0.45^{+0.43}_{-0.32}$	$-0.50^{+0.46}_{-0.33}$	$-0.45^{+0.40}_{-0.32}$	$0.09^{+0.37}_{-0.42}$	$-0.24^{+0.49}_{-0.45}$	$\mathbf{0.07^{+0.07}_{-0.05}}$	0.08
λ	$-0.22^{+0.17}_{-0.17}$	$-0.25^{+0.24}_{-0.22}$	$-0.13^{+0.21}_{-0.21}$	$-0.09^{+0.18}_{-0.19}$	$0.18^{+0.18}_{-0.19}$	$-0.16^{+0.26}_{-0.26}$	$0.75^{+0.17}_{-0.43}$	$\mathbf{0.24^{+0.05}_{-0.04}}$

anti-correlated. We identify two primary sources of bias which can have an effect on the estimated correlation. We note that both of these systematics are inducing positive covariance for anti-correlated quantities. We discuss and quantify each of these effects in the following.

The first potential systematic arises from a bias in the estimated scaling relation. A biased scaling relation induces a positive correlation; however, a significant change on the estimated the anti-correlation requires a substantial bias in the scaling relation. To illustrate this effect, we generate a set of simulated clusters and estimate the correlation coefficient under different scenarios. We, first, take the LoCuSS weak-lensing masses and assume a $M_{\rm gas}$– $M_{\rm WL}$ and L_K– $M_{\rm WL}$ relations. Draw a random $M_{\rm gas}$ and L_K from a multivariate log-normal distribution with a variable correlation coefficient and 20% intrinsic scatter. Then, residuals are measured by assuming a biased scaling relation. Finally, the correlation coefficient is estimated according to Equation 7.2. For each input correlation coefficient, 1,000 realizations of LoCuSS-like cluster sample are generated. Figure 7.3 illustrates the shift in the estimated

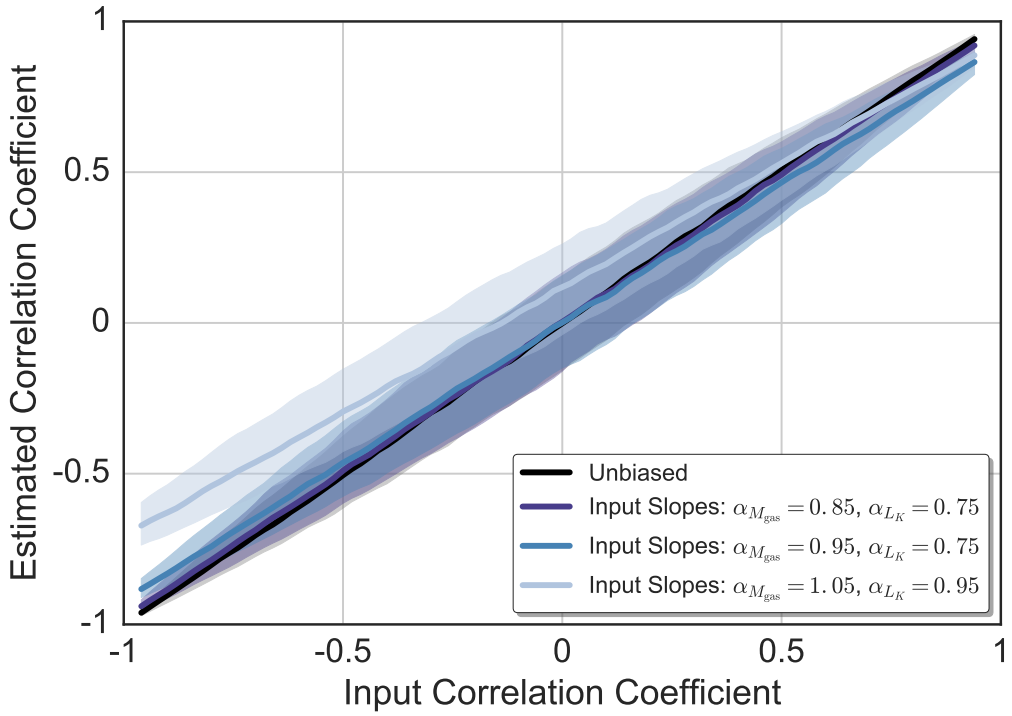

Figure 7.3: The effect of bias in the scaling relation on the estimated property correlation. This illustrates the estimated property correlation for a set of simulated clusters, with the same mass as the LoCuSS cluster sample. To estimate the correlation coefficient a fixed slope of 0.75 for both $M_{\rm gass}$–M and L_K–M relation is assumed, while the input slope takes different value specified in the legend. The shaded areas are 68% confidence intervals derived from 1,000 realizations for each input correlation coefficient.

correlation coefficient as a function of input correlation coefficient. The shift in the estimated correlation coefficient is modest even for a very large bias in the assumed scaling relation. A biased scaling relation is typically a indicator of unaccounted sample selection. Because the sample selection is already taken into account in our analysis, we do not expect a large bias in the estimated scaling relation for this sample.

The second potential systematic arises from the additional, uncalibrated scatter exists for a measured quantity with respect to the intrinsic halo quantity that we are after. Although correcting for this effect is relatively straightforward, this requires a prior estimate of this scatter conditioned on the halo quantity, for exam-

ple $\mathrm{Var}(L_K|M_{\rm star})$. Figure 7.4 illustrates the effect of this additional scatter for a LoCuSS-like cluster sample. To make a realization, we, first, take the LoCuSS weak-lensing masses and assume a $M_{\rm gas}$– $M_{\rm WL}$ and a $M_{\rm star}$– $M_{\rm WL}$ relations with unity slope. Then, a random realization of $M_{\rm gas,true}$ and $M_{\rm star,true}$ is drawn from a multivariate log-normal distribution with a variable input correlation coefficient and 20% scatter. Then, each intrinsic halo quantity, $M_{\rm gas,true}$ and $M_{\rm star,true}$, is further perturbed with an additional scatter to get observed quantities, $X = L_{K,\rm obs}$ and $Y = M_{\rm gas,obs}$. Finally, the property correlation of this realization is estimated according to Equation 7.2. For each input correlation coefficient, 1,000 realizations are generated to get an estimate of the bias uncertainties. Figure 7.4 illustrates the bias in the estimated property correlation for a noisy measurement of true halo quantities, i.e. X and Y. This additional scatter washes out the (anti)-correlation signal, unless this additional scatter is calibrated and corrected for.

As an illuminative example, the systematic difference between the posterior estimate of correlation coefficient for $\{M_{\rm gas}, L_K\}$ pair and $\{M_{\rm gas}, \lambda\}$ pair could be understood via the "additional scatter" effect. We assume that λ is a nosier measurement of L_K with additional scatter of $\sim 20\%$. If we further assume that the value of correlation coefficient between L_K and $M_{\rm gas}$ about fixed halo mass is ~ -0.5 and the intrinsic scatter of both $M_{\rm gas}$ and L_K are $\sim 10\%$, then the expected value of correlation coefficient for λ and $M_{\rm gas}$ would be about -0.22. This simple calculation is in excellent agreement with our findings, and can describe the systematic shift in the posterior distributions (See Figure 7.2).

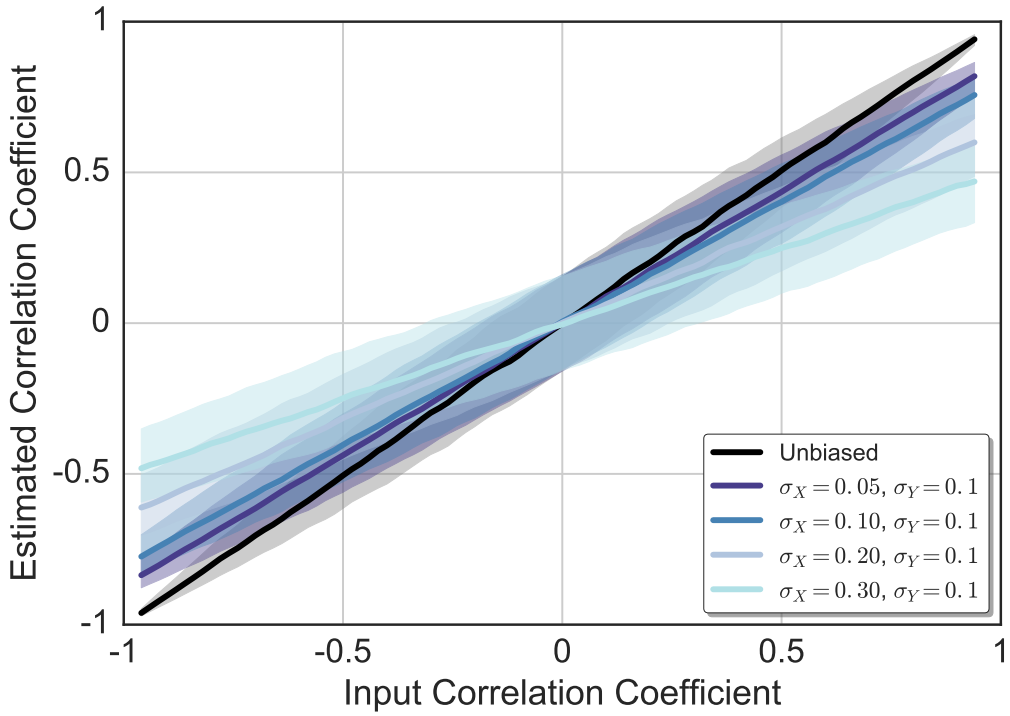

Figure 7.4: The effect of additional scatter between intrinsic quantities and the measured quantities on the estimated property correlation. This figure illustrates the estimated correlation coefficient for a set of simulated cluster samples, with the same mass as the LoCuSS cluster sample. To estimate the property correlation of two observed quantities, we generate a realization of intrinsic quantities with an input property correlation, and then add additional uncalibrated scatter to the intrinsic quantities to get measured quantities. Finally, the property correlation of measured quantities are estimated. These additional, uncalibrated scatters are specified in the legend. The shaded areas are 68% confidence intervals derived from 1,000 realizations for each input correlation coefficient.

CHAPTER VIII

Conclusion

Galaxy clusters, the most massive collapsed objects in the Universe, are recognized as a powerful probe of the cosmological parameters (Weinberg et al. 2013; Huterer & Shafer 2018). These cosmic giants can be employed to study both the growth of structure and the expansion history of the Universe. To achieve this goal, the past generation of cluster samples have successfully delivered competitive cosmological constraints and practically illustrated that a population of clusters can produce complementary results. Now, the next generation surveys are in the process of obtaining larger and deeper cluster samples to further improve our understanding of the physics of the Universe. One such ongoing survey is the Dark Energy Survey (DES, Dark Energy Survey Collaboration et al. 2016), which is expected to gather information on tens of thousands of these systems. While such large samples of clusters provide ample opportunities for discovery, accurate and precise statistical modeling of these systems are a significant challenge, which is the subject of this dissertation. The developed models and techniques in this dissertation are a significant step forward in the modeling of multi-wavelength cluster samples.

One popular approach in constraining the cosmological parameters is to compare the population statistics of galaxy clusters, as a function their observable properties, with theoretical predictions. To calculate the conditional population statistics of clusters, Chapter II proposes a new statistical approach. This population model enables fast and accurate computation of cluster space density as a function of their observable properties, which is a key ingredient of cluster cosmology analysis. The space density of clusters as a function of their observable properties and redshift is a quantity which can readily be derived from a cluster sample. However, theory predicts the halo mass function (HMF), the space density of dark matter halos as a function of their mass and redshift. Bridging the gap between theoretical predictions and measured quantities demands two fundamental elements: (1) Modeling HMF as a function cosmological parameters, (2) and mapping cluster observable properties to the underlying halo mass.

Since the pioneering work of Press & Schechter (1974), many forms of HMF have been put forward in the literature. Today, there are HMF which are calibrated, with better than one-percent precision, against large suites of cosmological simulations. The current modeling challenge rests in the second element mentioned above, modeling and calibrating the underlying mass of a cluster sample. This dissertation aims to address this challenge by adopting a multi-wavelength cluster analysis approach. Simulations have also played a pivotal role in assessing the accuracy and precision of the mass calibration models. As of today, the mass calibration of a cluster sample remains as the dominant source of uncertainty in any cluster cosmology study.

Zwicky (1937) was the first to estimate a cluster's mass from luminous matter. Since then several independent mass measurement techniques have been put forward – e.g., weak-lensing, dynamical, and hydrostatic techniques – and applied to various

cluster samples. Adding to this stream of work, I take multiple approaches to model and infer the mass scale of various cluster samples. In Chapter IV and Chapter V, I present a novel mass calibration technique based on the ensemble of kinematics of cluster member galaxies. This model is then applied to the cluster samples derived from an optical, SDSS, and an X-ray, XMM-XXL, survey. To calibrate the mass of a multi-wavelength cluster sample derived from the LoCuSS survey, I develop a new inference algorithm. This model, which is described in Chapter VI, determines the mass–property relations for nine observables by employing weak-lensing measurements of 41 LoCuSS-selected clusters.

The stacked technique developed in Chapter IV, similar to other stacking techniques, comes with an inherent drawback. These techniques are capable of estimating only the expected conditional halo mass and cannot capture the probability distribution of individual cluster's mass. To determine this probability distribution, an independent inference algorithm is required, which is the subject of a new project I am currently pursuing. It is typically assumed that the probability distribution of halos' mass has a log-normal form, which is widely used in the literature. In Chapter III, I employ halos derived from suites of cosmological simulations to assess the accuracy of this assumption. To extend this work, Chapter VI presents a new set of estimates for the scatter about the mean relation, which defines the width of the log-normal distribution, for nine cluster observables. The uncertainty in the halo mass scatter conditioned on cluster observables is a primary source of systematic uncertainty in cosmological analysis with clusters. In collaboration with the members of the DES cluster working group, I am developing an algorithm which employs multi-wavelength cluster data to infer this scatter parameter.

Another essential element of a multi-wavelength cluster analysis is the property covariance between two intrinsic properties of a halo population. In Chapter III, I explore this property covariance for a set of halos derived from hydrodynamical simulations. An outcome of this study leads to new prediction that two phase of baryonic content of halos – stellar mass and gas mass – are anti-correlated at fixed halo mass. Chapter VII of this dissertation employs the LoCuSS cluster sample to empirically estimate this quantity for a broad set of cluster observables. The results establish the first observational estimate of the degree of anti-correlation between the stellar and gas content of halos and confirms the predictions made in Chapter III. Furthermore, the results of this analysis open up a new, promising research direction which has an impact on the better understanding of how the astrophysics affect the baryonic content halos. Due to the importance of the property covariance on multi-wavelength cluster cosmology analysis, an important future direction would be to better constrain this quantity via future observations and to incorporate this covariance into inference algorithms.

Inference models should be vetted with synthetic data before being applied to observational data. In Chapter III of this dissertation, I utilize a set of halos derived from hydrodynamical simulations to validate the proposed halo population model introduced in Chapter II. The results of this validation suggest that while the log-normal model is an accurate description of halo properties conditioned on halo mass, the commonly used power-law model with constant slope is an insufficient approximation. This study suggests that low-order polynomial should be sufficient to capture the scale- and time-dependent behaviors of the local slope and covariance of halo mass-halo properties scaling relation.

In a separate effort, I employed clusters derived from a light-cone produced by N-body simulations to validate the stacked dynamical mass calibration technique proposed in Chapter IV. As a third application of synthetic data for validation, the hierarchical Bayesian model developed in Chapter VI is validated with many realizations of synthetic LoCuSS-like datasets. As a future direction, I am currently developing a technology which generates synthetic X-ray emission maps from groups and clusters of galaxies to study the appropriate parameterization of the selection function of X-ray surveys. This new direction will enable cosmological analysis with cluster samples derived from X-ray surveys.

Narrowing the uncertainty on inferred cosmological parameters, ultimate goal of any cosmology analysis, demands a better understanding of systematic effects. Currently, the stacked weak–lensing method is the primary choice for the cluster mass calibration of the emerging DES cluster sample. Given the widespread use of this method, it is of interest to model and understand its potential systematics. A recent analysis of the weak–lensing method has identified the following effects as the primary sources of systematic uncertainty: (1) cluster's orientation, (2) contamination due to the correlated and uncorrelated projected structures, (3) the misidentified cluster centers, and (4) the intrinsic halo mass scatter about the mean relation. A future direction would be to model and calibrate each of the above systematics with DES-like cluster catalogs synthesized from cosmological simulations. The results of this analysis would support the science goal of constraining the dark energy equation of state with the DES cluster sample.

Ultimately, a full calibration of halo mass–multi-wavelength cluster observables should describe the full complexity of the mapping between halos and clusters. This requires full knowledge of astrophysics, the detection algorithms, and measurement

uncertainties. Building such knowledge requires a good understanding of the effects of error covariance, intrinsic covariance between observables, cluster environment, and sample selection, while the detection algorithms may alter measured properties due to induce cluster orientation, projection, mis-centering, blending, and fragmentation. Accurate and precise modeling of all these phenomena are the big challenge of cluster cosmology. Last but not least, the work developed in this dissertation provides an important step forward in the modeling of multi-wavelength cluster samples and represents a critical step in the quest to unleash the full statistical power of future cluster surveys.

Bibliography

Aarseth S. J., Turner E. L., Gott III J. R., 1979, ApJ, 228, 664

Abernethy J., et al., 2016, arXiv preprint arXiv:1610.00580

Adami C., et al., 2011, A&A, 526, A18

Adami C., et al., 2017

Adams F. C., Fatuzzo M., 1996, ApJ, 464, 256

Ahn C. P., et al., 2014, ApJS, 211, 17

Akrami M., 2011, ArXiv e-prints,

Akritas M. G., Bershady M. A., 1996, ApJ, 470, 706

Alam S., et al., 2015, ApJS, 219, 12

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

Aminrazavi M., 2013, The wine of wisdom: the life, poetry and philosophy of Omar Khayyam. Oneworld Publications

Amodeo S., et al., 2017, ApJ, 844, 101

Andreon S., 2010, MNRAS, 407, 263

Andreon S., 2012, A&A, 548, A83

Andreon S., Wang J., Trinchieri G., Moretti A., Serra A. L., 2017, ArXiv e-prints,

Areán D., Farahi A., Pando Zayas L. A., Landea I. S., Scardicchio A., 2014, Phys. Rev. D, 89, 106003

Areán D., Farahi A., Pando Zayas L. A., Salazar Landea I., Scardicchio A., 2015, Journal of High Energy Physics, 7, 46

Arnaud M., Pointecouteau E., Pratt G. W., 2005, A&A, 441, 893

Arnaud M., Pratt G. W., Piffaretti R., Böhringer H., Croston J. H., Pointecouteau E., 2010, A&A, 517, A92

Azarian M. K., et al., 2010, Missouri Journal of Mathematical Sciences, 22, 64

Bahé Y. M., McCarthy I. G., King L. J., 2012, MNRAS, 421, 1073

Ballay U., 1990, Arabica, 37, 389

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

Barnes D. J., Kay S. T., Henson M. A., McCarthy I. G., Schaye J., Jenkins A., 2017, MNRAS, 465, 213

Barsanti S., Girardi M., Biviano A., Borgani S., Annunziatella M., Nonino M., 2016, A&A, 595, A73

Bayliss M. B., et al., 2017, ApJ, 837, 88

Becker M. R., Kravtsov A. V., 2011, ApJ, 740, 25

Becker M. R., et al., 2007, ApJ, 669, 905

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

Behroozi P. S., Wechsler R. H., Conroy C., 2013b, ApJ, 770, 57

Benson A. J., et al., 2013, MNRAS, 428, 1774

Bertschinger E., 1985, ApJS, 58, 39

Bhavsar S. P., Aarseth S. J., Gott III J. R., 1981, ApJ, 246, 656

Biffi V., Sembolini F., De Petris M., Valdarnini R., Yepes G., Gottlöber S., 2014, MNRAS, 439, 588

Bird C. M., 1994, AJ, 107, 1637

Biviano A., Girardi M., 2003, ApJ, 585, 205

Biviano A., Murante G., Borgani S., Diaferio A., Dolag K., Girardi M., 2006, A&A, 456, 23

Blanchard A., Wachter K., Evrard A. E., Silk J., 1992, ApJ, 391, 1

Bleem L. E., Stalder B., Brodwin M., Busha M. T., Gladders M. D., High F. W., Rest A., Wechsler R. H., 2015, ApJS, 216, 20

Böhringer H., et al., 2004, A&A, 425, 367

Bond G. P., 1848, An Account of the Nebula in Andromeda. publisher not identified

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

Borgani S., Kravtsov A., 2011, Advanced Science Letters, 4, 204

Borgani S., Fabjan D., Tornatore L., Schindler S., Dolag K., Diaferio A., 2008, Space Sci. Rev., 134, 379

Bourdin H., Mazzotta P., Kozmanyan A., Jones C., Vikhlinin A., 2017, ApJ, 843, 72

Briel U. G., Henry J. P., Boehringer H., 1992, A&A, 259, L31

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

Busha M. T., Evrard A. E., Adams F. C., Wechsler R. H., 2005, MNRAS, 363, L11

Busha M. T., Wechsler R. H., Becker M. R., Erickson B., Evrard A. E., 2013, in American Astronomical Society Meeting Abstracts 221.

Carlberg R. G., 1994, ApJ, 433, 468

Carroll S. M., 1997, ArXiv General Relativity and Quantum Cosmology e-prints,

Cavaliere A., Fusco-Femiano R., 1976, A&A, 49, 137

Cavaliere A., Fusco-Femiano R., 1978, A&A, 70, 677

Cen R., Ostriker J. P., 2000, ApJ, 538, 83

Centrella J., Melott A. L., 1983, Nature, 305, 196

Chang C., et al., 2015, ApJ, 801, 73

Chiappetti L., et al., 2017

Chiu I., et al., 2016, MNRAS, 455, 258

Chiu I., et al., 2017, ArXiv e-prints,

Chojnacki A., Dai C., Farahi A., Shi G., Webb J., Zhang D. T., Abernethy J., Schwartz E., 2017, in Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp 1407–1416

Cohn J. D., Evrard A. E., White M., Croton D., Ellingson E., 2007, MNRAS, 382, 1738

Coziol R., Andernach H., Caretta C. A., Alamo-Martínez K. A., Tago E., 2009, AJ, 137, 4795

Croton D. J., et al., 2006, MNRAS, 365, 11

Cuesta A. J., Prada F., Klypin A., Moles M., 2008, MNRAS, 389, 385

Cunha C., 2009, Phys. Rev. D, 79, 063009

Czakon N. G., et al., 2015, ApJ, 806, 18

DESI Collaboration 2016, ArXiv e-prints [arXiv:1611.00036],

Dark Energy Survey Collaboration et al., 2016, MNRAS, 460, 1270

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

De Lucia G., Blaizot J., 2007, MNRAS, 375, 2

Diaferio A., 1999, MNRAS, 309, 610

Dietrich J. P., et al., 2014, MNRAS, 443, 1713

Doi M., et al., 2010, AJ, 139, 1628

Dolag K., Borgani S., Schindler S., Diaferio A., Bykov A. M., 2008, Space Sci. Rev., 134, 229

Dong F., Pierpaoli E., Gunn J. E., Wechsler R. H., 2008, ApJ, 676, 868

Duarte M., Mamon G. A., 2015, MNRAS, 453, 3848

Durret F., et al., 2011, A&A, 535, A65

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

Ebeling H., Edge A. C., Bohringer H., Allen S. W., Crawford C. S., Fabian A. C., Voges W., Huchra J. P., 1998, MNRAS, 301, 881

Ebeling H., Edge A. C., Allen S. W., Crawford C. S., Fabian A. C., Huchra J. P., 2000, MNRAS, 318, 333

Eckmiller H. J., Hudson D. S., Reiprich T. H., 2011, A&A, 535, A105

Efstathiou G., Davis M., White S. D. M., Frenk C. S., 1985, ApJS, 57, 241

Erben T., et al., 2013, MNRAS, 433, 2545

Erickson B. M. S., Cunha C. E., Evrard A. E., 2011, Phys. Rev. D, 84, 103506

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

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

Evrard A. E., 1997, MNRAS, 292, 289

Evrard A. E., Metzler C. A., Navarro J. F., 1996, ApJ, 469, 494

Evrard A. E., et al., 2002, ApJ, 573, 7

Evrard A. E., et al., 2008, ApJ, 672, 122

Evrard A. E., Arnault P., Huterer D., Farahi A., 2014, MNRAS, 441, 3562

Fabjan D., Borgani S., Rasia E., Bonafede A., Dolag K., Murante G., Tornatore L., 2011, MNRAS, 416, 801

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

Faltenbacher A., Diemand J., 2006, MNRAS, 369, 1698

Farahi A., Benson A. J., 2013, MNRAS, 433, 3428

Farahi A., Pando Zayas L. A., 2014, Physics Letters B, 734, 31

Farahi A., Evrard A. E., Rozo E., Rykoff E. S., Wechsler R. H., 2016, MNRAS, 460, 3900

Farahi A., Evrard A. E., McCarthy I., Barnes D. J., Kay S. T., 2017a, ArXiv e-prints,

Farahi A., et al., 2017b, ArXiv e-prints,

Gamow G., 1948, Nature, 162, 680

Gardner J., Koutra D., Mroueh J., Pang V., Farahi A., Krassenstein S., Webb J., 2017, arXiv preprint arXiv:1710.06839

Gaskin J. A., et al., 2015, in UV, X-Ray, and Gamma-Ray Space Instrumentation for Astronomy XIX. p. 96010J, doi:10.1117/12.2190837

Gelman A., Carlin J. B., Stern H. S., Rubin D. B., 2014, Bayesian data analysis. Vol. 2, Chapman & Hall/CRC Boca Raton, FL, USA

Gerke B. F., et al., 2005, ApJ, 625, 6

Gerke B. F., et al., 2012, ApJ, 751, 50

Gifford D., Miller C., Kern N., 2013, ApJ, 773, 116

Giles P. A., et al., 2016, A&A, 592, A3

Giles P. A., et al., 2017, MNRAS, 465, 858

Giodini S., Lovisari L., Pointecouteau E., Ettori S., Reiprich T. H., Hoekstra H., 2013, Space Science Reviews, 177, 247

Gladders M. D., Yee H. K. C., 2005, ApJS, 157, 1

Gonzalez A. H., Zaritsky D., Zabludoff A. I., 2007, ApJ, 666, 147

Gonzalez A. H., Sivanandam S., Zabludoff A. I., Zaritsky D., 2013, ApJ, 778, 14

Gough B., 2009, GNU scientific library reference manual. Network Theory Ltd.

Guglielmo V., et al., 2017, ArXiv e-prints [arXiv:1710.04667],

Gunn J. E., Gott III J. R., 1972, ApJ, 176, 1

Gunn J. E., et al., 1998, AJ, 116, 3040

Guo H., et al., 2015a, MNRAS, 446, 578

Guo H., et al., 2015b, MNRAS, 453, 4368

Guzzo L., et al., 2014, A&A, 566, A108

Hafez I., Stephenson F. R., Orchiston W., 2011, in , Highlighting the History of Astronomy in the Asia-Pacific Region. Springer, pp 121–138

Hahn O., Martizzi D., Wu H.-Y., Evrard A. E., Teyssier R., Wechsler R. H., 2017a, MNRAS, 470, 166

Hahn O., Martizzi D., Wu H.-Y., Evrard A. E., Teyssier R., Wechsler R. H., 2017b, MNRAS, 470, 166

Haines C. P., et al., 2009, ApJ, 704, 126

Haines C. P., et al., 2015, ApJ, 806, 101

Hao J., et al., 2009, ApJ, 702, 745

Hearin A. P., Watson D. F., 2013, MNRAS, 435, 1313

Hey T., Tansley S., Tolle K. M., et al., 2009, The fourth paradigm: data-intensive scientific discovery. Vol. 1, Microsoft research Redmond, WA

Heymans C., et al., 2012, MNRAS, 427, 146

Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, MNRAS, 449, 685

Hopkins A. M., et al., 2013, MNRAS, 430, 2047

Huterer D., Shafer D. L., 2018, Reports on Progress in Physics, 81, 016901

Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372

Juin J. B., Yvon D., Réfrégier A., Yèche C., 2007, A&A, 465, 57

Kaiser N., 1986, MNRAS, 222, 323

Kasun S. F., Evrard A. E., 2005, ApJ, 629, 781

Katz N., Weinberg D. H., Hernquist L., 1996, ApJS, 105, 19

Kelly B. C., 2007, ApJ, 665, 1489

Kennedy E. S., 1947, Isis, 38, 56

Kennedy E. S., 1950, Isis, 41, 180

Kennedy E. S., 1951, Journal of the American Oriental Society, 71, 13

Kennedy E. S., 1952, Isis, 43, 42

Kennedy E. S., 1966, Isis, 57, 365

Kettula K., et al., 2013, ApJ, 778, 74

Kettula K., et al., 2015, MNRAS, 451, 1460

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

Koester B. P., et al., 2007, ApJ, 660, 239

Kravtsov A. V., 2013, ApJL, 764, L31

Kravtsov A. V., Borgani S., 2012, ARAA, 50, 353

Kravtsov A. V., Vikhlinin A., Nagai D., 2006, ApJ, 650, 128

Laganá T. F., Zhang Y.-Y., Reiprich T. H., Schneider P., 2011, ApJ, 743, 13

Larson R. B., 1973, MNRAS, 161, 133

Lau E. T., Nagai D., Kravtsov A. V., 2010, ApJ, 708, 1419

Lauer T. R., Postman M., Strauss M. A., Graves G. J., Chisari N. E., 2014, ApJ, 797, 82

Lawrence A., et al., 2007, MNRAS, 379, 1599

Le Brun A. M. C., McCarthy I. G., Schaye J., Ponman T. J., 2017, MNRAS, 466, 4442

Le Fèvre O., et al., 2005, A&A, 439, 845

Le Fèvre O., et al., 2015, A&A, 576, A79

Licitra R., Mei S., Raichoor A., Erben T., Hildebrandt H., 2016, MNRAS, 455, 3020

Lieu M., et al., 2016, A&A, 592, A4

Lin Y.-T., Mohr J. J., Stanford S. A., 2003, ApJ, 591, 749

Lin Y.-T., Mohr J. J., Stanford S. A., 2004, ApJ, 610, 745

Lin Y.-T., Mohr J. J., Gonzalez A. H., Stanford S. A., 2006, ApJL, 650, L99

Lindberg D. C., Shank M. H., 2013, The Cambridge History of Science. The Cambridge History of Science Vol. 2, Cambridge University Press, doi:10.1017/CHO9780511974007

Liske J., et al., 2015, MNRAS, 452, 2087

Lovisari L., Reiprich T. H., Schellenberger G., 2015, A&A, 573, A118

Mamon G. A., Biviano A., Murante G., 2010, A&A, 520, A30

Mannucci F., Basile F., Poggianti B. M., Cimatti A., Daddi E., Pozzetti L., Vanzi L., 2001, MNRAS, 326, 745

Mantz A. B., 2016, MNRAS, 457, 1279

Mantz A., Allen S. W., Ebeling H., Rapetti D., Drlica-Wagner A., 2010, MNRAS, 406, 1773

Mantz A. B., Allen S. W., Morris R. G., Rapetti D. A., Applegate D. E., Kelly P. L., von der Linden A., Schmidt R. W., 2014, MNRAS, 440, 2077

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

Mantz A. B., Allen S. W., Morris R. G., Schmidt R. W., 2016a, MNRAS, 456, 4020

Mantz A. B., et al., 2016b, MNRAS, 463, 3582

Marrone D. P., et al., 2012, ApJ, 754, 119

Martel H., Robichaud F., Barai P., 2014, ApJ, 786, 79

Martino R., Mazzotta P., Bourdin H., Smith G. P., Bartalucci I., Marrone D. P., Finoguenov A., Okabe N., 2014, MNRAS, 443, 2342

Mathews W. G., 1978, ApJ, 219, 413

Maughan B. J., 2014, MNRAS, 437, 1171

Maughan B. J., Giles P. A., Rines K. J., Diaferio A., Geller M. J., Van Der Pyl N., Bonamente M., 2016, MNRAS, 461, 4182

Mayers J. A., et al., 2018, ArXiv e-prints,

McCarthy I. G., Schaye J., Bower R. G., Ponman T. J., Booth C. M., Dalla Vecchia C., Springel V., 2011, MNRAS, 412, 1965

McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936

Melchior P., et al., 2017, MNRAS, 469, 4899

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

Merloni A., et al., 2012, ArXiv e-prints [arXiv:1209.3114],

Mickaelian A., 2016, Astronomy Reports, 60, 857

Milkeraitis M., van Waerbeke L., Heymans C., Hildebrandt H., Dietrich J. P., Erben T., 2010, MNRAS, 406, 673

Miller C. J., Stark A., Gifford D., Kern N., 2016, ApJ, 822, 41

More S., Diemer B., Kravtsov A. V., 2015, ApJ, 810, 36

Mulroy S. L., et al., 2014, MNRAS, 443, 3309

Mulroy S. L., McGee S. L., Gillman S., Smith G. P., Haines C. P., Démoclès J., Okabe N., Egami E., 2017, MNRAS, 472, 3246

Munari E., Biviano A., Borgani S., Murante G., Fabjan D., 2013, MNRAS, 430, 2638

Murphy D. N. A., Geach J. E., Bower R. G., 2012, MNRAS, 420, 1861

Murray S. G., Power C., Robotham A. S. G., 2013, Astronomy and Computing, 3, 23

Mushotzky R. F., Serlemitsos P. J., Boldt E. A., Holt S. S., Smith B. W., 1978, ApJ, 225, 21

Nagai D., Kravtsov A. V., Vikhlinin A., 2007, ApJ, 668, 1

Nastasi A., et al., 2014, A&A, 564, A17

Navarro J. F., Frenk C. S., White S. D. M., 1995, MNRAS, 275, 720

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

Neugebauer O., 2012, A history of ancient mathematical astronomy. Vol. 1, Springer Science & Business Media

Nord B., Stanek R., Rasia E., Evrard A. E., 2008, MNRAS, 383, L10

Oguri M., 2014, MNRAS, 444, 147

Oguri M., Hamana T., 2011, MNRAS, 414, 1851

Okabe N., Smith G. P., 2016, MNRAS, 461, 3794

Okabe N., Zhang Y.-Y., Finoguenov A., Takada M., Smith G. P., Umetsu K., Futamase T., 2010, ApJ, 721, 875

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

Old L., Gray M. E., Pearce F. R., 2013, MNRAS, 434, 2606

Ortiz-Gil A., Guzzo L., Schuecker P., Böhringer H., Collins C. A., 2004, MNRAS, 348, 325

Ostriker J. P., Choi E., Ciotti L., Novak G. S., Proga D., 2010, ApJ, 722, 642

Pacaud F., et al., 2006, MNRAS, 372, 578

Pacaud F., et al., 2016, A&A, 592, A2

Paranjape A., Sheth R. K., Desjacques V., 2013, MNRAS, 431, 1503

Patil A., Huard D., Fonnesbeck C. J., 2010, Journal of statistical software, 35, 1

Penna-Lima M., Bartlett J. G., Rozo E., Melin J.-B., Merten J., Evrard A. E., Postman M., Rykoff E., 2017, A&A, 604, A89

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 1149

Perlmutter S., et al., 1999, ApJ, 517, 565

Pierre M., et al., 2016, A&A, 592, A1

Pillepich A., et al., 2018, MNRAS, 475, 648

Plagge T. J., et al., 2013, ApJ, 770, 112

Planck Collaboration et al., 2014, A&A, 571, A16

Planck Collaboration et al., 2016a, A&A, 594, A8

Planck Collaboration et al., 2016b, A&A, 594, A27

Planelles S., Borgani S., Dolag K., Ettori S., Fabjan D., Murante G., Tornatore L., 2013, MNRAS, 431, 1487

Pound R. V., Rebka G. A., 1960, Phys. Rev. Lett., 4, 337

Pratt G. W., Croston J. H., Arnaud M., Böhringer H., 2009, A&A, 498, 361

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

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical recipes in FORTRAN. The art of scientific computing

Ragep F. J., 2007, History of Science, 45, 65

Ragone-Figueroa C., Plionis M., Merchán M., Gottlöber S., Yepes G., 2010, MNRAS, 407, 581

Rasia E., et al., 2006, MNRAS, 369, 2013

Reed D. S., Bower R., Frenk C. S., Jenkins A., Theuns T., 2007, MNRAS, 374, 2

Richstone D. O., 1975, ApJ, 200, 535

Riess A. G., et al., 1998, AJ, 116, 1009

Rines K., Diaferio A., 2010, AJ, 139, 580

Rines K., Diaferio A., Natarajan P., 2007, ApJ, 657, 183

Robotham A. S. G., et al., 2011, MNRAS, 416, 2640

Rozo E., Rykoff E. S., 2014, ApJ, 783, 80

Rozo E., et al., 2009, ApJ, 703, 601

Rozo E., et al., 2010, ApJ, 708, 645

Rozo E., Rykoff E. S., Becker M., Reddick R. M., Wechsler R. H., 2015, MNRAS, 453, 38

Rubin V. C., Ford Jr. W. K., 1970, ApJ, 159, 379

Rykoff E. S., et al., 2012, ApJ, 746, 178

Rykoff E. S., et al., 2014, ApJ, 785, 104

Rykoff E. S., et al., 2016, ApJS, 224, 1

Sachs R. K., Wolfe A. M., 1967, ApJ, 147, 73

Sadibekova T., Pierre M., Clerc N., Faccioli L., Gastaud R., Le Fevre J.-P., Rozo E., Rykoff E., 2014, A&A, 571, A87

Sarazin C. L., 1988, X-ray emission from clusters of galaxies

Saro A., Mohr J. J., Bazin G., Dolag K., 2013, ApJ, 772, 47

Saro A., et al., 2017, MNRAS, 468, 3347

Schaye J., et al., 2010, MNRAS, 402, 1536

Schellenberger G., Reiprich T. H., 2017, MNRAS, 469, 3738

Schellenberger G., Reiprich T. H., Lovisari L., Nevalainen J., David L., 2015, A&A, 575, A30

Sembolini F., Yepes G., De Petris M., Gottlöber S., Lamagna L., Comis B., 2013, MNRAS, 429, 323

Shaw L. D., Holder G. P., Dudley J., 2010, ApJ, 716, 281

Simet M., McClintock T., Mandelbaum R., Rozo E., Rykoff E., Sheldon E., Wechsler R. H., 2017, MNRAS, 466, 3103

Skibba R. A., van den Bosch F. C., Yang X., More S., Mo H., Fontanot F., 2011, MNRAS, 410, 417

Skielboe A., Wojtak R., Pedersen K., Rozo E., Rykoff E. S., 2012, ApJL, 758, L16

Soares-Santos M., et al., 2011, ApJ, 727, 45

Springel V., et al., 2005, Nature, 435, 629

Springel V., et al., 2018, MNRAS, 475, 676

Stanek R., Rasia E., Evrard A. E., Pearce F., Gazzola L., 2010, ApJ, 715, 1508

Stanford S. A., Gonzalez A. H., Brodwin M., Gettings D. P., Eisenhardt P. R. M., Stern D., Wylezalek D., 2014, ApJS, 213, 25

Sun M., Voit G. M., Donahue M., Jones C., Forman W., Vikhlinin A., 2009, ApJ, 693, 1142

Tinker J. L., Norberg P., Weinberg D. H., Warren M. S., 2007, ApJ, 659, 877

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

Tormen G., Bouchet F. R., White S. D. M., 1997, MNRAS, 286, 865

Tremaine S., et al., 2002, ApJ, 574, 740

Truong N., et al., 2018, MNRAS, 474, 4089

Turner E. L., Aarseth S. J., Blanchard N. T., Mathieu R. D., Gott III J. R., 1979, ApJ, 228, 684

Valotti A., Pierre M., Farahi A., Evrard A., Faccioli L., Sauvageot J.-L., Clerc N., Pacaud F., 2017, ArXiv e-prints,

Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, ApJ, 640, 691

Vikhlinin A., et al., 2009, ApJ, 692, 1060

Vogelsberger M., et al., 2014, Nature, 509, 177

Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, PhysRep, 530, 87

White S. D. M., 1976, MNRAS, 174, 19

White S. D. M., Navarro J. F., Evrard A. E., Frenk C. S., 1993, Nature, 366, 429

Will C. M., 2014, Living Reviews in Relativity, 17, 4

Wilson S., et al., 2016, MNRAS, 463, 413

Wojtak R., Łokas E. L., Mamon G. A., Gottlöber S., Prada F., Moles M., 2007, A&A, 466, 437

Wu H.-Y., Hahn O., Evrard A. E., Wechsler R. H., Dolag K., 2013, MNRAS, 436, 460

Wu H.-Y., Evrard A. E., Hahn O., Martizzi D., Teyssier R., Wechsler R. H., 2015, MNRAS, 452, 1982

Xue Y.-J., Wu X.-P., 2000, ApJ, 538, 65

York D. G., et al., 2000, AJ, 120, 1579

Yoshikawa K., Jing Y. P., Börner G., 2003, ApJ, 590, 654

Zemp M., Gnedin O. Y., Gnedin N. Y., Kravtsov A. V., 2011, ApJS, 197, 30

Zhao H.-H., Li C.-K., Chen Y., Jia S.-M., Song L.-M., 2015, ApJ, 799, 47

Zou S., Maughan B. J., Giles P. A., Vikhlinin A., Pacaud F., Burenin R., Hornstrup A., 2016, MNRAS, 463, 820

Zu Y., Mandelbaum R., 2015, MNRAS, 454, 1161

Zwicky F., 1937, ApJ, 86, 217

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

van den Bosch F. C., Norberg P., Mo H. J., Yang X., 2004, MNRAS, 352, 1302