ABSTRACT

INVESTIGATIONS OF GALAXY CLUSTERS USING GRAVITATIONAL LENSING

Matthew P. Wiesner, Ph.D.
Department of Physics
Northern Illinois University, 2014
Huan Lin and Michael Fortner, Director

In this dissertation, we discuss the properties of galaxy clusters that have been determined using strong and weak gravitational lensing. A galaxy cluster is a collection of galaxies that are bound together by the force of gravity, while gravitational lensing is the bending of light by gravity. Strong lensing is the formation of arcs or rings of light surrounding clusters and weak lensing is a change in the apparent shapes of many galaxies. In this work we examine the properties of several samples of galaxy clusters using gravitational lensing. In Chapter 1 we introduce astrophysical theory of galaxy clusters and gravitational lensing. In Chapter 2 we examine evidence from our data that galaxy clusters are more concentrated than cosmology would predict. In Chapter 3 we investigate whether our assumptions about the number of galaxies in our clusters was valid by examining new data. In Chapter 4 we describe a determination of a relationship between mass and number of galaxies in a cluster at higher redshift than has been found before. In Chapter 5 we describe a model of the mass distribution in one of the ten lensing systems discovered by our group at Fermilab. Finally in Chapter 6 we summarize our conclusions.

NORTHERN ILLINOIS UNIVERSITY

DE KALB, ILLINOIS

AUGUST 2014

INVESTIGATIONS OF GALAXY CLUSTERS USING

GRAVITATIONAL LENSING

BY

MATTHEW P. WIESNER

A DISSERTATION SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS

Dissertation Director:
    Huan Lin and Michael Fortner

ACKNOWLEDGEMENTS

This dissertation is the first dissertation in astrophysics ever completed in the 119 year history of Northern Illinois University. I owe a great debt to many people who helped me to accomplish this.

I cannot overemphasize my debt to Dr. Huan Lin of Fermilab. He took me on as a student though he had no responsibility whatever to do so. He committed countless hours to training me, assisting me, teaching me astrophysics and reviewing my work. He went out of his way to see that I would receive proper training in how to take and analyze astronomical data and was always available when I needed help. Dr. Huan Lin made it possible for me to complete a doctorate in astrophysics, and I thank him most kindly for that.

Dr. David Hedin advised me and supported me constantly from the beginning to the end of my time at NIU. He helped me to connect with the Fermilab Center for Particle Astrophysics, he encouraged me to run the NIU Observatory and to teach Astronomy and he helped me find my postdoc. I am indebted to Dr. Hedin.

Dr. Michael Fortner was kind enough to take me on as a student and to advise me on how to progress to completion of my Ph.D. He was absolutely critical in my obtaining the Dissertation Completion Fellowship and in my publishing my first-author paper.

Dr. Tom Diehl has advised me and helped me throughout my time at Fermilab. He also made it possible for me to work as an observer for the Dark Energy Survey.

Dr. Marcelle Soares-Santos is an important collaborator and friend, and her cluster catalogs are central to some of my analyses.

DEDICATION

To my wife, Tiffany Nohelani, who has led me up the mountain. *Ad victoriam semper simul.*

To my father, who reminded me that the future is as filled with possibilities as the sky is filled with stars.

To my mother, who taught me the ways of science when I was still small.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

AN INTRODUCTION TO GALAXY CLUSTERS AND GRAVITATIONAL LENSING

Dimidium facti qui coepit habet: sapere aude, incipe. (He who has begun is half done: dare to know, begin.)–Horace

1.1   Introduction

Nobel prize winning physicist David Gross defines wisdom in this way [36]:

$$Wisdom = \frac{Knowledge}{Ignorance} \tag{1.1}$$

where ignorance is defined as awareness of a lack of knowledge. Our knowledge of the universe has increased tremendously over the last century, with groundbreaking discoveries throughout physics and astronomy. But as our knowledge of the universe has grown, so has our ignorance. For example as our understanding of galaxy clusters has increased, many new questions have arisen about their development and structure. As our knowledge of the structure of the universe has grown, we have become aware of dark matter and dark energy. By 2014 we have a rich understanding of the universe but it is accompanied by the knowledge that 95% of the universe is made of materials we do not really understand. Thus as our knowledge and our ignorance grow in physics, our quest for wisdom continues.

In this dissertation I attempt to increase the wisdom of humanity a little bit by studying clusters of galaxies and the phenomenon of gravitational lensing. Galaxy clusters hold information about the formation of structure in the universe as well as about the presence and abundance of dark matter. Gravitational lensing tells us both about the distribution of matter in massive objects like clusters and about the structure of source galaxies that are lensed. We will use clusters and lensing to infer how concentrated galaxy clusters are, to see how galaxies are distributed in clusters, to determine the mass distribution in one cluster and to find the overall mass of many clusters at greater distances than have been thus studied before.

In Chapter 1 we present the basic astrophysics behind galaxy clusters and gravitational lensing. In Chapter 2 we discuss the overconcentration problem, the idea that galaxy clusters may be more concentrated in their cores than had been expected. In Chapter 3 we discuss how the number of galaxies in a cluster changes with distance from the center. In Chapter 4 we present a new determination of the relation between the mass of galaxy clusters and the number of galaxies in the clusters. In Chapter 5 we focus on strong gravitational lensing observed in one of the clusters we studied, presenting a full lens and source model developed using a new Bayesian algorithm. Finally in Chapter 6 we summarize our conclusions. We begin with an overview of some important concepts in observational astronomy.

1.2   Concepts in Observational Astronomy

To begin, we summarize several important ideas in observational astronomy that will come up later in this thesis. First we discuss the concepts of magnitude, color and airmass.

Astronomical **magnitude** (symbolized as m) describes the brightness or luminosity of a celestial object. Apparent magnitude is defined as

$$m - m_{ref} = -2.5 \times log_{10} \frac{I}{I_{ref}} \tag{1.2}$$

I means intensity and the subscript *ref* means reference object, as magnitude is always defined with respect to some standard. Apparent magnitude describes flux received from an object. A related quantity called absolute magnitude describes the luminosity of an object by quantifying flux received at a particular distance from a source of radiation.

To an astronomer, **color** means the difference between magnitudes in two different wavelength bands, that is through two different filters. We frequently use *g-r* and *r-i* colors, although there are many other possible colors. The g, r and i filters are filters designed for the Sloan Digital Sky Survey (also including u at wavelengths shorter than g and z at wavelengths longer than i. Then considering Eq. 1.2, we find that g-r (where g and r are magnitudes measured in g and r filters) color is defined as

$$g - r = 2.5 \times log_{10} \left(\frac{I_r}{I_g} \frac{I_{g\_ref}}{I_{r\_ref}} \right) \tag{1.3}$$

Thus color is the logarithm of the ratio of fluxes in the two bands.

Airmass is the quantity used in astronomy to measure how high an object is in the sky. It is named such because it describes how much air light from a celestial object must pass through on the way to an observer. Airmass provides a limitation to astronomical observing, as images of an object at high airmass (very low in the sky) will be of very poor quality. The simplest definition of airmass is

$$X = \sec Z \tag{1.4}$$

In this equation, Z is the zenith angle of the object, that is the angle between the zenith (the point 90° above the horizon) and the object. An object at the zenith will have an airmass of 1 and airmass will increase as the object gets lower in the sky.

1.3   Cosmology

1.3.1   Redshift and Expansion

One of the most important discoveries in cosmology was Edwin Hubble's discovery of the expansion of the universe, leading to the conclusion that radial velocity of distant galaxies was directly related to galaxy redshift (Hubble's Law). Redshift is defined as

$$z = \frac{\lambda_{obs} - \lambda_{em}}{\lambda_{em}} \tag{1.5}$$

where λ_{obs} is the observed wavelength and λ_{em} is the emitted wavelength. Thus redshift measures the fractional change in wavelength of light emitted from some source. We are here concerned with the cosmological redshift due to the expansion of the universe, which can be defined as

$$z = \frac{1}{a} - 1 \tag{1.6}$$

where a is the dimensionless scale factor, which describes the rate of expansion of the universe. This is related to the scale factor $R(t)$ and the scale factor at time=0 R_0 by

$$a = \frac{R(t)}{R_0} \tag{1.7}$$

Most often the expansion of the universe is described by the Hubble parameter, which is related to the scale factor by

$$H = \frac{\dot{a}}{a} \qquad (1.8)$$

The Hubble parameter is usually stated as a function of redshift, as $H(z)$. The Hubble parameter at z=0 is called H_0. The general expression for $H(z)$ proceeds from Friedmann's equations (Equations 1.13 and 1.14), and is stated as

$$H(z) = H_0(1+z)[\Omega_{M,0}(1+z) + \Omega_{rad,0}(1+z^2) + \frac{\Omega_{\Lambda,0}}{(1+z^2)} + 1 - \Omega_0]^{1/2} \qquad (1.9)$$

In this equation, z is the redshift, Ω_M is the matter density parameter, Ω_Λ is the vacuum energy density parameter, Ω_{rel} is the relativistic particle (i.e., photon, neutrino) density parameter and Ω_0 is the total density parameter. The subscript 0 denotes the value at this time, as the values of these parameters have changed throughout time. We follow current cosmology to take $\Omega_{rad} = 0$ and $\Omega_0 = 1$, corresponding to a universe with zero curvature. In general, density parameters are represented as:

$$\Omega_x = \frac{\rho_x}{\rho_{crit}} \qquad (1.10)$$

where the subscript x denotes the quantity at hand (matter, vacuum energy, etc.).

Other important cosmological parameters include σ_8 and w. The quantity σ_8 is a special case of the quantity σ_R, which is the mean value of the fractional density perturbation over a sphere of co-moving radius R/a_0 [90], where a_0 is the scale factor at $z = 0$. Note that σ_8 is σ_R evaluated at $z = 0$ and $R = 8h^{-1}Mpc$. Thus σ_8 is one measure of the power spectrum at the current time, that is, a measure of clustering

and structure in the universe. The quantity w comes from the equation of state of the universe:

$$p = w \times \rho \tag{1.11}$$

where p is the overall pressure in the universe and ρ is the energy density of the universe. Thus w is the the ratio of p/ρ.

1.3.2 FRLW Metric and the Friedmann Equations

In General Relativity, the metric is a tensor that describes the fundamental properties of a spacetime, including the geometry, concepts of distance and time, and more. If a metric is to be consistent with GR, it must be able to provide a solution to Einstein's field equations:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + g_{\mu\nu}\Lambda = \frac{8\pi G}{c^4}T_{\mu\nu} \tag{1.12}$$

In Equation 1.12, $R_{\mu\nu}$ is the Ricci tensor, R is the Ricci scalar, $g_{\mu\nu}$ is the metric tensor, G is Newton's gravitational constant, Λ is the cosmological constant and $T_{\mu\nu}$ is the energy-momentum tensor. (The speed of light (c) is usually taken as equal to 1 to simplify.) There are a number of metrics that can provide solutions to Einstein's field equations; one that describes expanding spacetime is the Friedmann-Robertson-LeMaître-Walker (FRLW) metric. The FRLW metric describes a universe that is homogeneous and isotropic but changing in time [18]:

$$ds^2 = -dt^2 + a^2(t)\left[\frac{dr^2}{1 - \kappa r^2} + r^2 d\Omega^2\right] \tag{1.13}$$

where

$$\kappa = \frac{k}{R_0^2} \tag{1.14}$$

and k describes the curvature of the spacetime. Using Equation 1.13 in Equation 1.12, we can obtain the Friedmann equations, which describe the evolution of the universe. The first Friedmann equation is [18]

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{\kappa}{a^2} \tag{1.15}$$

where ρ describes the energy density of spacetime. The second Friedmann equation is

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) \tag{1.16}$$

where $\ddot{a}$ is the acceleration of expansion of the universe and p is the pressure of all components of spacetime.

1.3.3 Distance in Cosmology

Distance is an important concept in cosmology but the meaning of distance is sensitive to cosmology as distance measurements are done in an expanding universe with the possibility of curvature. We use two main definitions of cosmological distance, **luminosity distance** and **angular diameter distance**.

Luminosity distance is defined as [19]

$$d_L^2 \equiv \frac{L}{4\pi F} \tag{1.17}$$

where L is the luminosity of an object and F is the observed flux. Using the comoving coordinate to define the flux expected, d_L can be stated as

$$d_L(z) = \frac{c}{H_0}(1+z)I(z) \tag{1.18}$$

where

$$I(z) \equiv \int_0^z \frac{dz'}{\sqrt{\Omega_{M,0}(1+z')^3 + \Omega_{rel,0}(1+z')^4 + \Omega_{\Lambda,0} + (1-\Omega_0)(1+z')^2}} \tag{1.19}$$

Equation 1.18 is valid for $\Omega_0 = 1$.

Angular diameter distance is defined as

$$d_A \equiv \frac{D}{\theta} \tag{1.20}$$

where D is an observed object's linear diameter and θ is the object's angular diameter on the sky. It can be shown [19] that

$$d_A = \frac{d_L}{(1+z)^2} \tag{1.21}$$

1.4   Galaxy Clusters

1.4.1   Properties of Galaxy Clusters

Galaxy clusters are the largest gravitationally bound structures in the universe; the only larger structures are the clusters of clusters of galaxies that form the cosmic web of the universe. Some galaxy clusters contain ten or twenty galaxies and some

contain hundreds or even thousands of galaxies. A small galaxy cluster with fewer than 50 galaxies is often referred to as a group. The number of galaxies in a cluster is referred to as cluster **richness**. Galaxy clusters can have radii on the order of a megaparsec or more, where 1 Mpc is 3.26 million light years.

There are a number of galaxy clusters in our own cosmic neighborhood. Our own Milky Way galaxy is part of a small cluster of about 35 galaxies called the Local Group. The Virgo cluster is a relatively nearby cluster of spiral and elliptical galaxies that may contain up to 2000 galaxies. The Virgo cluster is part of the Virgo supercluster, a cluster of galaxy clusters of which the Local Group is a part. The Coma cluster [55] is a rich cluster in Coma Berenices that contains over 1000 galaxies.

A galaxy cluster is typically characterized by a collection of luminous elliptical galaxies surrounding one or several brightest cluster galaxies (BCGs). BCGs are typically giant elliptical (gE), D or cD galaxies (these are the largest types of galaxies known); they typically make up a significant fraction of the baryonic mass of the cluster. Galaxy clusters are believed to have begun to form early in the history of the universe and as such are often populated by early-type galaxies, that is ellipticals, of similar age and evolution (and thus color).

Elliptical galaxies are actually triaxial spheroidal systems. Elliptical galaxies are characterized by low rates of star formation and the lack of significant amounts of dust and gas which could be used in star formation. It is believed that as early galaxies were forming after the Big Bang the elliptical galaxies experienced very rapid star formation. This rapid star formation would have exerted pressure to prevent a collapse into a disk and would have used up much of the available star-forming material. Thus elliptical galaxies are observed to lack a significant disk component.
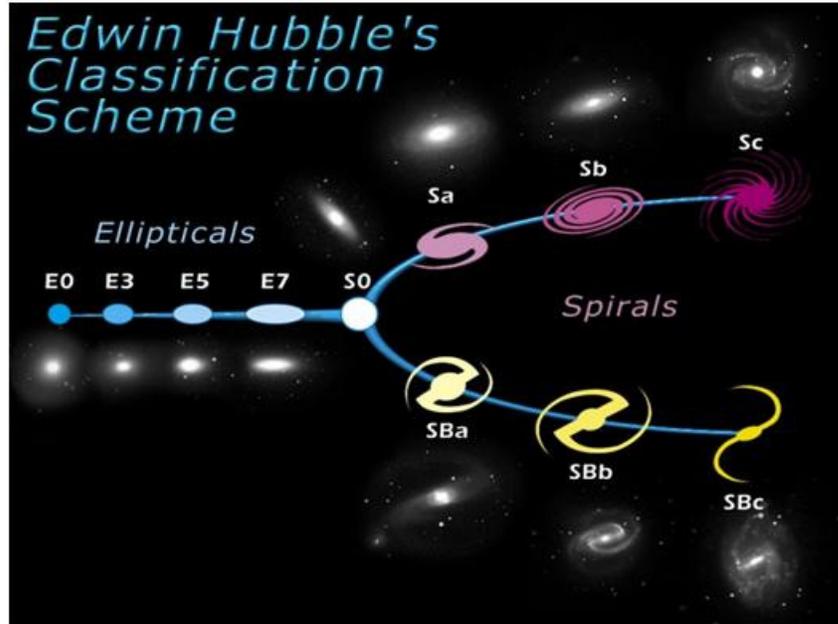
Figure 1.1: The Hubble sequence of galaxies, including ellipticals (on the left) and spirals (on the right). Lenticular (S0) galaxies form a transition type between elliptical and spiral galaxies. (NASA/STScI)

In the Hubble sequence (see Figure 1.1) elliptical galaxies are rated by their ellipticity:

$$e = 1 - \frac{b}{a} \tag{1.22}$$

Here b is the apparent minor axis of the elliptical profile and a is the apparent major axis. Elliptical galaxies are rated from E0 to E7 in the Hubble sequence, where E0 corresponds to $e = 0$ and E7 is $e = 0.7$. S0 galaxies share with ellipticals the lack of spiral arms but they have a disk like a spiral. S0 or lenticular galaxies are thus transitional types between ellipticals and spirals.

Galaxy clusters typically have masses from $10^{13} - 10^{15} h^{-1} M_\odot$, but the galaxies themselves make up a small portion of this total mass, typically about 1%. About 9% of the cluster mass comes from the intra-cluster medium (ICM), hot gas found in between cluster galaxies. But about 90% of the cluster mass is in the form of
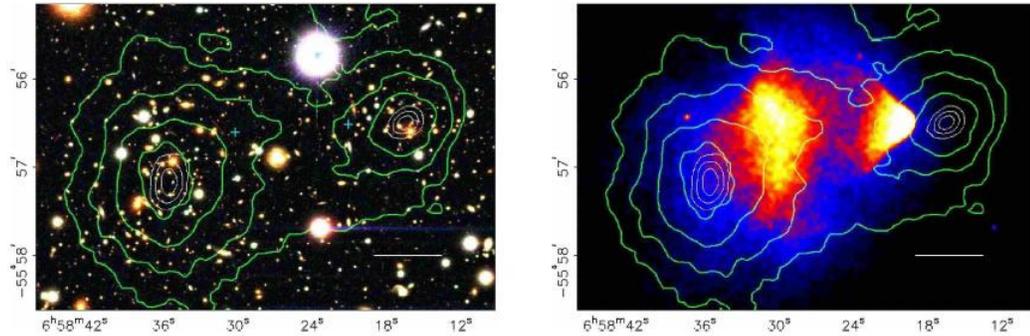
Figure 1.2: Images of the Bullet cluster from Clowe et al. (2006) [20]. The left image is an optical image from the Magellan telescope, while the right image is an x-ray image from the Chandra x-ray telescope. The contours indicate the weak lensing shear.

dark matter. The ICM can be mapped out using x-ray observations and the dark matter can be mapped out using weak lensing (see Section 1.7). The Bullet Cluster [20] is a well-studied system in which two galaxy clusters are interacting. Through a combination of x-ray observations and weak lensing, both the ICM and the dark matter can be represented (see Figure 1.2).

The galaxies and luminous gas are the most easily discernable evidence of a galaxy cluster, but these objects make up only a fraction of the overall cluster mass. As can be seen in Figure 1.2, the dark matter does not exactly follow the distribution of luminous matter. Since dark matter is weakly interacting, it does not collapse into galaxies and stars or radiate like baryonic matter. Thus dark matter distributions are fundamentally different than baryonic matter distributions. Some models of dark matter distributions have been made, including the NFW profile (see Section 1.7.2). The distribution of baryonic matter (e.g. galaxies) in a dark matter halo can be described by a Halo Occupation Distribution (HOD) function.

An HOD function describes three parameters in the halo model: (1) the number of galaxies per halo, (2) the spatial distribution of galaxies in a halo and (3) the

distribution of galaxy velocities [82]. The total number of galaxies in a halo as a function of mass is given as [82]:

$$\langle N \rangle_M = 1 + \left(\frac{M}{M_1} \right)^{\alpha} \tag{1.23}$$

where M is the halo mass. Spatial distribution of galaxies is given by the galaxy power spectrum. The monopole and quadrupole moments of the power spectrum are given by [82]:

$$P_0(k) = \left(1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2 \right) P_R(k) \tag{1.24}$$

and

$$P_2(k) = \left(\frac{4}{3}\beta + \frac{4}{7}\beta^2 \right) P_R(k) \tag{1.25}$$

where $P_R(k)$ is the real-space power spectrum and $\beta = \Omega_m^{0.6}/b_g$. The quantity b_g is the galaxy bias, which runs between $0.922 - 1.358$. Distribution of galaxies velocities is given by velocity dispersion, which is given as [82]

$$\sigma_v(M) = \alpha_v \left(\frac{GM}{2r_{200}} \right)^{1/2} \tag{1.26}$$

For $\alpha_v = 1$, this corresponds to the velocity dispersion for a singular isothermal sphere. The quantity r_{200} is the radius of a region of space within which the average density is 200 times the critical density of the universe; it is further discussed in § 2.3.2.2. The mass contained within r_{200} is called M_{200}; the number of galaxies within r_{200} is N_{200}; and the concentration of mass within r_{200} is c_{200}.

1.4.2   Galaxy Cluster Finding

Since the advent of CCD imaging, galaxy cluster-finding is done by using a computing algorithm based on some known properties of galaxy clusters. These properties might be the presence of a BCG and galaxies of similar color to that BCG[33, 9, 38, 13, 73]; it might be an overdensity in a local region of space as found in a Voronoi Tessellation [79]; it might be evidence of the Sunyaev-Zeldovich effect in the cosmic microwave background (see § 1.5.1) [94, 69, 40]. Once a cluster is found it is necessary to identify which galaxies are cluster members and which are simply along the line of sight and are really either in front of or behind the cluster. Such a determination can be done by measuring the redshift. Redshift is proportional to distance, and so if a galaxy is at similar redshift to the other galaxies in the cluster, then we can conclude it is a cluster member. Measurement of the redshift can be done using spectroscopy, measuring the spectrum of each of the galaxies believed to be in the cluster. However, spectroscopy is observationally expensive, requiring many hours of observing time to measure redshifts for all cluster galaxies.

Thus other methods are used to decide if galaxies are cluster members or not, including the red sequence method [33]. The red sequence method considers the fact that most galaxies in a cluster formed at about the same time and will be at similar evolutionary stages. Thus most galaxies in a cluster would be expected to have similar colors.

Galaxy cluster catalogs are described in terms of cluster *purity* and *completeness*. Purity describes what fraction of the objects in the catalog are actually clusters while completeness describes the fraction of existent clusters within defined redshift and magnitude limits that were found. In other words, purity tells us about how many

non-cluster interlopers exist in our cluster sample and completeness tells us how many clusters we missed.

1.5 Galaxy Cluster Cosmology

1.5.1 <u>Introduction to Cluster Cosmology</u>

Observational cosmology can be done using a number of probes, including Type Ia supernovae, baryon acoustic oscillations, cosmic shear, cosmic microwave background, galaxy clusters and more. Each probe has advantages and limitations associated with it. Galaxy clusters have the advantage that measuring their distribution in space and time tracks both the distribution of mass (related to Ω_M) and the rate of clustering in the universe (related to Ω_Λ). However one limitation to cluster cosmology is that to constrain cosmology with clusters it is necessary to understand the mass distribution of galaxy clusters. That is, we need to know the masses of the clusters we are studying. This mass distribution is called the **mass function** and describes the number of halos as a function of mass and redshift, $n(M, z)$. The general form of the mass function is [3]

$$\frac{dn}{dlnM} = \frac{\overline{\rho}_m}{M} \left[\frac{dln\sigma}{dlnM} \right] f(\sigma) \tag{1.27}$$

In Eq. 1.27, $\overline{\rho}_m = \Omega_m \rho_{crit}$ and $f(\sigma)$ is a model-dependent function of the filtered perturbation spectrum [3]. Press-Schechter theory gives a form of Eq. 1.27 as [26]:

$$\frac{dn(M, z)}{dM} = \sqrt{\frac{2}{\pi}} \frac{\rho_m \delta_c}{3M^2\sigma} e^{-\delta_c^2/2\sigma^2} \left[-\frac{R}{\sigma} \frac{d\sigma}{dR} \right] \tag{1.28}$$

Here δ_c is the critical overdensity above which objects collapse into gravitationally bound structures, σ is the RMS of the smoothed density field, and M is the total mass contained within a volume of space of radius R. Note that the resultant number density function will be a function of M, the cluster mass as well as redshift, z. In order to constrain the mass of galaxy clusters, it is necessary to find some way to determine cluster mass. There are four main ways to determine cluster mass: using x-ray temperatures, Sunyaev-Zeldovich distortion, optical masses and weak gravitational lensing [26].

The **x-ray temperature** method can be used by studying the hot intracluster medium (ICM) in a galaxy cluster. The temperature of the gas can be determined by finding the cutoff frequency of the x-ray radiation emitted. Then the temperature can be related to the mass of the cluster. Allen et al. (2011) give the formula for x-ray masses as

$$M(r) = -\frac{rkT}{G\mu m_p}\left[\frac{d\ln n_e}{d\ln r} + \frac{d\ln T}{d\ln r}\right] \tag{1.29}$$

where $M(r)$ is the mass within radius r, $T(r)$ is the temperature of the ICM, $n_e(r)$ is the electron density, k is Boltzmann's constant and μm_p is the mean molecular weight [3]. G is Newton's gravitational constant. **Sunyaev-Zeldovich distortion** tracks how photons from the CMB interact with the intracluster gas. CMB photons scatter off electrons in the hot gas through inverse Compton scattering. This causes a bias in frequency because the low-energy CMB photons gain energy from the hot electrons in the gas. The distortion is measured as a lack of photons at low energy and an excess of photons at higher energy. The amplitude of the distortion can be related to cluster mass. **Optical masses** can be found by counting galaxies using

imaging and finding galaxy velocity dispersion using spectroscopy. Optical mass can be stated as

$$M(r) = -\frac{r\sigma_r^2}{G} \left[\frac{d \ln \sigma_r^2}{d \ln r} + \frac{d \ln \nu}{d \ln r} + 2\beta \right] \qquad (1.30)$$

where $\nu(r)$ is the galaxy number density, $\sigma_r(r)$ is the 3-dimensional velocity dispersion and β is the velocity anisotropy parameter [3]. **Weak lensing** is a method of mass determination that is independent of gas temperature. Rather it depends on how the mass of the cluster causes path changes in photons traveling from more distant objects. Weak lensing will be discussed at greater length in § 1.7.

1.5.2 Examples of Cluster Cosmology

Rozo et al. (2010) [71] describe using the maxBCG cluster catalog found from the Sloan Digital Sky Survey to put constraints on the cosmological parameters Ω_M and σ_8 (see Figure 2.7). They do this by finding cluster abundances (the number of clusters in each richness and redshift bin) and by finding cluster masses using weak lensing [47]. The expected total number of clusters in the maxBCG catalog is given as:

$$\langle N \rangle = \int dM \ dz \ \frac{dn}{dM} \frac{dV}{dz} \langle \psi | M \rangle \langle \phi | z \rangle \qquad (1.31)$$

The function dn is given by an expression like Eq. 1.28 and dV/dz is the co-moving volume per unit redshift. $< \psi | M >$ is the probability that a halo of mass M falls within the richness bin defined by ψ [71]. $< \phi | z >$ is the probability that a halo at redshift z falls within the redshift bin defined by ϕ. Combining their results with Wilkinson Microwave Anisotropy Probe Year 5 cosmology they obtain $\sigma_8 = 0.807 \pm 0.020$ and $\Omega_M = 0.265 \pm 0.016$ (one-sigma errors).
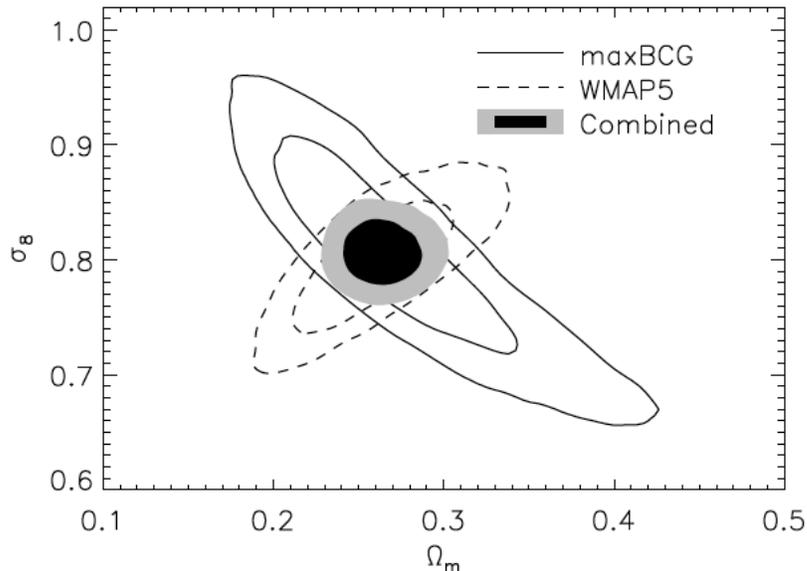
Figure 1.3: Figure 5 from Rozo et al. (2010) [71]. These are 1 and 2σ constraints on σ_8 and Ω_M from maxBCG clusters and WMAP Year 5 cosmology.

Vikhlinin et al. (2009) [85] present Chandra Space Telescope observations of 86 galaxy clusters found by ROSAT, 37 clusters at $< z >= 0.55$ and 49 clusters at $z \approx 0.05$. These observations allow improved determinations of cluster mass functions at low and high redshift. These data allow constraints on cosmological parameters, including w, Ω_M, and σ_8. They find $w_0 = -1.14 \pm 0.21$, $\Omega_M h = 0.184 \pm 0.035$ and $\sigma_8 = 0.803 \pm 0.0105$. In Figure 1.4 we show their Figure 3 which shows constraints on $\Omega_M h$ and σ_8 derived from their cluster analyses. In Figure 1.5 we show their Figure 2 which shows the derived cluster mass functions fit with theoretical curves based on different cosmologies. The left plot uses $\Omega_\Lambda = 0.75$ while the right plot uses $\Omega_\Lambda = 0$. It is evident that the left plot is a much better fit, suggesting that the standard cosmology of a non-zero Ω_Λ is correct. Vikhlinin et al. also show constraints on cosmological parameters related to dark energy, w_0 (w at the present time) and Ω_Λ at the present time, which they call Ω_X. In Figure
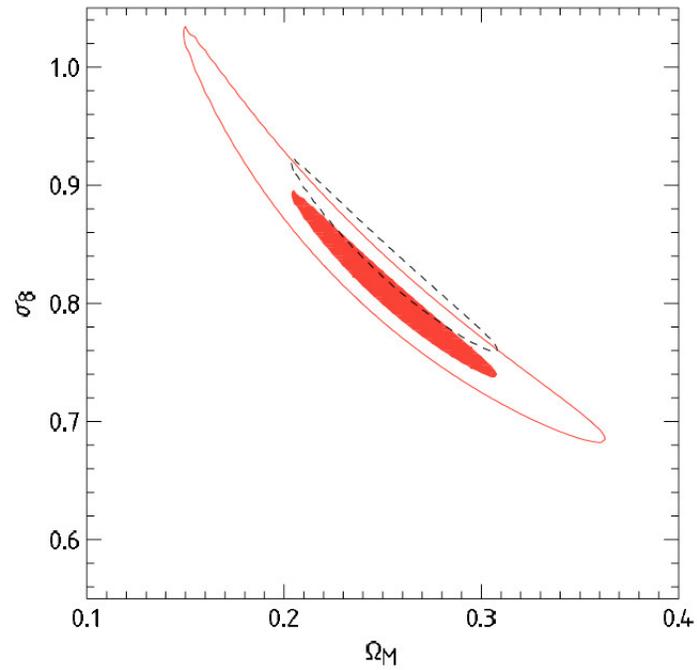
Figure 1.4: Constraints on the σ_8 and Ω_M parameters in a flat ΛCDM cosmology from the total (both low- and high-redshift) cluster sample from Vikhlinin et al. (2009) [85]. The inner shaded region represents 1σ and the outer contour is 2σ. The dashed contour shows how the 1σ region would be changed if the cluster mass is changed by the scale of the systematic errors ($\approx 9\%$).
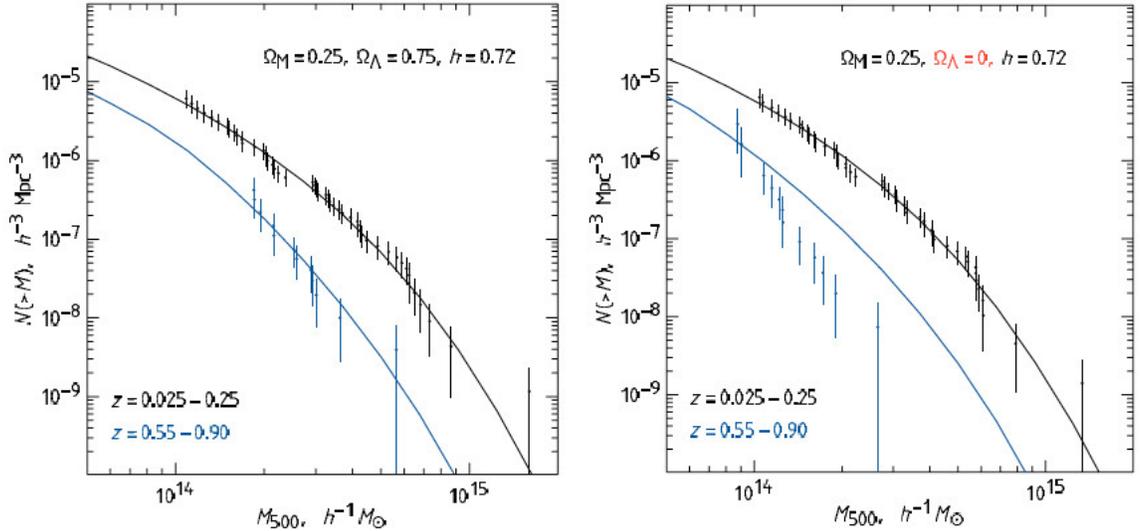
Figure 1.5: Cluster mass functions for different cosmologies from Vikhlinin et al. (2009) [85]. The left side shows the measured data along with the theoretical predictions for $\Omega_M = 0.25$ and $\Omega_\Lambda = 0.75$. The right side shows the same thing for $\Omega_M = 0.25$ and $\Omega_\Lambda = 0$.

1.6 we present their Figure 10, which shows constraints on these two parameters found from multiple cosmological methods. BAO is baryon acoustic oscillations, SN Ia is Type Ia supernovae and WMAP is the Wilkinson Microwave Anisotropy Probe space-based cosmic microwave background experiment. The inner red regions include their cluster cosmology. With all of these probes combined, they find $w_0 = -0.991 \pm 0.045 \pm 0.04_{systematic}$ and $\Omega_X = 0.740 \pm 0.012$.

Finally, Zu et al. (2014) [99] use weak gravitational lensing (see § 1.6) around maxBCG clusters (see § 2.3.1) to put constraints on σ_8 and Ω_M. Combining their results with WMAP7 cosmology they find $\Omega_M = 0.298 \pm 0.020$ and $\sigma_8 = 0.831 \pm 0.020$. In Figure 1.7 we present their Figure 9 in which they plot 1σ and 2σ constraints on these two parameters and also plot the combination of their analysis and WMAP7.
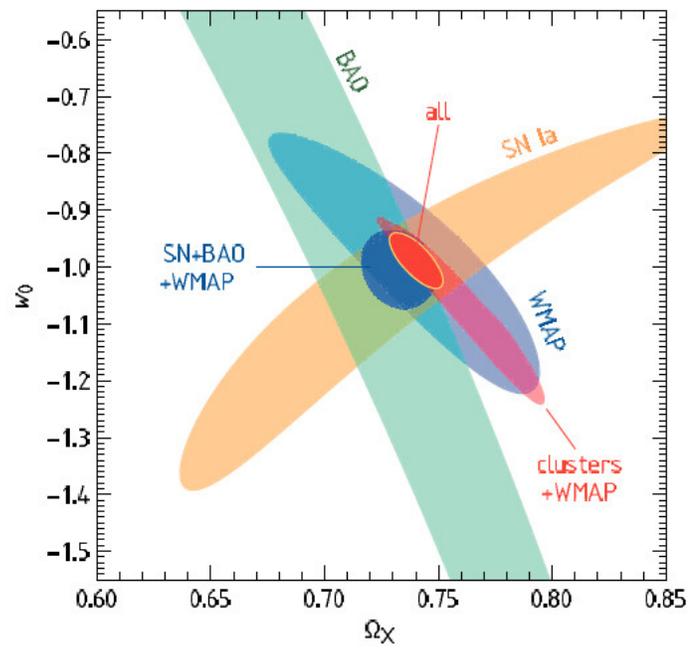
Figure 1.6: Constraints on dark energy parameters from supernovae, baryon acoustic oscillations, cosmic microwave background (WMAP) and galaxy clusters. This is Figure 10 from Vikhlinin et al. (2009) [85].
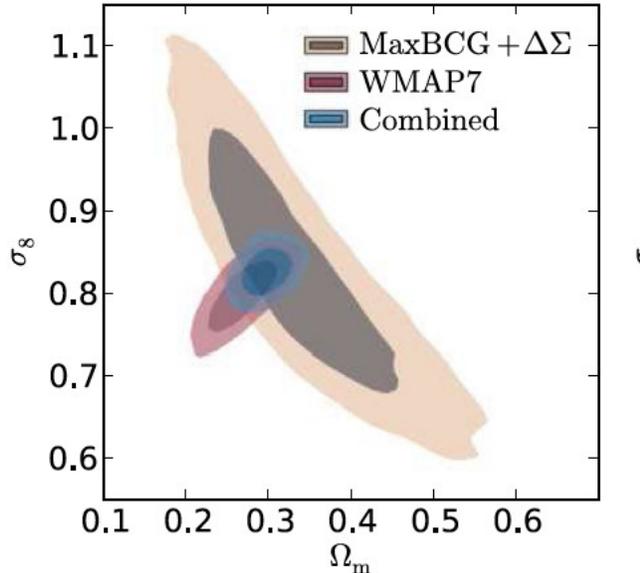
Figure 1.7: Constraints on σ_8 and Ω_M from Zu et al. (2014) [99]. The largest brown areas are 1 and 2σ constraints from maxBCG clusters and weak lensing. The smaller reddish areas are from WMAP7. The smallest blue areas are combined constraints.

To conclude, we see that cosmology can be well constrained by analyses of galaxy clusters combined with some method of determining cluster mass. In Chapter 4 we present our own relations that relate cluster mass to cluster richness, relations that will improve future cosmological measurements at higher redshift. These relations are based on measurements of mass using weak gravitational lensing, which is discussed in the next section.

1.6   Gravitational Lensing

Gravitational lensing is the deflection of light by mass. The theory of General Relativity predicts that all objects will follow geodesics as they travel through the universe; photons, traveling at the speed of light, follow null geodesics. These

geodesics are determined by the curvature of the universe in the path of the photons. Large concentrations of mass, like galaxy clusters, cause significant curvature, thus deflecting the photons significantly from their original paths. There are three major types of gravitational lensing: *strong lensing*, *weak lensing* and *microlensing*.

Strong lensing is a significant bending of light from a distant object, large enough that the resultant image can be tens of arcseconds across. (An arcsecond is $\frac{1}{3600}$ of a degree). This occurs when light from a distant object passes near to a massive object and then follows a straight line path to the observer. Strong lensing is what we observe when we see extended arcs in galaxy clusters or multiple images of quasars. When the distant object is an extended source (like a galaxy), we can see arcs or rings of light as the light from different parts of the source is lensed to different locations. When the distant object is a point source (like a quasar), we cannot see an arc or ring, but can only see multiple images. Strong lensing is only observed when the source galaxy (the one whose image is lensed) and the lensing galaxy (the one that bends the light) are aligned with the line of sight of an observer. In Figure 1.8 we include a Hubble Space Telescope image of galaxy cluster Abell 1689 showing multiple arcs produced by strong gravitational lensing.

Weak lensing is a more subtle bending of light from distant objects. (In Section 1.7 we will define the distinction between strong and weak lensing more rigorously.) Weak lensing cannot be observed directly in an image, but rather is a statistical effect which is only observed after analysis of the shapes of many galaxies. Thus weak lensing can be somewhat more challenging than strong lensing to observe, but weak lensing has the advantage that it is not dependent on a chance alignment of two objects, as strong lensing is. In Figure 1.9 we present a cartoon showing a field of background galaxies with no lensing (left) and with tangential shear (see § 1.7.3)
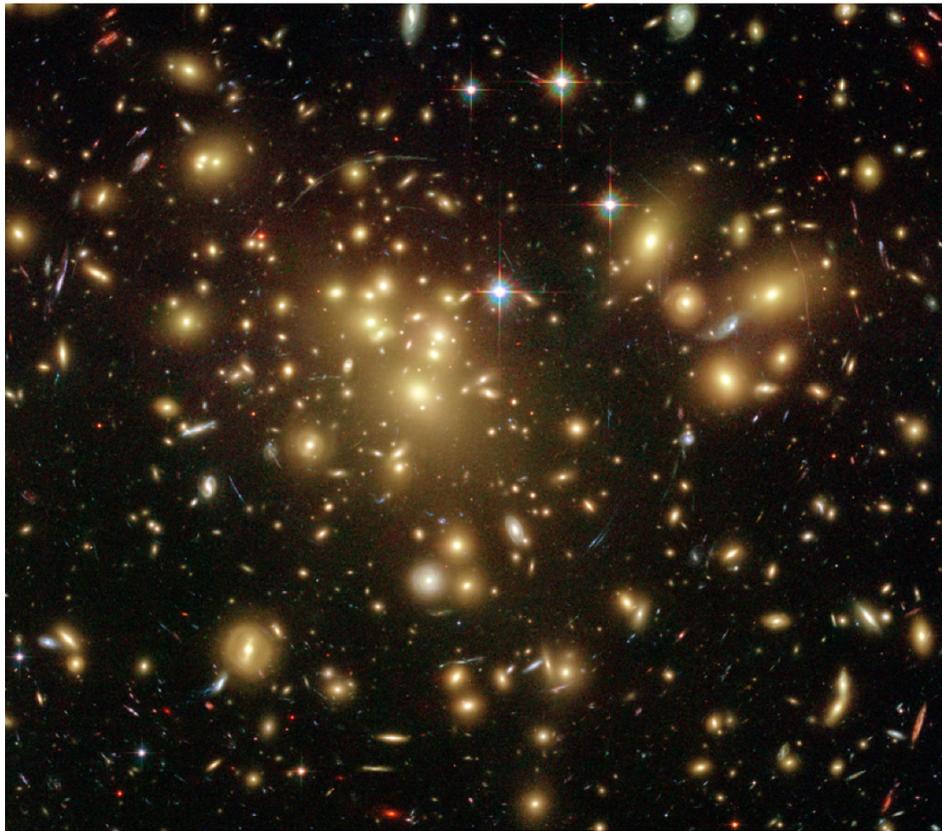
Figure 1.8: An image of galaxy cluster Abell 1689 showing arcs produced by strong gravitational lensing [61].
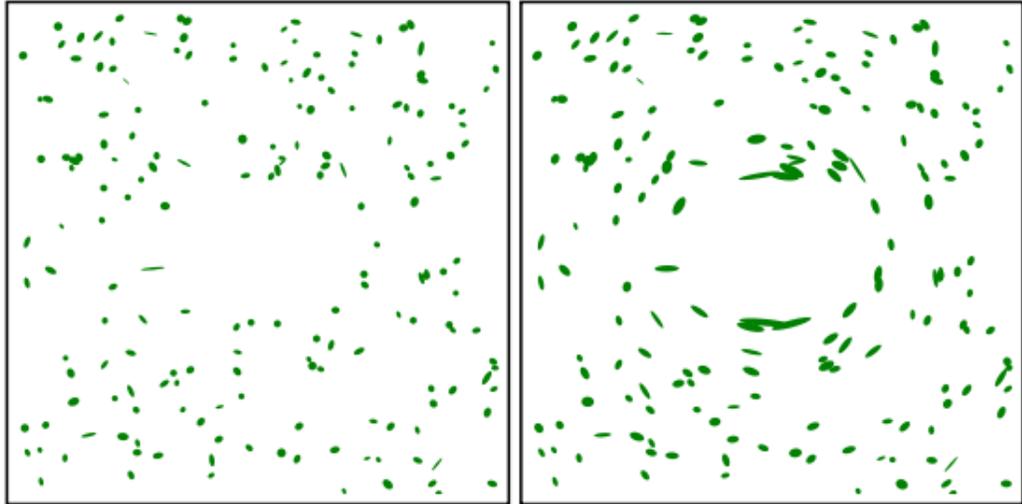
Figure 1.9: A cartoon showing a field of source galaxies with no weak lensing (left) and with tangential shear produced by weak lensing (right).

produced by weak lensing (right). Of course this would not be visible in an image, but rather has the effect shown in extreme form to demonstrate what is occurring.

Microlensing is the bending of light by a much less massive object, often by a single star. Because the lensing effect in microlensing is so small, on the order of microarcseconds, neither arcs nor changes in ellipticity are observed. Rather, microlensing is observed as a fluctuation in the luminosity of an object due to magnification from lensing. What is observed is a light curve that varies somewhat like a variable star or an eclipsing binary system, but with characteristics in the light curve that identify an event as microlensing. Microlensing has proven very useful in searches for extrasolar planets, and several exoplanet searches have used primarily this technique.
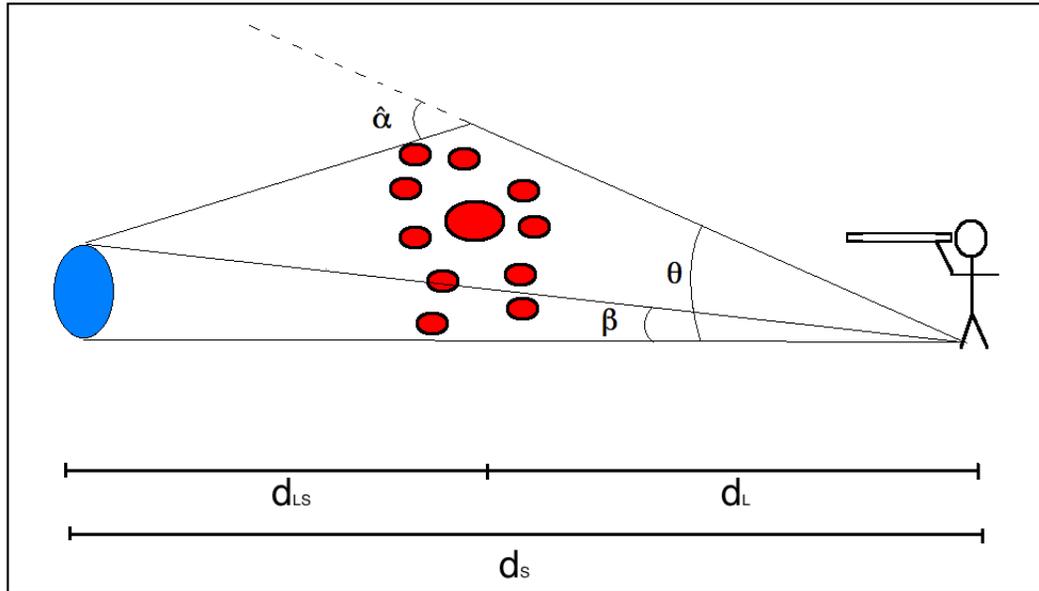
Figure 1.10: A sketch of a sample lensing system showing the basic parameters associated with lensing analysis.

1.7   Gravitational Lensing Theory

1.7.1   Lensing by a Point Mass

The simplest case in lensing theory is the deflection of light by a mass that is concentrated at a point. This is clearly an idealization, but for lensing by quasars or stars, this is an important approximation. Later in Section 1.7.2 we discuss lensing by an extended mass distribution. In Figure 1.10, we present the basic quantities associated with lensing, $\hat{\alpha}, \beta$ and θ. These are referred to respectively as actual deflection angle, angle between lens and source and angle between lens and image. The angular diameter distances in Figure 1.10 are d_L, the distance from the observer to the lens, d_S, the distance from the observer to the source galaxy and d_{LS}, the

distance from the lens to the source galaxy. The actual deflection angle is related to the reduced deflection angle by [19]

$$\hat{\alpha} = \frac{d_S}{d_{LS}}\alpha \tag{1.32}$$

Also the angular distance θ is related to the linear distance from the lens called the impact parameter b by

$$\theta = \frac{b}{d_L} \tag{1.33}$$

It may be noted from the figure that the various angular distances are related by a simple relation,

$$\alpha = \theta - \beta \tag{1.34}$$

This is called the lens equation. We can find the actual deflection angle by taking the gradient of the gravitational potential. The result for a point mass is

$$\hat{\alpha} = \frac{4GM}{c^2 b} \tag{1.35}$$

where G is Newton's gravitational constant, M is the mass of the lens, c is the speed of light and b is the impact parameter, the distance from the photon to the lens at closest approach. Combining this with the lens equation (Equation 1.34 we can say that

$$\beta = \theta - \frac{d_{LS}}{d_S d_L}\frac{4GM}{c^2\theta} \tag{1.36}$$

In the special case where the source and lens are aligned, then $\beta = 0$ and then θ has a special value, called the Einstein radius:

$$\theta_E = \sqrt{\frac{4GM}{c^2}\frac{d_{LS}}{d_L d_S}} \tag{1.37}$$

The Einstein radius is a measure of the size of a gravitational lens even in situations when the source and lens are not perfectly aligned. For strong lensing, the Einstein radius is tens of arcseconds, while with microlensing the Einstein radius is on the order of milliarcseconds, far too small to be directly observed. If we solve for the mass in Equation 1.37 we find that the mass contained within the Einstein radius of a gravitational lens is

$$M_E = \sqrt{\frac{c^2}{4G} \frac{d_L d_S}{d_{LS}} {\theta_E}^2} \tag{1.38}$$

Equation 1.34 can be rewritten in terms of θ_E as

$$\beta = \theta - \frac{\theta_E^2}{\theta} \tag{1.39}$$

This is a quadratic equation which has two solutions:

$$\theta_\pm = \frac{1}{2} \left(\beta \pm \sqrt{\beta^2 + 4\theta_E^2} \right) \tag{1.40}$$

Thus a point mass will almost always produce two images. As the source comes closer to the optic axis of the lens, the two images get closer to merging together into a single Einstein ring.

One of the properties of gravitational lenses is that they cause magnification, meaning that the image of a distant galaxy can appear brighter than would normally occur at a certain distance. Magnification arises because gravitational light deflection preserves surface brightness (by Liouville's theorem). The magnification is related to a change in apparent size of the source galaxy by [62]

$$magnification = \frac{image\ area}{source\ area} \tag{1.41}$$

For a circularly symmetric lens magnification μ is given by [62]:

$$\mu = \frac{\theta}{\beta}\frac{d\theta}{d\beta}$$

(1.42)

which for a point mass becomes

$$\mu_{\pm} = \left[1 - \left(\frac{\theta_E}{\theta_{\pm}}\right)^4\right]^{-1}$$

(1.43)

1.7.2   Lensing by an Extended Mass Distribution

For extended mass distributions like galaxies and galaxy clusters we need to consider not just the amount of mass M but also its distribution in space. The gravitational lensing observed will depend on the mass present in the lensing object and its distribution. Since most of the mass in a galaxy or galaxy cluster is contained in dark matter, these are dark matter mass profiles. Following classical mechanics, one of the most important quantities to specify for a mass distribution is the mass density $\rho(r)$. Most directly applicable to lensing is the related quantity $\Sigma(r)$ which is related to $\rho(r)$ by

$$\Sigma(r) = \int \rho(\vec{\xi}, z)dz$$

(1.44)

where $\vec{\xi}$ is a two dimensional vector in the lens plane [62] and z is distance perpendicular to the lens plane. Σ_{crit} is the critical surface mass density, which is

$$\Sigma_{crit} = \frac{c^2 d_S}{4\pi G d_L d_{LS}}$$

(1.45)

When $\Sigma = \Sigma_{crit}$, then $\beta = 0$ and $\theta = \alpha$ so we see an Einstein ring.

	Einstein Cross	Cusp Caustic	Fold Caustic
Source Plane			
Image Plane			
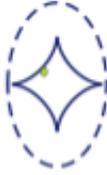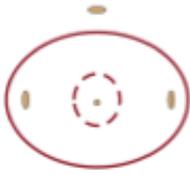
Figure 1.11: A demonstration of the lensing arcs that would be produced by any of several source galaxy positions with respect to the caustics [35].

A cartoon of lensing by an extended mass distribution is shown in Figure 1.11 [35]. This figure shows a distant source galaxy in the source plane in one of several positions. It also shows a less distant elliptical mass distribution in the image plane with one of several corresponding strong lensing arcs. Thus in this figure, the top row represents the source plane (where the more distant source galaxy is) and the bottom row represents the image plane (where the less distant galaxy cluster is). The elliptical structures in the image plane are called **critical curves** and the structures in the source plane are called **caustics**. The critical curves are the set of points where the magnification of an image would approach infinity. These curves mapped back into the source plane are the caustics. For an elliptical mass distribution the central caustic appears as this astroid shape. When an object crosses the caustics, the number of images changes.

We will utilize several mass distributions:

1. The singular isothermal sphere (SIS)

2. The singular isothermal ellipsoid (SIE)

3. The NFW profile

4. The pseudo-Jaffe profile

1.7.2.1   <u>SIS</u>

Binney and Tremaine [12] give an equation for the mass distribution of a galaxy, assuming the galaxy can be treated as a spherically symmetric system of stars behaving like an ideal gas at equilibrium [53]. This is their equation 4-115b:

$$\frac{d}{dr}\left(r^2\frac{d\ln\rho}{dr}\right) = -\frac{Gm}{k_BT}4\pi r^2\rho \tag{1.46}$$

where r is distance from the center, ρ is the density, m is the mass per particle, k_B is Boltzmann's constant and T is the temperature. The simplest solution to this equation is the singular isothermal sphere, which is a mass distribution defined as [53]:

$$\rho(r) = \frac{\sigma_v^2}{2\pi Gr^2} \tag{1.47}$$

where σ_v is the velocity dispersion in the lensing region. *Singular* means that this expression goes to infinity at $r = 0$ and *isothermal* means that the the system is at dynamic equilibrium. If we express the three-dimensional distribution instead as a

surface mass density, we obtain this equation for surface mass density as a function of position in an SIS:

$$\Sigma(r) = \frac{\sigma_v^2}{2G}\frac{1}{r} \tag{1.48}$$

1.7.2.2   SIE

A slightly more complex model takes account of the fact that the central lensing galaxy is often elliptical. We can obtain an expression for the surface mass density of a singular isothermal ellipsoid by replacing the term r with the square root of the product of the major (a) and minor (b) axes $\sqrt{ab}$. We then obtain [53]:

$$\Sigma(r) = \frac{\sqrt{\frac{a}{b}}\sigma_v^2}{2G}\frac{1}{a} \tag{1.49}$$

1.7.2.3   NFW

The Navarro-Frenk-White mass profile [63] was developed using N-body simulations of dissipationless dark matter particles. The NFW profile has proven to be a very good fit to the matter distribution in a dark matter halo. The NFW profile is:

$$\rho(r) = \frac{\delta_c \rho_{crit}}{(r/r_s)(1 + r/r_s)^2} \tag{1.50}$$

In this equation, ρ_{crit} is the critical density of the universe, δ_c is the characteristic overdensity of the halo, r is the distance from the center of the halo and r_s is the scale radius. The scale radius is the radius at which the decrease in density as a

function of radius changes from a $1/r$ relation to a $1/r^3$ relation. The critical density, ρ_{crit}, is defined as

$$\rho_{crit}(z) = \frac{3H(z)^2}{8\pi G} \tag{1.51}$$

where $H(z)$ is the Hubble parameter (see Equation 1.1) and G is Newton's gravitational constant. The characteristic overdensity of the halo is

$$\delta_c = \frac{200}{3} \frac{c^3}{\ln(1+c) - c/(1+c)} \tag{1.52}$$

In this equation, c is the concentration parameter, which is a measure of dark matter halo density in the inner regions of the cluster. Finally, scale radius is defined as $r_s = r_\Delta/c$ where r_Δ is the virial radius, the radius of a spherical region of space within which the density of the universe is $\Delta\rho_{crit}$.

What value is used for Δ varies among authors. In this work we will use $\Delta = 200$, so r_{200} is the radius of a region of space within which the average density is $200\rho_{crit}$. Then the total mass contained within this region is called M_{200} and the total number of galaxies in this region is N_{200}.

Finally the surface mass density for an NFW profile can be found by solving Eq. 1.44. Wright and Brainerd (2000) [95] present the solution as

$$\Sigma_{NFW}(x) = \begin{cases} \frac{2r_s\delta_c\rho_{crit}}{(x^2-1)} \left[1 - \frac{2}{\sqrt{1-x^2}} \operatorname{arctanh} \sqrt{\frac{1-x}{1+x}}\right] & x < 1 \\ \frac{2r_s\delta_c\rho_{crit}}{3} & x = 1 \\ \frac{2r_s\delta_c\rho_{crit}}{(x^2-1} \left[1 - \frac{2}{\sqrt{x^2-1}} \arctan \sqrt{\frac{x-1}{1+x}}\right] & x > 1 \end{cases}$$

where $x = r/r_s$.

1.7.2.4 Pseudo-Jaffe

The Jaffe profile (proposed by Walter Jaffe) is [45]

$$\rho(r) = C\frac{1}{r^2}\left(\frac{1}{a+r}\right)^2 \tag{1.54}$$

where C is a constant and a is the break radius, which describes how quickly mass density decreases with radius. The Jaffe profile was proposed because it can reproduce the de Vaucouleurs profile, which is a description of how luminosity changes as a function of radius in many galaxies. The de Vaucouleurs profile is (see also § 5.2)

$$I(r) = I_e \exp\left[-7.669\left[\left(\frac{r}{r_e}\right)^{1/4} - 1\right]\right] \tag{1.55}$$

where I is the observed light intensity, r is radius, r_e is the radius of an isophote containing half the light and I_e is the intensity at r_e. The pseudo-Jaffe profile is produced by adding a core radius, s [49]:

$$\rho(r) = C\frac{1}{(r^2 + s^2)}\frac{1}{(r^2 + a^2)} \tag{1.56}$$

The surface mass density of the pseudo-Jaffe profile is [49]

$$\Sigma(\xi) = \frac{\theta_E}{2}\left[\frac{1}{\sqrt{s^2 + \xi^2}} - \frac{1}{\sqrt{a^2 + \xi^2}}\right] \tag{1.57}$$

where $\xi^2 = x^2 + \frac{y^2}{q^2}$, x and y are image plane coordinates and q is the projected axis ratio.

1.7.3   Convergence and Shear

The image of the source object is changed by its interaction with the lensing mass. How the image changes is described by the parameters called *convergence* and *shear*. Lensing is considered strong lensing when $\kappa \approx 1$. When *kappa* and γ are $\ll 1$ then the lensing is considered weak lensing. Conceptually, convergence affects the size of the image but not the shape. Conversely, shear affects the shape of the image but not the size (see Figure 1.12). Typically an observed image has undergone a combination of convergence and shear. Note that convergence cannot be directly measured in weak lensing since galaxies may come in any range of sizes, so a change in galaxy size is not detectable. However shear can be measured by looking for a large number of galaxies that tend to a particular distortion of shape [41].

Convergence, κ, is the ratio of the surface mass density at a point to critical surface mass density:

$$\kappa(\theta) = \frac{\Sigma(\theta)}{\Sigma_{crit}} \tag{1.58}$$

where θ describes position in the lens plane. Convergence can also be defined as [95]

$$\kappa(\theta) = \frac{1}{2}\left(\frac{\partial^2 \psi}{\partial \theta_1^2} + \frac{\partial^2 \Psi}{\partial \theta_2^2}\right) \tag{1.59}$$

where θ_1 and θ_2 are lens plane coordinates (effectively x and y in the lens plane). In Equation 1.59 ψ is the lensing potential, defined as

$$\psi(\vec{\theta}) = \frac{d_{LS}}{d_L d_S}\frac{2}{c^2}\int \Phi(d_L\vec{\theta}, z)dz) \tag{1.60}$$
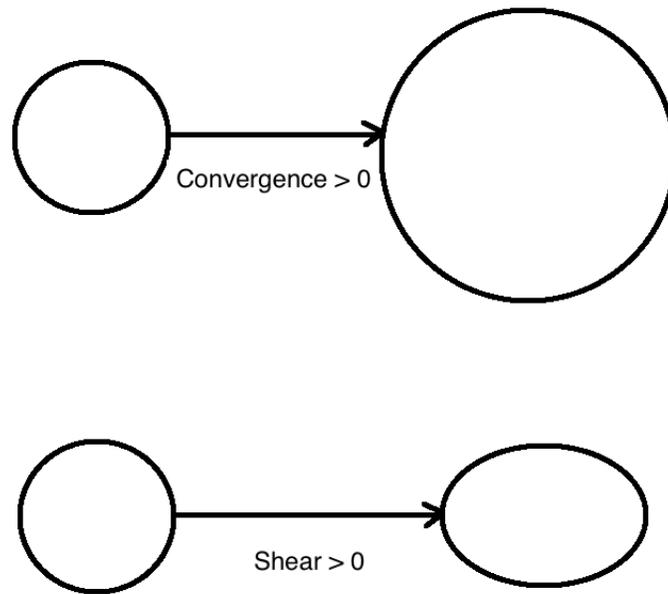
Figure 1.12: A comparison of the effects of convergence and shear on an image of an imagined spherical galaxy.

where c is the speed of light and Φ is the gravitational potential. Φ is a function of the mass distribution, and will depend on what mass model is used.

Shear is defined as a complex number:

$$\gamma = \gamma_1 + i\gamma_2 \qquad (1.61)$$

As usual for complex numbers, the magnitude of the shear is given by

$$\gamma = \sqrt{\gamma_1{}^2 + \gamma_2{}^2} \qquad (1.62)$$

and the two components of shear can be given in terms of the magnitude and a phase angle ϕ:

$$\gamma_1 = \gamma cos(2\phi)$$
$$\gamma_2 = \gamma sin(2\phi)$$

(1.63)

In terms of the lensing potential, the components of shear are equal to [60]

$$\gamma_1 = \frac{1}{2} \left(\frac{\delta^2 \Psi}{\delta \theta_1^2} - \frac{\delta^2 \Psi}{\delta \theta_2^2} \right)$$
$$\gamma_2 = \frac{\delta^2 \Psi}{\delta \theta_1 \delta \theta_2}$$

(1.64)

We seek to measure a change in shape of galaxies between the source (β) plane and the lens or image (θ) plane [41]. The mapping between source plane and the image plane is described by the lensing Jacobian. The lensing Jacobian is valid for all forms of lensing, but in weak lensing the Jacobian has the property that it is almost equal to the unit matrix, corresponding to very small Einstein radii. The Jacobian can be written as

$$\mathcal{A} = \frac{\delta \beta}{\delta \theta} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}$$

(1.65)

Thus a source galaxy that is circular and of unit radius will, after lensing involving convergence and shear, become an ellipse with major (a) and minor (b) axes [62]:

$$a = (1 - \kappa - \gamma)^{-1}$$
$$b = (1 - \kappa + \gamma)^{-1}$$

(1.66)

The inverse mapping from the source plane to the image plane is given by the magnification tensor, $\mathcal{M}$, which is the inverse of $\mathcal{A}$. The scalar magnification μ

describes the change in brightness of the galaxy in the lens plane. The magnification of the source galaxy is given by

$$\mu = det\mathcal{M} = \frac{1}{det\mathcal{A}} = \frac{1}{[(1-\kappa)^2 - \gamma^2]} \tag{1.67}$$

If the coordinates of the source galaxy position are (x_s, y_s) and the coordinates of the image position are (x, y), then the Jacobian can be written as

$$A = \begin{pmatrix} \frac{\delta x_s}{\delta x} & \frac{\delta x_s}{\delta y} \\[2mm] \frac{\delta y_s}{\delta x} & \frac{\delta y_s}{\delta y} \end{pmatrix} \tag{1.68}$$

The actual shape of a galaxy in an image is quantified by the ellipticity, e. There are several equivalent definitions of ellipticity that are used in lensing. They are [50]

$$e = \frac{1 - (b/a)^2}{1 + (b/a)^2} \tag{1.69}$$

or

$$e = \frac{1 - (b/a)}{1 + b/a} \tag{1.70}$$

where b is the minor axis and a is the major axis. We utilize the first definition in our analyses. Ellipticity has two components, e_1 and e_2, which are added in quadrature to give e. The components of ellipticity can be related to the second order (or quadrupole) moments of an image. First order moments describe the centroid of an object, while second order moments describe shape. First order moments can be given as [41]

$$\begin{aligned} x &= \int I(x,y)x dx dy \\ \overline{y} &= \int I(x,y)y dx dy \end{aligned} \tag{1.71}$$

for observed flux $I(x, y)$ and positions x and y. Quadrupole moments are defined as [26]

$$q_{ij} \equiv d^2\theta I_{obs}(\theta)\theta_i\theta_j \tag{1.72}$$

where I_{obs} is the observed flux. For dimensions x and y we obtain

$$q_{xx} = \frac{\int x^2 w(x, y)I(x, y)dxdy}{\int w(x, y)I(x, y)dxdy} \tag{1.73}$$

$$q_{yy} = \frac{\int y^2 w(x, y)I(x, y)dxdy}{\int w(x, y)I(x, y)dxdy} \tag{1.74}$$

$$q_{xy} = \frac{\int xy w(x, y)I(x, y)dxdy}{\int w(x, y)I(x, y)dxdy} \tag{1.75}$$

In this equation $I(x, y)$ is the number of pixel counts (related to object flux) as a function of position. The quantity $w(x, y)$ is a Gaussian weighting function [17]. Then e_1 and e_2 are defined as

$$\begin{aligned} e_1 &= \frac{q_{xx}-q_{yy}}{q_{xx}+q_{yy}} \\ e_2 &= \frac{2q_{xy}}{q_{xx}+q_{yy}} \end{aligned} \tag{1.76}$$

Note that a circular image has $q_{xx} = q_{yy}$ and $q_{xy} = 0$ [26]. Finally, we introduce tangential (e_T) and cross (e_X) ellipticity components, which are defined as

$$\begin{aligned} e_T &= e_1\cos(2\phi) - e_2\sin(2\phi) \\ e_X &= e_1\sin(2\phi) + e_2\cos(2\phi) \end{aligned} \tag{1.77}$$

The angle ϕ is the position angle, which is defined west of north. Mathematically ϕ is defined by

$$\phi = \arctan\left[\frac{-(RA_{gal} - RA_{BCG}) \times \cos(DEC_{BCG})}{DEC_{gal} - DEC_{BCG}}\right] \tag{1.78}$$

RA_{gal} and DEC_{gal} are the coordinates of one of the many source galaxies and RA_{BCG} and DEC_{BCG} are the coordinates of the BCG, taken as the center of the cluster. Ellipticity is related to shear by a factor R called *responsivity* by

$$\gamma = e/R \qquad (1.79)$$

Responsivity is a multiplicative factor that describes how a particular observing site and system will produce changes in ellipticity in response to a particular shear; it is one method of estimating the effect of the PSF on the shear (see § 1.7.4). So finally we have tangential and cross (or orthotangential) shear:

$$\gamma_T = \frac{e_1 \cos(2\phi) - e_2 \sin(2\phi)}{R} \qquad (1.80)$$

$$\gamma_X = \frac{e_1 \sin(2\phi) + e_2 \cos(2\phi)}{R} \qquad (1.81)$$

Tangential shear is produced by gravitational lensing. Orthotangential shear is not. Thus cross shear can be used as a check on measurements as if a lensing signal is observed in cross shear, this is evidence of an error. This will be discussed at greater length in Chapter 4.

1.7.4   Decoupling Shear from Systematics

Finding the actual change of shape of a number of source galaxies is not as simple as measuring the shapes and calculating shear. There are a number of systematics that must be considered before shapes can be measured. In Figure 1.13 Heymanns (2002) shows what these systematics are. For a galaxy, (1) long after light leaves the galaxy, the light has its path changed by the gravitational field of a massive object

Galaxies: Intrinsic galaxy shapes to measured image:

Intrinsic galaxy
(shape unknown)

Gravitational lensing
causes a *shear (g)*

Atmosphere and telescope
cause a convolution

Detectors measure
a pixelated image

Image also
contains noise
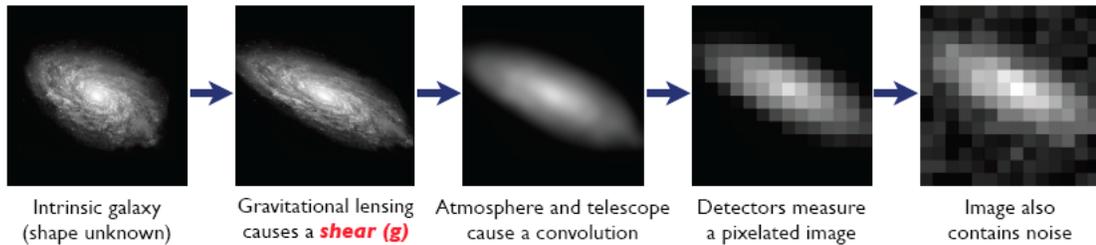
Stars: Point sources to star images:

Intrinsic star
(point source)

Atmosphere and telescope
cause a convolution

Detectors measure
a pixelated image

Image also
contains noise
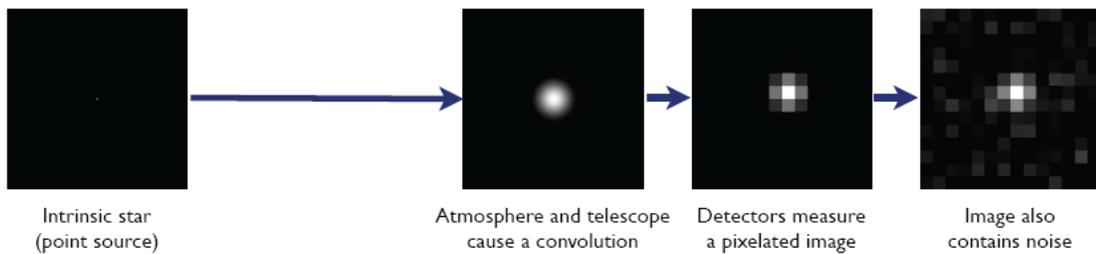
Figure 1.13: A graphic from Heymanns (2002) [41] showing the effects on galaxy shape.

(weak lensing shear). Then (2) the light is distorted by a point spread function (PSF) caused by the atmosphere and the telescope. Then (3) the detector causes pixelization of the image. Finally (4) there is also noise in the image, and so the signal to noise (S/N) ratio must be considered.

For stars, similar systematics occur. However it is somewhat simpler because stars are point sources. As shown the bottom left of Figure 1.13 a perfectly resolved star is a point of light. However once the point source is convolved with the PSF of the atmosphere and telescope, it is a disk of finite size, usually described by a Moffat profile (see Equation 5.4). Then like before the image is pixelated and there is some amount of noise.

Overall, the steps from taking data at a telescope to measuring mass of a set of clusters using weak lensing are [41]:

1. Take data

2. Reduce data

3. Detect objects in the data

4. Separate stars and galaxies

5. Model the PSF using the stars

6. Remove the effect of the PSF on the galaxy image

7. Measure the shear from the PSF corrected image

8. Fit the shear profile to a mass model

Data reduction involves a number of steps, including overscan subtraction, bias subtraction, flat-fielding, bad pixel masking and more. This is discussed in more detail in § 3.3. Object detection can be done using several routines; one of the most powerful is called `SourceExtractor` or `SExtractor`. This routine is described in § 3.3.2. Star-galaxy separation is done by identifying some feature that distinguishes stars from galaxies. In § 2.3.1 we describe doing this separation by finding magnitudes measured in different apertures. This can also be done by using the `SExtractor` parameters *CLASS_STAR* or *SPREAD_MODEL*. Modeling the PSF can be done by taking bright stars (we used $18 \leq i \leq 20$) and finding their second order moments. Removing the effect of the PSF involves several steps. One method is called the Kaiser-Squires-Broadhurst (KSB) method; it deconvolves the PSF from the shear by subtracting the PSF from the measured moments in Fourier space. KSB uses the assumption that the PSF can be described by a small but highly anisotropic distortion convolved with a large circularly symmetric seeing disk [41, 48]. The Hirata-Seljak method [43] also removes the PSF from the shear, but takes account

of the fact that the PSF is not Gaussian. Hirata-Seljak also corrects for shear se-
lection bias. After the effect of the PSF is removed, the shear is measured for each
source galaxy-cluster pair using Equation 1.80 and then the shear is averaged for
a particular radius bin. Finally the shear signal is fit to a mass model (we use an
NFW profile) by minimizing the χ^2 [17]:

$$\chi^2 = \sum_{i=1}^{N} \frac{[\gamma_i - \gamma_{NFW}(r_i; M_{200}; c_{200})]^2}{\sigma_\gamma^2} \tag{1.82}$$

Here γ_i is the shear measured in a particular radius bin (e.g. between radius of 0.1
and 0.2 megaparsecs from the BCG) and γ_{NFW} is the shear predicted by the NFW
model. The parameter σ_γ^2 is the standard deviation of the galaxy shears in that bin.
This will be described in more detail in Chapter 4. But first we consider strong
lensing observed in 10 galaxy clusters.

CHAPTER 2

A CONFIRMATION OF THE OVERCONCENTRATION

PROBLEM

Philosophy is written in this grand book the universe, which stands continually open to our gaze.–Galileo Galilei

2.1   Introduction

In this chapter, we describe studies of ten galaxy clusters that exhibit strong gravitational lensing. All of these systems were found in searches conducted by the Sloan Bright Arcs Survey (SBAS). These clusters range in redshift from $0.2595 - 0.5580$ and the gravitationally lensed source galaxies range in range from $0.6596 - 2.9437$.

In § 2.2 we describe how the clusters were found and how we took follow-up data on them. In § 2.3 we describe how we measured the properties of the clusters, including measurements of richness, mass, velocity dispersion and concentration. In § 2.4 we describe measurements of the properties of the strong lensing systems, including Einstein radius, lens mass and lens velocity dispersion. Finally in § 2.5 we describe analysis of these measurements including a determination that some of our clusters are more concentrated than would be expected.

Table 2.1: The coordinates and redshifts of the ten systems in this paper. Note that source redshift has not been determined for SDSS J1439+3250 so we present a range of possible source redshifts.

System	R.A. (deg)	Decl. (deg)	Lens z	Source z
SDSS J0900+2234	135.01128	22.567767	0.4890	2.0325
SDSS J0901+1814	135.34312	18.242326	0.3459	2.2558
SDSS J0957+0509	149.41318	5.1589174	0.4469	1.8230
SDSS J1038+4849	159.67974	48.821613	0.4256	0.966
SDSS J1209+2640	182.34866	26.679633	0.5580	1.018
SDSS J1318+3942	199.54798	39.707469	0.4751	2.9437
SDSS J1343+4155	205.88702	41.917659	0.4135	2.0927
SDSS J1439+3250	219.98542	32.840162	0.4176	1.0-2.5
SDSS J1511+4713	227.82802	47.227949	0.4517	0.985
SDSS J1537+6556	234.30478	65.939313	0.2595	0.6596

2.2   The Sample of Strong Lenses

The redshifts of all ten of the clusters are given in Table 2.1 along with their celestial coordinates (right ascension and declination). All of these systems are among the 19 strong lenses so far found and confirmed by the SBAS. The SBAS is a collaboration of scientists at Fermilab that has searched the data from the Sloan Digital Sky Survey (SDSS) to find strong lensing systems [2, 58, 25, 55, 56, 91]. The SDSS is an endeavour, begun in 2000, to map more than 25% of the sky and to obtain spectra for more than one million objects. The SBAS searched the SDSS data using two search algorithms, searching for: (1) blue objects that were within 10″ of luminous red galaxies (LRGs) [54] and (2) interacting galaxies [1, 2]. Blue objects surrounding red galaxies could be lensing systems as the more distant galaxies are often galaxies with high rates of star formation, making them appear blue. Galaxy clusters, often associated with strong lensing, are typically populated

by red elliptical galaxies. Objects classified as interacting galaxies can actually be a lensing arc surrounding a lensing galaxy, thus why the second method searched for these objects. Four of our systems were found using the first method and six of them were found using the second method.

2.2.1   Details on the Data

The SBAS took new images of the ten lensing systems at Kitt Peak National Observatory, using the 3.5-m Wisconsin-Indiana-Yale-NOAO (WIYN) telescope. These observations were conducted on February 26 and 27, 2009. Images were taken using the Mini-Mosaic camera, which uses two CCDs, each of dimensions 2048×4096 pixels. Images were taken using three filters, SDSS g, r and i filters. A collage of color images of sections of these images showing the strong lenses is provided in Figure 2.1.

The group took new images in order to obtain data with improved seeing, finer pixel scale and fainter magnitude limits than the images available from the SDSS. Astronomical *seeing* describes the image resolution at a particular observatory on a particular night. Seeing can be measured by analyzing the point spread function (PSF) of a star, where the PSF describes how the object's magnitude changes as a function of radius. Typically seeing is given as the full-width half-max (FWHM) of the PSF. *Pixel scale* describes how much of the sky is fit onto a single pixel in an image. The smaller the pixel scale, the better small details will be seen in an image. For example, a pixel scale of $0.2''$ would mean that $0.2''$ of the sky fits on a single pixel and it would take $18,000$ pixels to cover one full degree of sky. *Magnitude limit* describes the faintest objects that can be well observed in an image. Magnitude

Figure 2.1: A collage of the ten lensing systems. The bottom row is (left to right): SDSS J0901+1814, SDSS J1038+4849, SDSS J0900+2234, SDSS J1209+2640 and SDSS J0957+0509. The top row is (left to right): SDSS J1511+4713, SDSS J1537+6556, SDSS J1439+3250, SDSS J1318+3942 and SDSS J1343+4155. These images were produced by combining the stacked images in g, r and i filters. Each image has dimensions of $49'' \times 49''$.

limits can be measured by plotting a histogram of magnitudes and observing when the histogram begins to turn over from a maximum.

In the SDSS Data Release 7 (DR7), the median seeing is $1.4''$, the pixel scale is $0.396''$ and the magnitude limits are 22.2 in g and r filters and 21.3 in i-band. In our data seeing ranged from $0.49''$ for SDSS J1318+3042 to $1.54''$ for SDSS J1209+2640. Median seeing for our data was $0.74''$ for February 26 and $0.75''$ for February 27, significantly better than SDSS seeing. Pixel scale for the Mini-Mosaic camera that the group used at WIYN is $0.14''$, less than half that of the SDSS camera. Finally, we measured magnitude limits for our data to be approximately 24 in g and r-band and 23 in i-band. For each data image, the exposure time was 450-s and two exposures were taken for each field in each filter. Later the exposures were stacked, leading to a total exposure time of 900-s (15 minutes) per field in each filter.

The data were reduced using the NOAO Image Reduction and Analysis Facility [83]. The data were flat-fielded using both dome flats taken on site and superflats produced from data images. Cosmic rays were removed using the IRAF task LA-Cosmic [84]. Next, the two images in each filter were WCS-corrected and stacked, again in IRAF. Object magnitudes were measured in several measurement apertures using SExtractor [11]. Finally, the instrumental magnitudes measured by SExtractor were converted to calibrated magnitudes. This was done by finding the model magnitudes of stars in the SDSS DR7 Catalog Archive Server that also appeared in the WIYN data and finding the offset in magnitudes in the g, r, and i bands. The median offset in each filter for each field was then added to the SExtractor magnitudes (again using `MAG_AUTO`).

2.3   Galaxy Cluster Properties

2.3.1   Identifying Cluster Galaxies

We first sought to characterize richness of the clusters in terms of N_{gals}, the number of cluster members within 1 h^{-1} Mpc of the BCG [37] by using the maxBCG method. The maxBCG method [51] uses three primary features of galaxy clusters to facilitate the detection of clusters in survey data. First, galaxies in a cluster tend to be close together near the center and to become more separated from one another toward the outskirts of the cluster. Second, galaxies in a cluster tend to closely follow a sequence in a color-magnitude diagram; this is referred to as the E/S0 ridgeline, where E and S0 refer to galaxy types in the Hubble classification. Finally, galaxy clusters typically contain a central BCG, which is defined as the

brightest galaxy in the cluster. In all of the clusters in our sample, one or two BCGs can be seen near the center of the cluster surrounded by lensing arcs. While the dark matter halo dominates the lensing potential, the BCG contributes to the lensing potential as well since it comprises a large fraction of the baryonic matter in the cluster. Typically the BCG would be expected to have a color similar to that of the other cluster galaxies and it is also expected to be almost at rest with respect to the halo of the cluster.

Considering these properties of cluster galaxies, we searched the SExtractor catalog files for objects that:

1. Were classified as galaxies, not stars

2. Were within 1 h^{-1} Mpc of the central BCG

3. Had the characteristic $r - i$ color of the E/S0 ridgeline

4. Had the characteristic $g - r$ color of the E/S0 ridgeline

5. Met a particular magnitude limit (0.4L*)

In order to separate galaxies from stars, we compared two different SExtractor magnitudes, `MAG_AUTO` and `MAG_APER`. `MAG_AUTO` is the flux measured above background in a variable-size elliptical aperture. `MAG_APER` uses a circular aperture of fixed size to determine magnitude; we used a diameter of 2.0″. The difference `MAG_APER − MAG_AUTO` (henceforth Δm), can be used to identify the galaxies: stars stand out from galaxies because stars typically have a nearly identical shape while galaxies generally do not. Thus for stars the fixed aperture of `MAG_APER` will measure a fairly constant fraction of the light that the variable aperture of `MAG_AUTO` will measure. Therefore, the difference between the measurements (Δm) will be mostly
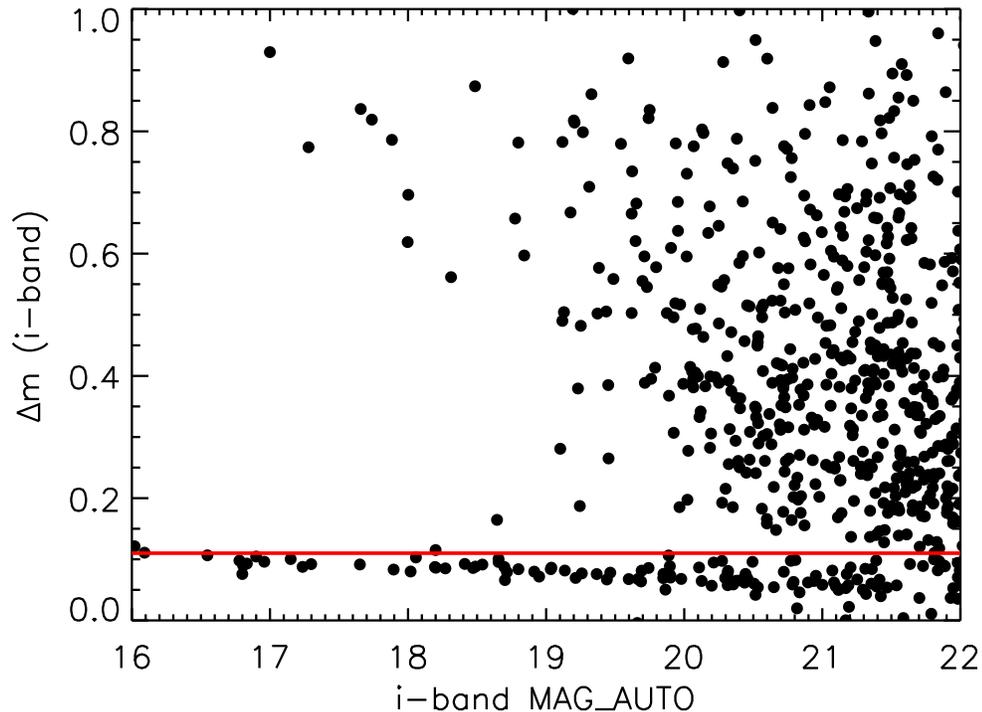
Figure 2.2: A sample plot of Δm vs. i-band `MAG_AUTO` for SDSS J1439+3250. Recall that Δm is $2''$ `MAG_APER` − `MAG_AUTO`. The horizontal red line is the star-galaxy cutoff we used, meaning objects with Δm$\leq$0.11 were cut as stars.

constant for stars, but not for galaxies. We used this fact to find stars by plotting Δm vs. `MAG_AUTO` (see Figure 2.2). In this plot, stars will be found on a mostly horizontal line of nearly constant Δm value; this line is referred to as the stellar locus.

We also tried using the SExtractor parameter `CLASS_STAR` for star-galaxy separation by requiring $0 \leq$ `CLASS_STAR` $\leq$ `0.9` (1 is highly star-like and 0 is highly galaxy-like in this parameter) and remeasuring N_{gals} with this requirement. We chose this cutoff because when we plotted `CLASS_STAR` against i-band magnitude (`MAG_AUTO`), we found a tight stellar sequence within 0.1 of `CLASS_STAR` = 1. We found that the mean difference in N_{gals} values was 0.3, which corresponds to a mean
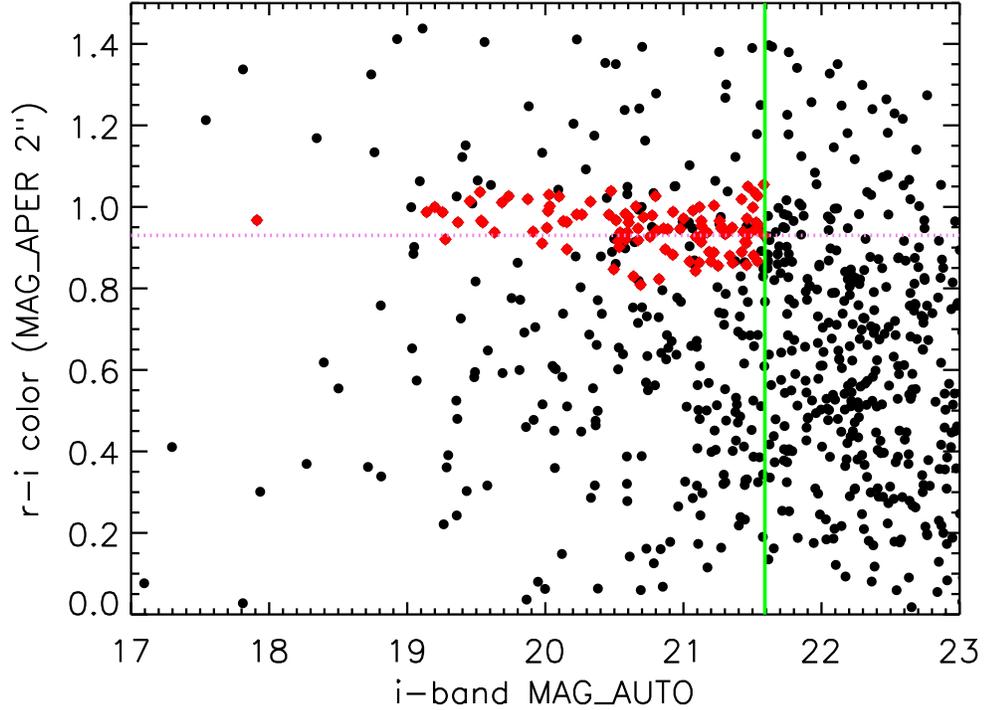
Figure 2.3: The $r - i$ color-magnitude diagram for SDSS J1209+2640. The black dots denote the galaxies, the red diamonds denote the cluster galaxies, the vertical green line shows the value of $0.4L^*$ and the horizontal violet dotted line represents the red sequence $r - i$ color. The objects plotted are galaxies within $1\ h^{-1}$ Mpc of the BCG.

percent difference of 1.7%. Thus we conclude that the Δm cut method is equivalent to using `CLASS_STAR`.

In order to select galaxies that are members of the cluster, we used the red sequence method [33, 51]. This approach involves plotting a color-magnitude diagram of the $g - r$ and $r - i$ colors of the galaxies vs. their i-band magnitude, looking for a nearly horizontal line of galaxies of similar color. Galaxies in a cluster are at similar redshifts and will be largely coeval, leading them to have similar colors. Thus the galaxies that populate the red sequence are likely to be cluster members. For each cluster we identified the $g - r$ and $r - i$ color of the red sequence on the plots. A sample color-magnitude diagram is shown in Figure 2.3.

We also used a second method to check our identification of the red sequence color. For both $g - r$ and $r - i$ colors, we made a histogram of the colors of the galaxies within 1 h^{-1} Mpc of the BCG and found the distribution near the red sequence color we had previously identified. We then fit this section of the histogram with a Gaussian profile and found the mean color of the red sequence galaxies.

Ultimately we used the first method (color-magnitude diagrams) to obtain a reasonable range of values for the colors of the red sequence and we used the second method (histograms) to determine final values for the colors. When we made color cuts, we only allowed galaxies that were within 2σ of the $r - i$ and $g - r$ colors, where σ was defined as:

$$\sigma = \sqrt{(\sigma_{intrinsic})^2 + (\sigma_{color})^2} \tag{2.1}$$

Here $\sigma_{intrinsic}$ is the intrinsic scatter in the red sequence color in the absence of measurement errors, which we took to be 0.06 for $r - i$ and 0.05 for $g - r$ [51]. σ_{color} is the color measurement error found by adding the SExtractor aperture magnitude measurement errors in quadrature.

Finally we cut any galaxies that had a magnitude dimmer than $0.4L^*$, where L^* is defined as the luminosity at which the luminosity function [74] changes from a power law to an exponential relation. In the maxBCG algorithm $0.4L^*$ is used as a limiting magnitude [52], and so we adopt this as our magnitude limit as well. We referred to a table of $0.4L^*$ [5] as a function of z to make cuts, allowing only galaxies brighter than $0.4L^*$ in i-band. All values used for cluster galaxy cuts are provided in Table 2.2.

Table 2.2: A summary of the values of limits used for richness measurements. Δm is the magnitude measured in i-band in $2''$ `MAG_APER` minus the magnitude in the same band measured in `MAG_AUTO`. Δm was used for star-galaxy separation. The $g - r$ and $r - i$ colors are based on measurements in the $2''$ aperture. Finally, the magnitude at $0.4L^*$ was found in the i-band.

System	Δm	g-r color	r-i color	0.4L* Magnitude
SDSS J0900+2234	0.56	1.83	0.73	21.20
SDSS J0901+1814	0.22	1.72	0.52	20.26
SDSS J0957+0509	0.15	1.78	0.71	21.26
SDSS J1038+4849	0.07	1.72	0.62	20.84
SDSS J1209+2640	0.34	1.79	0.93	21.59
SDSS J1318+3942	0.06	1.73	0.73	21.15
SDSS J1343+4155	0.16	1.75	0.54	20.71
SDSS J1439+3250	0.11	1.74	0.67	20.78
SDSS J1511+4713	0.17	1.78	0.75	20.97
SDSS J1537+6556	0.14	1.50	0.52	19.38

2.3.2 Cluster Properties

2.3.2.1 Area Corrections

We applied the four cuts described in § 2.3.1 to measure N_{gals}. However we found that for several of the ten systems, regions of the cluster were not in the image. The reason for this is that when we took the data, our primary focus was on the strong lensing arcs, which were near the center in all of our images. In order to address this problem and still obtain accurate values for N_{gals}, we extrapolated values for N_{gals} in the area off the CCD. In order to do this, we divided the 1 h^{-1} Mpc aperture into six annuli with constantly increasing radii, as shown in Figure 2.4.
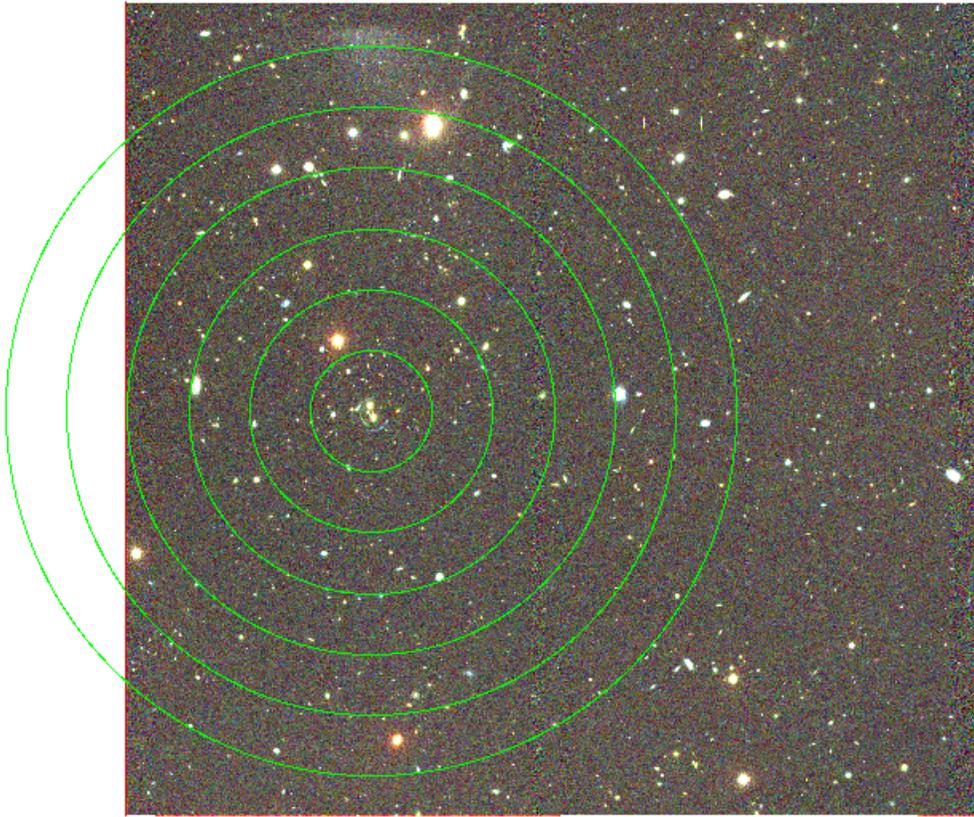
Figure 2.4: An image of SDSS J1038+4849, with a circular region of radius 1 h^{-1} Mpc centered on the BCG. We have divided the aperture into six annuli in order to apply Equation 2.3 for area corrections.

We assumed that the number of galaxies in each annulus should only be a function of radius; this would suggest that the number of galaxies per area should be a constant in each annulus. Mathematically,

$$\frac{N_{total}}{N_{on\_CCD}} = \frac{A_{ann}}{A_{ann\_on\_CCD}} \tag{2.2}$$

so

$$N_{total} = N_{on\_CCD} \left(\frac{A_{ann}}{A_{ann\_on\_CCD}} \right) \tag{2.3}$$

where N_{total} means the total number of galaxies in each annulus, N_{on_CCD} means the number of galaxies actually found in the image in each annulus, A_{ann} means the area of the annulus and $A_{ann\_on\_CCD}$ means the area of the annulus that was on the CCD.

We checked the accuracy of Equation 2.3 using the SDSS data. We measured N_{gals} twice, once covering the full 1 h^{-1} Mpc (taking this as true N_{gals}) and once covering only as much of the 1 h^{-1} Mpc as was on the CCD in the WIYN data. We then used Equation 2.3 to predict the final values of N_{gals} based on the measurements with the WIYN area cuts. Finally we compared the predicted values for N_{gals} to the measured (true) values and found them to be similar. We plot the two sets of N_{gals} against each other in Figure 2.5. Note that the points follow the $y = x$ line very closely, indicating that the measured and extrapolated values are quite similar and suggesting that the richness extrapolation works well. The typical fractional error in the extrapolated values is 0.06.
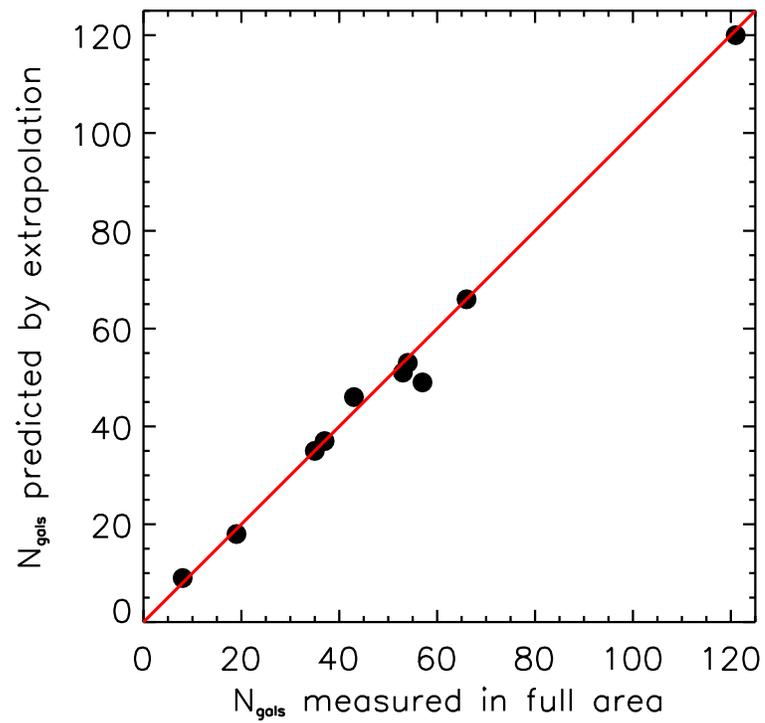
Figure 2.5: This plot is a test of the accuracy of the N_{gals} extrapolation described in Equation 2.3. Here we plot N_{gals} values measured in SDSS data, with the measured values on the x-axis and the predicted values on the y-axis. The red line is the $y = x$ line. Since the data closely follow the $y = x$ line, we conclude that the predictions from the extrapolation are quite accurate.

2.3.2.2   Richness Measurements

We next found the richness, N_{200} [37], the number of galaxies in a spherical region within which the density was $200\rho_{crit}$, where ρ_{crit} is the critical density of the universe. The radius of this spherical region of space is termed r_{200}. Hansen et al. [37] give r_{200} as:

$$r_{200} = 0.156(N_{gals})^{0.6}h^{-1}Mpc \qquad (2.4)$$

We used the area-corrected values for N_{gals} when calculating r_{200}. In order to find N_{200} we again applied the four cuts discussed in § 2.3.1, this time using r_{200} as the distance cut rather than 1 h^{-1} Mpc. Finally, once we found N_{200}, we again applied the area corrections using Equation 2.3.

We used the variable elliptical aperture of `MAG_AUTO` and the circular $2''$ and $3''$ diameter apertures using `MAG_APER` in order to determine object magnitudes and thus colors. We used $2''$ and $3''$ because both were significantly larger than the seeing FWHM, for which the median value was about $0.75''$. The differences in colors measured in different apertures were usually small, on the order of 0.05 magnitudes, but could be up to 0.2 magnitudes. Since identification of a cluster galaxy depends on color, there was a resulting variation in richness values for different apertures. We determined that the $2''$ aperture had the highest signal to noise by comparing the measurement errors of the $g - r$ and $r - i$ colors to see in which aperture the errors were typically lowest. We found that the $2''$ aperture typically had the lowest error value; therefore we used the colors and thus richness values in the $2''$ aperture for richness measurements. However, we considered the variation in richness values to determine the error in richness: we took the standard deviation of the three values for N_{200} for each cluster and used these values for the uncertainty in N_{200}.

2.3.2.3   <u>Cluster Mass</u>

We define M_{200} to be the mass contained within a spherical region of radius r_{200} [47]. An empirical relation between mass and richness is found in Johnston et al. [47] using a large sample of maxBCG clusters from the SDSS:

$$M_{200}(N_{200}) = M_{200|20} \left(\frac{N_{200}}{20} \right)^{\alpha_N} \tag{2.5}$$

In this equation $M_{200|20} = (8.8 \pm 0.4_{stat} \pm 1.1_{sys}) \times 10^{13} h^{-1} M_{\odot}$ and $\alpha_N = 1.28 \pm 0.04$. Equation 2.5 was found empirically using data from the SDSS, using mean redshift of $z = 0.25$.

The error in M_{200} values was considered in Rozo et al. [70]. In that paper, the logarithmic scatter in mass at fixed richness is given as:

$$\sigma_{lnM|N} = 0.45^{+0.20}_{-0.18} \tag{2.6}$$

We thus can approximate the uncertainty in the mass itself as:

$$\Delta M = 0.45 M_{200} \tag{2.7}$$

We also propagate error from the uncertainty in values of N_{200} through equation 2.5. Our final values for error on M_{200} were found by adding the uncertainty in the mass and the propagated error in quadrature. The propagated fractional errors had a median value of 0.13 while the scatter described by Equation 2.7 had a value of 0.45. The combined fractional errors had a median value of 0.47, with the scatter in mass dominating the errors.

2.3.2.4   Velocity Dispersion

Becker et al. [8] give an empirical relationship for velocity dispersion as a function of richness found from redshifts of cluster members in the maxBCG cluster sample:

$$\langle \ln \sigma_v \rangle = A + B \ln \frac{N_{200}}{25} \tag{2.8}$$

The constants A and B are referred to as mean normalization and mean slope, respectively. They are given as $A = 6.17 \pm 0.04$ and $B = 0.436 \pm 0.015$. Becker et al. [8] also found a relation for the scatter, S, in the velocity dispersion. The scatter is defined to be the standard deviation in $\ln \sigma_v$:

$$S^2 = C + D \ln \frac{N_{200}}{25} \tag{2.9}$$

where $C = 0.096 \pm 0.014$ and $D = -0.0241 \pm 0.0050$. We used this relation to calculate the errors on the velocity dispersion values, defining the errors as one standard deviation. We also propagated the error on N_{200} through Equation 2.8 and added these errors in quadrature to the errors found from Equation 2.9. Again the propagated errors are minimal: The median fractional error on the velocity dispersions from the propagated error on N_{200} is 0.08, while the median fractional error from Equation 2.9 is 0.31, leading to an overall median fractional error of 0.33.

2.3.2.5   Errors on Richness and Mass

In order to better constrain the error on our richness measurements, we also measured colors and richnesses for the 10 systems using the SDSS data. We found

Table 2.3: A comparison of N_{gals} values measured in different SExtractor apertures and in SDSS data. These values for N_{gals} have not been area corrected with Eq. 2.3. For the SDSS values, any area which is not on the CCD in the WIYN data is excluded from consideration. Note that for SDSS J1537+6556 much of the WIYN area is outside the SDSS footprint, so the SDSS value is biased low.

System	N_{gals} (MAG_AUTO)	N_{gals} ($2''$ MAG_APER)	N_{gals} ($3''$ MAG_APER)	$N_{gals}(Sloan)$
SDSS J0900+2234	23	28	29	56
SDSS J0901+1814	8	14	8	11
SDSS J0957+0509	15	28	26	63
SDSS J1038+4849	16	15	17	32
SDSS J1209+2640	85	101	98	190
SDSS J1318+3942	21	23	23	39
SDSS J1343+4155	26	25	32	46
SDSS J1439+3250	48	55	51	82
SDSS J1511+4713	22	29	29	54
SDSS J1537+6556	14	20	18	8

that richness values from the SDSS are typically much higher than those found in this paper; the mean ratio of N_{gals}(SDSS) to N_{gals}($2''$ MAG_APER) is 1.75 (for WIYN N_{gals} before area corrections, using only cluster area found both in WIYN and SDSS data (see Table 2.3).

These differences apparently arise because there is a larger error in magnitudes measured in the SDSS than in the data used here. This allows some objects to be counted as cluster members in the SDSS that are not counted as cluster members in the WIYN data. Note that in Figure 2.6, a color-color diagram for SDSS J1318+3942, more cluster members are found in SDSS data, but those objects are much more scattered in color-color space and many are not true cluster members. On the other hand, fewer objects are found in the WIYN data, but these objects form a much tighter red sequence and are more likely to be genuine cluster members.
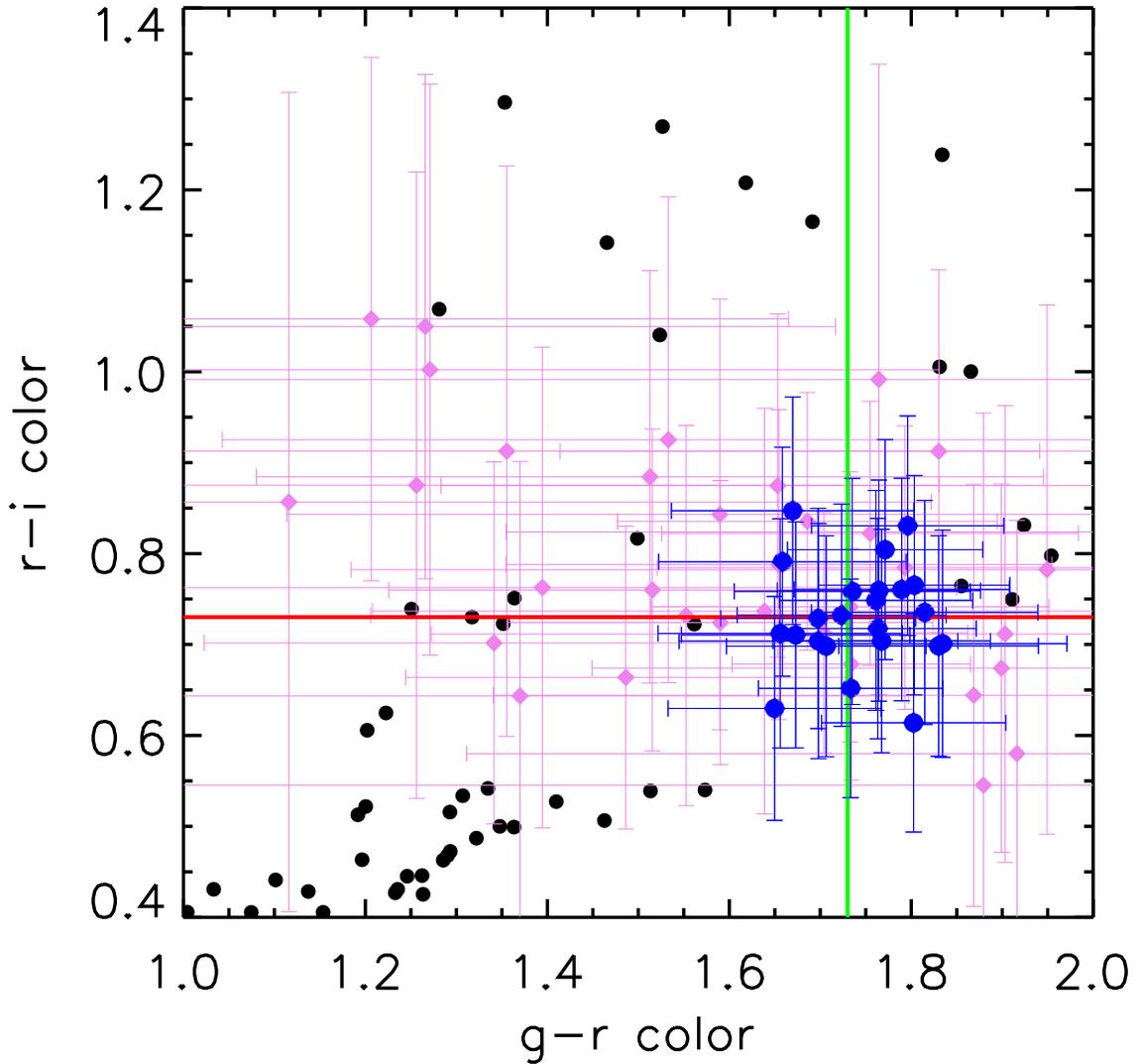
Figure 2.6: A comparison of cluster members found in our data (2″ MAG_APER) and in the SDSS data for SDSS J1318+3942. The larger blue circles represent cluster members found in our data and the smaller violet diamonds are cluster members found in the SDSS data. The smallest black circles are all galaxies within 1 h^{-1} Mpc of the BCG that are brighter than $0.4L^*$ in our data but do not meet the color cuts to be considered cluster members. The error bars represent 2σ, with σ defined by Eq. 2.1. The solid lines mark the cluster red sequence colors for the WIYN data: the vertical green line marks the g-r color and the horizontal red line marks the r-i color.

We also include Figure 2.7, in which we show the deviation of each cluster galaxy's color from the measured color of the E/S0 ridgeline; we plot this vs. SDSS i-band magnitude for all ten clusters. We found g-r and r-i colors for objects considered to be cluster galaxies within 1 h^{-1} Mpc of the BCG in WIYN data and in SDSS data and compared them to the characteristic red sequence colors of the respective clusters. We also found the errors in colors for both sets of data using Equation 2.1 to find σ. We used magnitude errors reported by SExtractor for WIYN data and errors on model magnitudes for SDSS data. The error bars shown represent 2σ. It can be seen in Figure 2.7 that the differences between the measured color and the cluster color are much larger in the SDSS data than in WIYN data but the errors are larger for SDSS data as well. Due to these larger errors in SDSS data, there is a higher likelihood that objects with larger color deviations will still be counted as cluster members.

The differences in richness values between WIYN and SDSS data persist even at bright magnitudes. We measured values for N_{gals} at an i-band magnitude of 19.38, which is the value for $0.4L*$ corresponding to $z = 0.25$. We found that the mean ratio of N_{gals}(SDSS) to $N_{gals}(2'' \texttt{MAG\_APER})$ is 1.63, meaning that SDSS values are typically about 60% higher than WIYN values. Thus we find that in general for these ten clusters richness values measured in our data do not closely match values measured in the SDSS data.

However, since the mass-richness relation (Equation 2.5) is calibrated from SDSS data, if we use WIYN richness values with this equation, we would expect the masses to be biased to be too low. Therefore, we determined it would be necessary to scale our measured richness values up to match SDSS values. To do that, we we first found all objects that were counted as cluster galaxies (N_{gals}) only in WIYN (not in SDSS) and then found the opposite, objects counted as cluster galaxies only in
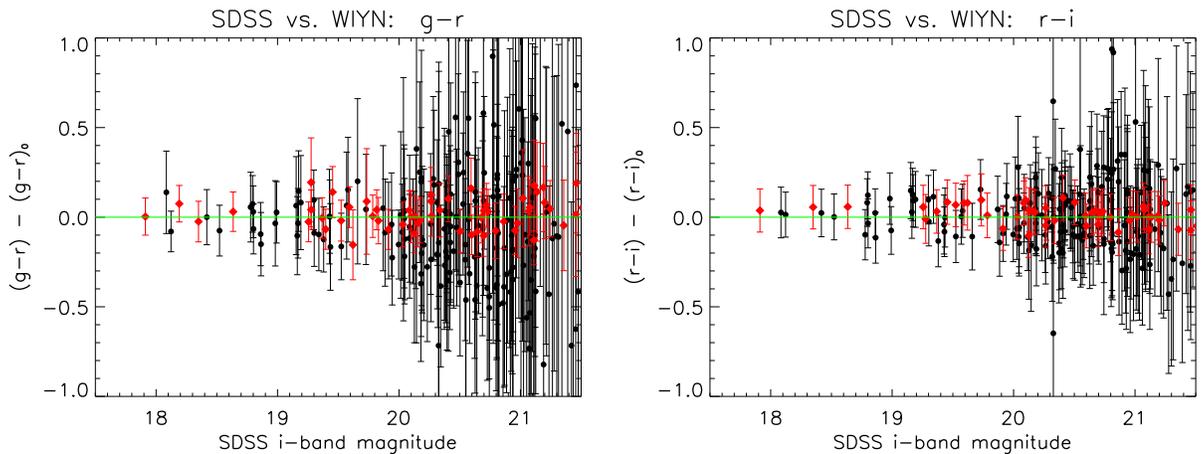
Figure 2.7: A plot of color difference versus SDSS i-band magnitude for all cluster members in both SDSS and WIYN data. Color difference is defined as the difference between the actual $r - i$ or $g - r$ color of each cluster galaxy and the measured red sequence color for that cluster. Cluster galaxies in each of the ten clusters are plotted together here. The red diamonds denote WIYN data points and the black circles denote SDSS data points. The error bars represent 2σ, where σ is defined by Eq. 2.1. For SDSS measurements, color is found from SDSS model magnitudes and the red sequence colors were measured in SDSS data. For WIYN measurements, color is found from $2''$ MAG_APER magnitudes and red sequence colors were measured in WIYN data. Note that WIYN data points are found much closer to the central line that represents color difference of 0, while SDSS points can be found further away. To be counted as cluster members, points must be within 2σ of the cluster red sequence colors, but 2σ is larger for the SDSS points.

SDSS but not in WIYN. We then also found the galaxies counted as cluster galaxies in both WIYN and SDSS. Our goal was to constrain the amount that SDSS was overcounting galaxies. To do that we found the ratio

$$C = \frac{N_1 + N_2}{N_1} = 1 + \frac{N_2}{N_1} \tag{2.10}$$

where N_1 represents the number of cluster members found in both WIYN data and SDSS data and N_2 represents the number of cluster members found only in SDSS data. Since we expect the numbers of galaxies in each magnitude bin to be a Poisson distribution, the standard deviation on N_1 and N_2 would be simply the square root of each. Then the fractional error on Equation 2.10 would be

$$\sigma_C = \frac{N_2}{N_1} \sqrt{\frac{1}{N_2} + \frac{1}{N_1}} \tag{2.11}$$

We then plotted C against binned WIYN i-band (`MAG_AUTO`) model magnitude. The result is shown in Figure 2.8.

We fit the data with a linear relation using IDL routine `FITEXY`, which applies a linear fit including error bars. The final relation found was

$$C = (0.222 \pm 0.116)m_{i\_WIYN} + (-2.84 \pm 2.29) \tag{2.12}$$

The magnitude m_{i_WIYN} is WIYN i-band magnitude from `MAG_AUTO`. When this equation is evaluated at i-band $m = 19.38$, the value for $0.4L*$ at the mean SDSS redshift of 0.25, then $C = 1.47$. We took this as the correction factor for our richness values.
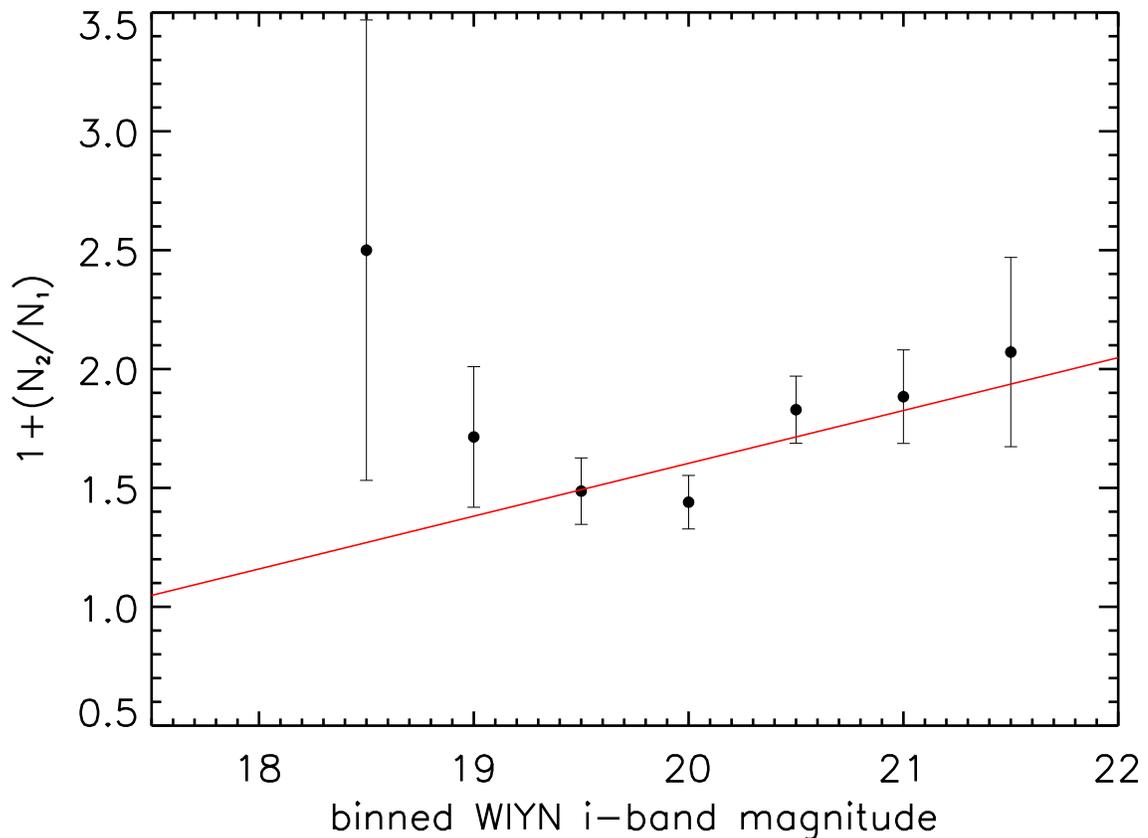
Figure 2.8: A plot comparing objects counted as cluster galaxies only in SDSS data and in both WIYN and SDSS data. Here N_1 is the number of cluster members found in both SDSS and WIYN in that magnitude bin and N_2 is the number of cluster members found only in SDSS. We plot the ratio $1 + \frac{N_2}{N_1}$ (which we refer to in the text as C) on the y-axis and the magnitude bin on the x-axis, where magnitude bins are 0.5 magnitude in size. The red line is a linear best fit, found using IDL routine `FITEXY`. The equation of that line is $C = (0.222 \pm 0.116)m_{i\_WIYN} + (-2.84 \pm 2.29)$, where m_{i_WIYN} represents magnitude in i-band `MAG_AUTO`.

Table 2.4: A summary of the quantities measured for the ten galaxy clusters. These are all based on colors measured in the 2″ aperture. The N_{gals} and N_{200} values are area-corrected using Eq. 2.3 but are not scaled up. All the other values are based on these area-corrected but not scaled richness values.

System	N_{gals}	r_{200} $(h^{-1}Mpc)$	N_{200}	M_{200} $(10^{14}h^{-1}M_\odot)$	σ_v (km/s)	c_{200}
SDSS J0900+2234	28	1.15	30 ± 4.1	1.48 ± 0.715	518^{+196}_{-153}	$8.27^{+14.9}_{-3.32}$
SDSS J0901+1814	15	0.792	11 ± 0.58	0.409 ± 0.186	334^{+137}_{-98}	$19.0^{+91.0}_{-9.00}$
SDSS J0957+0509	29	1.18	36 ± 3.4	1.87 ± 0.8708	561^{+200}_{-153}	$7.49^{+9.81}_{-2.78}$
SDSS J1038+4849	16	0.823	15 ± 0.62	0.609 ± 0.276	383^{+150}_{-108}	$34.9^{+18800}_{-21.3}$
SDSS J1209+2640	101	2.49	214 ± 11.5	18.3 ± 8.32	1219^{+293}_{-240}	$3.64^{+2.81}_{-1.21}$
SDSS J1318+3942	24	1.050	25 ± 4.2	1.17 ± 0.583	478^{+191}_{-150}	$9.9^{+31.8}_{-4.39}$
SDSS J1343+4155	28	1.15	29 ± 1.1	1.42 ± 0.641	510^{+182}_{-135}	$14.3^{+118}_{-7.26}$
SDSS J1439+3250	59	1.80	105 ± 18	7.35 ± 3.69	894^{+296}_{-250}	$3.20^{+2.41}_{-1.03}$
SDSS J1511+4713	31	1.22	40 ± 2.9	2.14 ± 0.981	587^{+203}_{-154}	$7.69^{+6.51}_{-2.45}$
SDSS J1537+6556	22	0.997	22 ± 2.9	0.994 ± 0.477	452^{+177}_{-136}	$18.8^{+49.2}_{-7.84}$

We measured N_{gals} and corrected these values for missing area in WIYN using Equation 2.3. Then we included the above correction factor when calculating r_{200}, letting

$$r_{200} = 0.156(N_{gals\_SDSS})^{0.6} = 0.156(1.47N_{gals\_WIYN})^{0.6} \qquad (2.13)$$

We remeasured N_{200} using the new value for r_{200} and corrected for missing area. Finally we scaled these new N_{200} values by multiplying them by the same scale factor of 1.47. We used these scaled values of N_{200} to find M_{200}, velocity dispersion and concentration parameter. We give values for all quantities found without the scale factor in Table 2.4 and we give the values found with the scale factor in Table 2.5.

We find the scaled values for N_{200} are on average 1.7 times bigger than the unscaled values. This leads the new values for M_{200} (those found from the scaled richness

Table 2.5: A summary of the quantities measured for the ten galaxy clusters. These are all based on colors measured in the $2''$ aperture. The N_{gals} and N_{200} values are area-corrected using Eq. 2.3 and are scaled up using Eq. 2.10.

System	N_{gals}	r_{200} $(h^{-1}Mpc)$	N_{200}	M_{200} $(10^{14}h^{-1}M_{\odot})$	σ_v (km/s)	c_{200}
SDSS J0900+2234	28	1.45	53 ±7.6	3.046 ± 1.48	662^{+233}_{-187}	$5.13^{+5.36}_{-1.82}$
SDSS J0901+1814	15	0.996	22 ± 2.4	0.993 ± 0.468	452^{+174}_{-132}	$9.63^{+12.7}_{-3.52}$
SDSS J0957+0509	29	1.48	57 ± 5.1	3.38 ± 1.57	686^{+226}_{-176}	$5.15^{+4.68}_{-1.72}$
SDSS J1038+4849	16	1.036	25 ± 1.8	1.17 ± 0.536	477^{+177}_{-132}	$16.8^{+73.3}_{-7.80}$
SDSS J1209+2640	101	3.13	317 ± 17	30.2 ± 13.7	1446^{+307}_{-258}	$2.69^{+1.90}_{-0.890}$
SDSS J1318+3942	24	1.32	44 ± 5.0	2.41 ± 1.14	612^{+215}_{-168}	$5.83^{+7.30}_{-2.17}$
SDSS J1343+4155	28	1.45	43 ± 1.6	2.31 ± 1.046	603^{+203}_{-153}	$9.11^{+22.0}_{-3.94}$
SDSS J1439+3250	59	2.27	158 ± 28	12.4 ± 6.26	1069^{+331}_{-288}	$2.36^{+1.72}_{-0.790}$
SDSS J1511+4713	31	1.54	70 ± 5.6	4.40 ± 2.030	751^{+237}_{-185}	$5.21^{+3.44}_{-1.52}$
SDSS J1537+6556	22	1.25	41 ± 9.6	2.21 ± 1.20	594^{+243}_{-204}	$10.0^{+14.7}_{-3.88}$

values) to be 2.0 times larger than the previous values. Also new values for velocity dispersion are 1.3 times larger than previous values, while new values for concentration parameter are all smaller, on average 0.63 times the previous values (see § 2.5.2).

2.3.2.6 Comparison of Results

Several other groups have measured cluster masses or related quantities for some of our clusters. Oguri et al. (2012) [66] present combined strong and weak lensing analyses for 28 clusters, including 4 of the clusters discussed in this paper. This allowed us to compare our results for M_{200} to their results for these four systems. As Oguri et al. (2012) [66] present values for M_{vir}, we converted these to M_{200} values using the method described in Appendix A of Johnston et al. [47] (see §2.5.1).

Table 2.6: M_{200} values from other papers for several of our systems Oguri et al. (2012) [66], Bayliss et al. [7] and Drabek et al. [27]. Note that the values from Oguri et al. (2012) [66] have been converted from M_{vir} to M_{200} using the process detailed in the appendix of Johnston et al. [47]. To convert the errors, we simply converted the upper and lower errors on M_{vir} given in Oguri et al. [66] using the same method. All M_{200} values have the units $10^{14}h^{-1}M_{\odot}$. Though there are significant differences in the M_{200} values for some clusters, overall our values seem consistent with those of the other groups. See Figure 2.9.

System	M_{200}(this paper)	M_{200}(Oguri)	M_{200}(Bayliss)	M_{200}(Drabek)
SDSS J0957+0509	3.38 ± 1.57	$1.17^{+0.77}_{-0.55}$	$8.01^{+4.98}_{-6.40}$	-
SDSS J1038+4849	1.17 ± 0.536	$0.681^{+0.48}_{-0.11}$	$2.06^{+1.18}_{-0.36}$	-
SDSS J1209+2640	30.2 ± 13.7	$5.50^{+1.67}_{-1.32}$	$16.8^{+6.43}_{-11.0}$	-
SDSS J1343+4155	2.31 ± 1.046	$3.34^{+1.38}_{-1.11}$	$8.13^{+4.76}_{-6.87}$	6.60 ± 3.20
SDSS J1439+3250	12.4 ± 6.26	-	-	4.73 ± 2.84

Bayliss et al. [7] provided velocity dispersions for 4 of our clusters. We used the relation between cluster mass and galaxy velocity dispersion given in Evrard et al. [31] to find M_{200}:

$$b_v^{\frac{1}{\alpha}} M_{200c} = 10^{15} M_{\odot} \frac{1}{h(z)} \left(\frac{\sigma_{gal}}{\sigma_{15}} \right)^{\frac{1}{\alpha}} \tag{2.14}$$

Here $h(z)$ is the Hubble parameter, $b_v = \sigma_{gal}/\sigma_{DM}$ is the velocity bias (we assume $b_v = 1$), σ_{gal} is the galaxy velocity dispersion, σ_{DM} is the dark matter velocity dispersion, $\sigma_{15} = 1084 \pm 13$ km s^{-1}, and $\alpha = 0.3359 \pm 0.0045$. Drabek et al. [27] present masses for two clusters, SDSS J1343+4155 and SDSS J1439+3250, based on spectroscopy of a sample of galaxies in these clusters. We summarize all the values of M_{200} found by these groups in Table 2.6.

In Figure 2.9, we plot the M_{200} values from the three other papers against our M_{200} values; the dotted line in the plot is the $y = x$ line. We find that our values are reasonable in light of the findings of other groups as when we plot our values against those from other groups, the points are all scattered around the $y = x$ line.
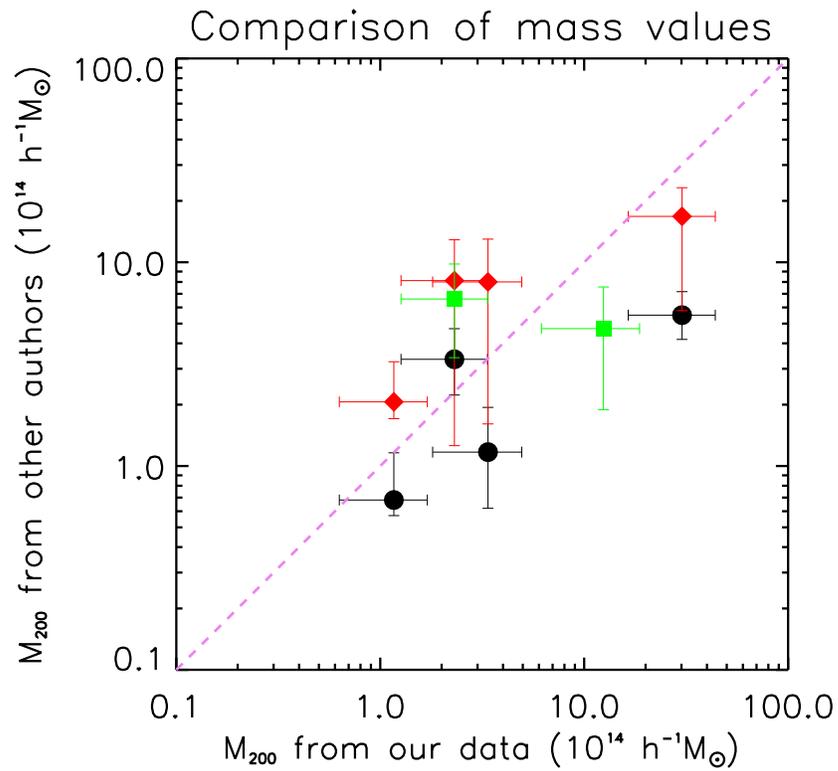
Figure 2.9: A comparison of M_{200} values in this paper and in other papers. The black circles represent mass values in Oguri et al. (2012) [66], the red diamonds represent mass values in Bayliss et al. [7] and the green squares represent mass values in Drabek et al. [27]. The dotted violet line is the $y = x$ line.

2.4 Strong Lensing Properties

In a strong lensing system, if the source galaxy and the galaxy cluster are perfectly aligned, then the image formed will be a perfect ring, or Einstein ring. The radius of this ring is referred to as the Einstein radius. The Einstein radius for a symmetric mass distribution treated as a thin sheet is given by Narayan and Bartelmann [62]:

$$\theta_E = \sqrt{\frac{4GM}{c^2}\frac{D_{ds}}{D_d D_s}} \qquad (2.15)$$

where D_d, D_s, and D_{ds} are angular diameter distances to the lens, to the source, and from lens to source, respectively, c is the speed of light, G is the gravitational constant, and M is the mass contained within the Einstein radius. We estimated the Einstein radius of each of the clusters directly by fitting a circle to the visible arc and measuring the radius of that circle. The values found here are all very similar to those presented in the SBAS discovery papers, with a median difference of 2.5%.

In order to try to quantify the uncertainty in our measurements, we measured the Einstein radii for all the objects again several months after the first measurement without referencing previous data. In all cases the differences between the original and new measurements were between $0.03''$ and $0.6''$. Since this represents up to 10% of the value of θ_E, we estimated the uncertainty in θ_E as 10%.

We note however that this method of estimating Einstein radius can lead to large systematic errors, so we also compared our values for Einstein radii to values from other groups. West et al. [91] present strong lensing models for three of our systems and [66] present models for four of our systems. Both groups have measurements for SDSS J1343+4155, so we compared values for a total of six systems. We provide measured Einstein radii from these papers in Table 2.7. For SDSS J0900+2234

Table 2.7: A comparison of values for Einstein radius measured in this paper, in West et al. [91] and in Oguri et al. (2012) [66].

System	θ_E(arcsec) (this paper)	θ_E(arcsec) (West et al.)	θ_E(arcsec) (Oguri et al.)
SDSS J0900+2234	8.0 ± 2.7	8.32	-
SDSS J0901+1814	6.9 ± 2.3	6.35	-
SDSS J0957+0509	8.2 ± 2.7	-	$5.2^{+0.5}_{-0.5}$
SDSS J1038+4849	8.6 ± 2.9	-	$12.6^{+1.3}_{-1.6}$
SDSS J1209+2640	11 ± 3.7	-	$8.8^{+0.9}_{-0.9}$
SDSS J1343+4155	13 ± 4.3	7.05	$5.4^{+2.5}_{-1.6}$

and SDSS J0901+1814, our estimates are almost exactly the same as the values in [91]. However for the other four systems, the scatter (standard deviation) in values is larger, between $2.1''$ and $4.0''$. We account for this error by calculating the fractional error in the values for θ_E and then finding the median value of the fractional errors for each of the six systems. The median value of the fractional errors is 0.32, or 32%, which we added in quadrature to the 10% errors to find final error values.

Solving Equation(2.15) for the mass, we obtain:

$$M = \theta_E^2 \frac{c^2}{4G} \frac{D_d D_s}{D_{ds}} \tag{2.16}$$

Using the redshifts listed in Table 2.1 for the galaxy clusters and the source galaxies, we calculated the angular diameter distances. We then used the Einstein radii we had measured to calculate the masses of the lenses.

Table 2.8: A summary of the properties measured for the ten strong lensing systems. Rescaled Einstein radii are Einstein radii projected to fiducial redshifts, of $z_{lens} = 0.433$ and $z_{source} = 1.65$. See Equation 2.21. Source redshift has not yet been determined for the arc in SDSS J1439+3250 and so we can only present a range of redshifts, leading to a range of values for mass and velocity dispersion.

System	θ_E (arcsec)	$M_{lens}(10^{12}h^{-1}\mathbf{M_\odot})$	$\sigma_v(\mathbf{km/s})$	$\theta_E(rescaled)$
SDSS J0900+2234	8.0 ± 2.7	11 ± 7.3	648 ± 108	7.9 ± 2.7
SDSS J0901+1814	6.9 ± 2.3	5.5 ± 3.7	564 ± 93.9	5.9 ± 2.0
SDSS J0957+0509	8.2 ± 2.7	12 ± 8.0	680 ± 113	8.5 ± 2.8
SDSS J1038+4849	8.6 ± 2.9	15 ± 10.0	780 ± 130	11 ± 3.8
SDSS J1209+2640	11 ± 3.7	36 ± 24.0	691 ± 115	19 ± 6.2
SDSS J1318+3942	9.1 ± 3.0	12 ± 8.0	336 ± 55.9	8.2 ± 2.7
SDSS J1343+4155	13 ± 4.3	24 ± 16.0	804 ± 134	12 ± 3.9
SDSS J1439+3250	7.4 ± 2.5	7.4 ± 4.9 -10.0 ± 6.7	596 ± 99.2 - 708 ± 118	7.1 ± 2.4
SDSS J1511+4713	5.4 ± 1.8	6.3 ± 4.2	631 ± 105	7.3 ± 2.4
SDSS J1537+6556	8.5 ± 2.8	8.7 ± 5.8	715 ± 119	9.4 ± 3.1

Finally, we calculated the velocity dispersions of the regions of the clusters inside θ_E assuming the mass distribution was well fit by a singular isothermal sphere (SIS). We used the following equation, from [62]:

$$\sigma_v = \sqrt{\frac{\theta_E c^2 D_s}{4\pi D_{ds}}} \tag{2.17}$$

All values measured for the strong lenses are presented in Table 2.8.

In Figure 2.10 we compare the velocity dispersions found from lensing to those found from richness measurements. Note that these velocity dispersions measure different things: the velocity dispersion from lensing describes the velocity dispersion inside θ_E and the velocity dispersion from N_{200} describes the velocity dispersion within the much larger distance r_{200}. We see in Figure 2.10 that many of the clusters are found along the $y = x$ line, several are found above it and several
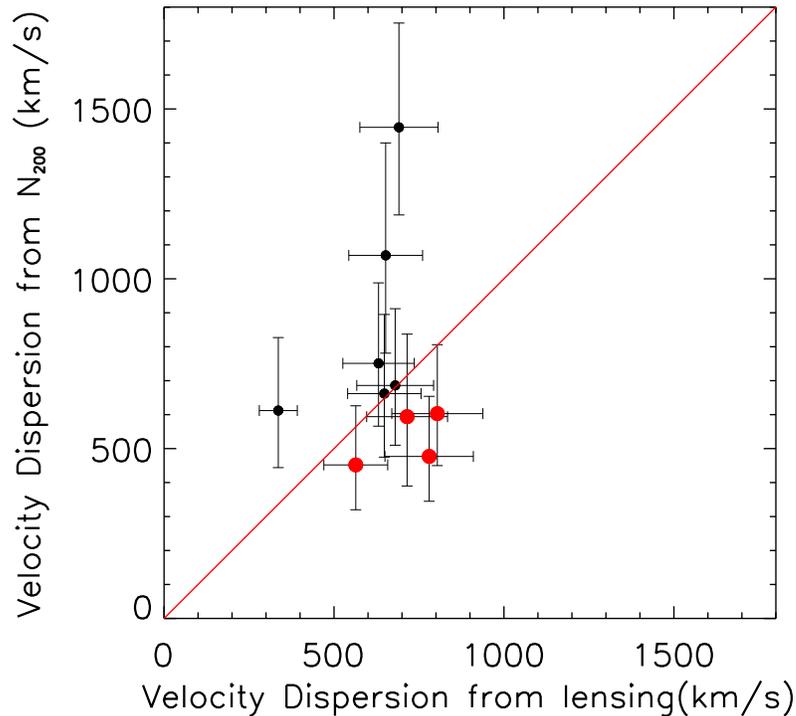
Figure 2.10: A comparison of velocity dispersions found from N_{200} and found from Einstein radii. The line shown has the equation $y = x$. The clusters on or above the $y = x$ line are all higher mass clusters. The lowest mass clusters are shown as red points.

are found below it. For the clusters found along the $y = x$ line, we see that the velocity dispersions are similar within the two different radii, θ_E and r_{200}, which suggests that these systems are largely isothermal since velocity dispersions are similar at different radii. For the systems found above the $y = x$ line, the velocity dispersion at large radii is much larger than at small radii, indicating that much of the mass is found at larger distance from the BCG, suggesting a low value for c_{200}. Several systems (shown as red points) are found below the $y = x$ line, suggesting that the velocity dispersion within θ_E is larger than that found within r_{200}. This indicates that for several clusters there is more mass within the smaller radius and suggests that the concentration parameter is large. Our highest mass clusters are

found above the $y = x$ line (suggesting lower concentration parameter), while our lower mass clusters are found below the $y = x$ line (suggesting higher concentration parameter). This would agree with what we discuss in the next section, that our highest mass clusters are not overconcentrated but our lowest mass clusters seem to be.

2.5 Applications to Cosmology

2.5.1 <u>An Overconcentration Problem?</u>

Several recent papers [64, 34, 32, 66] have presented evidence that galaxy clusters that exhibit strong lensing have higher concentration parameters than ΛCDM would predict. The most recent considerations [32, 66] suggest that this overconcentration is most significant at cluster masses less than $10^{14}h^{-1}M_\odot$. Overconcentration can be illustrated by comparing Einstein radii to M_{200} [34]. Since Einstein radii are dependent on both cluster mass and cluster concentration parameter, such a comparison will yield larger Einstein radii than would be expected for particular M_{200} values.

Considering this, we have compared Einstein radius to M_{200} for our ten systems. One complication in making this comparison is that Einstein radius is a function of redshift. Since all of our systems have different redshifts for both lens and source, in order to compare them, we needed to scale them to a single, constant redshift for lens and source. We chose both the lens and source redshifts (we refer to them henceforth as fiducial redshifts) by taking the mean of the ten lens redshifts and the mean of the ten source redshifts. Our fiducial redshifts are $z_d = 0.433$ for the lens and $z_s = 1.65$ for the source.

To scale Einstein radii to the fiducial redshifts, we needed to find a scale factor k that would satisfy:

$$\theta_{E\_scaled}(z_{d\_fiducial}, z_{s\_fiducial}) = k \times \theta_E\_measured(z_d, z_s) \tag{2.18}$$

We note that Equation 2.17 can be rearranged as

$$\theta_E = \frac{4\pi\sigma_v^2}{c^2}\frac{D_{ds}}{D_s} \tag{2.19}$$

Since σ_v is proportional to the mass and does not depend on redshift, θ_E scales with redshift according to the ratio D_{ds}/D_s. Thus solving Equation 2.18 for k we obtain:

$$k = \frac{\theta_E\_scaled}{\theta_E\_measured} = \frac{\frac{4\pi\sigma_v^2}{c^2}\frac{D_{ds\_fiducial}}{D_{s\_fiducial}}}{\frac{4\pi\sigma_v^2}{c^2}\frac{D_{ds}}{D_s}} \tag{2.20}$$

and since σ_v does not scale with redshift, it cancels. Then

$$k = \frac{D_{ds\_fiducial}/D_{s\_fiducial}}{D_{ds}/D_s} \tag{2.21}$$

We applied Equation 2.21 to find the scale factor k for each cluster and then scaled each Einstein radius to the fiducial values.

In order to compare the relation between Einstein radius and M_{200} for our data to the relation that ΛCDM would predict, we refer to the models presented in Duffy et al. (2008) [28] and Oguri et al. (2012) [66] which predict concentration as a function of cluster mass. Duffy et al. present relations for any general galaxy clusters while Oguri et al. present relations for lensing-selected clusters, that is those that were

found because of gravitational lensing (like the clusters discussed in this chapter).
Concentration parameter, c_Δ, is defined as

$$c_\Delta = \frac{r_\Delta}{r_s} \tag{2.22}$$

The r_s term is the scale radius, a term in the Navarro-Frenk-White (NFW) model of
dark matter halo density [63] (see § 1.7.2). The quantity Δ is the virial overdensity.
In this paper we use $\Delta = 200$, but Oguri et al. (2009) [65] use $\Delta = vir$, where
the virial overdensity is the local overdensity that would cause halo collapse; it is a
function of redshift.

Duffy et al. (2008) [28] present a relation for c_{vir} in general clusters, citing results
obtained from N-body simulations conducted using WMAP5 cosmology:

$$\bar{c}_{vir}(sim) = \frac{7.85}{(1+z)^{0.71}} \left(\frac{M_{vir}}{2.78 \times 10^{12} M_\odot} \right)^{-0.081} \tag{2.23}$$

We consider this relation at $z = 0.45$, for consistency with the lensing-selected
relation below. Oguri et al. (2012) [66] present a relation for c_{vir} in lensing-selected
clusters, using ray tracing to estimate the effect of lensing bias:

$$\bar{c}_{vir}(z = 0.45) \approx 6.3 \left(\frac{M_{vir}}{5 \times 10^{14} h^{-1} M_\odot} \right)^{-0.2} \tag{2.24}$$

In order to compare our data to these predictions, we chose a range of values of
M_{vir} and used Equations 2.23 and 2.24 to find the corresponding values for c_{vir}. We
then used the relations in Johnston et al. [47] and Hu and Kravtsov [44] to convert
from c_{vir} and M_{vir} to c_{200} and M_{200}. Finally we used the range of values for M_{200}
and the predicted values for c_{200} to find predicted values for Einstein radius (θ_E) by

using the NFW profile (see Equation 2.25 below). We plotted the relations between M_{200} and θ_E as the general and lensing-selected predictions in Figures 2.11 and 2.12.

To find a predicted Einstein radius we used the NFW density profile, expressed as

$$\rho(r) = \frac{\rho_s}{(r/r_s)(1 + r/r_s)^2} \tag{2.25}$$

where r is the distance from the center of the cluster, ρ_s is a characteristic density, and r_s is the scale radius, given by $r_s = r_{200}/c_{200}$. We implemented Equation 13 in Wright and Brainerd [95], an equation that describes surface mass density $\overline{\Sigma}_{NFW}$ in the NFW model. The Einstein radius θ_E is given implicitly by the solution of [62]:

$$\overline{\Sigma}_{NFW}\left(\frac{\theta_E}{r_s}\right) = \Sigma_{crit} \tag{2.26}$$

where the critical surface mass density Σ_{crit} is

$$\Sigma_{crit} = \frac{c^2}{4\pi G}\frac{D_s}{D_d D_{ds}} \tag{2.27}$$

Thus we found Einstein radius by solving for $\overline{\Sigma}_{NFW}$ and using that to find θ_E.

2.5.2   Consideration of the Overconcentration Problem

The final result of our analysis is shown in Figures 2.11 and 2.12. Figure 2.11 shows the relation between M_{200} and θ_E for our measured values of M_{200} while Figure 2.12 shows the relation for the new M_{200} values that come from the scaled-up richness values. We consider Figure 2.12 to be more reliable as it uses richness

values scaled to correspond with values from SDSS data, which was used to calibrate the mass-richness relation.

In Figure 2.11 there is a noticeable disagreement between our data and the predicted relations. It can also be seen that the lower-mass clusters disagree more while the higher-mass clusters fit the predictions better, as found by other authors. However in Figure 2.12, we see that all clusters are shifted to higher masses by an average factor of 2.0. In the plot of the scaled values, we see that many of the clusters now closely follow the lensing-selected prediction. There are still four clusters that do not fit the predicted relations. These clusters are SDSS J0901+1814, SDSS J1038+4849, SDSS J1343+4155 and SDSS J1537+6556, which are the lowest mass clusters in our sample. SDSS J1318+3942, which is also among the lowest mass clusters, is found close to the predicted line, but still slightly above it.

We determined values for c_{200} for our clusters by using our measured values for M_{200} and θ_E in Equations 2.25 and 2.26; values are listed in Table 2.4. We estimated errors on c_{200} by varying M_{200} and θ_E to the maximum and minimum values allowed by their respective error bars. Maximum values for c_{200} were found with minimum M_{200} and maximum θ_E while minimum values for c_{200} were found with the opposite. For smaller values of M_{200}, this led to very large upper error bars on c_{200} as a very high concentration parameter would then be required to achieve the large Einstein radius.

Our measurements of c_{200} follow the trends noted earlier: for many of the clusters, our measured values of c_{200} are within the range of predictions, but for the lowest mass clusters measured values of c_{200} are higher than predictions. The average value for c_{200} predicted for our scaled values of M_{200} by Equation 2.23 (for general clusters) is 3.4 while the average value predicted by Equation 2.24 (for lensing-selected clusters) is 5.7. The average of our ten measured values of c_{200} is 7.3, which is slightly
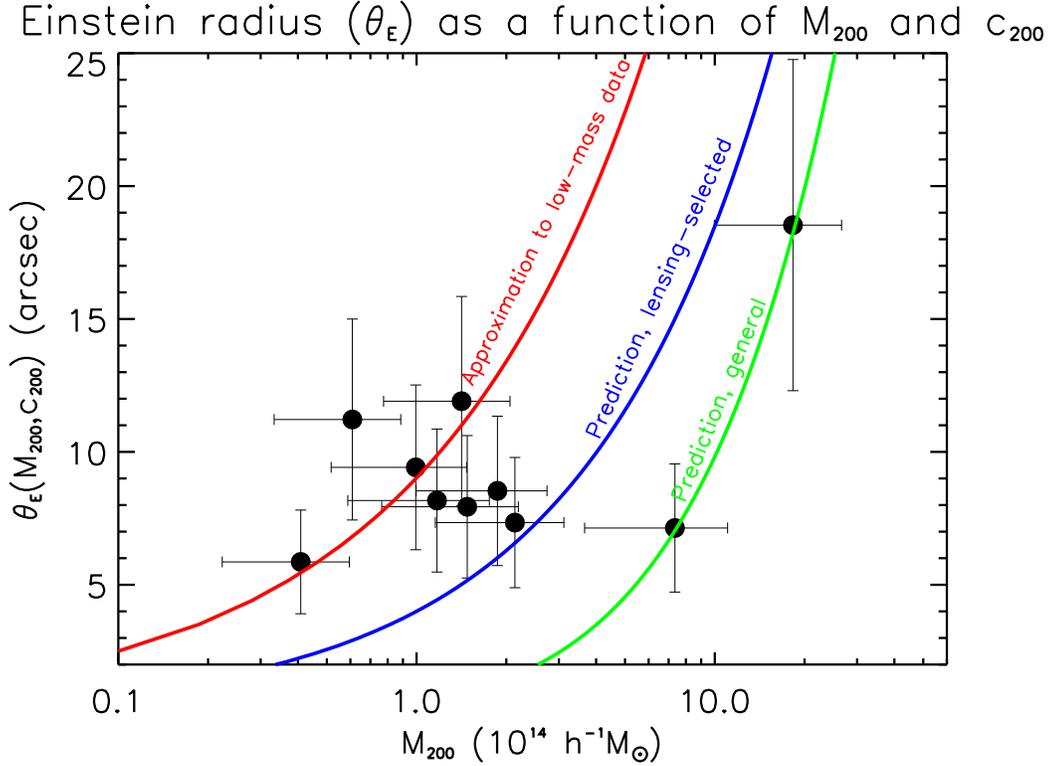
Figure 2.11: A plot of Einstein radius versus M_{200} for unscaled M_{200} values, with Einstein radii scaled to fiducial redshifts. The theoretical lines come from a prediction of Einstein radii for given M_{200} and c_{200} values found by using an NFW [63] fit to the mass and concentration. The general clusters line was found using predicted c_{200} values found from Equation 2.23 and the lensing-selected clusters line was found using predicted c_{200} values found from Equation 2.24. The average c_{200} for general clusters is 3.6 and for lensing-selected clusters it is 6.4. Both Equations took $z = 0.45$. The approximate fit line was found by multiplying the values of c_{vir} resulting from Eq. 2.24 by 1.9. We tried different factors to multiply c_{vir} until the resultant line went approximately through the low mass data points.
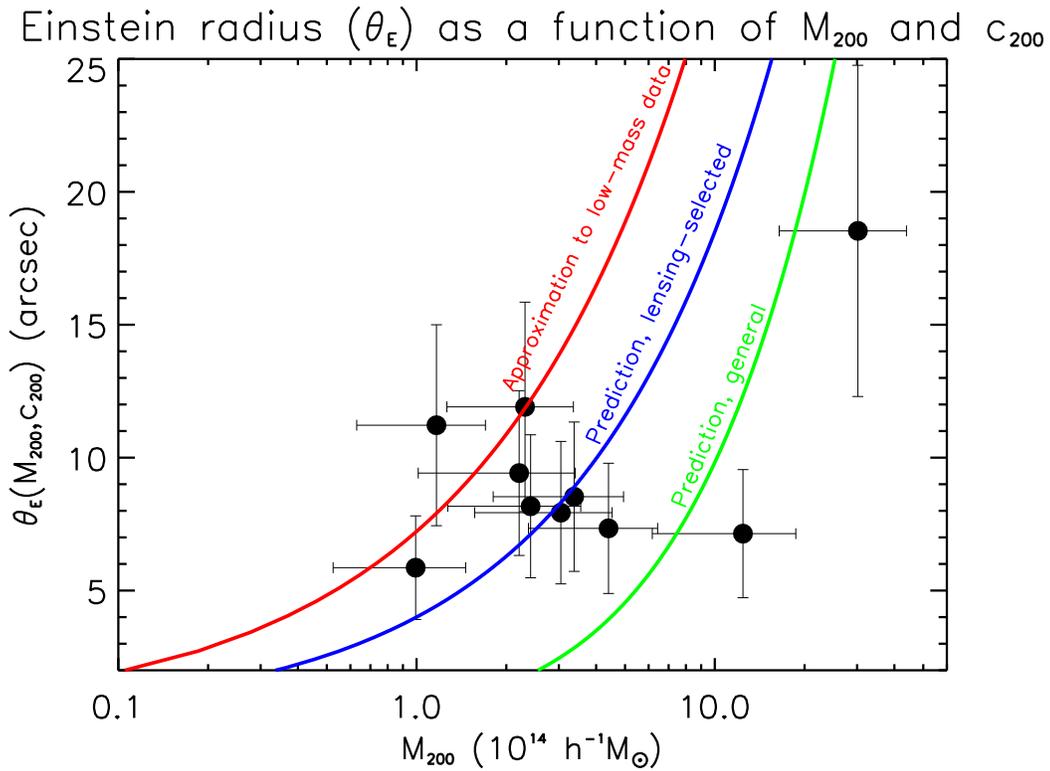
Figure 2.12: The same plot as Figure 2.11 but using M_{200} values found by scaling richness values up using Eq. 2.12. The average c_{200} for general clusters is 3.4 and for lensing-selected clusters it is 5.7. To find the approximate fit to the low mass data, we multiplied all c_{vir} values from Eq. 2.24 by 1.5. Note that with scaled richness values, the points all move closer to the predicted values.

larger than the lensing-selected prediction. However for our four lowest mass clusters the average c_{200} value is 11.6, much larger than the lensing-selected prediction. The four clusters we identify as overconcentrated above have the following values for c_{200}: for SDSS J0901+1814, $c_{200} = 9.6^{+13}_{-3.5}$; for SDSS J1038+4840, $c_{200} = 17^{+73}_{-7.8}$; for SDSS J1343+4155, $c_{200} = 9.1^{+22}_{-3.9}$; and for SDSS J1537+6556, $c_{200} = 11^{+15}_{-3.9}$. These clusters have respectively M_{200} of 0.99, 1.2, 2.3 and $2.2 \times 10^{14} h^{-1} M_\odot$, which are the lowest masses in our sample.

Concentration parameters (c_{vir}) based on strong and weak lensing measurements are provided in Oguri et al. (2012) [66] for two of these four clusters. We convert these to c_{200} using the method discussed in §2.5.1. For SDSS J1038+4840, $c_{200} = 33.8^{+0.00}_{-18.3}$ and for SDSS J1343+4155, $c_{200} = 4.25^{+1.38}_{-0.790}$. Thus for SDSS J1038+4840, the second lowest mass cluster in our sample, both sets of measurements find this cluster to be significantly overconcentrated. For SDSS J1343+4155 the evidence for overconcentration is not as strong.

In Figures 2.13 and 2.14 we consider the mass-concentration relation, comparing $\log(c_{200})$ to $\log(M_{200})$. Figure 2.13 is the mass-concentration relation for our measured values of c_{200} and M_{200} without scaling and Figure 2.14 is this relation using scaled richness values. We also include three lines in Figures 2.13 and 2.14: the blue solid line is the prediction from Oguri et al. (2012) [66] for lensing-selected clusters, the green solid line is the best-fit to the data in the Oguri paper (Equation 26 in [66]) and the red dotted line is the best fit to our data. Equation 26 in [66] is:

$$\overline{c_{vir}} = (7.7 \pm 0.6) \left(\frac{M_{vir}}{5 \times 10^{14} h^{-1} M_\odot} \right)^{-0.59 \pm 0.12} \tag{2.28}$$

We used the same method as discussed in §2.5.1 to add the prediction and best fit from Oguri et al. (2012) [66] to Figures 2.13 and 2.14. For the predicted line, we

applied Equation 2.24 and for the best fit from Oguri et al. (2012) [66] we applied Equation 2.28. In Figure 2.13 the slope is $\alpha = 0.45 \pm 0.30$ while in Figure 2.14 $\alpha = 0.45 \pm 0.23$. Note that the error bars are larger on c_{200} in Figure 2.13; this is because when calculating error bars, the minimum M_{200} was small and maximum θ_E was large, leading to very large values for c_{200}. Fedeli et al. (2011) [32] suggests that for clusters that are not overconcentrated, α should be no larger than 0.2. At 1σ, our lowest value of α is 0.15 for unscaled values and 0.22 for scaled values. Both of these values are consistent with clusters that are not overconcentrated, again suggesting that most of our clusters are not overconcentrated. Prada et al. (2011) [68] suggest in their Figure 12 that $\log(c_{200})$ should be less than about 0.8 at $z = 0.5$. This is again consistent with most of our clusters, although not for the lowest mass clusters. Note in Figure 2.14 that the four lowest mass clusters have values of $\log(c_{200})$ above 1.0 which suggest that these clusters are overconcentrated.

We find in Figure 2.13 that our data points are mostly above the predicted line, suggesting many of our clusters are overconcentrated. However when we use the more reliable scaled values in Figure 2.14 we find that most of the clusters are found near the predicted line, but the lowest mass clusters (the four identified above) remain above the prediction. This again confirms our previous statement that most of our clusters do not appear to be overconcentrated, but there is evidence for overconcentration at lower cluster masses.

Thus for most of our clusters, ΛCDM seems to match their observed properties. But for our several clusters showing evidence of overconcentration, what does the overconcentration problem suggest is happening in galaxy clusters? It seems to suggest that clusters are collapsing more than ΛCDM would predict [16, 32, 66]. The dark matter halo associated with a galaxy cluster is expected to have undergone an adiabatic collapse during the formation of the cluster. The baryonic matter in the
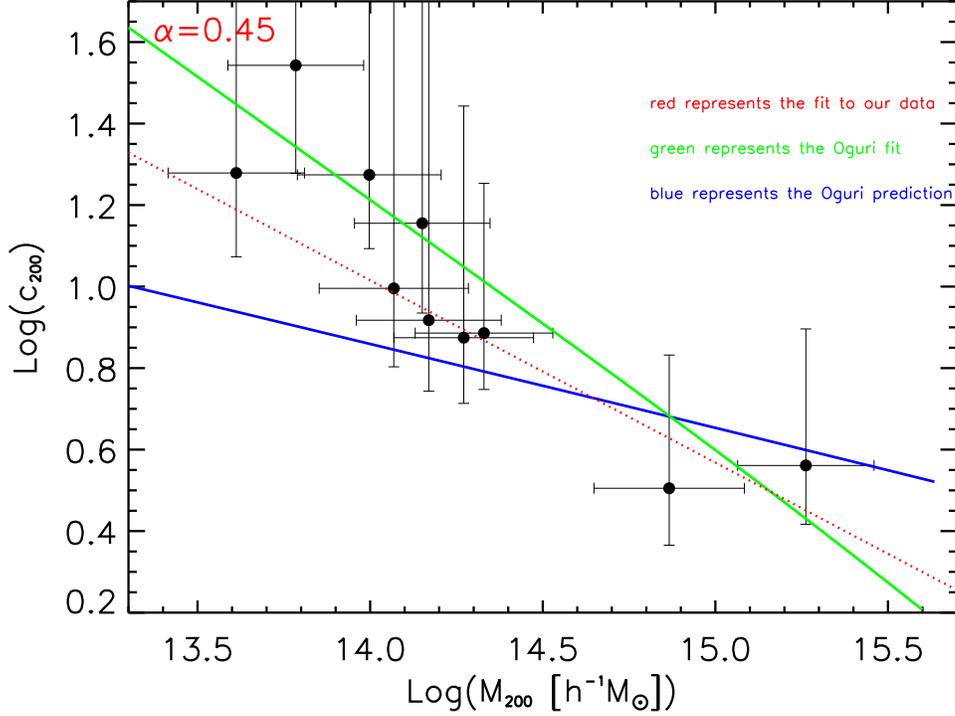
Figure 2.13: A plot of the logarithm of the concentration parameter c_{200} versus logarithm of M_{200}. This is modeled on Figure 1 in [32]. The red line is the best fit to the data and has a slope $\alpha = 0.45 \pm 0.30$. The solid green line is the fit to the data in [66] (their Equation 26). The solid blue line is the Oguri prediction for lensing-selected clusters (Equation 2.24). The large vertical error bars arise on the low-mass clusters due to how c_{200} changes as a function of M_{200} and θ_E. We found the upper vertical error bars on c_{200} by setting M_{200} and θ_E to their minimum and maximum values, respectively. When M_{200} is very small, a very large value for c_{200} is required to achieve the large value for Einstein radius.
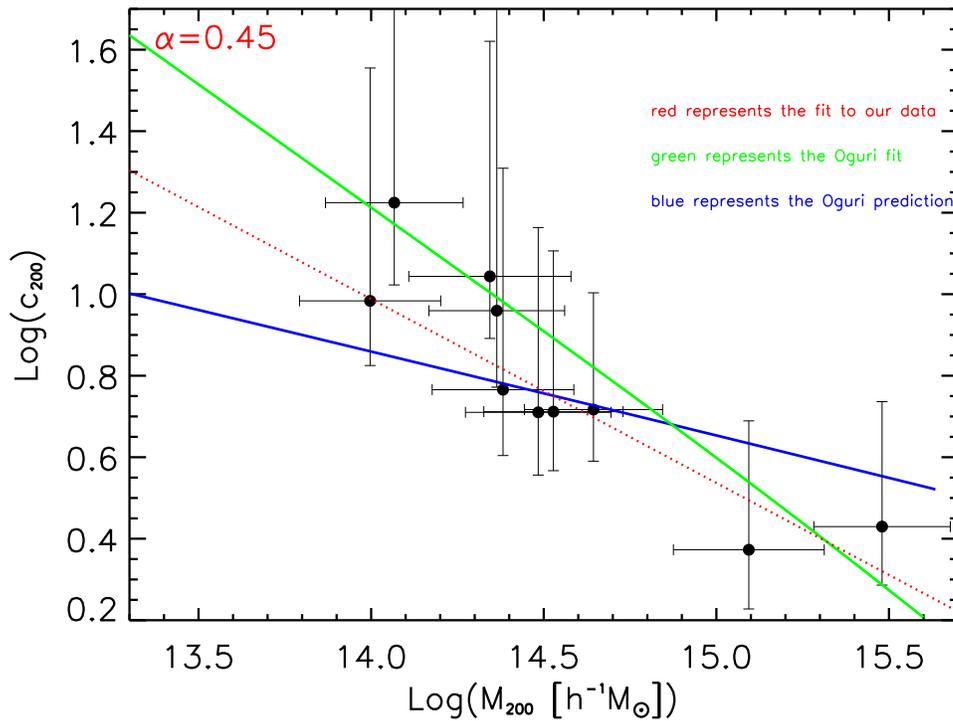
Figure 2.14: The same plot as Figure 2.13 but made with M_{200} and c_{200} values corresponding to richness values scaled up using Eq. 2.12. The slope of the best fit line is $\alpha = 0.45 \pm 0.23$. Note that the values for M_{200} have been shifted to the right and thus many of the points fit the predicted relations now. However the lowest mass points still do not match the predictions.

cluster (concentrated in the BCG) would also have collapsed. The baryonic matter would likely have dragged the dark matter along with it, augmenting the collapse of the halo. Since we find some clusters to be more concentrated than expected, it may be that the halo collapsed more than expected due to the contribution of the baryons. It is suggested [32, 66] that the overconcentration is most significant in lower-mass clusters because in these clusters the BCG makes up a larger percentage of the overall cluster mass. Thus the baryons would contribute to the halo collapse more in a lower-mass cluster than in a higher-mass cluster.

CHAPTER 3

INVESTIGATING THE RICHNESS EXTRAPOLATION

There are more things in heaven and earth, Horatio, than are dreamt of in your philosophy.–from *Hamlet* by William Shakespeare

3.1   Introduction

As we mentioned in Chapter 2, we measured richness for the ten clusters by counting all of the galaxies meeting certain criteria. However we found that for several of the ten systems, some regions of the cluster were not in the image. Thus we used Equation 2.3 in order to extrapolate the total values for N_{gals} and N_{200}. In this section we describe how we took further images of the clusters in order to completely image all cluster galaxies, to obtain more reliable richness values and to assess the accuracy of the richness extrapolation method.

In § 3.2 we describe how we planned and completed the new observations at Apache Point Observatory. In § 3.3 we describe the data reductions and calibrations done to prepare the data images for science analysis. In § 3.4 we detail how we identified the characteristic colors of the clusters and measured the additional cluster galaxies not found in the original imaging. In § 3.5 we present a final determination of the accuracy of the richness extrapolation method. Finally in § 3.6 we present new imaging of a foreground galaxy cluster found in the APO images.

3.2   Data Acquisition and Reduction

3.2.1   Planning the Observations

We proposed to Apache Point Observatory for 4 half-nights on the APO 3.5-m telescope. APO assigns observing in half-nights, so one group observes for the first half of the night and a different group observes for the second half. The proposal was accepted, but we only obtained 3 half nights, on February 20, March 16 and March 18 of 2012. For the first two nights, we would be observing for the first half of the night (sunset to about midnight) and for the last night we would have a second half (about midnight to sunrise). Unfortunately on March 18 the observatory was closed due to snow on the mountain, so we were only able to take data on two of the three half-nights.

Observing at APO is usually done remotely, as long as at least one of the observers is trained for remote observing. As H. Lin and S. Allam had been trained, we observed remotely from Fermilab. Remote observing is done through the program TUI, or Telescope User Interface, which was written by Russell Owen at the University of Washington. TUI produces a graphical user interface through which all elements of telescope function can be controlled. One technician works at the Observatory during the night to address any issues that cannot be resolved remotely.

The APO 3.5-m telescope is a Ritchey-Chretien reflector on an alt-azimith mount. A Ritchey-Chretien telescope is a Cassegrain telescope that uses a hyperbolic mirror for both the primary and secondary mirrors in order to minimize spherical abberations. An alt-azimuth mount moves along two perpendicular axes, one vertical and one horizontal. This is as opposed to an equatorial mount, in

which one of the axes is aligned with the celestial equator. There are a number of instruments which can be used on this telescope, including several spectrometers, a spectrograph and an optical imaging camera called SPIcam (Seaver Prototype Imaging camera). We used SPIcam for these observations.

SPIcam uses a single CCD of dimensions 2048x2048 pixels and has pixel scale of $0.14''$ per pixel. However since this is usually much better than typical seeing, the CCD is typically operated in binned mode, giving pixel scale of $0.28''$ per pixel. SPIcam has a field of view of $4.7'$ on each side. The original images of the clusters taken at the WIYN telescope used the Mini-Mosaic imager which has 2 CCDs each with field of view of $9.6'$. Thus SPIcam provides a much smaller view of the sky and more images would be required in order to observe the entire cluster.

In order to plan our observations, we placed circular regions files corresponding to an area with a radius of r_{200} on the WIYN images in astronomical image viewer DS9. (A "region" in DS9 is a shape that can be made any size that can be superimposed on an image. We used both circular and square regions to help plan our observing.) We then tried to cover as much of the area off the CCD with square regions $4.7'$ on a side, corresponding to the size of SPIcam. With some of the images, we could fit most of the missing area with 1 or 2 pointings of the APO telescope, while with others we needed up to 9. In Figure 3.1, we show an image of SDSS J0901+1814 and SDSS J1439+3250, showing how we used square regions to fit the missing area.
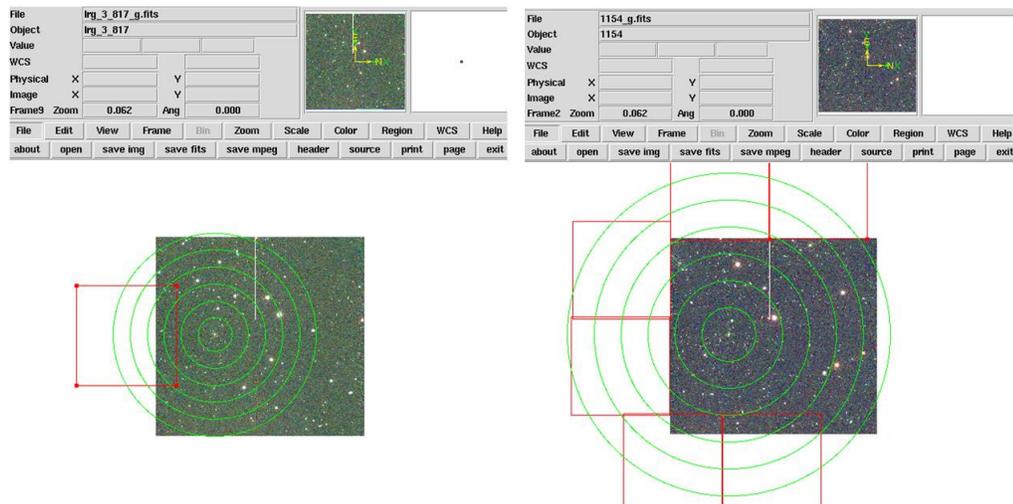
Figure 3.1: An image of SDSS J0901+1814 and SDSS J1439+3250 showing how we designed pointings for APO followup of the WIYN images. Note that we could mostly cover the missing area of SDSS J0901+1814 with only one pointing at APO while we required six pointings to mostly cover the missing area of SDSS J1439+3250. (SDSS J0901+1814 was imaged, but observations of SDSS J1439+3250 did not ultimately fit into the two half nights we had.) There are six circles presented as these are the six annuli used for Equation 2.3.

3.2.2 Details of the Observations

We determined that we could cover most of the missing area in the WIYN images in the 3 half-nights we were allotted by using 29 pointings of the telescope, assuming that we could complete one pointing every 30 minutes. A pointing in this case consisted of slewing the telescope to a particular set of celestial coordinates and then taking images in three filters, SDSS g, r and i filters. We needed to use three different filters as information on g-r and r-i colors is critical to using the red-sequence method for cluster galaxy identification. In each filter, we took two 210 second exposures, dithering by $15''$ between each of the two exposures. Dithering (moving the telescope slightly) is often done in astronomical observations in order to identify cosmic rays and to avoid placing celestial objects permanantly on a bad column on the CCD. During image processing, the dithered exposures are stacked so bad columns or cosmic rays can be rejected in comparisons between the two images, and the total exposure time is $210 \, s \times 2 = 420$ seconds.

At the beginning of each half night we always took calibration images, including bias frames, dome flats and sky flats. Bias frames are 0 second integrations, meaning the CCD is not exposed to any light, but it is read out. This allows the observer to identify noise that occurs during CCD readout. We took 10 bias frames and then co-added them before using them for data reduction. We took 5 dome flats in each of the three filters and then the same number of sky flats. A dome flat is an image taken by pointing the telescope at a screen illuminated by a highly stable and uniform light source in the observatory dome. A sky flat is an image of the twilight sky, which is fairly evenly illuminated by the recently-set Sun.

We used the centers of the square regions to determine the right ascension and declination of the point where each planned pointing of the telescope would be centered. We proceeded to image each of the fields in order, beginning with the fields with the lowest airmass (see § 1.2). Due to several limiting factors, we only were able to complete imaging of 15 of the planned 29 fields. The limiting factors included the following: (1) as previously mentioned, we lost the March 18 observations; (2) airmasses for all fields were too high to begin observing any of them right at the end of twilight on March 16 and so we repeated a previous observation; (3) on the first night, I miscalculated the center of one of the pointings, requiring us to redo it; and (4) we had an occasional slight overrun on the planned 30 minutes per pointing. The 15 fields observed are listed in Table 3.1. We also list median seeing in this table; astronomical **seeing** describes the quality of the images taken on a particular night. Mathematically, seeing is the full-width half-max (FWHM) of the point-spread function (PSF) of stars in an image.

3.3   Data Reduction

Before science analysis can be completed on CCD images, they must be reduced and calibrated. As described at greater length in my Master's thesis [93], data reduction and calibration consists primarily of the following steps:

1: Basic reductions

2: WCS correction

3: Bad pixel removal

4: Image stacking

Table 3.1: The observing parameters for the data taken at APO. The number given after the name indicates the image number, as for some clusters several different fields were imaged. RA and DEC are of the center of each field. RA is given in the form (hours:min:sec) and DEC in (deg:min:sec). Stated values of airmass and seeing are median values for each image. The upper objects are those observed on February 20 and the lower ones are those observed on March 16.

System	RA	DEC	Airmass	Seeing ($''$)
SDSS J0900+2234-1	09:00:02.992	+22:30:54.74	1.62	1.24
SDSS J0901+1814-1	09:01:22.013	+18:11:00.71	1.34	0.965
SDSS J0957+0509-1	09:57:47.383	+05:05:02.00	1.50	1.11
SDSS J0957+0509-2	09:57:29.853	+05:05:01.11	1.35	0.955
SDSS J1038+4849-1	10:38:43.035	+48:45:12.72	1.47	1.03
SDSS J1209+2640-1	12:09:52.658	+26:40:22.47	1.61	1.31
SDSS J1209+2640-2	12:09:48.246	+26:35:36.95	1.43	1.08
SDSS J1209+2640-3	12:09:27.287	+26:34:52.67	1.29	1.39
SDSS J1209+2640-4	12:09:06.168	+26:34:53.18	1.18	1.25
SDSS J0900+2234-1	09:00:02.992	+22:30:54.74	1.088	1.050
SDSS J1209+2640-5	12:08:53.778	+26:39:12.05	1.61	1.60
SDSS J1209+2640-6	12:08:53.959	+26:43:57.72	1.31	1.12
SDSS J1209+2640-7	12:09:08.579	+26:48:02.63	1.19	1.15
SDSS J1209+2640-8	12:09:30.411	+26:48:06.57	1.11	1.20
SDSS J1209+2640-9	12:09:51.871	+26:45:12.67	1.059	1.34
SDSS J1318+3942-1	13:18:11.757	+39:39:19.93	1.10	1.18

5:   Magnitude calibration

We will describe each of these steps in brief.

3.3.1   Basic reductions

Basic reductions used on the APO data included overscan removal, bias subtraction and flat fielding. Basic reductions were done using the NOAO Image Reduction and Analysis Facility (IRAF). Overscan is a small section of the CCD that contains no science data but does provide information on the noise or bias level produced during readout of the CCD. The bias level is the values measured in each pixel in the absence of real data. The constant overscan value is subtracted from the whole image and then the overscan region is removed.

A bias frame is a readout of the CCD with zero integration time and the shutter closed. Its purpose is to identify pixel-to-pixel variations in the bias level of the CCD. In other words, different pixels may have different bias levels during the readout of the CCD; since the bias level varies spatially across the image, the removal of the constant value of the overscan may leave some pixels with more bias. Thus we take bias frames to identify the variations in response to readout. It is usual to take a number of bias frames and then take the average of them before subtracting them from the data images. This is done as a single bias frame may have random variations in it which would introduce more noise into the data images. Such random variations would be eliminated when the average is taken; thus a master bias or zero frame is ultimately subtracted from the data image.

Flat fielding is the use of uniformly illuminated images to remove pixel-to-pixel variations in images. In other words, flat fielding helps to ensure that every pixel

in the image will give the same output measurement for the same input flux. We combined each of the dome flats taken in the same filter and each of the sky flats taken in the same filter and produced master dome flats and master sky flats in each filter. We then divided each data image by the master flats in that filter.

3.3.2 WCS Correction

World Coordinate System (WCS) correction is a correction applied across an image to ensure that the coordinate measurements put into the images at the telescope closely match standard catalogs. If there are any differences, then coordinates in data images are corrected. We completed this step using the program called SCAMP [10]. SCAMP is one of a series of astronomical analysis programs written by E. Bertin that are called *Astromatic* programs. The Astromatic programs each have a different main function, but their operation is very similar in that the programs are invoked from the command line while referring to a configuration file.

In order to run SCAMP, we needed to first run another Astromatic program called SourceExtractor (SExtractor). This program identifies the sources (stars and galaxies) in an image along with their locations, magnitudes and more. To run SCAMP, we first had to run SExtractor to produce FITS_LDAC files. FITS stands for Flexible Image Transport System and is the typical format for astronomical images; LDAC stands for Leiden Data Analysis Center, and is a data format developed by E. Bertin and others especially for use in large astronomical surveys. SCAMP ultimately finds a solution to update all positions to correspond to standard catalogs and updates image headers with this information.

3.3.3 Bad Pixel Removal

Two types of pixels need to be identified and removed, cosmic rays and bad pixels. Cosmic ray removal was done using the program LACosmic [84]. LACosmic uses a Laplacian algorithm to find and remove cosmic rays based on the characteristic way cosmic ray hits transition from background to maximum measured flux. LACosmic is provided in several different codes, but we used the implementation in IRAF Command Language (CL).

CCDs often have some number of bad pixels, pixels that do not respond linearly to input flux. Typically bad pixels in data images are removed by using a bad pixel map to identify them during processing. We made a bad pixel mask by taking flat-field images of two different exposure lengths. On March 16 we took dome flats in the r-band, integrating for 6 seconds a total of 5 times and then combining the 5 exposures. We did this again while integrating for only 3 seconds, and combining each of these 5 exposures. We then took the ratio of the short exposure flat-field to the long exposure flat-field. This method works because in the ratio image, good pixels will all have almost exactly the same value since good pixels will measure more flux in the long exposure than in the short exposure. Bad pixels will not have the same ratio as the good pixels and can be identified on this basis. Finally we used the IRAF task `CCDMASK` to convert the ratio image to a bad pixel mask.

There were some other bad pixels which were not found using this method, and so these were masked out by hand, by making a list of the coordinates of bad pixels. The images made from these pixels were combined with the images made using the ratio of flat fields to make a single bad pixel map, one for each of the g, r and i-filters. These bad pixel maps were fed into the stacking program.

3.3.4 Image Stacking

After bad pixels and cosmic rays were removed, the two 210 second dithered exposures in each filter could be stacked. Stacking was done with the Astromatic program called SWARP [10]. SWARP takes the input data images in FITS format and makes necessary shifts and interpolations to combine the images. The output is a single image in each filter with an equivalent exposure time of 420 seconds.

SExtractor is then run on the final stacked images to identify sources and measure their magnitudes and other quantities. This time SExtractor is run to produce ASCII output. Each object is listed by SExtractor in the output data file along with all the other quantities that it had been set to measure. However thus far the magnitudes that SExtractor outputs are not yet actual object magnitudes but are only instrumental magnitudes.

3.3.5 Magnitude Calibration

Instrumental magnitudes are magnitudes measured based on fluxes observed in a particular instrument. These fluxes have to be compared to the actual fluxes from some standard objects in order to convert instrumental magnitudes to calibrated magnitudes. This is often done by observing a standard star during the observing session, that is a star whose actual magnitude is well known. We instead converted our instrumental magnitudes measured by SExtractor to calibrated magnitudes by referring to magnitudes measured in the SDSS. We did this by querying the SDSS Catalog Archive Server (CAS) at http://cas.sdss.org for stars in the fields we observed. We then found the magnitudes measured for these stars by SExtractor and

the magnitudes measured for the same stars in the SDSS and found the median differences. We were able to match stars between the APO and SDSS data by using the IDL program CLOSE_MATCH_RA_DEC [46] which compares celestial coordinates to make a list of matching objects. Median differences were calculated between the magnitudes measured in APO data and those measured in the SDSS for each field in each filter.

A sample plot of magnitude differences is shown in Figure 3.2 and a full list of offsets is provided in Table 3.2. Errors reported in Table 3.2 were found by calculating Poisson errors (σ_o is the standard deviation of the differences between Sloan and APO magnitudes and N_o is the number of data points):

$$error = \frac{\sigma_o}{\sqrt{N_o}} \tag{3.1}$$

Once magnitude offsets were found, they were added to magnitudes measured in APO data, converting them from instrumental magnitudes to calibrated magnitudes. All magnitudes reported henceforth in this chapter include these offsets.

We finally also include in Figure 3.3 an image of SDSS J1209+2640 with circles of radius r_{200}. The two images show WIYN imaging alone and WIYN imaging with APO imaging superimposed, indicating how much of the cluster was covered. Both WIYN and APO images are color composites made of images taken in g, r and i filters. We note that there are still small gaps in imaging; we quantify the area of these gaps in Section 3.5.1. We also note that in several of the APO images there are small dark regions near the center of the images: this was due to condensation on the camera during February 20 observations; we saw no indication that this would significantly bias our analyses and so did not pursue further efforts to remove them.
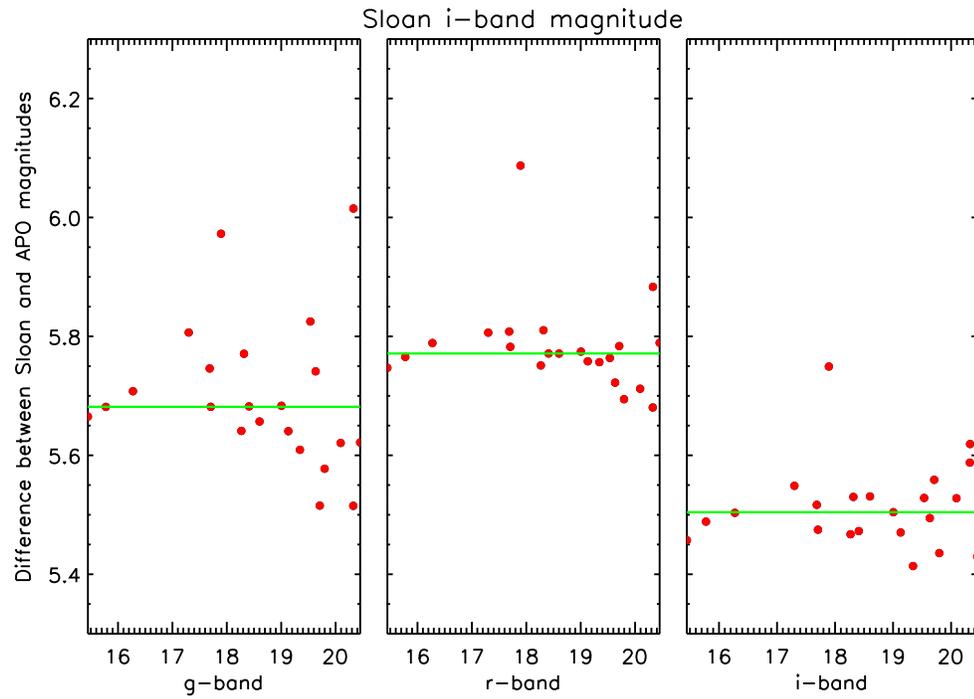
Figure 3.2: A plot of magnitude offsets (Sloan magnitude - APO instrumental magnitude) for SDSS J0900+2234.
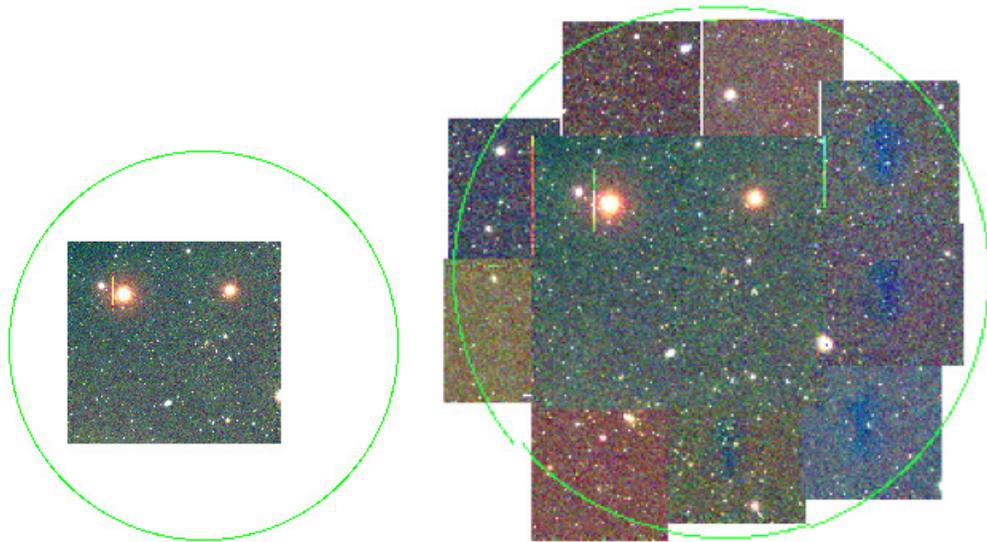
Figure 3.3: An image of SDSS J1209+2640 before (left) and after (right) APO imaging. In the "after" image we superimpose APO imaging onto the WIYN image. The large circle is a circle of radius r_{200}.

Table 3.2: The offsets added to SExtractor magnitudes in each band to convert them to calibrated magnitudes. These were found from SDSS stars.

System	Offset in g-band	Offset in r-band	Offset in i-band
SDSS J0900+2234-1	5.443 ± 0.036	5.560 ± 0.023	5.380 ± 0.018
SDSS J0901+1814-1	5.450 ± 0.026	5.584 ± 0.014	5.359 ± 0.013
SDSS J0957+0509-1	5.452 ± 0.028	5.582 ± 0.018	5.374 ± 0.016
SDSS J0957+0509-2	5.511 ± 0.091	5.596 ± 0.030	5.373 ± 0.011
SDSS J1038+4849-1	5.484 ± 0.26	5.605 ± 0.19	5.345 ± 0.19
SDSS J1209+2640-1	5.414 ± 0.026	5.546 ± 0.016	5.343 ±0.040
SDSS J1209+2640-2	5.524 ± 0.011	5.600 ± 0.016	5.411 ± 0.013
SDSS J1209+2640-3	5.424 ± 0.035	5.561 ± 0.025	5.389 ± 0.018
SDSS J1209+2640-4	5.538 ± 0.017	5.602 ± 0.014	5.393 ± 0.0094
SDSS J0900+2234-1	5.682 ± 0.027	5.771 ± 0.017	5.504 ± 0.016
SDSS J1209+2640-5	5.587 ± 0.069	5.723 ± 0.019	5.499 ± 0.018
SDSS J1209+2640-6	5.734 ± 0.16	5.762 ± 0.24	5.521 ± 0.30
SDSS J1209+2640-7	5.747 ± 0.12	5.807 ± 0.058	5.537 ± 0.026
SDSS J1209+2640-8	5.693 ± 0.029	5.740 ± 0.014	5.487 ± 0.020
SDSS J1209+2640-9	5.749 ± 0.044	5.765 ± 0.022	5.510 ± 0.017
SDSS J1318+3942-1	5.712 ± 0.031	5.738 0.019	5.502 ± 0.025

3.4 Measuring Richness

Once we had catalogs listing calibrated object magnitudes, we could apply the machinery developed for the WIYN images (see Chapter 2) to identify cluster galaxies found in these fields. We wrote an IDL program to apply (as in Chapter 2) 5 criteria:

1: r-i color of galaxy is less than 2σ different from cluster r-i color

2: g-r color of galaxy is less than 2σ different from cluster g-r color

3: galaxy is within radius of 1 $h^{-1}Mpc$ or r_{200}

4: galaxy is no dimmer than 0.4L*

5: galaxy meets cuts to be considered a galaxy and not a star

3.4.1 <u>Finding APO Red Sequence Colors</u>

Ordinarily to find the red sequence r-i and g-r colors for galaxy clusters we would make a color-magnitude plot and look for the emergence of a line of galaxies of constant color. We could not do that with the APO data because only small sections of the clusters were present in the images, since by design we only imaged the outer regions of the clusters that were not imaged in the WIYN data. We also could not simply use colors previously identified in WIYN data, as colors have some dependence on observing site and instruments. Finally, we could not always directly find the difference in magnitudes between WIYN and APO data as for some images

there is almost no overlap between WIYN and APO. Thus we began with SDSS cluster data, which covered all of the clusters we investigated here.

We had identified red sequence colors in SDSS data for all ten WIYN clusters. We had also identified red sequence colors in WIYN; see Table 3.3 for these values. In order to convert from SDSS colors to WIYN colors to APO colors, we used the following equation (c always stands for color):

$$c_{APO} = c_{SDSS} + (c_{APO} - c_{SDSS}) + (c_{WIYN} - c_{SDSS}) \qquad (3.2)$$

We use this equation in order to convert from SDSS colors to WIYN colors to APO colors, allowing us to use the complete coverage of SDSS but still make a comparison between WIYN and APO. The differences between WIYN colors and SDSS colors, i.e. $(c_{WIYN} - c_{SDSS})$), were found by comparing the red sequence colors from the two different data sets for the clusters. The quantities called $(c_{APO} - c_{SDSS})$ were found by matching galaxies brighter than $i = 21.5$ in SDSS and APO data and then finding the median difference in both g-r and r-i colors. Matching was done using the routine *CLOSE_MATCH_RA_DEC.pro*. A plot of color difference vs. SDSS color was made while calculating the median color difference; a sample plot is shown in Figure 3.4. Final values for APO colors are shown in Table 3.4.

3.4.2   Finding Cluster Galaxies in APO Images

We finally wrote a program called *APO_richness.pro* in IDL that read in cluster centers of all clusters, the size of 1 h^{-1} Mpc or r_{200} in degrees, the value for 0.4L*, the APO g-r and r-i colors for each image, and so on. It then applied the cuts discussed at the beginning of this section to identify which galaxies were cluster

Table 3.3: WIYN and Sloan colors for the clusters observed at APO and differences between colors. Note that c_{WIYN}-c_{SDSS} means the difference between colors in WIYN data and SDSS data.

Cluster	WIYN g-r	WIYN r-i	SDSS g-r	SDSS r-i	g-r $c_{WIYN}-$ c_{SDSS}	r-i $c_{WIYN}-$ c_{SDSS}
SDSS J0900+2234	1.83	0.73	1.60	0.83	0.23	-0.10
SDSS J0901+1814	1.72	0.52	1.67	0.53	0.05	-0.01
SDSS J0957+0509	1.78	0.71	1.51	0.74	0.27	-0.03
SDSS J1038+4849	1.72	0.62	1.64	0.69	0.08	-0.07
SDSS J1209+2640	1.79	0.93	1.65	0.95	0.14	-0.02
SDSS J1318+3942	1.73	0.73	1.65	0.78	0.08	-0.05
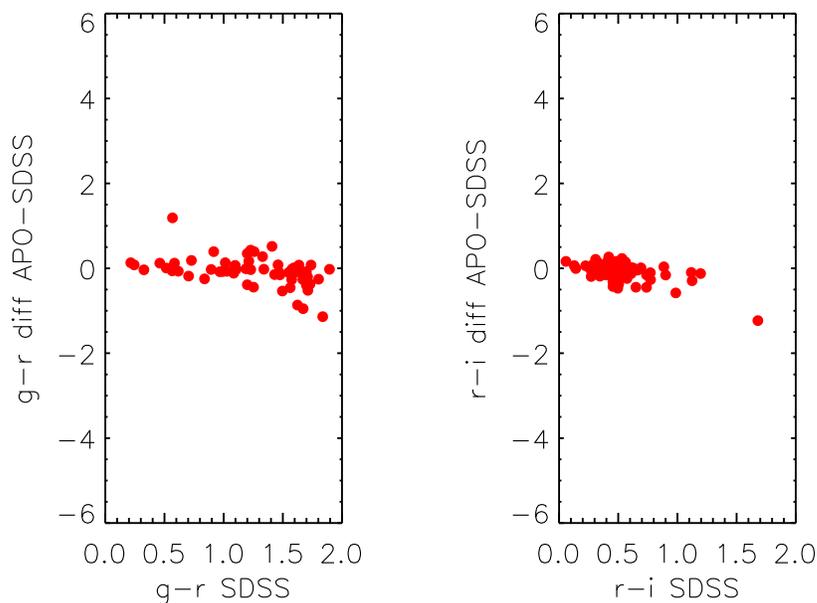
Figure 3.4: A plot of SDSS color vs. color difference ($c_{APO} - c_{SDSS}$) for SDSS J1209+2640, Image 1.

Table 3.4: Differences between APO and Sloan galaxy magnitudes and measured APO colors.

Cluster	g-r $c_{APO} - c_{SDSS}$	r-i $c_{APO} - c_{SDSS}$	APO g-r	APO r-i
SDSS J0900+2234-1	-0.0041 ± 0.033	-0.047 ± 0.027	1.83	0.73
SDSS J0901+1814-1	-0.13 ± 0.024	-0.00026 ± 0.017	1.59	0.52
SDSS J0957+0509-1	-0.014 ± 0.023	0.0035 ± 0.021	1.77	0.71
SDSS J0957+0509-2	-0.039 ± 0.028	-0.014 ± 0.023	1.74	0.70
SDSS J1038+4849-1	-0.013 ± 0.041	0.033 ± 0.018	1.71	0.65
SDSS J1209+2640-1	-0.032 ± 0.026	-0.019 ± 0.021	1.76	0.91
SDSS J1209+2640-2	0.044 ± 0.027	-0.032 ± 0.020	1.83	0.90
SDSS J1209+2640-3	-0.065 ± 0.040	-0.048 ± 0.031	1.73	0.88
SDSS J1209+2640-4	-0.017 ± 0.032	-0.029 ± 0.021	1.77	0.90
SDSS J0900+2234-1	0.0047 ± 0.043	0.0012 ± 0.028	1.83	0.68
SDSS J1209+2640-5	0.0070 ± 0.025	-0.043 ± 0.020	1.80	0.89
SDSS J1209+2640-6	0.078 ± 0.022	-0.053 ± 0.015	1.87	0.88
SDSS J1209+2640-7	0.022 ± 0.023	-0.028 ± 0.017	1.81	0.90
SDSS J1209+2640-8	-0.0076 ± 0.027	-0.039 ± 0.017	1.78	0.89
SDSS J1209+2640-9	0.024 ± 0.019	-0.037 ± 0.019	1.81	0.89
SDSS J1318+3942-1	0.022 ± 0.029	-0.067 ± 0.024	1.75	0.66

galaxies and which ones were not. In addition this routine matched APO galaxies to galaxies identified as cluster galaxies in WIYN that overlapped with the APO images. Thus we were able to identify how effective the routine was at identifying cluster galaxies since we had some regions of the images in which we knew which galaxies should be considered cluster galaxies.

For the WIYN-APO overlap region of each APO image, there were three classes of galaxy:

1. Galaxies considered cluster galaxies in APO and in WIYN (**matched**)

2. Galaxies considered cluster galaxies in APO only (**APO only**)

3. Galaxies considered cluster galaxies in WIYN only (**missed**)
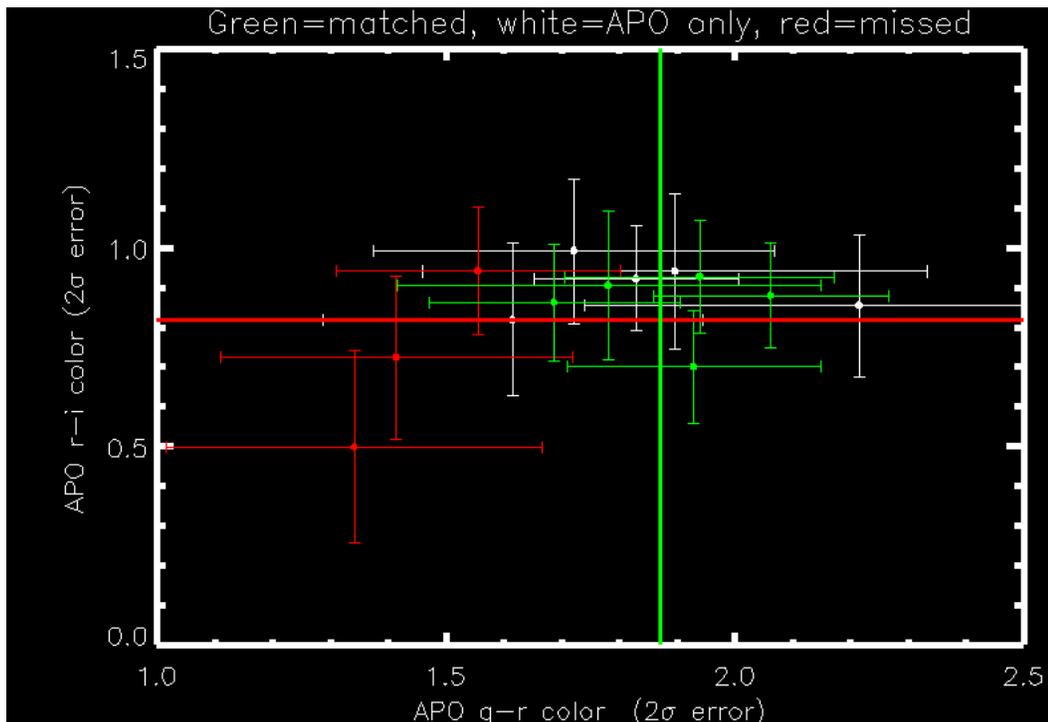
Figure 3.5: The output plot from *APO_richness.pro* for SDSS J1209+2640, Image 6. This shows galaxies matched with cluster galaxies in WIYN (green) considered cluster galaxies in APO only (white) and considered cluster galaxies in WIYN only (red). The green vertical line and red horizontal line are the cluster g-r and r-i color, respectively.

For each cluster, the program made a plot of these three types of galaxies, shown in Figure 3.5. By varying the values of APO red sequence color up and down, we identified what cluster colors produced the highest rate of matched galaxies and the lowest possible rates of APO only or missed galaxies. Thus the final values for APO colors are slightly different than those given in Table 3.4. We list the numbers of each of the three types of galaxies as well as the final values for red sequence colors in Table 3.5.

Considering the total numbers in each category, we find that we missed as many galaxies as we overcounted in the overlap region, leading us to believe that the

Table 3.5: The numbers of galaxies that are (1) matched to WIYN cluster galaxies, (2) found in APO only or (3) found in WIYN only. We also provide here the final values for cluster colors, found by maximizing the number of matches between WIYN and APO cluster galaxies. The quantities Δg-r and Δr-i mean the difference between these final colors and the measured colors stated in Table 3.4 (i.e., these values-previous).

Cluster	Number Matched	Number APO only	Number Missed	APO g-r	APO r-i	Δg-r	Δ r-i
SDSS J0900+2234-1	5	3	5	1.84	0.76	0.01	0.03
SDSS J0901+1814-1	2	0	1	1.60	0.55	0.01	0.03
SDSS J0957+0509-1	1	0	0	1.76	0.71	-0.01	0
SDSS J0957+0509-2	1	0	1	1.73	0.70	-0.01	0
SDSS J1038+4849-1	0	0	0	1.69	0.65	-0.02	0
SDSS J1209+2640-1	2	0	1	1.73	0.91	-0.03	0
SDSS J1209+2640-2	1	2	1	2.03	0.83	0.20	-0.07
SDSS J1209+2640-3	0	0	0	1.76	0.90	0.03	0.02
SDSS J1209+2640-4	0	0	0	1.79	0.92	0.02	0.02
SDSS J0900+2234-1	1	1	1	1.83	0.73	0	0.05
SDSS J1209+2640-5	1	0	0	1.80	0.90	0	0.01
SDSS J1209+2640-6	5	5	3	1.87	0.82	0	-0.06
SDSS J1209+2640-7	3	2	2	1.80	0.91	-0.01	0.01
SDSS J1209+2640-8	8	2	0	1.77	1.02	-0.01	0.13
SDSS J1209+2640-9	0	0	1	1.81	0.89	0	0
SDSS J1318+3942-1	2	1	0	1.75	0.66	0	0
TOTAL	**32**	**16**	**16**	-	-		

method used to identify cluster galaxies would give a reasonable estimate in the region beyond the overlap.

So we finally applied all cuts to the galaxies in the APO images and excluded any galaxies that had previously been counted in WIYN. Thus we obtained richness values for APO data only, values of both N_{gals} and N_{200} (see Table 3.6).

3.5 The Accuracy of the Area Extrapolation

Finally we can use these data to consider the accuracy of the area extrapolation used in 2. We obtain final richness values for each cluster by adding the cluster galaxies found only in APO data to cluster galaxies found only in WIYN data. We can then compare these values to those obtained by using Equation 2.3 to extrapolate richness values (see Table 3.8).

However, before we make final conclusions, there are two more factors to consider: First, even with APO imaging, there are still small parts of the clusters for which no imaging was obtained; second, we must quantify uncertainties in our richness measurements.

3.5.1 How Much Area was Missed in APO?

In order to determine how much cluster area was still missed even with the new imaging, we used DS9 regions as discussed in Section 3.2.1. We superimposed square regions that matched the size of the APO images over circular regions representing areas of radius r_{200} or 1 h^{-1} Mpc and looked for any space that was not covered. If space was not imaged in either APO or WIYN, then we fit that area with a polygonal

Table 3.6: Final numbers of cluster galaxies found in APO data but not in WIYN data. By N_{gals} and N_{200}, we mean galaxies meeting criteria to be cluster galaxies and within 1 h^{-1} Mpc or r_{200}, respectively, but only present in APO images.

Cluster	N_{gals}	N_{200}
SDSS J0900+2234-1	0	0
SDSS J0901+1814-1	1	0
SDSS J0957+0509-1	1	1
SDSS J0957+0509-2	1	2
SDSS J1038+4849-1	0	0
SDSS J1209+2640-1	0	4
SDSS J1209+2640-2	0	1
SDSS J1209+2640-3	0	14
SDSS J1209+2640-4	0	6
SDSS J0900+2234-1	0	0
SDSS J1209+2640-5	0	10
SDSS J1209+2640-6	0	8
SDSS J1209+2640-7	0	4
SDSS J1209+2640-8	0	3
SDSS J1209+2640-9	0	4
SDSS J1318+3942-1	0	1
TOTAL	**3**	**58**
Total By Cluster		
SDSS J0900+2234	0	0
SDSS J0901+1814	1	0
SDSS J0957+0509	2	3
SDSS J1038+4849	0	0
SDSS J1209+2640	0	54
SDSS J1318+3942	0	1

Table 3.7: Comparison of values for N_{gals} and N_{200} from APO and WIYN data combined (*Measured*) and from WIYN only, using extrapolation (*Predicted*). The values given for ΔN_{gals} or ΔN_{200} are those from combined richness minus those from extrapolated richness (i.e., $N_{A+W} - N_{ex.}$). Thus negative values indicate that the extrapolated value was higher than observed data while positive values indicate the extrapolated value was lower than the observed data.

Cluster	Measured N_{gals}	Predicted N_{gals}	ΔN_{gals}	Measured N_{200}	Predicted N_{200}	ΔN_{200}
SDSS J0900+2234	28	28	0	29	30	-1
SDSS J0901+1814	15	15	0	11	11	0
SDSS 0957+0509	30	29	1	36	36	0
SDSS J1038+4849	15	16	-1	14	15	-1
SDSS J1209+2640	101	101	0	199	214	-15
SDSS J1318+3942	23	24	-1	25	25	0

region. The advantage of polygonal regions is that they can have as many vertices as necessary to fit a region of unusual size. An image of how we did this for SDSS J1209+2640 is shown in Figure 3.6. Note that in this figure we display both an area of radius r_{200} and the nine APO pointings (circle and squares, respectively). The solid green areas represent the area for which there is still no image information.

Once we fit polygonal regions to the areas off of the images, we used a Fortran routine from C. Peng [67] called *Fillpoly.f*. This routine will take the vertices of a polygonal DS9 region and fill in the area with pixels. *Fillpoly.f* was designed to create masks when using Galfit (which we describe in Chapter 4). We did this once for each individual piece of missing area and then counted pixels in each piece, finally summing them all. This gave us the total area not imaged in units of pixels. Finally we found r_{200} and $1\ h^{-1}$ Mpc in pixels and found the total number of pixels in the entire cluster area by using $A = \pi r^2$, with r in pixels. By comparing the number of pixels in the missing area to the number of pixels in the total area we were able to find the percentage of coverage of the cluster.
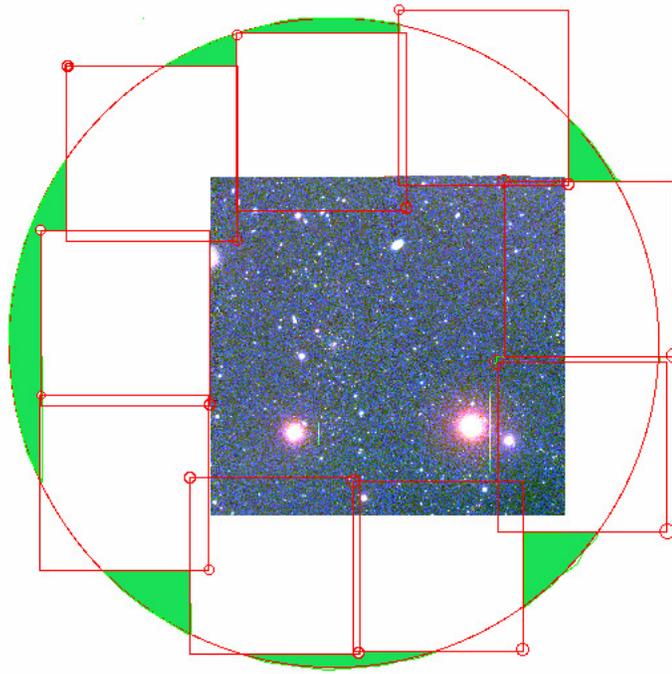
Figure 3.6: The WIYN image for SDSS J1209+2640 with several regions superimposed. The large red circle represents an area of radius r_{200} and the red squares represent the locations of the nine pointings for this cluster done at APO. The solid green areas are the areas for which there is still no image data.

3.5.2   <u>Consideration of Errors</u>

Finally we consider errors both in values *predicted* by the extrapolation equation and in values *measured* from WIYN and APO data. First, for WIYN data we consider the error in richness values in a slightly different way than in Section 2.3.2.2. Previously we found the error on richness by finding richness values in 3 different SExtractor magnitude apertures (`MAG_AUTO`, 2″`MAG_APER` and 3″`MAG_APER`). We then found the standard deviation in the 3 values and took this as the error on the richness value. When we extrapolated richness values using Equation 2.3, we found error by using this equation, assuming that error scales up with richness (subscript *ac* means area corrected, using Equation 2.3):

$$N_{200\_ac\_err} = N_{200\_ac} \frac{N_{200\_error}}{N_{200}} \tag{3.3}$$

Another approach (which we use here) would be to take the error in richness in two pieces, error in richness values *on* the CCDs and error in richness values extrapolated *off* the CCDs. For error in richness on CCDs we continue to use the standard deviation in measurements in different apertures, but we also add another element to track intrinsic scatter in richness, taking the square root of each data value. Thus

$$N_{gals\_on\_CCD\_err} = \sqrt{\sigma\left(Ngals\_values\right)^2 + \left(\sqrt{N_{gals}}\right)^2} \tag{3.4}$$

For the area off the CCD, we note that the number of galaxies found off the CCD in Equation 2.3 depends on the fraction of cluster area on the CCD, so the standard deviation does as well. Thus we write down the standard deviation as

(where subscript A represents *on the CCD*, subscript B represents *off*, T_i means total area of i_{th} annulus, A_i means area of annulus on CCD and N_{Ai} means number of cluster galaxies observed in WIYN data in i_{th} annulus):

$$\sigma_{N_{gals\_off\_CCD}} = \sqrt{\sum_{i=1}^{6} N_{Ai} \left(\frac{T_i - A_i}{A_i} \right)} \qquad (3.5)$$

We finally add in quadrature both the richness uncertainty for on the CCD (Eq. 3.4) and that for off the CCD (Eq. 3.5)). The error bars on richness predictions in Figure 3.7 and error values in Table 3.8 are produced in this way.

To find errors on richness values combining APO and WIYN measurements, we add in quadrature Poisson errors (square root of data value) for APO measurements to errors from Equation 3.4.

3.5.3   Final Values

In Table 3.8 we present the total imaging coverage of each cluster and also the total number of galaxies found from WIYN and APO data. To obtain these numbers we took the total number of galaxies counted in WIYN and in APO and scaled these numbers up to reflect the amount of area still not imaged, following the equation

$$\frac{N_{gals}(observed)}{Percent\ imaged} = \frac{N_{gals}(total)}{100\%} \qquad (3.6)$$

In Figure 3.7 we compare the total richness measurements taken from APO and WIYN data ("measured"; on x-axis) to predictions from Equation 2.3 ("predicted"; on y-axis). The green line is the y=x line and errors are 2σ, with σ being the total error found following the methods of Section 3.5.2. Note that the observed

Table 3.8: Total imaging coverage of each cluster and final results for measured and extrapolated richness.

Cluster	Total Coverage for N_{gals}	Total Coverage for N_{200}	Measured N_{gals}	Predicted N_{gals}	Measured N_{200}	Predicted N_{200}
SDSS J0900+2234	100%	99.4%	28 ± 6.2	28 ± 6.2	29 ± 7.1	30 ± 7.1
SDSS J0901+1814	96.5%	100%	16 ± 5.2	15 ± 5.1	11 ± 3.6	11 ± 3.6
SDSS J0957+0509	100%	100%	30 ± 8.9	29 ± 8.8	36 ± 6.8	36 ± 6.7
SDSS J1038+4849	99.1%	100%	15 ± 4.0	16 ± 4.0	14 ± 3.9	15 ± 3.9
SDSS J1209+2640	100%	93.5%	101 ± 13.2	101 ± 13.2	213 ± 16.1	214 ± 27
SDSS J1318+3942	99.5%	99.1%	23 ± 4.9	24 ± 5.0	25 ± 5.9	25 ± 5.8

final numbers closely follow predictions, with all data points falling on the y=x line. Thus we conclude that the extrapolation relation, Equation 2.3, is highly accurate as the measurements obtained using the APO follow-up observations match the predictions exquisitely.

3.6   New Imaging of a Foreground Cluster

In our analysis of images of SDSS J1209+2640, we noticed in Image 9 (taken on 16 Mar) that there was a collection of galaxies surrounding a bright elliptical galaxy, but that none of them were marked as cluster members for SDSS J1209+2640. Upon further investigation we determined that this was in fact a separate cluster that happens to overlap with SDSS J1209+2640 but is a foreground object. A search of the NASA/IPAC Extragalactic Database (NED) showed that this cluster was known, although as far as we can determine this is the best imaging that exists for this cluster. This cluster was part of two 2010 cluster catalogs based on SDSS Data Release 6, J. Hao's gmBCG cluster catalog [39] and Wen, Han &
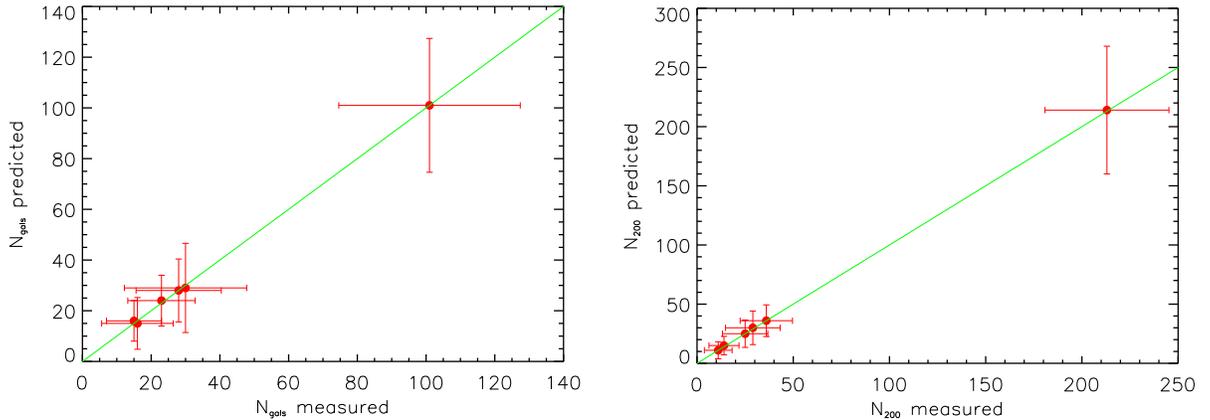
Figure 3.7: A comparison of richness values measured in APO data after corrections for missing area ("measured") against richness values predicted by WIYN measurements with area extrapolation ("predicted"). The green line is the y=x line.

Li's cluster catalog [92]. We refer to this cluster by the name from the first paper, SDSS J120947.0+264354. This paper gives a photometric redshift of this cluster as z=0.321209. Since the (spectroscopic) redshift of SDSS J1209+2640 is z=0.5580, it is clear that SDSS J120947.0+264354 is a foreground object, not associated with SDSS J1209+2640.

We found an image of this location in the SDSS, this is shown in Figure 3.8. At RA and DEC of (12:09:46.952, +26:43:55.89) there is what appears to be the BCG for this cluster. It shows evidence of being composed of two separate elliptical galaxies, although they blend together. Another slightly bluer galaxy is found very near the BCG(s). A number of smaller elliptical galaxies are clustered around the BCG(s).

In Figure 3.9 we show a section of the APO image of this cluster. Finally, in Figure 3.10 we present the WIYN image of SDSS J1209+2640 next to the APO image of SDSS J1209+2640, Image 9. The circles on the images represent an area of radius 1 h^{-1} Mpc. Note that the circle showing 1 h^{-1} Mpc is larger for SDSS
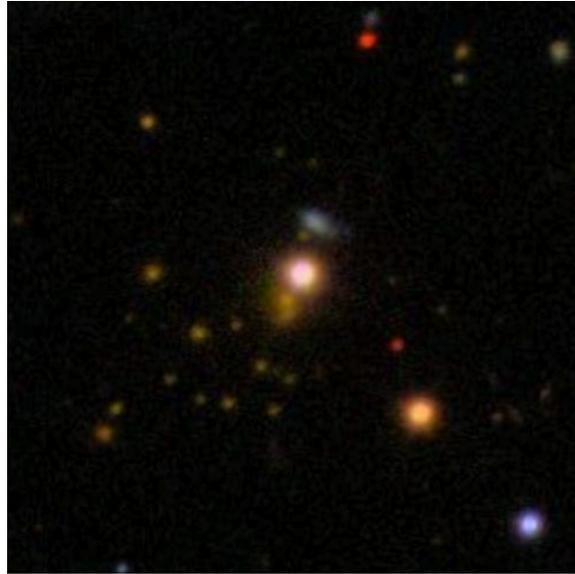
Figure 3.8: An image of SDSS J120947.0+264354 from the SDSS.

Figure 3.9: An image of SDSS J120947.0+264354 from the APO data, Image 9 of SDSS J1209+2640. This image has dimensions of $100''$ square. Note the BCG near the center, with a bluer galaxy above it.
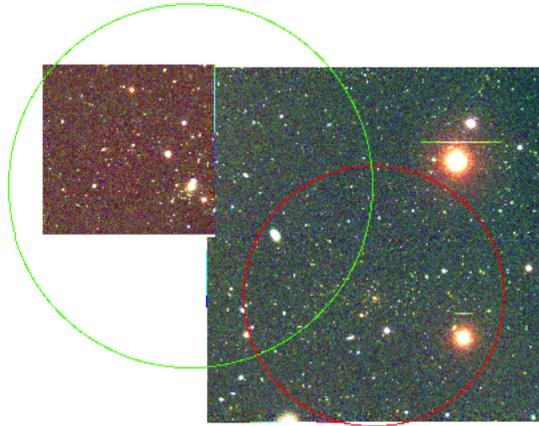
Figure 3.10: An image of SDSS J120947.0+264354 in the APO image of SDSS J1209+2640, Image 9 (upper left) placed next to the WIYN image of SDSS J1209+2640. The superimposed circles indicate an area of radius 1 h^{-1} Mpc.

J120947.0+264354 since this cluster is at lower redshift (and therefore closer to us).

CHAPTER 4

DEVELOPING A HIGHER-REDSHIFT

MASS-RICHNESS RELATION

What could be more beautiful than the heavens, which contain all beautiful things?–
Nicolaus Copernicus

4.1   Introduction

4.1.1   Mass-Richness Relations

As discussed in Chapter 1, when observing a galaxy cluster, a number of properties can be measured directly, including cluster richness (the number of galaxies in the cluster), the brightness of each of the cluster galaxies, the morphologies of cluster galaxies and the galaxy star-forming rate. However some important quantities cannot be measured directly but must be inferred from measurable properties. Mass is one such quantity. Thus instead we measure quantities that can be related to mass; these are called *mass proxies*. Cluster richness is commonly used as a mass proxy, but in order for it to give meaningful results, the relation between mass and richness must be calibrated. This calibration is referred to as a mass-richness relation.

Methods of determining cluster mass are discussed more in § 1.5.1. As discussed, methods of finding cluster mass include X-ray temperature, cluster velocity disper-

sion, frequency bias caused by the Sunyaev-Zeldovich effect and tangential shear caused by weak lensing. Weak lensing can cause a systematic change in the ellipticities of many galaxies behind the galaxy clusters (see Chapter 1); this observed change is called *shear*. As discussed in Chapter 1, only tangential shear is produced by weak lensing, so we expect no signal in orthotangential or cross shear. Since the observed shear is related to cluster mass, weak lensing can provide a measure of cluster mass. However often weak lensing nearby individual clusters does not provide the statistics for a significant measure of mass. Therefore we can use **stacked weak lensing**, where shear is measured for galaxies nearby multiple clusters within a particular range of richnesses and redshift and the shear signal is added together for each richness and redshift bin. The larger numbers of galaxies make it possible to obtain a reliable measurement of mass. Tangential shear is found for a combination of galaxy clusters and source galaxies, where source galaxies are galaxies nearby the cluster in projection, but at a larger redshift than the cluster. The source galaxies are the galaxies whose shape is measured, and the mass associated with the clusters is the cause of the tangential shear that describes the shape variation.

We observed stacked weak lensing shear as well as cluster richness in three data sets: the coadd of Stripe 82 of the Sloan Digital Sky Survey (SDSS), the Blanco Cosmology Survey (BCS) and the Dark Energy Survey (DES) Science Verification (SV) data. These are important results because all of these data sets reach to higher redshift (median ≈ 0.6) than previous surveys such as the Sloan Digital Sky Survey (median $z = 0.25$), where the best previous measurement of the mass-richness relation was done.

In § 4.1.2 we describe previous measurements of mass-richness relations at lower redshift. In § 4.2 we describe the properties of our three data sets. In § 4.3 we describe measurements of cluster richness. In § 4.4 we describe how stacked weak

lensing shear measurements were done and how mass was found from shear re-sults. In § 4.5 we describe systematics that introduced uncertainties into our mass measurements and how we compensated for them. In § 4.6 we describe null tests conducted to check that the weak lensing shear signal is real, and will not appear when there is no source of signal. In § 4.7 we present our results for mass-richness relations for the Stripe 82 coadd including each of the systematics individually and then altogether. In § 4.8 we present mass-redshift and mass-concentration relations for the Stripe 82 coadd. Finally in § 4.9 we present mass-richness relations for BCS and DES-SV data.

4.1.2   Previous Work

Johnston et al. (2007) [47] and Sheldon et al. (2007) [76] present average shear profiles from stacked weak lensing measurements around about 130,000 galaxy clus-ters at median redshift of 0.25 found in the SDSS. These clusters were taken from the maxBCG cluster catalog [52]. For richness, they measure N_{200}, the number of cluster galaxies inside radius r_{200}; r_{200} is given by our Equation 2.4. To fit weak lensing shear, they find the quantity $\Delta\Sigma$, which is [95]:

$$\Delta\Sigma = \overline{\Sigma(x)} - \Sigma(x) = \Sigma_{crit}\gamma \tag{4.1}$$

Here $x = r/r_s$, where r is the radius and r_s is the scale radius (see 1.7.2.3). $\overline{\Sigma(x)}$ is the mean surface mass density inside x and $\Sigma(x)$ is surface mass density at position x. Σ_{crit} is the critical surface mass density, defined in Chapter 1. Finally γ is the observed shear.

Johnston et al. fit the shear profiles with a mathematical expression that includes all of the major expected contributions to the shear. Their expression is

$$< \Delta\widehat{\Sigma(R)} >= \frac{M_0}{\pi R^2} + p_c \langle \Delta\Sigma_{NFW} \rangle(R) + (1-p_c)\langle \Delta\Sigma^2_{NFW} \rangle(R) + B\Delta\Sigma_l + \Delta\Sigma_{NL} \quad (4.2)$$

The brackets around $\Delta\widehat{\Sigma(R)}$ indicate that this quantity is averaged over the probability distribution of halo masses. The first term describes the mass of the central brightest cluster galaxy (BCG). The second term is the contribution from the dark matter halo, described by an NFW profile (see Section 4.2). The quantity p_c is the fraction of correctly centered halos; this needs to be considered because the BCG is taken to be the cluster center, but in some fraction of clusters, the BCG will not be at the center of the halo. Johnston et al. give

$$p_c(N_{200}) \equiv \frac{1}{1 + exp(-q)} \quad (4.3)$$

where

$$q = ln(1.13 + 0.92(N_{200}/20)) \quad (4.4)$$

Thus the second term includes the correctly centered halos and the third term includes the miscentered halos. In the third term the quantity $\Delta\Sigma^2_{NFW}$ is the NFW profile from the dark matter halos where the BCG is miscentered and $1 - p_c$ is the fraction of halos that contain a miscentered BCG. The fourth term describes the contribution from neighboring dark matter halos. $\Delta\Sigma_l$ is the shear signal produced by neighboring halos and B is the bias parameter:

$$B \equiv b(M_{200}, z)\Omega_m \sigma_8^2 D(z)^2 \quad (4.5)$$

The term $b(M_{200}, z)$ is the linear bias parameter and $D(z)$ is the linear growth factor, which is a function of the scale factor a. Finally the fifth term in Equation 4.2 is the contribution from non-linear shear. These different contributions to the measured shear signal have different strengths at different radii. Figure 8 in Johnston et al. (our Figure 4.1) shows the contribution of each quantity as a function of radius. After finding cluster masses (M_{200}) as a function of richness, Johnston et al. fit a power relation to their results. Their final mass-richness relation is

$$M_{200}(N_{200}) = M_{200|20}(N_{200}/20)^{\alpha_N} \tag{4.6}$$

where

$$M_{200|20} = (8.8 \pm 0.4_{stat} \pm 1.1_{sys}) \times 10^{13} h^{-1} M_\odot \tag{4.7}$$

and

$$\alpha_N = 1.28 \pm 0.04 \tag{4.8}$$

Simet et al. (2011) [77] present measurements of weak lensing shear in the SDSS Stripe 82 (see Section 4.2.1). They use the same maxBCG catalog as Johnston et al. but they use only clusters found in the Stripe 82 coadd. The clusters, being a subset of the clusters used in Johnston et al., have the same median redshift of ≈ 0.25, but more source galaxies can be observed as the Stripe 82 as the magnitude limits are $18 < i < 24$. Simet et al. consider many of the same systematics as Johnston et al., including halo miscentering. They consider miscentering of the BCG by creating a set of mock catalogs from their data. These catalogs preserve galaxy positions, shape errors, photometric redshifts and more but replace actual source galaxy shears with expected shears from a shear model. Since the shear model is produced by a halo model, they control how many of the halos have a miscentered BCG and can
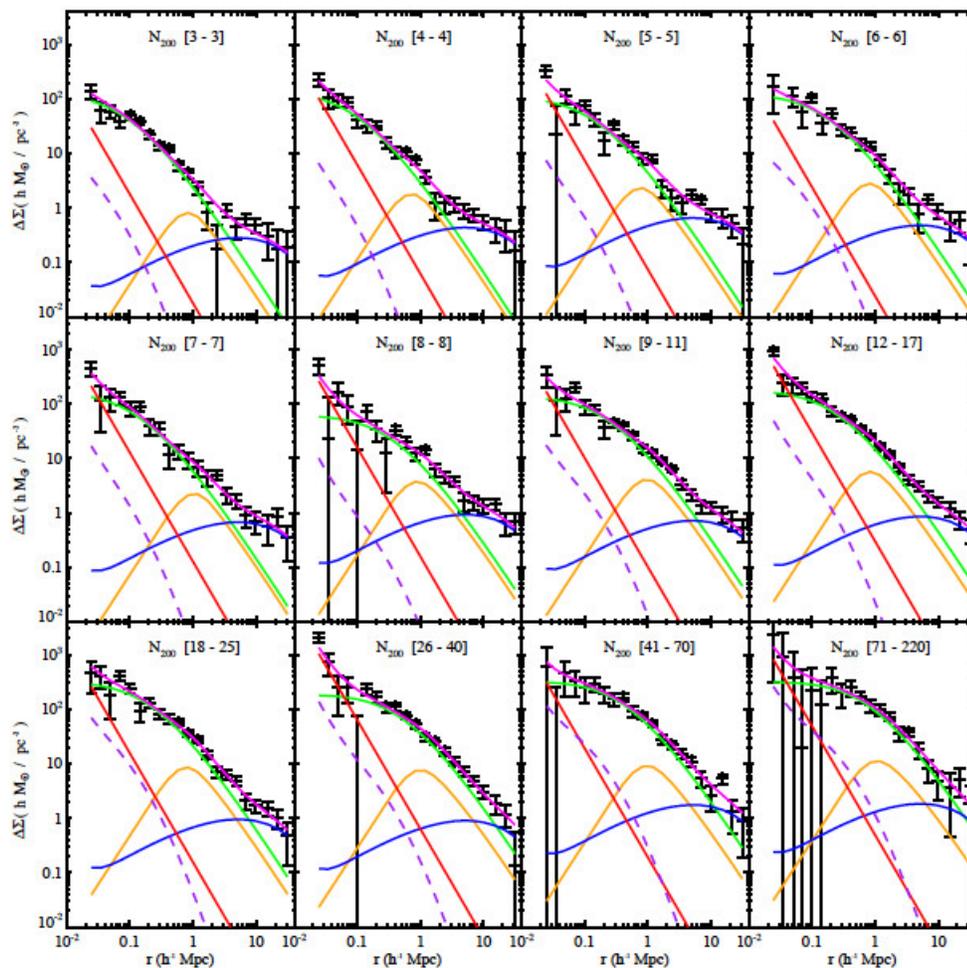
Figure 4.1: Johnston et al. (2007) [47] Figure 8, showing the contribution of different elements of their shear model (Equation 4.2) at different radii for 12 different N_{200} richness bins. Red is the central BCG (term 1), green is the NFW halo profile (term 2), orange is the miscentered halo component (term 3), blue is the neighboring halos (term 4) and purple is the non-linear contribution (term 5). Magenta curves show the sum of these components.

obtain information on the actual halo masses. They then fit an NFW profile to the mock data and find that the measured mass is underestimated due to miscentering of the BCG. They then find a relationship between the richness and the scale of miscentering:

$$\frac{M_{200,true}}{M_{200,mis}} = 1.44 \pm 0.17 \left(\frac{N_{200}}{20}\right)^{-0.21 \pm 0.18} \tag{4.9}$$

They also find a mass-richness relation that is a power law, with

$$M_{200|20} = (9.56 \pm 0.75) \times 10^{13} h^{-1} M_\odot \tag{4.10}$$

and

$$\alpha = 1.10 \pm 0.12 \tag{4.11}$$

4.2 The Data Sets

4.2.1 Stripe 82 of the Sloan Digital Sky Survey

The Sloan Digital Sky Survey (SDSS) began in 1998 and sought to image 10,000 deg^2 in the North Galactic Cap and to take spectroscopy of one million galaxies and one hundred thousand quasars in this same region [4]. The SDSS uses a 2.5-m telescope located at Apache Point Observatory. The SDSS uses a camera with 24 2048x2048 CCDs of pixel scale $0.396''$. Stripe 82 is a section of the footprint of the Sloan Digital Sky Survey (SDSS). The complete footprint of the SDSS is shown in Figure 4.2; note that Stripe 82 falls along the celestial equator between declinations of $-1.25°$ and $+1.25°$. Stripe 82 was imaged repeatedly during the fall when the North Galactic Cap was not observable. The main SDSS area reaches $r \approx 22.4$ and
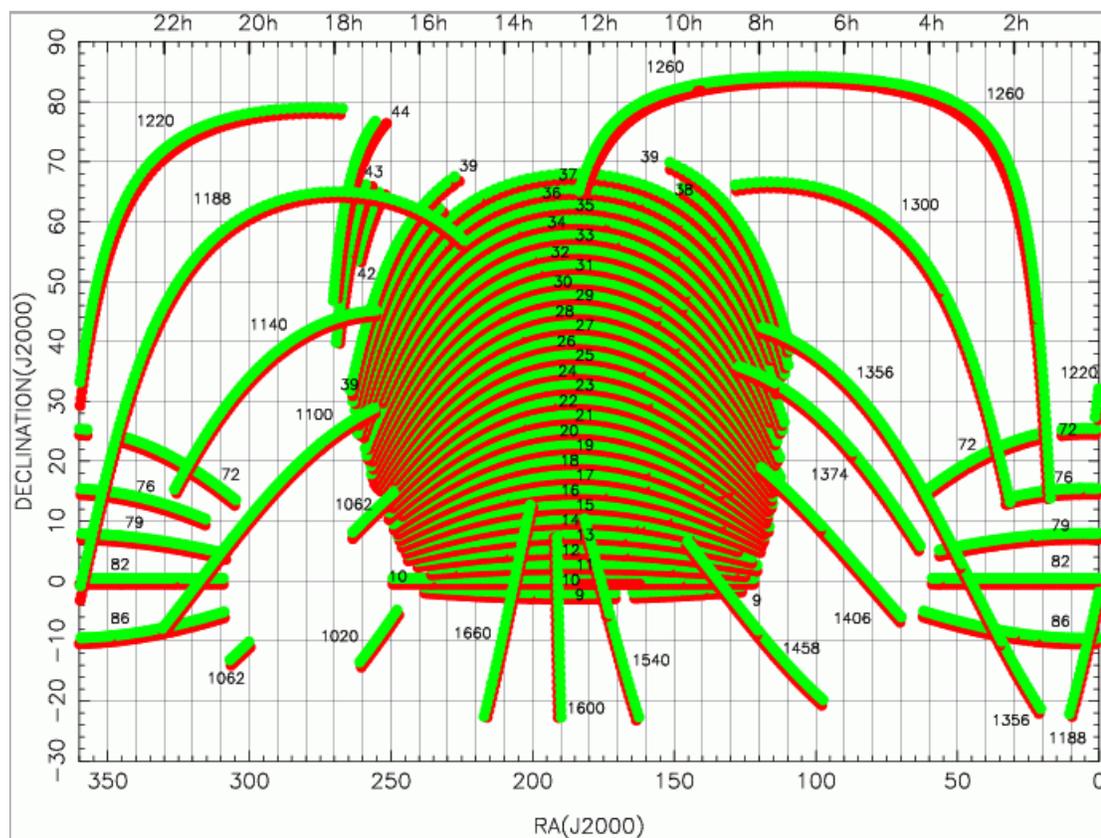
Figure 4.2: The Sloan Digital Sky Survey footprint (as of DR7) [78]. Note that each imaging section is referred to as a "stripe", which is made of two "strips". The red vs. green bands denote the 2 strips, which together make a stripe. Stripe 82 is found along the celestial equator (declination of 0).

median seeing of 1.4″, while Stripe 82 reaches to $r \approx 24.4$ and median seeing of 1.1″. This is possible as about 20 runs were taken of each field in the Stripe 82 coadd and these data were then coadded.

Source galaxies were selected from a catalog of luminous red galaxies provided by H. Lin. There are 5,875,133 objects in this catalog, ranging from $0.3 \leq z < 1.60$ and i-band magnitude $18 < i < 24$. All galaxies in this catalog include shape information measured by the SDSS pipeline.

4.2.2   The Blanco Cosmology Survey

The Blanco Cosmology Survey (BCS) was a survey conducted from 2005-2008 at the Cerro Tololo Inter American Observatory (CTIO) Victor M. Blanco Telescope using the Mosaic2 Imager [24]. The Mosaic2 imager was a prime focus camera that uses eight 2048x4096 CCDs. The field of view of the camera was 26.8 ′ per side, and the total camera area is about 0.4 deg^2. The BCS project was awarded 60 nights of observations [24] using the Blanco at CTIO. Images were taken using SDSS $griz$ filters.

There were two primary fields imaged in this survey, one at 5.5 hours right ascension and one at 23.5 hours; both fields were at declination about $-55°$. The total area of the survey is about 75 deg^2. These particular fields were chosen as they overlap with other surveys, including the South Pole Telescope (SPT), the Arcminute Cosmology Bolometer Array Receiver (ACBAR), the Atacama Pathfinder Experiment (APEX) and the Atacama Cosmology Telescope (ACT).

Source galaxies in BCS were found by running `SExtractor` on the original BCS PSF-homogenized images and matching information taken from this run to data

provided by the BCS collaboration. We found values for image moments, RA, DEC and `FLUX_RADIUS` by running `SExtractor`. We then matched these objects to data from the BCS collaboration on RA, DEC, magnitudes, `SExtractor` flags, `SPREAD_MODEL` and photometric redshift. We ultimately put all of these parameters together in two separate source galaxy catalogs, one for the 5 hr field and one for the 23 hr field. There are 1,794,195 galaxies in the BCS source galaxy catalog, with photometric redshifts in the range $0.006 < z < 2.96$ and i-band magnitude in the range $18 \leq i < 23.5$. We calculated ellipticities which we called *ME1CORR* and *ME2CORR* by using Equation 1.76.

4.2.3 The Dark Energy Survey–Science Verification

The Dark Energy Survey (DES) is a survey to learn more about the nature of dark energy by taking high-precision photometric data of 5000 square degrees of the southern sky. DES also uses the Blanco telescope, having constructed the 570 megapixel, Dark Energy Camera (DECam) and installed it on the Blanco in place of Mosaic2. DECam uses 62 2kx4k imaging CCDs. (See Figure 4.3 for an image of the dome (left) and the camera at the prime focus of the telescope with the author (right).) This survey will study dark energy using four major complementary methods: type Ia supernovae, baryon acoustic oscillations, weak gravitational lensing and galaxy cluster counting. DES saw first light in September 2012 and then continued to take data for small sections of its footprint during science verification (SV). I use about 100 square degrees of data of DES SV-A1 (area 1) for measurements of the mass-richness relation. DES began full survey operations in August 2013 and will continue for at least five years.
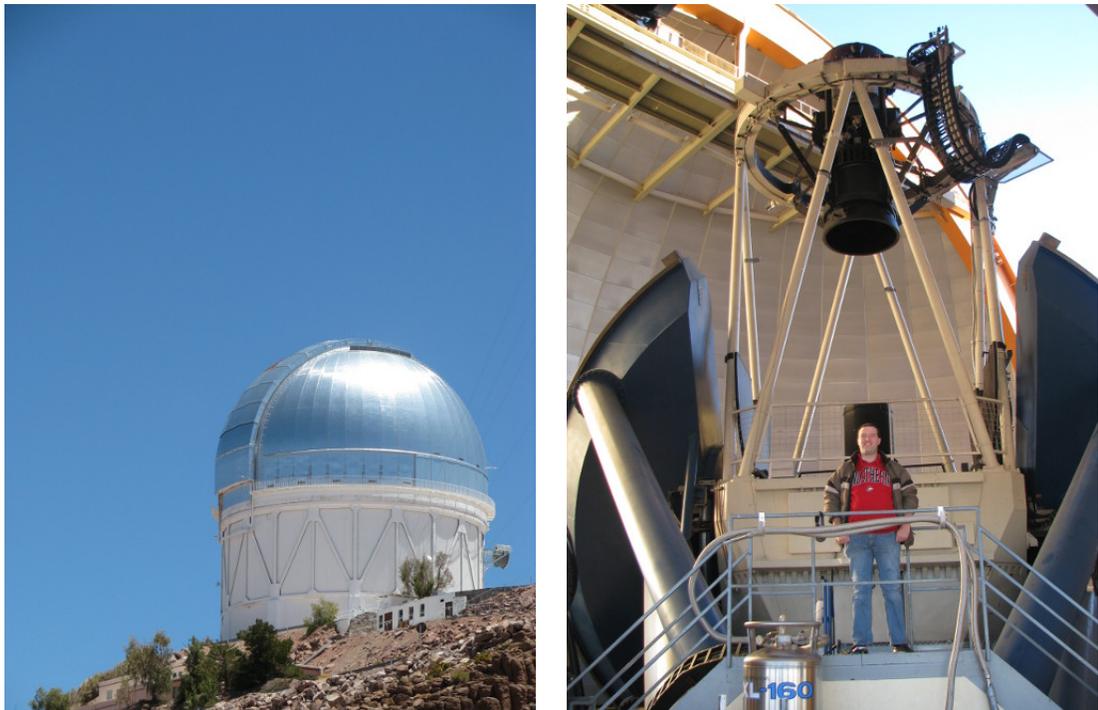
Figure 4.3: Left: An image of the dome of the Blanco telescope where DES is being conducted. Right: DECam and the author.

Figure 4.4: The highest richness cluster in the Stripe 82 coadd cluster catalog, with $N_{VT} = 127$. This is SDSS J0104+0004 (from the SDSS Image Server).

Source galaxies in DES-SV were found using a catalog of galaxies provided by H. Lin. There are 1,772,862 objects in this catalog with photometric redshifts in the range $0.01 < z \leq 2.0$ and i-band magnitude in the range $18 \leq i < 23.5$. All source galaxies included shape information found by SExtractor in the DES reduction pipeline.

4.3 Measurements of Cluster Richness

M. Soares-Santos provided us with a catalog of 19,376 clusters found in the Stripe 82 coadd using a Voronoi tessellation (VT) cluster finder. (The highest richness cluster in the sample is shown in Figure 4.4.) Voronoi tessellation is a method of relating distances between different locations in a large array of points

(called *seeds*). In a Voronoi tessellation, regions called Voronoi cells are defined in which all points within that cell are closer to that seed than to any other seed. In this method, the seeds would be galaxies. If the number and size of Voronoi cells exceed certain criteria, then it is declared to be a cluster. This was found to be a way of assembling a cluster sample with high purity and completeness while not depending on galaxy magnitude or color [79]. The total number of cluster members is defined by the number of objects within the overdense region in the VT method. Thus the richness measure is not N_{200} but is instead a new richness measure called N_{VT}. The Stripe 82 coadd VT catalog is biased towards lower richness clusters as the VT method is more likely to find these than the rarer high richness clusters.

We wanted to find an approximate relation between N_{200} and N_{VT} both to compare results for each and in order to use Equation 4.9, which depends on N_{200}. In order to do this, we found galaxy clusters listed both in the maxBCG catalog [52] and in the Stripe 82 VT catalog. We found 163 galaxy clusters in both catalogs and plotted the N_{200} from the maxBCG catalog against the N_{VT} from the VT catalog. The result is shown in Figure 4.5. It can be seen that there is not a clear relation between N_{VT} and N_{200}. Nevertheless we tried to find an approximate relation by fitting a power relation to the data (the power relation was a better fit than a simple linear fit). The equation we obtained relating the two richness measures was:

$$N_{200} = 3.88 \pm 0.105 (N_{VT})^{0.621 \pm 0.00689} \qquad (4.12)$$

Bleem et al. [13] present a sample of 763 clusters found in the BCS data. These clusters were found using a red sequence method presented by Gladders and Yee (2000) [33]; note that this is the same method behind the maxBCG method we use in Chapters 2 and 3. This method finds clusters by looking for overdensities of galaxies
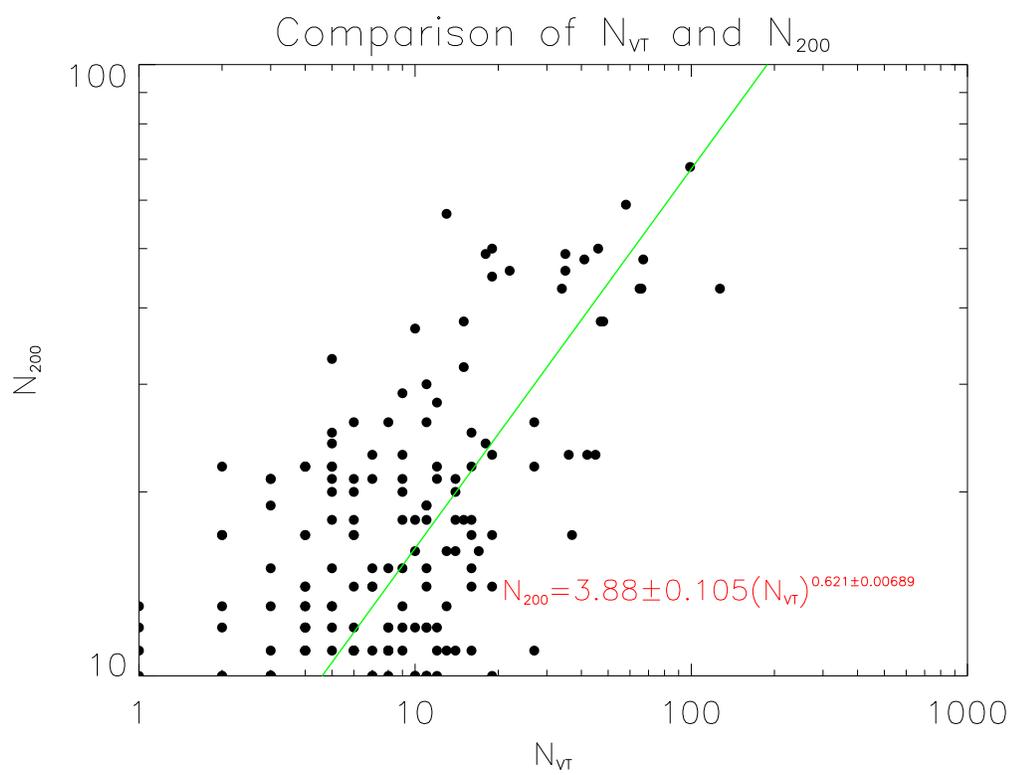
Figure 4.5: A plot of N_{VT} vs. N_{200} for 163 galaxy clusters.

in position, in color and in magnitude. That is, it looks for places where there are more galaxies than normal of a similar color. Richness was measured using the quantity λ, which is a statistic developed to minimize scatter in the mass-richness relation [72].

Clusters in the DES-SV data were also found by M. Soares-Santos using the Voronoi tessellation method. There are 27,230 clusters that were found in this preliminary DES-SV VT catalog. As in the Stripe 82 catalog, most of these clusters are found at low richness values. Due to the nature of DES as a higher redshift survey than SDSS, there are more clusters at higher redshift in the DES-SV cluster sample than in the Stripe 82 sample.

In order to compare λ to our other richness measurements, we found 206 clusters that overlapped between the DES-SV cluster sample and the 763 clusters from Bleem et al. Following the same method described above, we plotted N_{VT} as a function of λ, as seen in Figure 4.6. We found the relation between these variables to be

$$N_{VT} = 0.194 \pm 0.00935(\lambda)^{1.56 \pm 0.0143} \tag{4.13}$$

In Figure 4.7 we include a Hammer-Aitoff projection map of the whole sky showing the locations of our 3 data sets.

Finally we consider the richness and redshift ranges of the galaxy clusters. For the 19,706 Stripe 82 coadd clusters the redshift range is $0 \leq z \leq 0.98$ and the richness range is $1 \leq N_{VT} \leq 127$. We imposed a criterion that cluster needed to be at $z = 0.1$ or higher as below that significance of the N_{VT} detections was low. For the 763 BCS clusters, the redshift range is $0.13 \leq z \leq 0.75$ and the richness range is $10 \leq \lambda \leq 64.3$. For the 27,230 DES-SV clusters the redshift range is $0.12 \leq z \leq 1.68$ and the richness range is $1 \leq N_{VT} \leq 538$. In Figure 4.8 we show
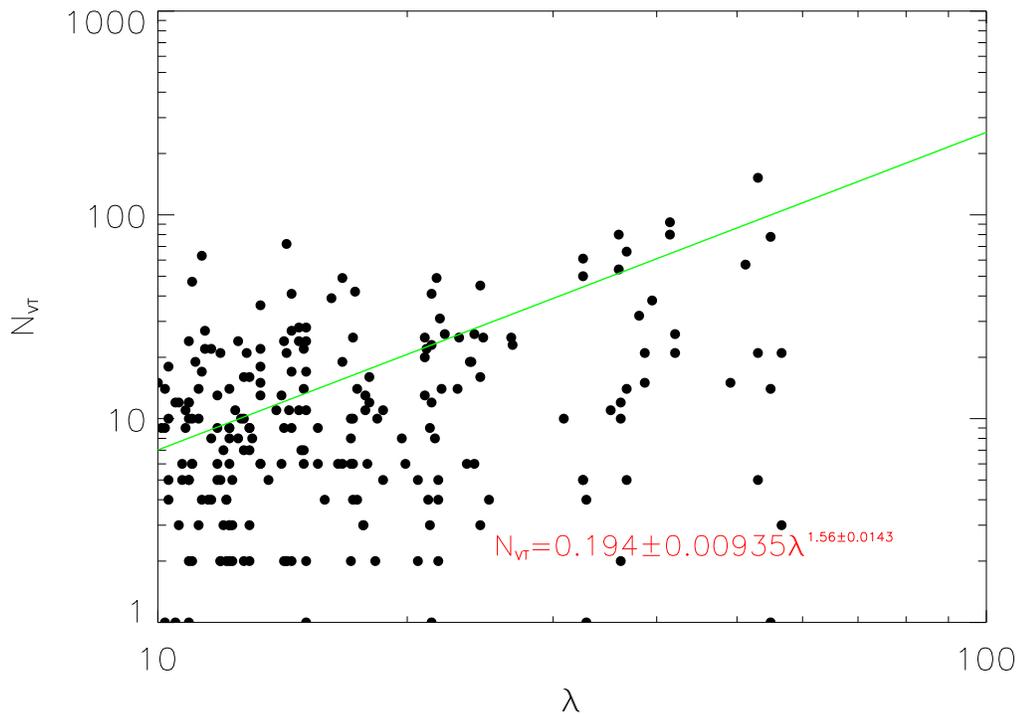
Figure 4.6: A plot of the richness parameter λ vs. N_{VT} for 206 galaxy clusters.

histograms of richness and redshift for all three cluster samples. Note that all three have the most clusters at redshifts above 0.5 while richnesses are dominated by low values. These samples were deliberately chosen as they extend to higher cluster redshift so that we could find the mass-richness relation at higher redshift.

4.4   Measurements of Cluster Mass

4.4.1   The Algorithm for Measurement of Shear

The quantity that can be directly observed in images of galaxy clusters is galaxy ellipticity. In order to find tangential shear, we use the relations described in Section
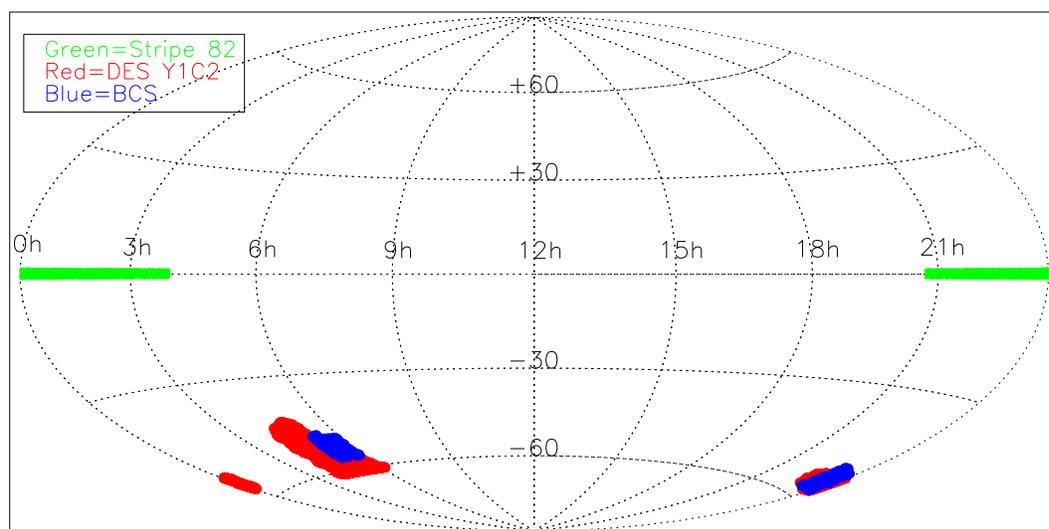
Figure 4.7: A plot of the whole sky showing the locations of the three datasets used in this chapter. DES Y1C2 is the DES-SV data; Y1C2 stands for Year 1, Catalog 2.
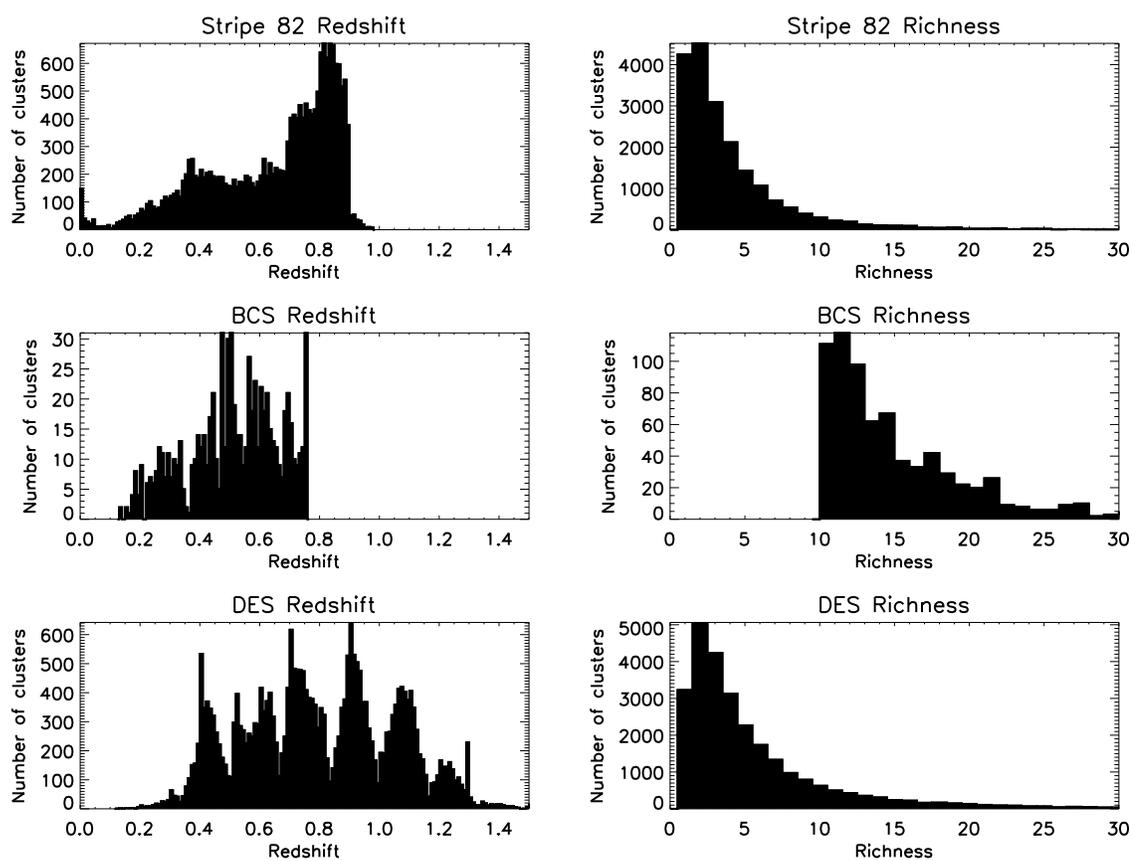
Figure 4.8: Histograms of redshift and richness for all three data sets.

1.7.3. Tangential shear is related to the mass and concentration of the lensing mass distribution. In order to find these quantities, we apply an NFW profile and find the best fit mass and concentration that produce the observed shear. We did this by using the equations given in Wright and Brainerd (2000) [95]. Weak lensing shear as a function of radius predicted by the NFW model is given as

$$\gamma_{NFW}(x) = \frac{\overline{\Sigma}_{NFW}(x) - \Sigma_{NFW}(x)}{\Sigma_{crit}} \tag{4.14}$$

Here x is a dimensionless radius equal to r/r_s, where r is the distance from the cluster center in the lens plane and r_s is the scale radius defined in Chapter 1. $\Sigma_{NFW}(x)$ is the surface mass density of the galaxy cluster's dark matter halo in the NFW model and $\overline{\Sigma}_{NFW}(x)$ is the mean surface mass density of the halo. Putting in Wright and Brainerd's Equations 11 and 13 for $\Sigma_{NFW}(x)$ and $\overline{\Sigma}_{NFW}(x)$ we obtain their Equation 14 for shear predicted by the NFW profile:

$$\gamma_{NFW}(x) = \begin{cases} \dfrac{r_s \delta_c \rho_c}{\Sigma_c} g_<(x) & x < 1 \\[2em] \dfrac{r_s \delta_c \rho_c}{\Sigma_c} [\frac{10}{3} + 4 \ln \frac{1}{2}] & x = 1 \\[2em] \dfrac{r_s \delta_c \rho_c}{\Sigma_c} g_>(x) & x > 1 \end{cases} \tag{4.15}$$

where

$$\delta_c = \frac{200}{3} \frac{c^3}{\ln(1+c) - c/(1+c)} \tag{4.16}$$

and

$$g_<(x) = \frac{8\,arctanh(\sqrt{(1-x)/(1+x)})}{x^2\sqrt{1-x^2}} + \frac{4}{x^2}\ln\left(\frac{x}{2}\right) - \frac{2}{(x^2-1)} + \frac{4\,arctanh(\sqrt{(1-x)/(1+x)})}{(x^2-1)\sqrt{1-x^2}}$$

(4.17)

and

$$g_>(x) = \frac{8\,arctan(\sqrt{(x-1)/(1+x)})}{x^2\sqrt{x^2-1}} + \frac{4}{x^2}\ln\left(\frac{x}{2}\right) - \frac{2}{(x^2-1)} + \frac{4\,arctan(\sqrt{(x-1)/(1+x)})}{(x^2-1)^{3/2}}$$

(4.18)

In all of these equations

$$x = \frac{r}{r_s}$$

(4.19)

and

$$r = D_L(\theta_1^2 + \theta_2^2)^{1/2}$$

(4.20)

The thetas above are image plane coordinates in the plane of the galaxy cluster.

In order to fit shear as a function of radius to an NFW profile, H. Lin wrote a function in IDL called *shearnfw*. This function implements Equation 4.15 by using the IDL routine *lmfit*. This routine takes in an independent variable, a dependent variable and an array of parameters to fit for. The *lmfit* routine applies a least squares fit to a non-linear function, in this case to the *shearnfw* function. The independent variable was the distance of a source galaxy from the center of the cluster, the dependent variable was the measured shear and the two parameters being fit for were M_{200} and c_{200}. The output of the function was a prediction of the tangential or orthotangential shear from the NFW model as a function of distance from the cluster center.

The *lmfit* function also read in partial derivatives of the function in order to do the fit. Two partial derivatives were output by *shearnfw* and input to *lmfit*, $\frac{\partial \gamma_{NFW}}{\partial M_{200}}$ and $\frac{\partial \gamma_{NFW}}{\partial c_{200}}$. These derivatives are implemented as:

$$\frac{\partial \gamma_{NFW}}{\partial M_{200}} = \frac{\partial g}{\partial r_s} \frac{\partial r_s}{\partial r_{200}} \frac{\partial r_{200}}{\partial M_{200}} \tag{4.21}$$

and

$$\frac{\partial \gamma_{NFW}}{\partial c_{200}} = \frac{\partial g}{\partial r_s} \frac{\partial r_s}{\partial c_{200}} + \frac{\partial \delta_c}{\partial c_{200}} \tag{4.22}$$

4.4.2   How the Algorithm was Applied

All of the above routines and others were assembled into an IDL program called `NFW.pro`. At the start of this program, data was read in from one of the three data sets, either the SDSS Stripe 82 coadd, BCS or DES. The data included information on a set of clusters and a set of source galaxies. Each cluster would have a large number of source galaxies nearby it (nearby in the same x-y plane, but at a higher redshift); these source galaxies would have their ellipticities measured and thus tangential shears calculated. All of these shears would be combined for clusters in a particular redshift and richness bin and they would then be run through the *shearnfw* fitting routine. We used 4 redshift bins:

- $0.1 < z \leq 0.4$ (low-z)

- $0.4 < z \leq 0.7$ (mid-z)

- $0.7 < z \leq 1.0$ (high-z)

- $0.1 < z \leq 1.0$ (all-z)

Table 4.1: Richness bins for shear measurements in the Stripe 82 coadd. When all bins have the same richness range we give only the value in the low-z bin, but these are the values for all 4 bins.

Bin Number	N_{VT} (Low-z)	N_{VT} (Mid-z)	N_{VT} (High-z)	N_{VT} (All-z)
1	1	–	–	–
2	2	–	–	–
3	3	–	–	–
4	4	–	–	–
5	5	–	–	–
6	6	–	–	–
7	7	–	–	–
8	8	–	–	–
9	9-10	–	–	–
10	11-12	–	–	–
11	13-15	–	–	–
12	16-20	–	–	–
13	21-30	–	–	–
14	31-40	–	–	–
15	41-50	41-50	41-140	41-50
16	51-60	51-60	none	51-60
17	61-70	61-70	none	61-70
18	71-80	71-127	none	71-80
19	81-127	none	none	81-127

We also used a number of richness bins; these varied by redshift bin and were chosen to maximize the number of clusters in each bin. The bins used for the Stripe 82 coadd are given in Table 4.1.

At the start of the program, several parameters were read in (see Table 4.2). The parameters *drad* and *nrad* were used to control the total distance from the BCG (in Mpc) to which the fitting routine would proceed. The maximum distance is *drad* $\times$ *nrad*. The maximum and minimum i-band magnitudes for source galaxies were controlled by *maglim1* and *maglim2*. The fitting routine could be run multiple

Table 4.2: A list of pre-set parameters and their values used in the program *NFW.pro.*

Name	Function	Value
drad	Radius step size (in Mpc)	0.1
nrad	Number of radius steps	30
maglim1	lower limit i-band magnitude for source galaxies	20.5
maglime2	upper limit i-band magnitude for source galaxies	23.5
niter	number of iterations of fitting	1
z_{fid}	fiducial lens redshift	0.55
z_{fids}	fiducial source redshift	0.75

times to improve the fit; this was controlled by *niter*. We set *niter* to 1 as the fitting results were not found to improve significantly with more iterations.

At the beginning of `NFW.pro`, a loop was begun that would run through a series of richness bins in a particular redshift range. For each individual cluster in a richness and redshift bin, the program would calculate angular diameter distance for the cluster. It would then cut any source galaxies that are outside the allowed magnitude limits, had photometric redshifts less than the cluster redshift plus 0.1, had ellipticities greater than 2 and were farther than $drad \times nrad$ (usually 3.0 $h^{-1}Mpc$) from the BCG. It then calculates tangential and orthotangential ellipticities of the source galaxies that passed all of these cuts using Equation 1.77. Finally ellipticities were converted to shears by dividing by responsivity.

For the Stripe 82 coadd, we used

$$R = 2(1 - \sigma_{SN}^2) = 1.73 \tag{4.23}$$

where σ_{SN} is the intrinsic galaxy shape noise in the SDSS, taken to be $\sigma_{SN} = 0.37$. *Shape noise* describes the intrinsic variation in galaxy shapes, independent of any measurement uncertainty. For the BCS and DES-SV data, we took $R = 1$. This

value of responsivity was found by H. Lin by measuring ellipticities of a number of source galaxies in images taken at the Blanco telescope using DECam and then measuring the ellipticities of a number of source galaxies in the same region of the sky using Hubble Space Telescope COSMOS data. Responsivity is related to the ratio of ellipticities in the two data sets. Responsivity was found to be 1 for DES-SV and BCS data.

For Stripe 82 data, PSF deconvolution was done by R. Reis using the Hirata-Seljak method [43]. For DES and BCS data a simpler estimate was done [17]: we identified the stars and found their second moments. We then subtracted these second order moments from those of the galaxies. This method was facilitated in BCS data as the data were PSF-homogenized.

We then calculated a scale factor to scale all cluster shears to the same redshift. This was done because weak lensing shear is a function of redshift, and while all clusters in a redshift/richness bin are at similar redshift, they are not at identical redshift. This scaling puts all clusters at the same redshift. This scaling factor is

$$correction = \frac{d_S d_{L\_fid} d_{LS\_fid}}{d_{S\_fid} d_L d_{LS}} \tag{4.24}$$

where d_S, d_L and d_{LS} are angular diameter distances to the source, the lens and from lens to source, respectively. The ones marked fiducial mean at the fiducial redshift, which is $z_{fiducial\_lens} = 0.55$ for the lens and $z_{fiducial\_source} = 0.75$ for the source. Thus all objects are scaled to these two redshifts. Equation 4.24 was obtained by taking the ratio of Σ_{crit} for the cluster redshift to the fiducial redshift. In other words, what we wanted to find was a correction, where

$$\gamma_{fiducial} = \gamma_{actual} \times correction \tag{4.25}$$

Then

$$correction = \frac{\gamma_{fiducial}}{\gamma_{actual}} \tag{4.26}$$

Since the magnitude of tangential shear is proportional to Σ_{crit}^{-1} and the other factors in the shear are not functions of redshift, we can say the following:

$$correction = \frac{\Sigma_{crit\_fiducial}^{-1}}{\Sigma_{crit\_actual}^{-1}} \tag{4.27}$$

since

$$\Sigma_{crit} = \frac{c^2 d_S}{4\pi G d_L d_{LS}} \tag{4.28}$$

we are led to Equation 4.24.

After calculating shears for all source galaxies in a particular redshift/richness bin and correcting them with Equation 4.24, the program prepares to fit the shears to an NFW profile. First it calculates $H(z)$ and ρ_{crit} at the fiducial redshift. It also needs to calculate Σ_{crit}^{-1} using redshifts of the source galaxies and of the cluster. This calculation is very important to the final results; we discuss analysis of systematics that affect Σ_{crit}^{-1} in § 4.5.3. Next the routine calculates the standard deviation in the tangential and orthotangential shear values using the unscaled shear values. We use unscaled shear values to avoid biasing by a large scale factor. This can be a problem when the source galaxy redshift is very close to the fiducial redshift. We scale the standard deviations after calculating them. We then calculate average tangential shear by using inverse variance weighting, again to overcome the effects of large scale factors. The equation we use is

$$\gamma_{ave} = \frac{\sum_{i=1}^{N} \left[\frac{\gamma_i}{\sigma_{shear}^2} \right]}{\frac{1}{\sum_{i=1}^{N} \left[\sigma_{shear}^2 \right]}} \tag{4.29}$$

for N source galaxies where σ^2_{shear} is the standard deviation on the values of tangential shear in that redshift and richness bin. Then the error on this average value is found as

$$\Delta\gamma_{ave} = \sqrt{\frac{1}{\sum\limits_{i=1}^{N} \frac{1}{\sigma^2_{shear}}}} \qquad (4.30)$$

Finally we take the average shear values found from Equation 4.29 and the predicted shear values found from Equation 4.15 and minimize the χ^2 function:

$$\chi^2 = \sum_{i=1}^{N} \frac{\left[\gamma_i - \gamma_{NFW}(r_i; M_{200}; c_{200})\right]^2}{\sigma^2_{\gamma}} \qquad (4.31)$$

The values of M_{200} and c_{200} that minimize the χ^2 are taken as the model fits for mass and concentration. The average values of tangential and orthotangential shear are plotted as a function of radius along with the predicted values of tangential and orthotangential shear. Two sample plots are shown in Figure 4.9, one for the low redshift $(0.1 - 0.4)$ bin with $N_{VT} = 6$ and one for the high redshift $(0.1 - 1.0)$ bin with $N_{VT} = 11 - 12$. Note that as would be expected, there are far more clusters (555) for the lower richness $N_{VT} = 6$ bin. For the higher richness $N_{VT} = 11 - 12$ bin there are only 90 clusters. However there are more source galaxies $(435, 759)$ in the higher richness bin than in the lower richness $(209, 784)$ bin. Note that in both cases there is good evidence for a tangential shear signal, but orthotangential shear is consistent with zero.

Finally in Tables 4.3, 4.4, 4.5 and 4.6 we present measurements of richness, stacked cluster mass and concentration as well as reduced χ^2 and the χ^2 probability
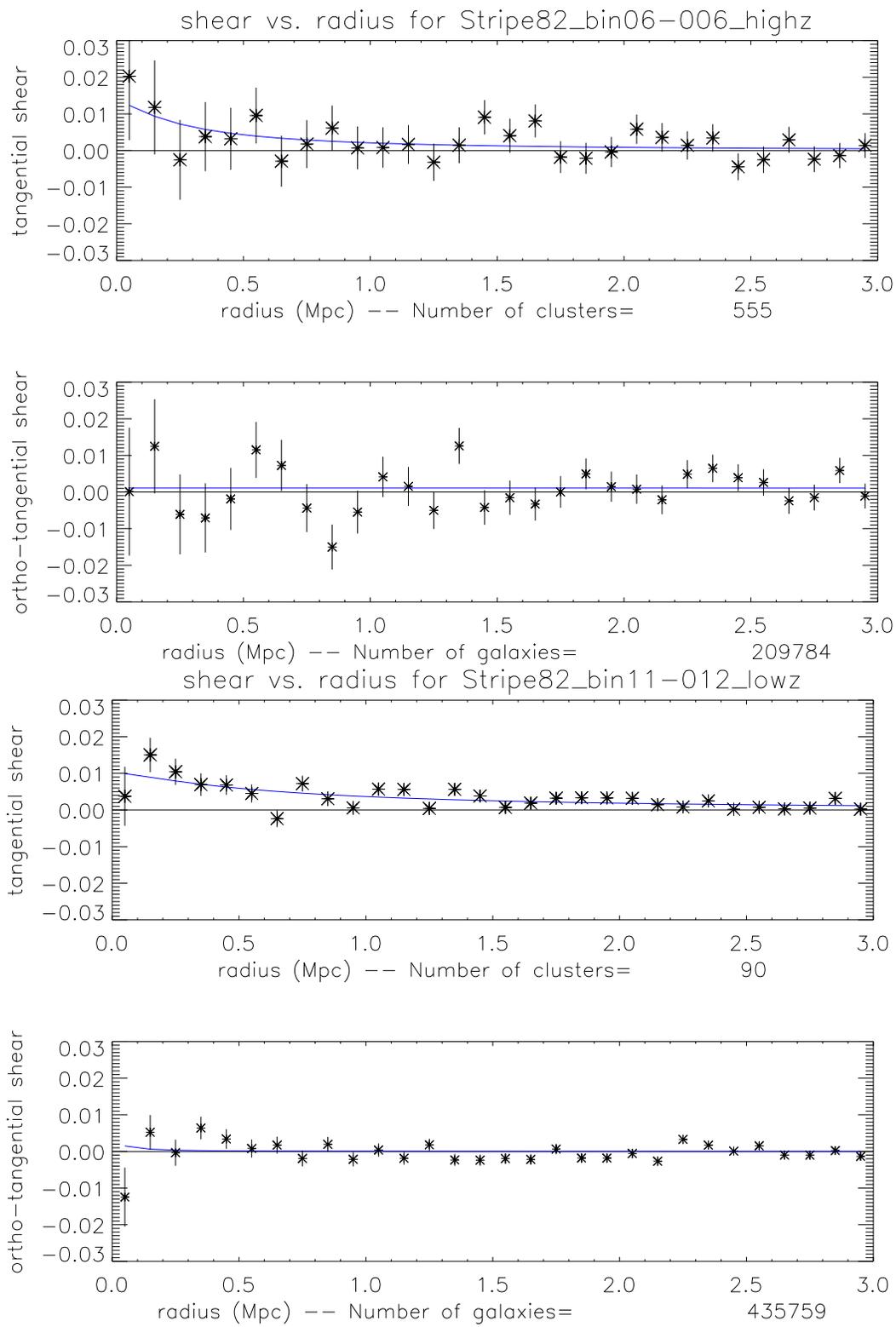
Figure 4.9: Upper: The shear profile for Stripe 82 coadd, high-z bin, $N_{VT} = 6$. Lower: The shear profile for Stripe 82 coadd, low-z bin, $N_{VT} = 11 - 12$.

for each richness and redshift bin. The value of N_{VT} given is the median of the N_{VT} values in that richness bin. The error on N_{VT} is the range of N_{VT} values, given by

$$\Delta N_{VT} = \frac{[max(N_{VT}) - median(N_{VT})] + [median(N_{VT}) - min(N_{VT})]}{2} \quad (4.32)$$

The error on the M_{200} values is the standard deviation on values of M_{200} output by the *lmfit* routine in IDL. The error on c_{200} is also output as the standard deviation from *lmfit*. The χ^2_{red} we report here is the reduced χ^2 for the null test of fitting the weak lensing shear to 0 (see Equation 1.82 for the definition of χ^2). Note that reduced χ^2 means χ^2/dof, where *dof* means degrees of freedom. The number of degrees of freedom is the number of data points minus the number of fit parameters (here 2, M_{200} and c_{200}). The χ_{red} we report here is not the same as the χ^2_{red} used for the fit to the NFW profile (Equation 4.31). What we are checking here is whether the tangential shear profile is consistent with zero. If this is a good fit (reduced $\chi^2 \approx 1$) then there is no signal. If this is a bad fit (reduced $\chi^2 > 1$), then we have cause to believe there is a tangential shear signal. The equation we use in the code is

$$\chi^2 = \left(\frac{\gamma_{ave} - 0}{\Delta\gamma_{ave}}\right)^2 \quad (4.33)$$

The χ^2 probability is equal to

$$Prob(\chi^2_{red} > x) = \int_x^\infty \chi^2_{red} \quad (4.34)$$

The resultant expresion is given as [6]

$$Prob(\chi^2, N) = \int_{\chi^2}^\infty \frac{2^{-N/2}}{\Gamma(N/2)} \chi^{N-2} e^{-\chi^2/2} d\chi'^2 \quad (4.35)$$

where $\Gamma(N/2)$ is the gamma function. This probability is the probability that a function which does genuinely describe a set of N data points would give a value of χ^2 as large, or larger, than the one you already have. Another way of saying this is that the χ^2 probability is the probability of the correct fit being found outside of the fit. Thus for a good fit the probability will be of order 0.5. We again implemented the χ^2 probability to test whether the tangential shear signal was consistent with 0. So a very small value of the χ^2 probability would indicate the shear signal not being consistent with 0 while a larger probability would indicate the shear signal is consistent with 0. Note that in Tables 4.3, 4.4 and 4.6 the probabilities are all quite small, which suggests that the fit to 0 is poor and there is measurable weak lensing signal. In Table 4.5 many of the probabilities are higher, which suggests the weak lensing signal is less clear for the high-z bin.

4.5 Systematics in Measurement of Cluster Mass in Stripe 82 Coadd

4.5.1 Central BCG

In our fitting of the mass and concentration, we fit an NFW profile to the dark matter halo of the clusters. In so doing, we did not include a term for the mass of the central BCG. Thus we wanted to try to estimate what effect the mass of the BCG would have on the shear profile. To do this we ran the `NFW.pro` program and excluded the central 0.1 h^{-1} Mpc. We did this as the mass of the BCG would usually affect only the innermost regions of the shear profile, as shown in Figure 4.1.

Table 4.3: Weak lensing shear fit results for low redshift clusters ($0.1 \leq z \leq 0.4$). This table includes values for richness and error, mass (M_{200}) and error, concentration (c_{200}) and error, null test (fit to 0) reduced χ^2 and corresponding probability of fit.

N_{VT}	M_{200}	c_{200} error	χ^2_{red}	**Prob**
1 ± 0	0.0147 ± 0.0100	4.87 ± 5.77	1.54	0.0306
2 ± 0	0.0505 ± 0.0145	2.52 ± 1.16	1.83	0.00368
3 ± 0	0.0773 ± 0.0222	1.56 ± 0.772	2.23	0.000126
4 ± 0	0.0729 ± 0.0330	0.890 ± 0.628	1.73	0.00769
5 ± 0	0.139 ± 0.0417	1.066 ± 0.541	2.39	0.000028
6 ± 0	0.125 ± 0.0389	2.28 ± 1.18	1.81	0.00421
7 ± 0	0.206 ± 0.0633	1.30 ± 0.704	1.99	0.000977
8 ± 0	0.0226 ± 0.0315	1.28 ± 2.88	2.13	0.000305
9 ± 0.500	0.0154 ± 0.0205	1.87 ± 4.17	2.40	0.000025
11 ± 0.500	0.500 ± 0.103	0.926 ± 0.350	4.17	0
14 ± 1.00	0.417 ± 0.116	0.633 ± 0.300	4.11	0
18 ± 2.00	0.331 ± 0.0954	0.993 ± 0.506	2.45	0.000016
23 ± 4.50	0.400 ± 0.159	0.444 ± 0.289	2.65	0.000002
35 ± 4.50	1.15 ± 0.275	1.069 ± 0.508	2.75	0.000001
46 ± 4.50	1.42 ± 0.326	1.45 ± 0.646	2.49	0.000011
54 ± 4.50	2.65 ± 0.842	0.825 ± 0.559	2.10	0.000386
65 ± 3.00	2.27 ± 0.710	1.075 ± 0.683	2.77	0.000001
78 ± 4.00	2.75 ± 1.28	1.089 ± 1.046	1.81	0.00413
127 ± 14.0	2.41 ± 1.51	0.615 ± 0.835	1.26	0.157

Table 4.4: Weak lensing shear fit results for mid-redshift clusters ($0.4 \leq z \leq 0.7$). This table includes values for richness and error, mass (M_{200}) and error, concentration (c_{200}) and error, null test (fit to 0) reduced χ^2 and corresponding probability of fit

N_{VT}	M_{200}	c_{200} error	χ^2_{red}	**Prob**
1 ± 0	0.0564 ± 0.0315	0.777 ± 0.635	1.48	0.0452
2 ± 0	0.00171 ± 0.00143	0.0000880 ± 0.279	1.63	0.0165
3 ± 0	0.0477 ± 0.0302	1.011 ± 0.994	1.24	0.168
4 ± 0	0.187 ± 0.0500	1.87 ± 0.831	1.92	0.00176
5 ± 0	0.114 ± 0.0607	1.092 ± 0.961	2.14	0.000269
6 ± 0	0.198 ± 0.115	0.444 ± 0.377	1.98	0.00108
7 ± 0	0.217 ± 0.126	0.591 ± 0.528	1.49	0.0413
8 ± 0	0.522 ± 0.217	0.425 ± 0.302	1.91	0.00194
9 ± 0.500	0.448 ± 0.132	1.17 ± 0.626	2.31	0.000061
11 ± 0.500	0.725 ± 0.201	1.24 ± 0.644	2.22	0.000140
14 ± 1.00	0.0241 ± 0.0705	0.827 ± 3.46	1.06	0.376
17 ± 2.00	0.703 ± 0.205	1.330 ± 0.720	1.98	0.00109
24 ± 4.50	0.923 ± 0.294	1.13 ± 0.689	1.64	0.0155
35 ± 4.00	1.38 ± 1.045	0.423 ± 0.673	1.64	0.0152
44 ± 3.00	0.680 ± 0.649	6.83 ± 7.01	1.61	0.0183
54 ± 2.00	1.74 ± 0.886	1.45 ± 1.470	1.38	0.0825
63 ± 4.00	2.42 ± 2.023	0.450 ± 0.920	1.069	0.364
91 ± 7.50	0.898 ± 1.89	0.617 ± 2.53	0.712	0.876

Table 4.5: Weak lensing shear fit results for high redshift clusters ($0.7 \leq z \leq 1.0$). This table includes values for richness and error, mass (M_{200}) and error, concentration (c_{200}) and error, null test (fit to 0) reduced χ^2 and corresponding probability of fit

N_{VT}	M_{200}	c_{200} error	χ^2_{red}	**Prob**
1 ± 0	0.00387 ± 0.0172	0.0232 ± 0.163	0.921	0.590
2 ± 0	0.0378 ± 0.0747	0.432 ± 1.038	0.937	0.564
3 ± 0	0.294 ± 0.136	0.00541 ± 0.253	1.49	0.0408
4 ± 0	0.0237 ± 0.0625	1.11 ± 4.31	1.22	0.189
5 ± 0	0.292 ± 0.174	3.49 ± 2.41	1.91	0.00192
6 ± 0	0.264 ± 0.204	1.72 ± 2.073	0.716	0.872
7 ± 0	0.307 ± 0.273	1.69 ± 2.37	0.553	0.977
8 ± 0	0.0818 ± 0.307	0.333 ± 1.58	0.892	0.636
9 ± 0.500	0.848 ± 0.442	1.28 ± 1.17	1.024	0.429
11 ± 0.500	0.351 ± 0.381	2.044 ± 3.33	0.814	0.753
14 ± 1.00	0.293 ± 0.608	0.0296 ± 0.438	1.056	0.383
17 ± 2.00	1.85 ± 1.12	1.74 ± 1.67	0.961	0.526
24 ± 4.50	2.86 ± 2.75	0.167 ± 0.644	1.35	0.0983
34 ± 4.50	4.74 ± 4.43	4.20 ± 4.42	1.18	0.233
44 ± 5.00	0.0138 ± 1.11	0.349 ± 30.1	1.20	0.209

Table 4.6: Weak lensing shear fit results for all redshift clusters ($0.1 \leq z \leq 1.0$). This table includes values for richness and error, mass (M_{200}) and error, concentration (c_{200}) and error, null test (fit to 0) reduced χ^2 and corresponding probability of fit

N_{VT}	M_{200}	c_{200} error	χ^2_{red}	**Prob**
1 ± 0	0.0263 ± 0.0131	0.861 ± 0.622	1.59	0.0207
2 ± 0	0.00610 ± 0.00693	0.970 ± 1.58	1.77	0.00576
3 ± 0	0.0633 ± 0.0211	0.783 ± 0.388	2.86	0
4 ± 0	0.00261 ± 0.00663	1.36 ± 5.73	2.86	0
5 ± 0	0.143 ± 0.0409	0.697 ± 0.307	3.29	0
6 ± 0	0.152 ± 0.0415	1.026 ± 0.474	3.031	0
7 ± 0	0.0128 ± 0.0175	1.65 ± 3.67	2.74	0.000001
8 ± 0	0.205 ± 0.0841	0.401 ± 0.240	2.72	0.000001
9 ± 0.500	0.296 ± 0.0537	1.39 ± 0.453	3.57	0
11 ± 0.500	0.546 ± 0.0884	1.051 ± 0.317	5.68	0
14 ± 1.00	0.305 ± 0.101	0.448 ± 0.231	4.30	0
17 ± 2.00	0.435 ± 0.0879	1.105 ± 0.405	3.83	0
24 ± 4.50	0.526 ± 0.142	0.562 ± 0.266	3.60	0
35 ± 4.50	1.22 ± 0.264	1.055 ± 0.453	3.082	0
45 ± 4.50	1.38 ± 0.312	1.35 ± 0.594	2.44	0.000018
54 ± 4.50	2.32 ± 0.617	1.031 ± 0.564	2.52	0.000008
65 ± 4.00	2.32 ± 0.672	0.932 ± 0.569	3.028	0
78 ± 4.00	2.70 ± 1.28	0.974 ± 0.974	1.72	0.00844
99 ± 18.0	2.028 ± 1.23	0.647 ± 0.833	0.936	0.566

In Figure 4.10 we present the same shear profiles as shown in Figure 4.9 but now excluding the central 0.1 h^{-1} Mpc.

In Figure 4.11 we compare the cluster masses found with and without the central 0.1 h^{-1} Mpc region of the clusters. The green line is the $y = x$ line. The points follow this line quite closely, with a few exceptions. Thus we conclude that the mass of the central BCG has minimal impact on the shear profiles.

4.5.2   Halo Miscentering

When we fit an NFW profile to the dark matter halos of the galaxy clusters as in Section 4.4.2, we take the BCG to be the center of the halo, measuring all distances with respect to this. However, as discussed in Section 4.1.2, this is will be incorrect some fraction of the time, as the BCG at times is not the center of the dark matter halo. If mass is measured at a location outside the center of a halo, the mass found for that halo will be less than it should be, since you are then treating the outer region of a halo as the center. As this will bias mass measurements, our analysis needs to include a consideration of the contribution from halos that are not centered on the BCG. In order to consider halo miscentering we follow the analysis of Simet et al. [77] in Equation 4.9.

We applied Equation 4.9 by taking the masses output by the fitting routine as $M_{200,mis}$ and multiplying them by 1.44. Then the true masses ($M_{200,true}$) are a function of N_{200}, following Equation 4.9. This process was done in our program which calculated the final mass-richness relation, `MASS_RICHNESS.pro`. Before we calculated $M_{200,true}$ we had to convert N_{VT} to N_{200} using Equation 4.12.
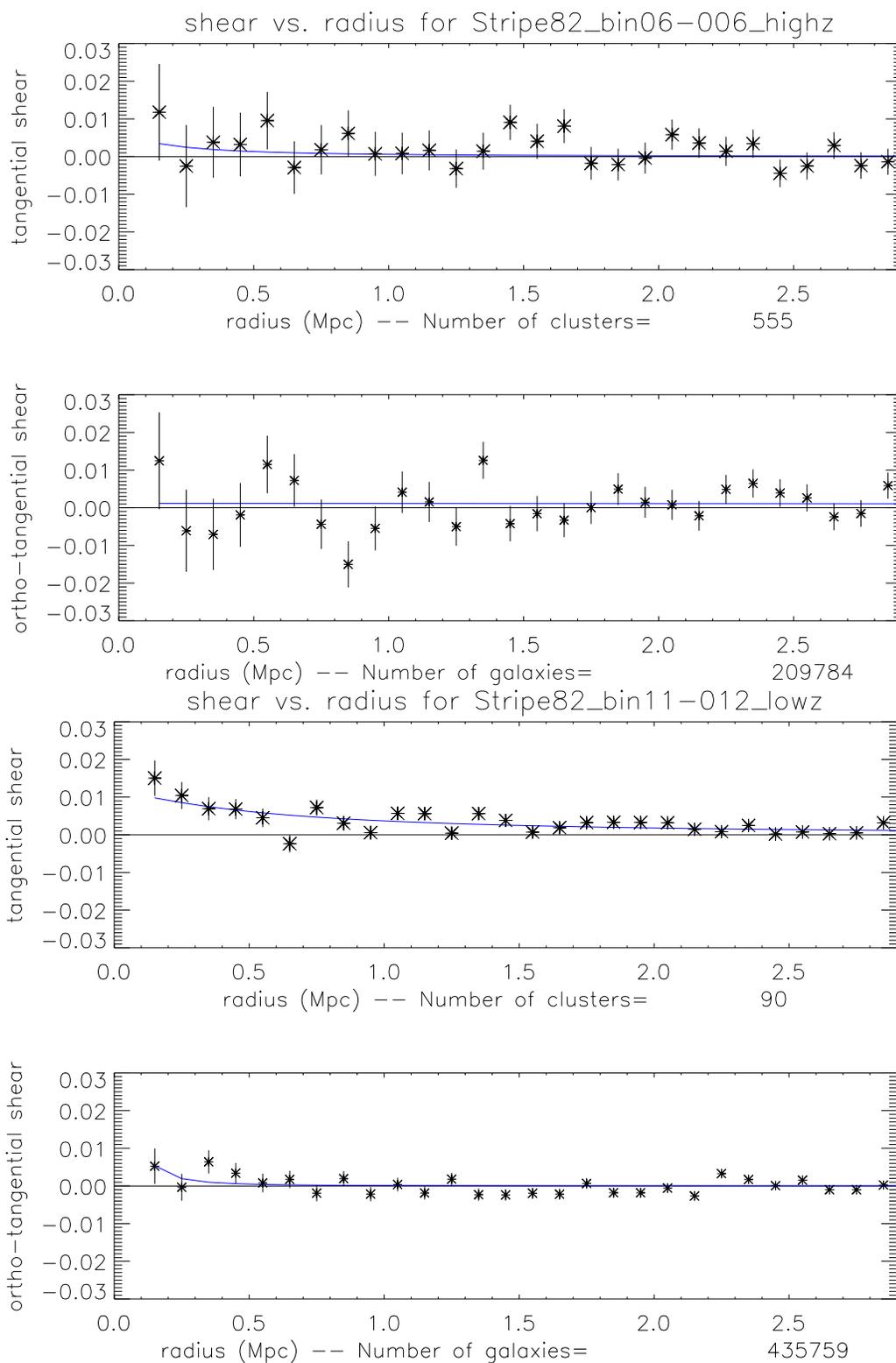
Figure 4.10: These are the same shear profiles as shown in Figure 4.9, but with the central $0.1h^{-1}$ Mpc excluded. Upper: The shear profile for Stripe 82 coadd, high-z bin, $N_{VT} = 6$. Lower: The shear profile for Stripe 82 coadd, low-z bin, $N_{VT} = 11 - 12$.
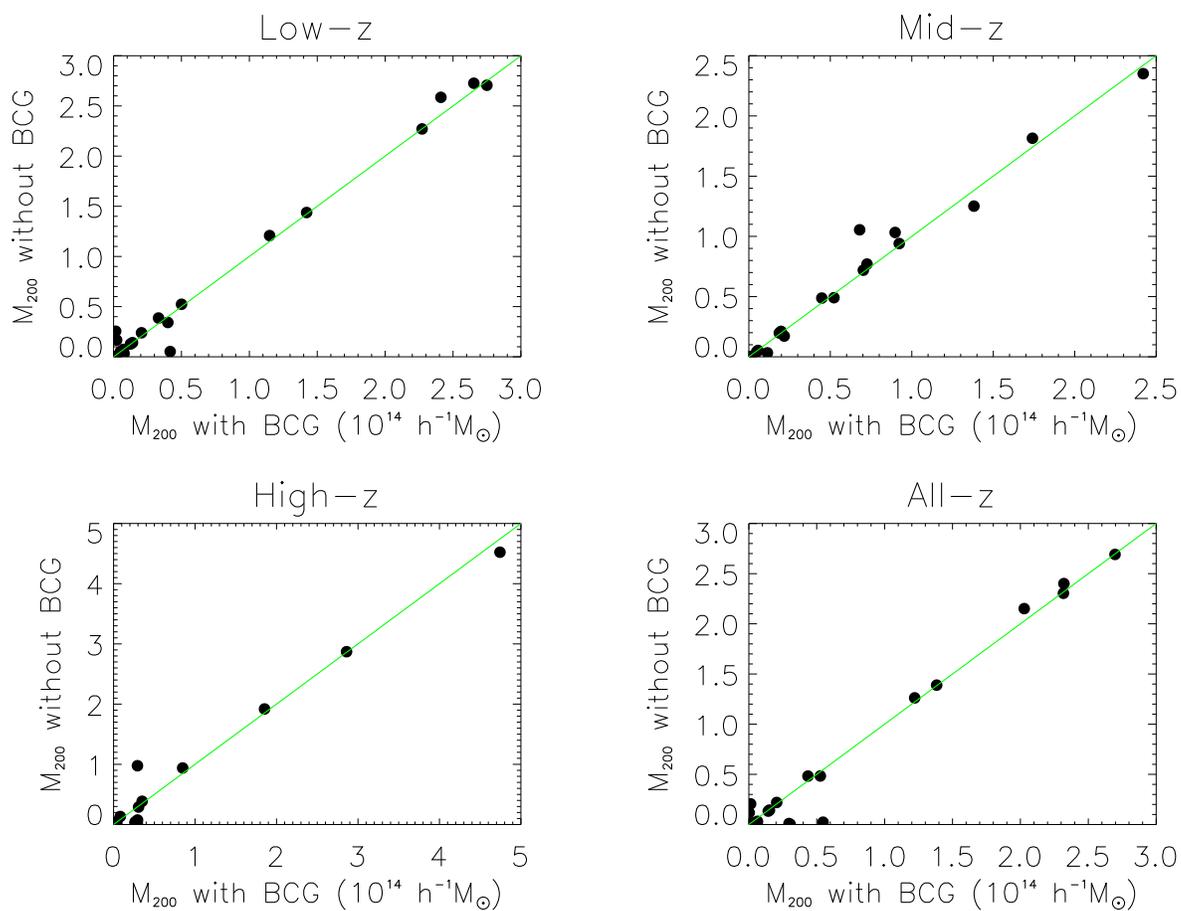
Figure 4.11: M_{200} with and without the central $0.1\ h^{-1}$ Mpc included in fitting. The green line is the $y = x$ line. Note that the points mostly follow this line.
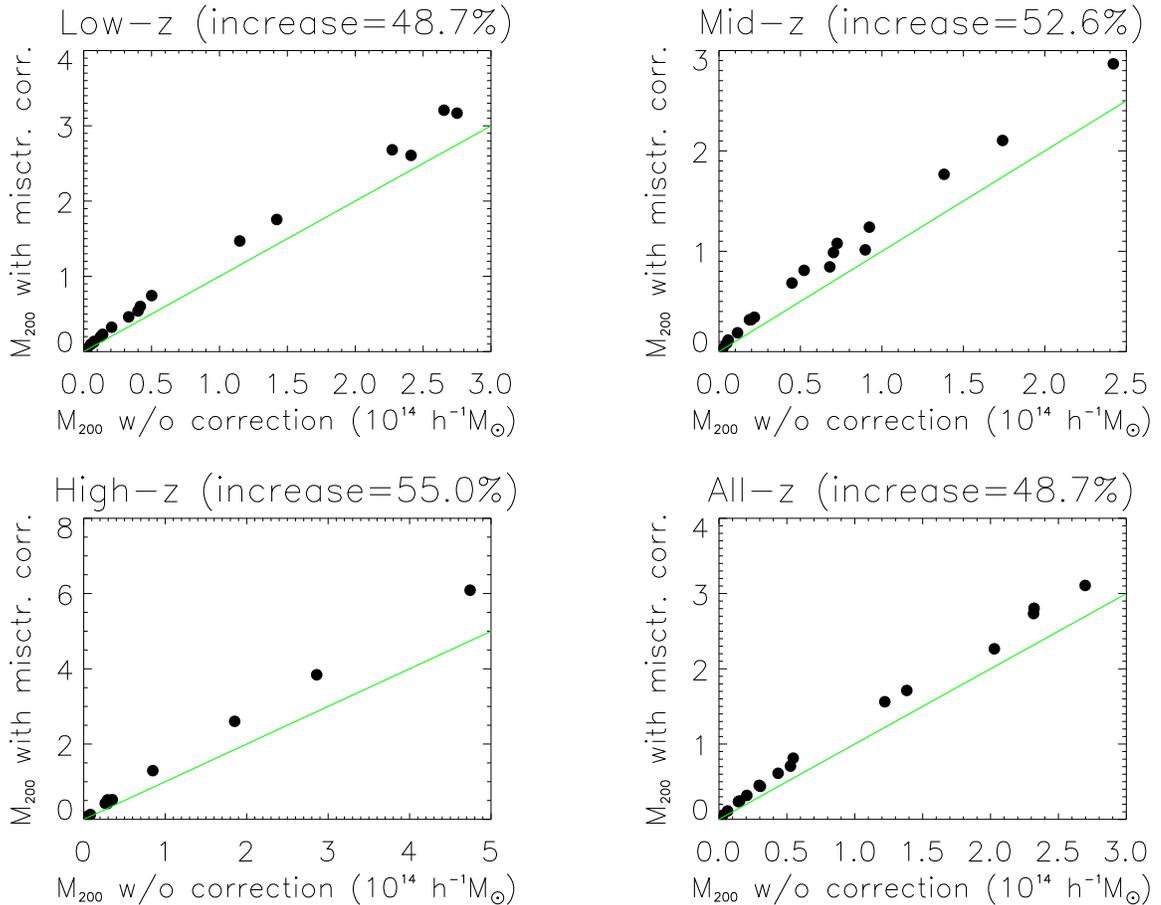
Figure 4.12: M_{200} values before and after miscentering correction. The green line is the $y = x$ line. Note that all masses are pushed upward by about 50%.

In Figure 4.12 we compare masses found without the miscentering correction to masses found after the miscentering correction. Note that all masses are systematically higher. For each of the redshift bins we found the ratio of miscentering-corrected mass to the uncorrected masses and then found the median of the ratios. This median is shown in Figure 4.12 as *increase*. Note that the percent increases are signficant and range from 48.7% to 52.6%. Thus we conclude that the miscentering correction causes masses to increase by about 50% and we conclude that miscentering corrections must be part of our final mass determinations.

4.5.3   <u>Errors on Photometric Redshifts</u>

All source galaxies in the sample had photometric redshifts measured for them. A photometric redshift (or photo-z) is a redshift measured by observing magnitudes in multiple filters and then using a neural network to estimate redshift based on this information and a small set of spectroscopic redshifts (spec-z). The relation between the magnitudes measured in different filters has a dependence on the redshift of an object, thus photometry can be used to estimate redshifts. Photometric redshifts are usually used in large surveys where it is implausible to find spectroscopic redshifts for most objects. The advantage to this method is that redshifts can be estimated for many objects, but the disadvantage is that errors on photometric redshifts are significantly larger than those for spectroscopic redshifts.

Our weak lensing mass measurements depend on photometric redshifts because tangential shear γ depends on Σ_{crit}^{-1}, which is a function of redshift. Thus we need to consider the errors which may be introduced into our measurements by uncertainties on photometric redshifts. To do this, we assembled a sample of spectroscopic redshifts to compare to.

We gathered public data from four data sets: The Stripe 82 coadd spectroscopic data; VVDS, DEEP2 and VIPERS. To access spectroscopy from the Stripe 82 coadd we used the SDSS-III CAS ($http : //skyserver.sdss3.org/CasJobs/$) to query for spectroscopic redshifts for objects within Stripe 82 that had i-band magnitudes $19 \leq i \leq 23.5$, type=3 (galaxies) and zwarning=0. The last parameter is a measure of the quality of the spectroscopic redshift. DEEP2 [23] is a project which uses the Keck telescopes at Mauna Kea. The DEEP2 project seeks to understand the evolution of galaxies and galaxy clusters by taking spectra of more than 50,000

Table 4.7: The number of objects in each spectroscopic sample.

Spectroscopic sample	Number of objects
DEEP2	7916
SDSS-III	21775
VIPERS	2093
VVDS	2396

galaxies out to redshift greater than 0.7. The VIMOS VLT Deep Survey (VVDS) [87] is a survey conducted using the VIMOS spectrograph on the Very Large Telescope (VLT). VVDS observed 3 fields that observed 35,526 galaxies of magnitude ranging between 17.5 and 24.75 and redshift between 0 and 6.7. The VIMOS Public Extragalactic Redshift Survey (VIPERS) [86] is a survey of high-redshift objects taken using the VIMOS spectrograph on the VLT. VIPERS observed two fields of objects chosen from the Canada-France-Hawaii Telescope Legacy Survey-Wide (CFHTLS-Wide) optical photometric catalog. In the first public data release of VIPERS there are 54,756 galaxies with spectroscopy. In each of these surveys we match the galaxy catalog with our own catalog of galaxies in Stripe 82. Our final catalog of objects with spectroscopic redshifts is a list of all objects from these 4 sets of data that overlap with objects in the Stripe 82 coadd. The total number of matched objects from each data set is given in Table 4.7, with 34,180 objects being matched altogether. After eliminating redundancies among the data sets, we found a total of 34,033 matched source galaxies with spectroscopic redshifts.

We first directly compared the photometric and the spectroscopic redshifts of these 34,033 objects, shown in Figure 4.13. We find fairly good agreement between spectroscopic and photometric redshifts, with this agreement decreasing at higher redshift, as would be expected. In order to better constrain the effect of photo-z
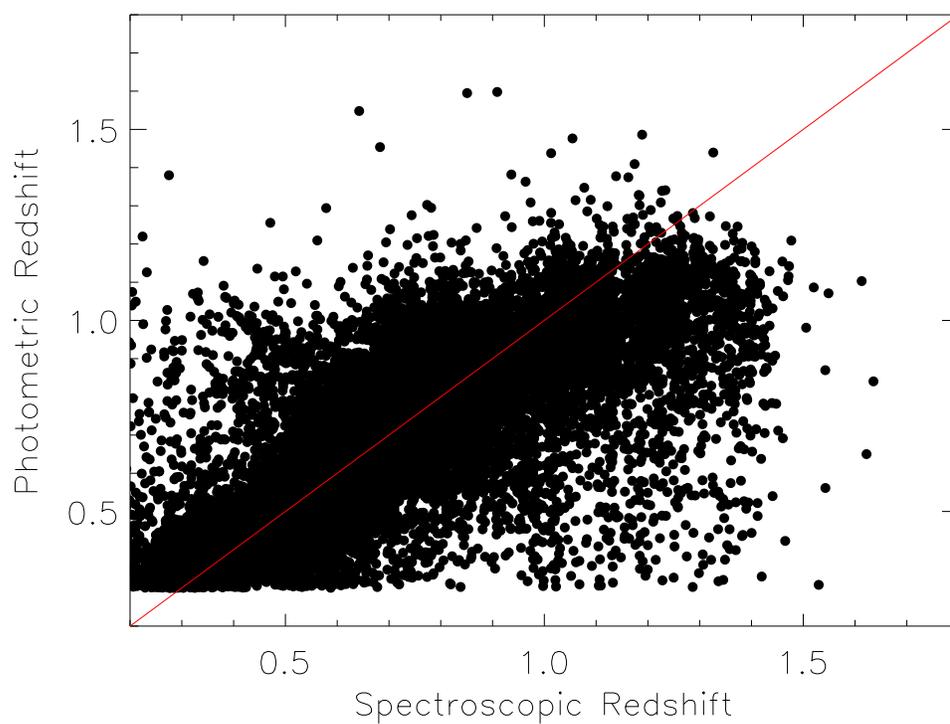
Figure 4.13: A comparison of photometric redshifts and spectroscopic redshifts for the 34,033 objects in the Stripe 82 coadd for which we also found spectroscopic redshifts. The red line is the $y = x$ line.

error, we analyzed how Σ_{crit}^{-1} changed with different measures of redshift. We looked at four different ways of finding Σ_{crit}^{-1} and then we compared them to see how much Σ_{crit}^{-1} varied based on the method:

1. Find Σ_{crit}^{-1} using the median photo-z of the source galaxies.

2. Find Σ_{crit}^{-1} for each source galaxy and then take the average. Use all photo-zs.

3. Find Σ_{crit}^{-1} for each galaxy that has a spec-z but use the photo-z for each of those galaxies. Photo-zs must be weighted.

4. Find Σ_{crit}^{-1} for each galaxy that has a spec-z but use the spec-z for each of those galaxies. Spec-zs must be weighted.

We discuss each of these methods individually next and then compare results. Note that Σ_{crit} is given by Equation 1.45.

4.5.3.1 $\underline{\Sigma_{crit}^{-1}}$ **from Median Photometric Redshift**

The original method (Method 1) by which we found Σ_{crit} was to find photometric redshift of the cluster and then find the median of all of the photometric redshifts of the source galaxies. These redshifts were used to find the angular diameter distances, then Σ_{crit} was calculated from them. Recall that galaxies were considered source galaxies if they were (1) within 3 h^{-1} Mpc of the BCG, (2) had i-band magnitude between 20.5 and 23.5, (3) had ellipticities less than 2 and (4) had $z_{source} > z_{cluster} + 0.1$. To test this method we produced a table of 1000 redshifts from $0.1-1.0$, treating these as cluster redshifts. We then took the first 1000 objects in the source galaxy catalog and looped through these to find source galaxies to correspond to each of

the "clusters"; we only required that the source galaxies have $20.5 \leq i \leq 23.5$ and that $z_{source} > z_{cluster} + 0.1$.

4.5.3.2 $\underline{\Sigma_{crit}^{-1}}$ from All Photometric Redshifts

For Method 2 (using all photometric redshifts) we used almost the same method as in the previous Section. However rather than finding the median source redshift, for each cluster we looped through all source galaxies and found Σ_{crit} for each. Then for each cluster we found a mean value of Σ_{crit}.

4.5.3.3 $\underline{\Sigma_{crit}^{-1}}$ from Photometric Redshift of Galaxies with Spectroscopic Redshifts

In Method 3 we began to use the 34,033 galaxies for which we had both photometric and spectroscopic redshifts. However since we were using this small spectroscopic sample to find values of Σ_{crit} for the full Stripe 82 coadd dataset, we had to weight the small spectroscopic sample to match the distribution of the much larger Stripe 82 sample. We utilized a routine in C++ called `calcweights.cpp` that was developed by C. Cunha [22] that uses a nearest-neighbor code to estimate a redshift distribution. The theory behind this code is described in Lima et al. (2008) [57].

This routine reads in two files, a training set that contains spectroscopic redshifts and observables for those galaxies with the spec-zs and a photometric redshift distribution with the observables for all galaxies in the photo-z distribution. The goal is to take a small data set with spectroscopic redshifts and match it to a photometric distribution by comparing to a set of observables. We used four observables:
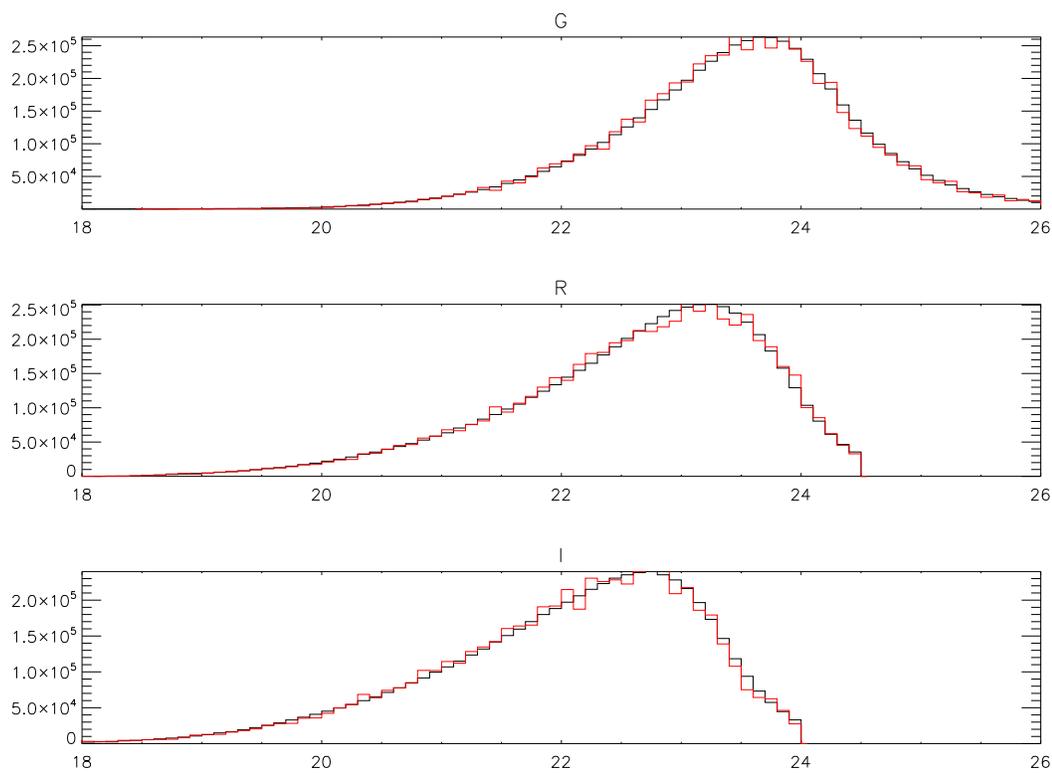
Figure 4.14: Magnitude distributions from the full Stripe 82 data and from the weighted spectroscopic sample. The black histograms are from the full Stripe 82 coadd photometric data while the red histograms are from the weighted spectroscopic sample.

g, r and i-band magnitude and photometric redshift. The routine allows you to vary the number of nearest neighbors used; we chose 10. The output of the routine is a set of weights; all the weights added together add to 1. The weight multiplied by the total number of photometric objects gives the number of times each galaxy with a spectroscopic redshift must be used to reproduce the distribution (e.g. on a magnitude histogram) of the photometric set. Figure 4.14 displays the magnitude histograms in g, r, and i-band for the entire Stripe 82 coadd source galaxy catalog. In it we also overlay the distribution reproduced by using the 34,033 galaxies with spectroscopic redshifts multiplied by the weights from this code.

Now that we had these weights we could find Σ_{crit} using Method 3. We produced a cluster sample like before, ranging in redshift from $0.1 - 1.0$. The source redshifts we used in Method 3 were the photometric redshifts that corresponded to the 34,033 galaxies with spectroscopic redshifts. Then for each cluster redshift we looped through all the (photometric) source redshifts, producing a value of $\Sigma_{crit} \times weight$ for each. Finally we found the mean by finding for N source galaxies:

$$\Sigma_{crit\_ave} = \frac{\sum\limits_{i=1}^{N} \Sigma_{crit\_i} \times weight_i}{\sum\limits_{i=1}^{N} weight_i} \tag{4.36}$$

4.5.3.4 Σ_{crit}^{-1} from Spectroscopic Redshifts

Finally in Method 4, we used almost exactly the same method as Method 3, but this time the source redshifts were the spectroscopic redshifts. Again we found $\Sigma_{crit} \times weight$ for each source galaxy for a particular cluster redshift and again we found the mean value of Σ_{crit} for that cluster redshift by using Equation 4.36.

4.5.3.5 Conclusion

Our final product in each case was a lookup table of values of Σ_{crit} for a set of cluster redshifts ranging from $0.1 - 1.0$. Finally we computed Σ_{crit}^{-1} as that is the quantity needed for lensing analysis. In Figure 4.15 we plot Σ_{crit}^{-1} as a function of redshift from $0.1 - 1.0$. In this plot the green line is from Method 1, the black line is from Method 2, the red line is from Method 3 and the blue line is from Method 4. Note that the original method of finding Σ_{crit}^{-1} (Method 1) deviates the most from
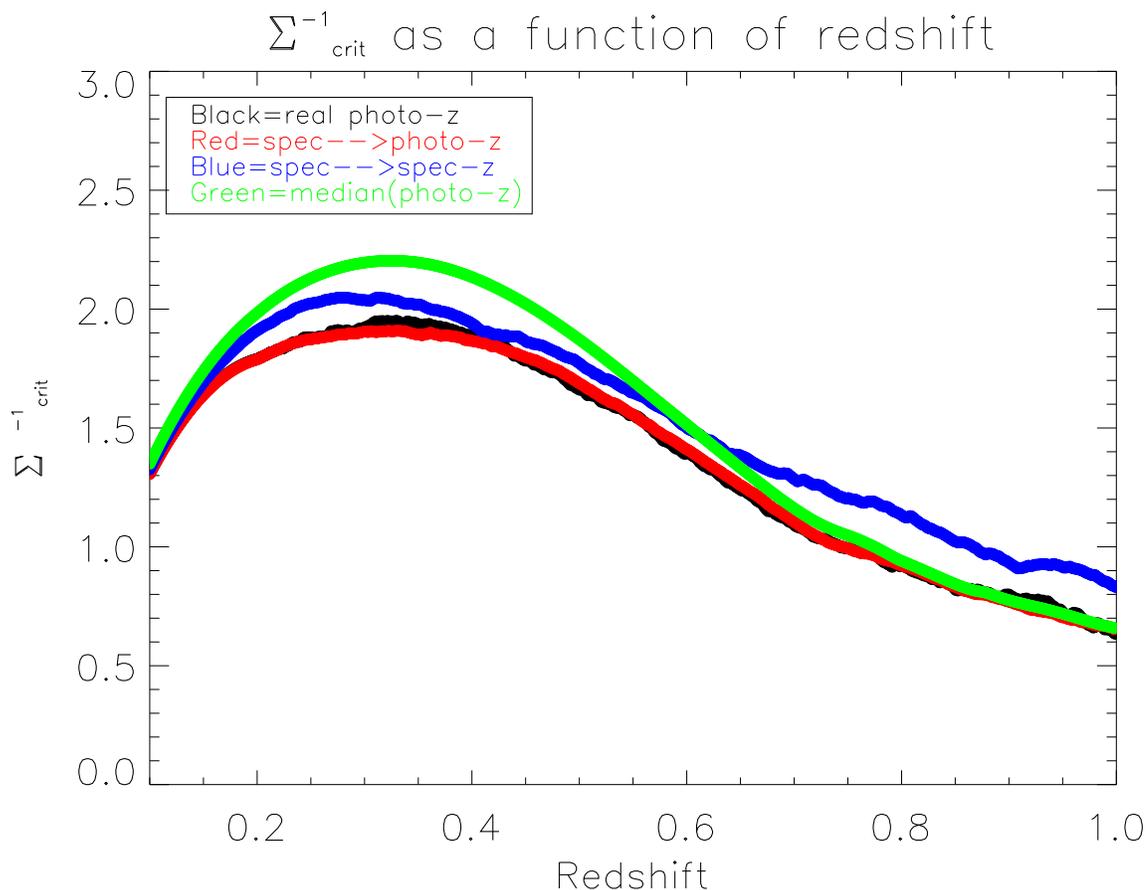
Figure 4.15: Σ_{crit}^{-1} as a function of redshift for each of the four different methods described in this section. The green line is from Method 1, the black line is from Method 2, the red line is from Method 3 and the blue line is from Method 4. All values of Σ_{crit}^{-1} are multiplied by 10^{-16} m^2/kg.

the other methods. The red and the black lines (Methods 3 and 2) are the closest, as they should be, since the red line describes the small photometric set projected to represent the full photometric set using the weights.

We finally conclude that there are significant differences in Σ_{crit}^{-1} depending on which method of finding it we use. We also conclude, from the closeness of the black and the red lines, that the weights code is highly effective at reproducing the photometric distribution from the spectroscopic distribution. Therefore we finally

will use the weighted distribution of spectroscopic redshifts (the blue line in Figure 4.15) to calculate Σ_{crit}^{-1}.

4.5.4 Foreground Galaxy Contamination

4.5.4.1 The Problem

Weak lensing shear was measured on a sample of galaxies that were more distant than each of the lensing clusters. In NFW.pro we measured shear only on galaxies that were at a redshift that was 0.1 greater than cluster redshift. The redshifts used for these cuts are photometric redshifts, and thus they have larger errors than if they were spectroscopic redshifts. These errors can allow some galaxies that are actually at lower redshift than the cluster (i.e., in front of the cluster) to be included in the sample of background source galaxies. This would contaminate the shear signal with galaxies that cannot exhibit shear. This is especially a problem nearby a galaxy cluster, as cluster galaxies can be misidentified as background galaxies. The contamination can be quantified by noting that average tangential shear (γ_{ave}) in a particular richness bin is

$$\gamma_{ave} = \frac{\sum_{i=1}^{N_{real}} \gamma_{T\_real} + \sum_{i=1}^{N_{fake}} \gamma_{T\_fake}}{N_{real} + N_{fake}} \tag{4.37}$$

where *real* and *fake* refer to source galaxies and misidentified foreground galaxies respectively and N means total number of galaxies found in that richness bin. We expect $\Sigma\gamma_{T\_fake}$ to sum to zero since this includes only shape noise and no actual shear. However, the N_{fake} term remains, and we need to find a correction factor $C(r)$

to account for it. What we want is to remove the N_{fake} term in the denominator. Thus we need to multiply the measured γ_{ave} by this term:

$$C(r) = \frac{N_{real} + N_{fake}}{N_{real}} \qquad (4.38)$$

But since we cannot directly measure N_{real}, we can instead measure a number density, the number of galaxies per area per cluster. Thus

$$C(r) = \frac{N_{real} + N_{fake}}{N_{real}} = \frac{n_{measured}(r)}{n_{random}} \qquad (4.39)$$

where n means number density. Thus $n_{measured}$ is the number of galaxies measured per area per cluster at various radii around clusters; this tells us about $N_{real} + N_{fake}$ because it accounts for all apparently source galaxies found nearby clusters (including misidentified cluster galaxies). The second term, n_{random}, is the number of galaxies per area per cluster measured either in a random part of the sky not near a cluster or far enough away from the cluster that the concentration of galaxies that marks a galaxy cluster does not affect the number density. We will use the latter in our corrections, but we will show that the two options (random part of sky or far from cluster) give the same results.

4.5.4.2 The Solution

Number density n is measured by finding the total number of galaxies per area per cluster in a given richness bin. To find this, we go to the output files from `NFW.pro` and find the total number of galaxies that have passed the cuts to be considered source galaxies in increments of $0.1h^{-1}$ Mpc of distance from the BCG. We then

divide this number by the product of the area of an annulus at that increment of distance from the BCG and the number of clusters in that richness bin. In other words we measure

$$n = \frac{\frac{N_{total}}{A_{annulus}}}{N_{clusters}} \tag{4.40}$$

where N_{total} is the total number of apparently source galaxies in that richness and redshift bin, $A_{annulus}$ is the area of the annulus and $N_{clusters}$ is the number of clusters in that bin. We initially used the same richness bins that we also used to measure cluster masses (presented in Table 4.1), however we found that these bins contained too few clusters at higher richness to allow us to have reasonable statistics for the number density. Thus we instead made a set of bins that combined the higher richness bins, as shown in Table 4.8.

When we first measured the number densities, we found that they all decreased with radius, even for the random points, as seen in the upper panels of Figure 4.16. This is unexpected behavior, as random points should overall have no dependence of number density on position. It was found that this behavior was due to the nature of Stripe 82 as a thin strip of observations (2.5 degrees wide in declination). For clusters that were near the edge of the images the number of nearby galaxies that could be measured would be limited to those that appeared on the images. As the distance from the cluster center increased, the number of galaxies per area would drop, not due to an actual decrease but simply because we had progressed outside the available imaging region. To address this issue, we reran `NFW.pro` while rejecting any clusters for which a circle of radius 3.0 h^{-1} Mpc would be off the imaging area. The result was as shown in the lower panels of Figure 4.16. Since we reduced the number of clusters used in this analysis in order to avoid biasing by clusters too close
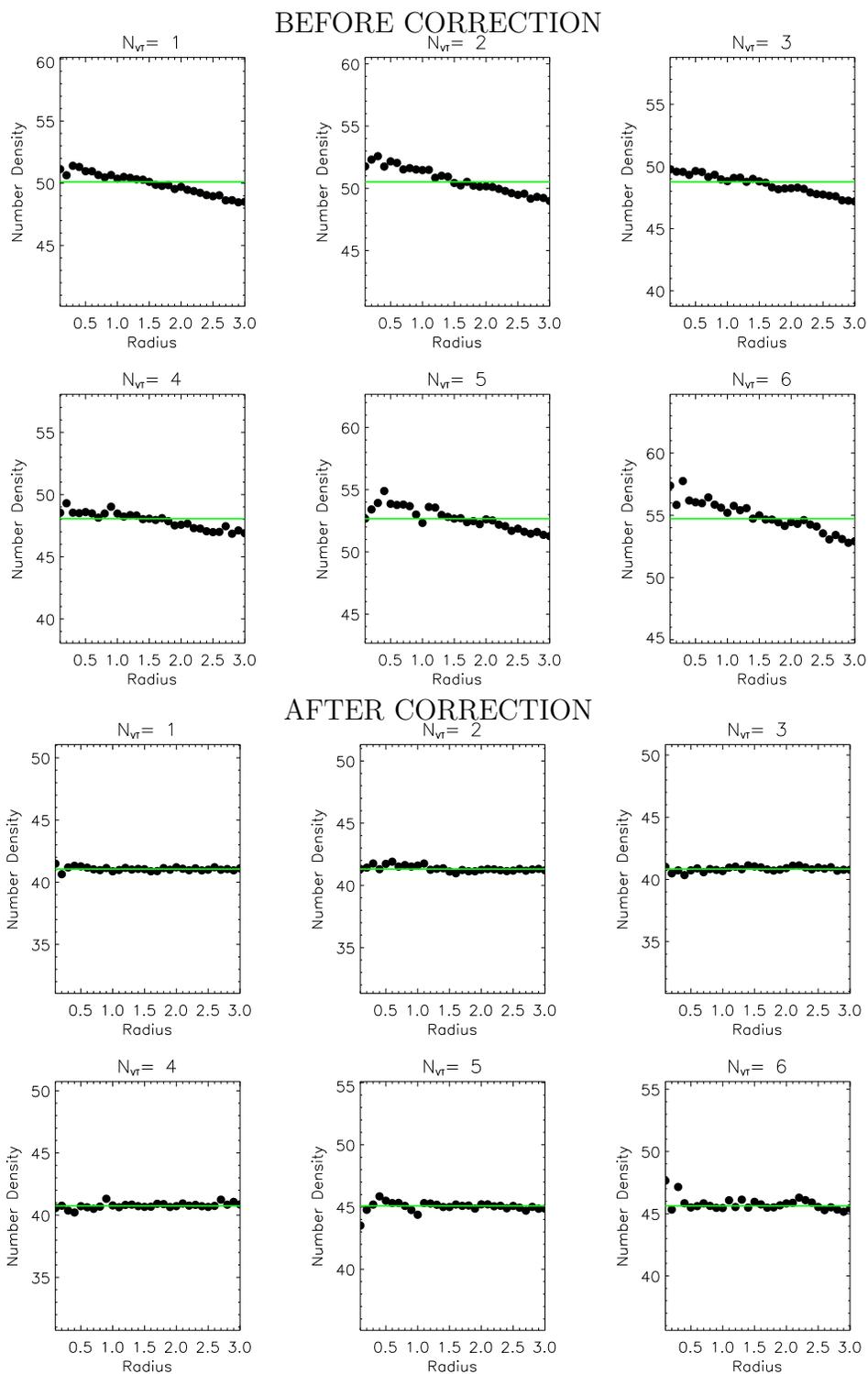
Figure 4.16: A plot of number density as a function of radius for the first six richness bins, done before (upper) and after (lower) removal of clusters that were close to the edges of Stripe 82. This is for $0.4 \leq 0.7$. The green line is the median number density for that richness bin. Note that in the lower panel, the data points form a much flatter horizontal line, as is expected.

Table 4.8: The bins used to measure galaxy number density, the number of clusters (N) in each bin and the number of clusters that passed cuts to be considered completely contained within the Stripe 82 data set. Low-z is $0.1 < z \leq 0.4$, mid-z is $0.4 < z \leq 0.7$, high-z is $0.7 < z \leq 1.0$ and all-z is $0.1 < z \leq 1.0$.

Bin number	N_{VT} range	N (low-z)	N (low-z, pass cut)	N (mid-z)	N (mid-z, pass cut)	N (high-z)	N (high-z, pass cut)	N (all-z)	N (all-z, pass cut)
1	1	543	285	1205	1042	2355	1990	4103	3399
2	2	619	289	1310	1119	2515	2212	4444	3787
3	3	428	222	924	808	1693	1461	3045	2601
4	4	275	143	654	586	1168	1029	2097	1799
5	5	238	132	411	359	772	677	1421	1213
6	6	179	95	318	269	555	467	1052	877
7	7	126	46	223	196	354	314	703	606
8	8	98	64	171	148	263	224	532	442
9	9-10	150	76	224	198	303	257	677	561
10	11-20	266	156	357	319	394	340	1017	854
11	21-127	123	89	86	76	62	53	271	216

to the edge of Stripe 82, we also present the final number of clusters that passed cuts to be completely on Stripe 82 in Table 4.8.

We found that number density for random points is about the same as the number density in the region of a cluster but far away from the center (see Figure 4.17). Because of this we chose to use the number density in the region of clusters but far away as the reference point for our foreground corrections. As the number density approached a constant by $1.4h^{-1}$ Mpc, we measured the number density between $1.4 - 3.0h^{-1}$ Mpc. Plots of number density for clusters for all redshift and richness bins may be seen in Figures A.1, A.2, A.3 and A.4 in Appendix A. The final foreground corrections as defined by Equation 4.39 are provided for the lowest radius ($0.1h^{-1}$ Mpc) in Table 4.9. We give the lowest radius as these are

Table 4.9: Foreground corrections at $0.1h^{-1}$ Mpc as a function of richness bin. Low-z is $0.1 < z \le 0.4$, mid-z is $0.4 < z \le 0.7$, high-z is $0.7 < z \le 1.0$ and all-z is $0.1 < z \le 1.0$.

Richness bin	Low-z Correction	Mid-z Correction	High-z Correction	All-z Correction
1	1.06	1.22	1.00	1.01
2	1.05	1.24	1.00	1.00
3	1.05	1.40	1.06	1.00
4	1.09	1.41	1.03	1.00
5	1.11	1.52	1.12	1.00
6	1.04	1.54	1.13	1.04
7	1.18	1.48	1.00	1.00
8	1.15	1.50	1.06	1.12
9-10	1.06	1.46	1.00	1.00
11-20	1.17	1.43	1.03	1.00
21-127	1.03	1.26	1.05	1.00

the largest corrections. The corrections are applied at for each radius bin (of size $0.1h^{-1}$ Mpc), but beyond the first several bins most corrections are 1 (meaning no change). We did not allow any of the final corrections to fall below 1 even if the value of the correction factor was less than the reference value at that point. We did this because a final correction less than 1 is not physically meaningful. All values of average tangential shear (γ_{ave}) were multiplied by the correction factors before fitting was done.

4.6   Null and Other Systematics Tests for Stripe 82 Coadd

A null test means a test that measurements where no signal should be expected do in fact show no signal. We conducted two null tests: First, we measured or-
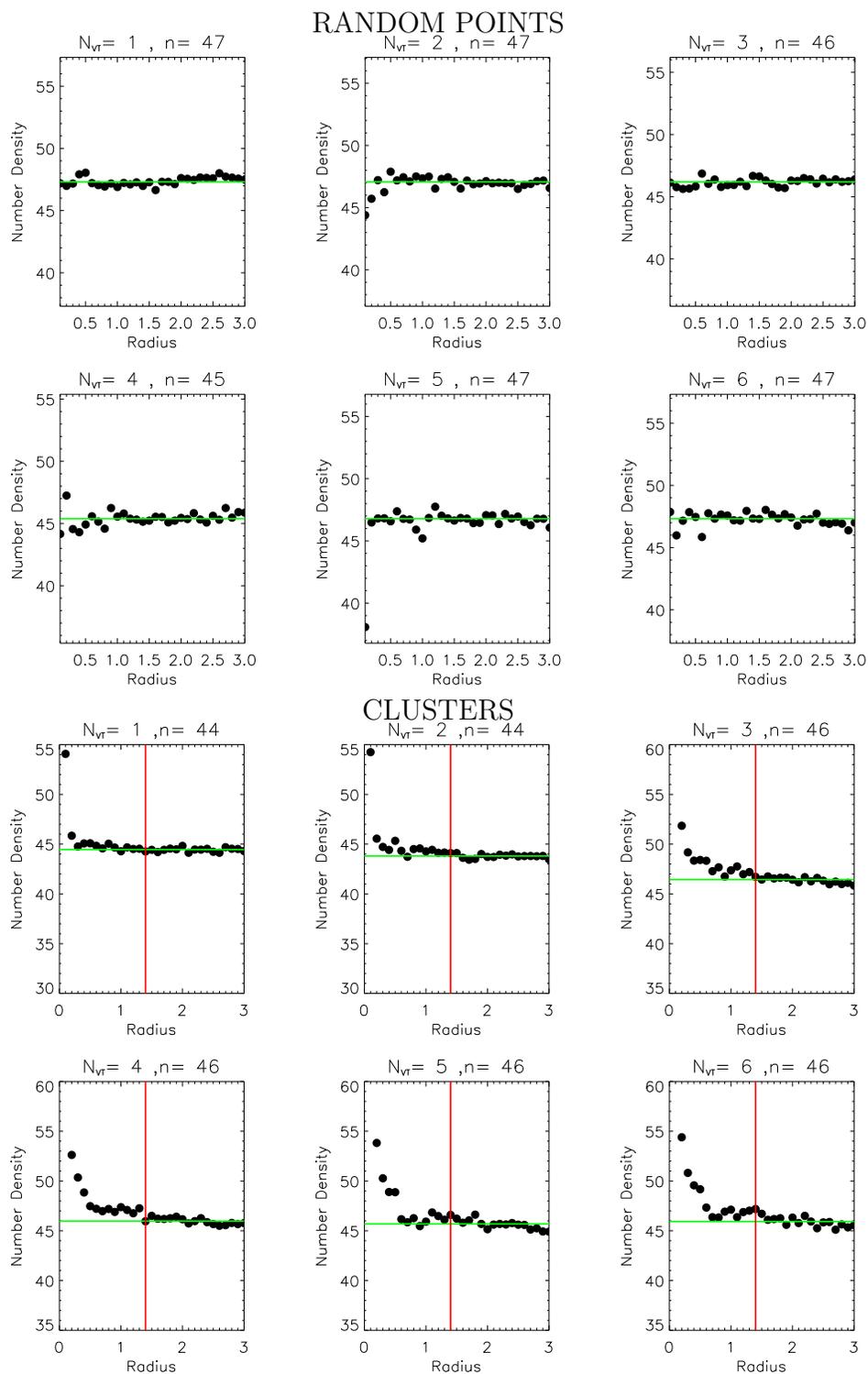
Figure 4.17: Plots comparing number density found for the first six richness bins for $0.4 \leq 0.7$ for random points (upper) and actual clusters (lower). The green line represents the median value of number density and the vertical red line represents the beginning radius at which the median number density was found for actual clusters.

thotangential shear for each richness and redshift bin, as orthotangential shear is not produced by lensing. Second, we measured weak lensing signal around random points in the sky rather than around galaxy clusters. Lastly we conducted one further systematics test: we checked to see if there was a dependence of shear on sky position in Stripe 82.

4.6.1   Orthotangential Shear Test

Every time we run `NFW.pro` we obtain shear profile plots for both tangential and orthotangential shear. For tangential shear we expect to see a curve up near the center of the profile (consistent with non-zero tangential shear) while for orthotangential shear we expect to see a line running along $x = 0$ (consistent with zero tangential shear). If we see significant deviations from this flat line for orthotangential shear, this can be evidence of error as weak lensing will not produce orthotangential shear. As seen in Figure 4.9 typically we can see a clear signal in tangential shear while the orthotangential shear is apparently consistent with zero.

In Figure 4.18 we compare the χ^2 probability calculated for the tangential shear to that calculated for orthotangential shear. Recall from Equation 4.35 that this quantity tells us the probability that the χ^2_{red} for a fit to 0 will exceed the given value of χ^2_{red}. For null data, like we expect orthotangential shear signal to be, this probability should be on the order of 0.5, since we expect the orthotangential shear to be well fit to 0. For real data this probability should be very small, since we would expect real shear measurements to not be consistent with 0. Note that in Figure 4.18 we find that χ^2 probability is almost always significantly higher for orthotangential shear than it is for tangential shear. For orthotangential shear, the
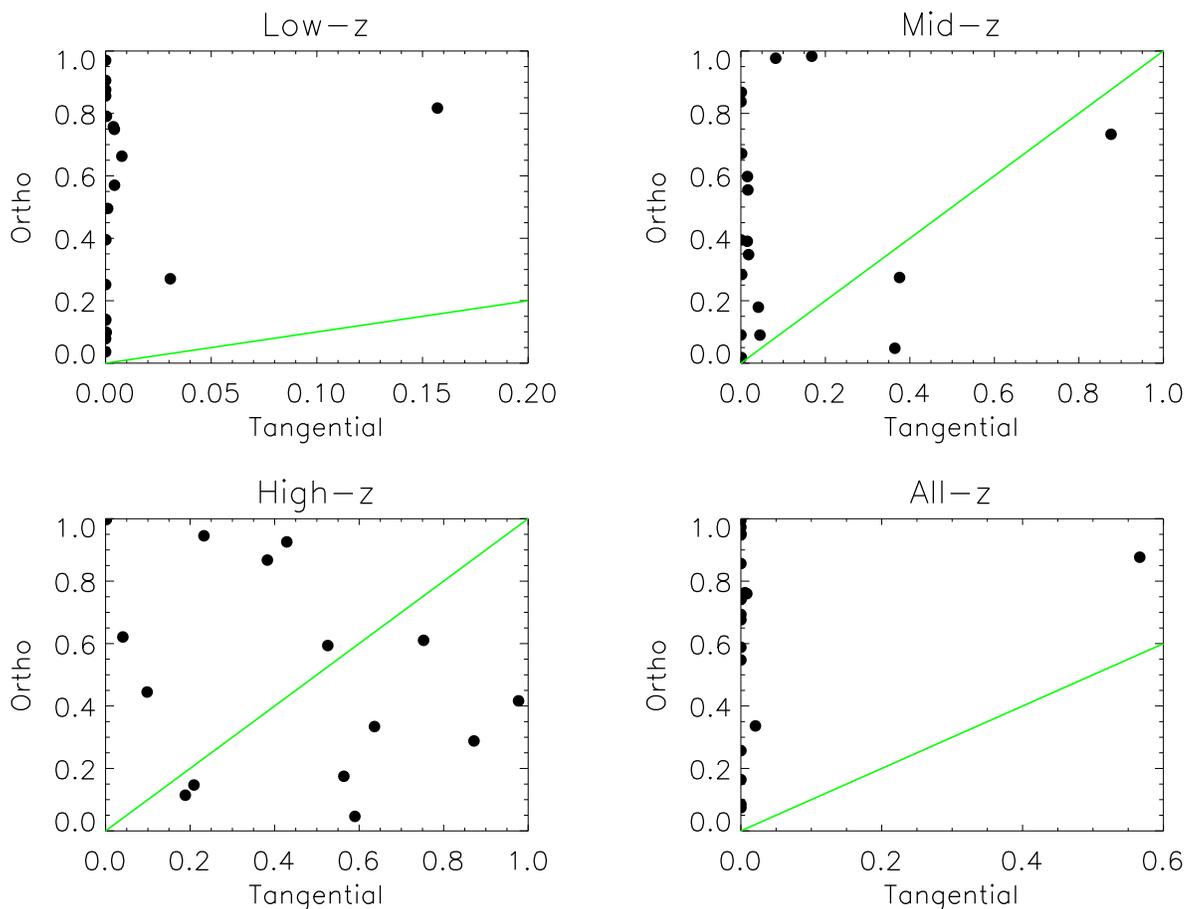
Figure 4.18: χ^2 probabilities compared for tangential and orthotangential shear. The green line is the $y = x$ line.

probability ranges from 0 to 1, and so on average is about 0.5, as expected. In most cases, the probabilities for tangential shear are near 0, meaning they are not well fit to a tangential shear of 0. Thus we conclude that in most cases, there is evidence for tangential shear signal, and not for orthotangential shear. The one notable exception is the high-z data. The high-z data has the lowest statistics of any of the bins, and so the tangential shear fits have higher error bars. The tangential shear signal is less obvious, with χ^2 probabilities from 0 to 1, while the orthotangential shear signal is unchanged.

4.6.2   Random Points Test

We would expect no tangential shear signal around random points in the sky while we would expect a non-zero tangential shear around clusters. If random points were to give signal similar to that measured near clusters, this would be evidence that there is a systematic problem in our measurements. In order to conduct this random points test we first generated a sample of random points. As there are 19,706 clusters in the complete Stripe 82 VT sample, we also generated 19,706 random points. There are 11,123 clusters in the low RA ($\leq 59.7°$) field and 8,583 clusters in the high RA ($> 310°$) field. To generate random points, we used the IDL function `RANDOMU`. The commands used for the low RA field were:

$$A = (1 + 58.7 * RANDOMU(seed, N))$$
$$DEC1 = (-1.25 + 1.25 * RANDOMU(seed, N/2)) \tag{4.41}$$
$$DEC2 = (1.25 - 1.25 * RANDOMU(seed, N/2))$$

The function $RANDOMU$ generates N pseudo random numbers in the range $0 < Y < 1.0$. The *seed* is an integer used to initialize the random number generator; we took it to be 1.0. Note that for the RA we multiplied the output from `RANDOMU` by 58.7 so that the maximum RA corresponds to the maximum RA in this section of the real data, 59.7°. The declination section is divided into two pieces, as there are both positive and negative declinations. This entire process was repeated for the high RA field. Finally we produced a catalog of these 19,706 random RA and DEC and kept the real cluster values of N_{VT} and z.

Once we generated the random points catalog, we reran `NFW.pro` using the random points as the clusters. We found that the tangential shear profiles generated

were now consistent with zero, as would be expected for random points. In Figure 4.19 we show the same shear profiles as shown in Figure 4.9, but now for the random points. Note that for the low-z N_{VT} $11 - 12$ bin, the tangential shear is clearly consistent with zero. In the high-z bin it is not as clearly consistent with zero, but note the larger error bars. This is as was noted before, that the high-z bin has the lowest statistics (since there are far fewer source galaxies available) and therefore the fits have larger error bars. Nonetheless if you compare Figure 4.19 to Figure 4.9, you will see the latter has a clear tangential shear signal while the former does not.

In Figure 4.20 we plot the reduced χ^2 values from tangential shear fits for both the actual data (upper) and for the random points (lower). Note that for the real data the reduced χ^2 values are mostly larger than 1, meaning that the real data is not consistent with a model of no shear signal. On the other hand, for the random points the reduced χ^2 values are mostly near 1, meaning that the random points are consistent with no tangential shear. This reinforces our claim that we have measured a significant weak lensing shear signal.

4.6.3   Dependence of Shear on Sky Position

The last test conducted was to check if shear depends on sky position, as we would not expect such a dependence. The way we did this was to divide the clusters in the Stripe 82 coadd into 6 equal sections in declination. These sections were:

1. $-0.75° \leq DEC < -0.5°$

2. $-0.5° \leq DEC < -0.25°$
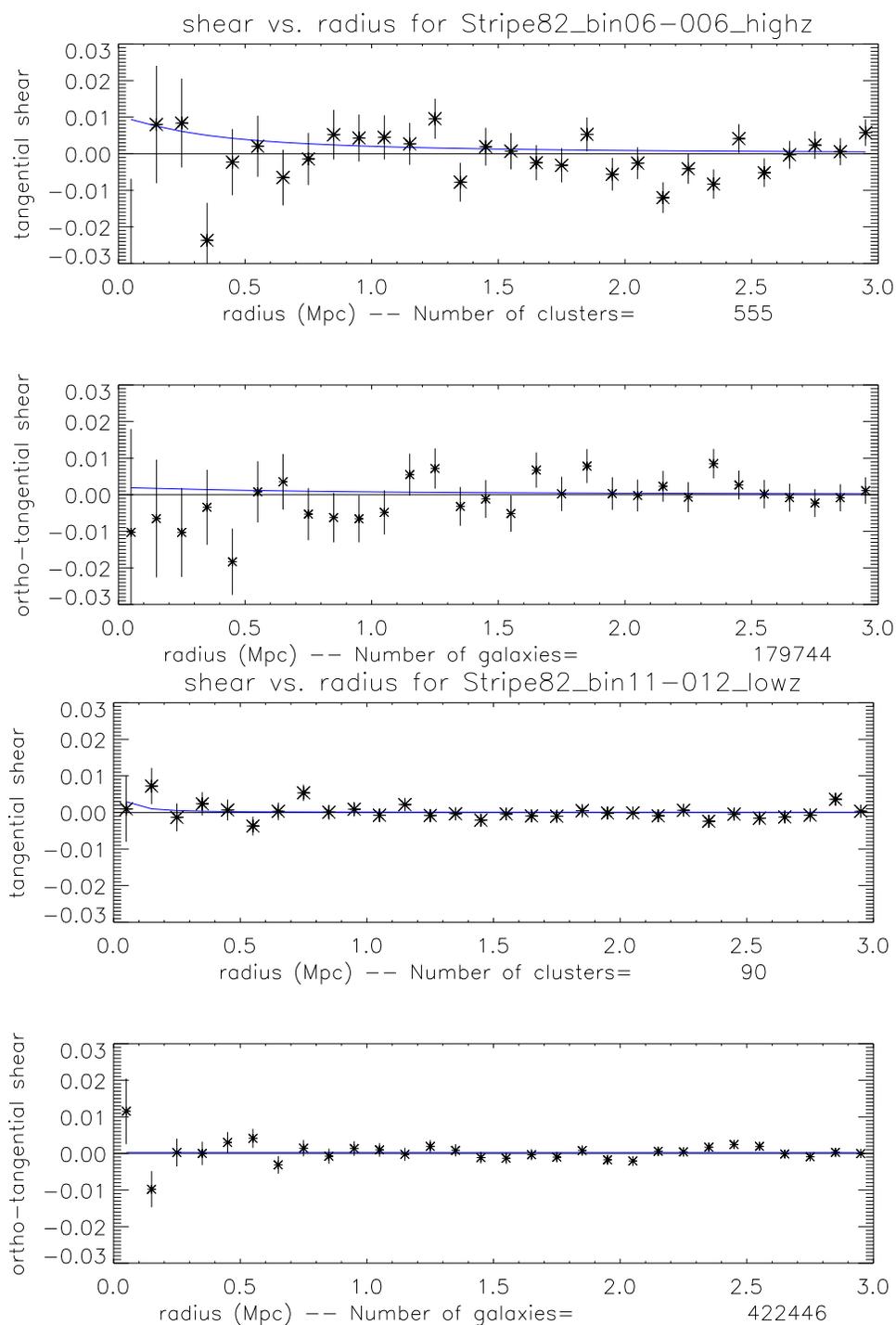
Figure 4.19: Shear profiles for random points. Note that for the low-z case the tangential shear is consistent with zero. For the high-z case, there is a fit to the points, but there is more scatter than there was in Figure 4.9.

REAL POINTS
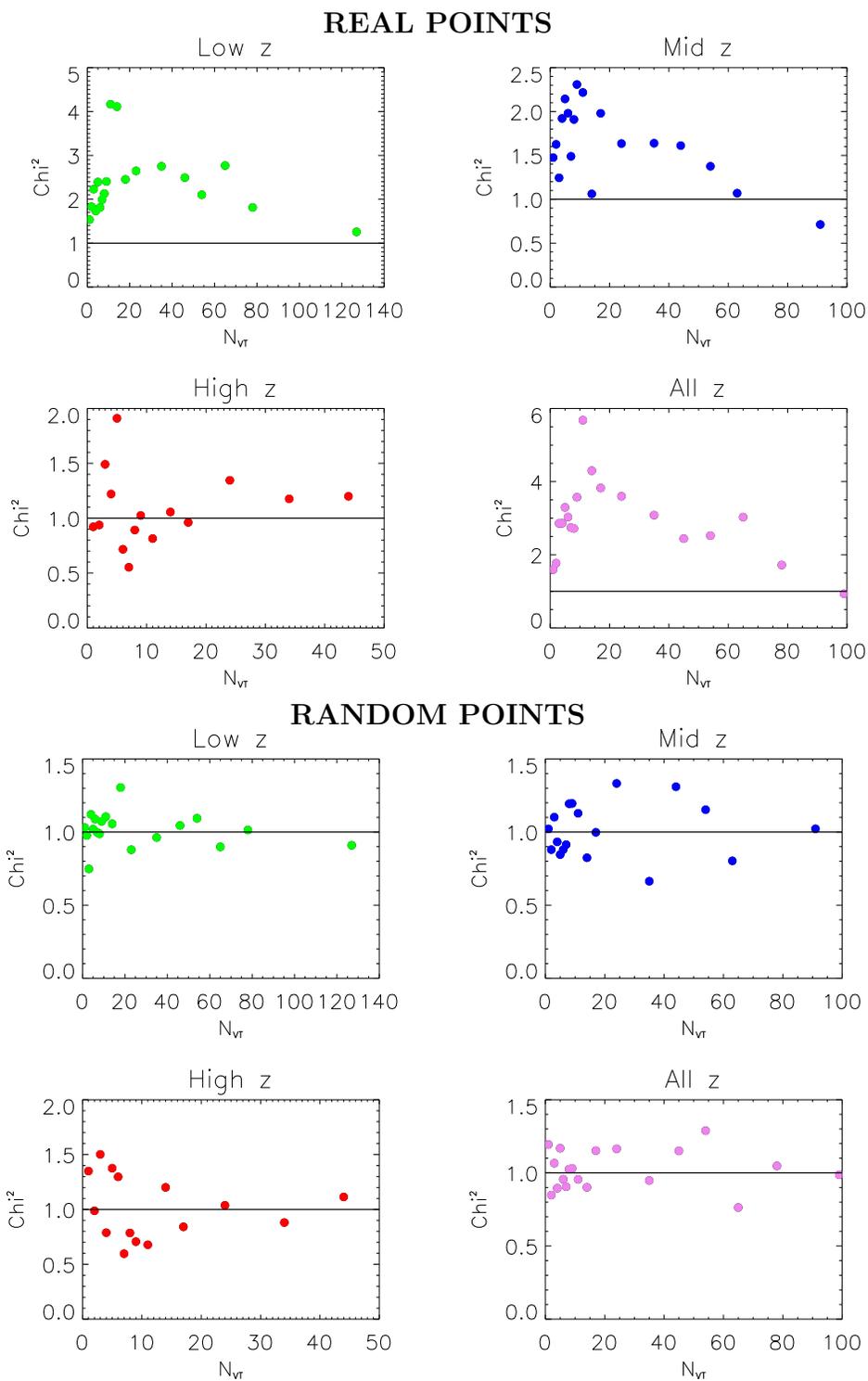
RANDOM POINTS

Figure 4.20: Reduced χ^2 values for tangential shear fits to clusters and for the same fits to random points, given compared to N_{VT}. Note that for the clusters, the reduced χ^2 values are well above 1, meaning they are not consistent with a model of no shear signal. For the random points, reduced χ^2 values are close to 1, suggesting no evidence of shear signal (as expected).

3. $-0.25° \leq DEC < 0°$

4. $0° \leq DEC < 0.25°$

5. $0.25° \leq DEC < 0.5°$

6. $0.5° \leq DEC < 0.75°$

We chose these sections instead of the full range from $-1.25° - 1.25°$ so that no clusters were right on the edge of the area (we noted in § 4.5.4 that this could be a problem). We then reran `NFW.pro` allowing in each case only clusters fitting within these regions.

In Figure 4.21 we plot the values of M_{200} found for each stripe of declination for the lowest four values of richness, $N_{VT} = 1, 2, 3$ and 4. We plot these richness bins as they have the most clusters and thus the best statistics. There is no obvious trend of variation in mass with position in the sky, thus we conclude that there is no evident dependence of shear on sky position.

4.7   Mass-Richness Relations for Stripe 82 Coadd

4.7.1   The Mass-Richness Code

In order to find a mass-richness relation by using the weak lensing shear results, we developed a routine called `MASS_RICHNESS.pro`. First we read in richness and mass results from `NFW.pro`. Then we used the IDL routine `FITEXY` to apply a linear fit to $\ln N_{VT}/20$ (as independent variable) and $\ln M_{200}$ (as dependent variable). `FITEXY` calculates a linear fit while including errors on x and y; for errors we used the
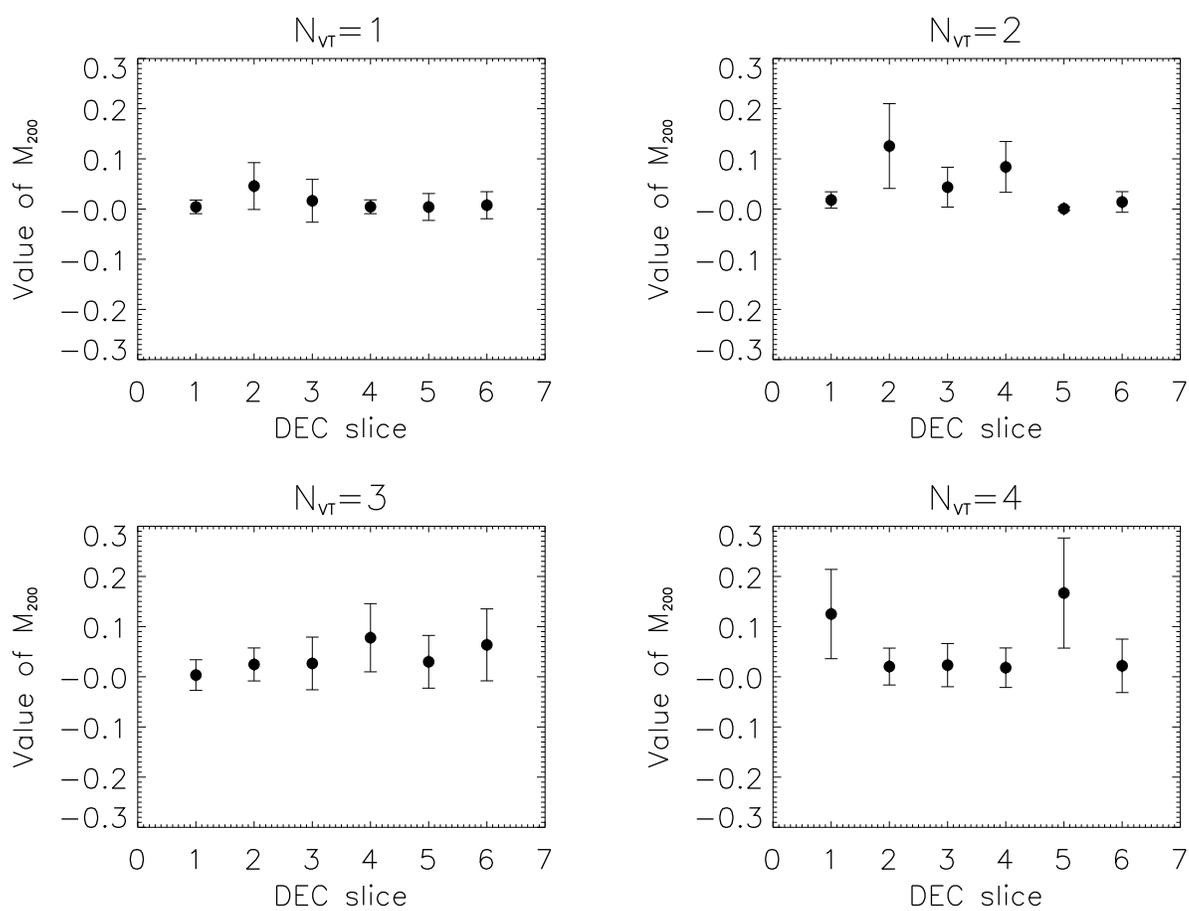
Figure 4.21: Mass as a function of declination for $N_{VT} = 1, 2, 3$ and 4. The declination slices are 0.25° wide, given by 4.6.3.

fractional error, $\Delta N_{VT}/N_{VT}$ and $\Delta M_{200}/M_{200}$. We used natural log for the fit as we want to find a power law relation, as in Section 4.1.2. Note that if we want to find this:

$$y = ax^b \tag{4.42}$$

then we need to find this in a linear fit:

$$\ln y = b \ln x + \ln a \tag{4.43}$$

Thus if we fit a linear relation to $\ln x$ and $\ln y$, then the power b will be equal to the slope and the coefficient a will be e raised to the power of the y-intercept. We also found values of reduced χ^2 and χ^2 probability for each fit. Before running MASS_RICHNESS.pro we tested it by running it on the M_{200} and N_{200} data from Johnston et al. [47]. We found that we exactly reproduced their mass-richness relation, so we were confident that the routine functioned well. In each case we found mass-richness relations for four separate redshift bins, $0.1 - 0.4$, $0.4 - 0.7$, $0.7 - 1.0$ and $0.1 - 1.0$. Note that all mass-richness relations are given in the form:

$$M_{200} = M_{200|20} \left(\frac{N_{VT}}{20} \right)^{\alpha} \tag{4.44}$$

where α is the parameter we called b in Equation 4.42, the *slope of the power law.* This parameter tells us how quickly mass increases with an increase in richness. $M_{200|20}$ is the *mass coefficient* of M_{200} when $N_{VT} = 20$, and corresponds to the parameter a in Equation 4.42.

4.7.2 Initial Mass-Richness Relations for the Stripe 82 Coadd

Our initial mass-richness relations for the Stripe 82 coadd were found with no corrections for the systematics discussed in Section 4.5. We ran `NFW.pro` using the four redshift bins and then ran `MASS_RICHNESS.pro` on the results. The four mass-richness fits are shown in Figure 4.23 and our numerical results are shown in Table 4.10. In Figure 4.23 we note that the error bars are smallest for the low-z and all-z bins, while they are largest for the high-z bin. The violet line is the fit to the all-z relation; we overplot each relation with this line to compare. We note that the slope of the power relation is significantly higher for the high-z relation than for the others. The reduced χ^2 value is the best for the low-z bin.

We also note that in all plots there are at least several points well out of line with the others. For the low-z bin, there are two points that are well out of line with the others. These are the points for $N_{VT} = 8$ and $N_{VT} = 9 - 10$. In Figure 4.22 we present the shear profiles for these points along with that for $N_{VT} = 4$. Note that there are 98 clusters for the $N_{VT} = 8$ bin and 150 clusters for the $N_{VT} = 9 - 10$ bin while there are 275 clusters for the $N_{VT} = 4$ bin. The fits are less good for the first two larger richness bins, and there is some evidence of an orthotangential shear signal. For the $N_{VT} = 4$ bin orthotangential shear is 0 and the tangential shear signal is more clear. Thus we conclude that these two points are simply outliers because of poor statistics.
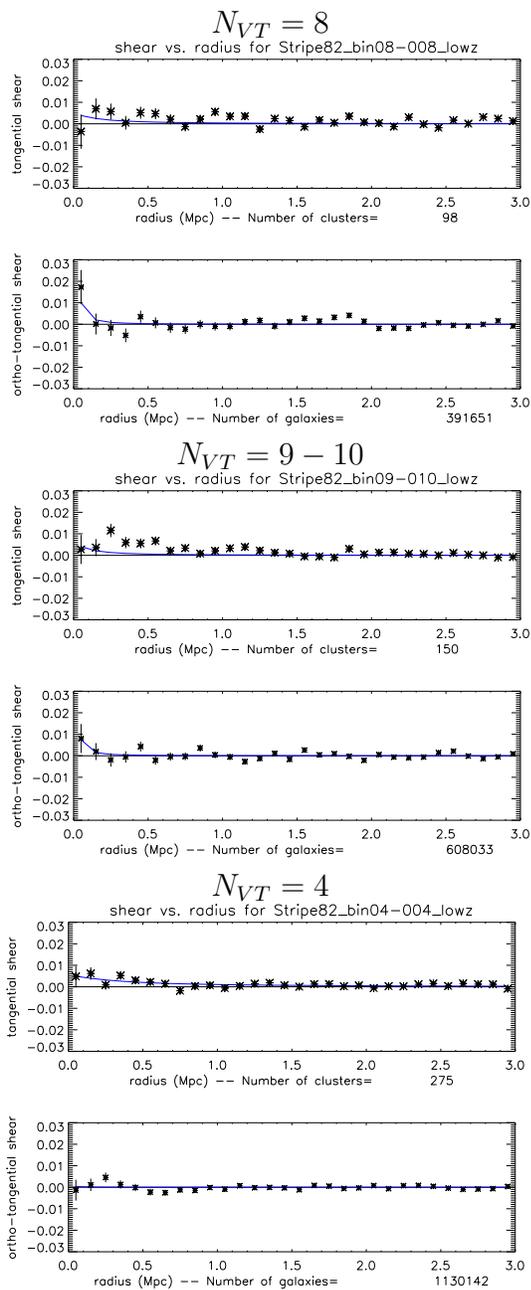
Figure 4.22: Shear profiles for several richness bins in the low-redshift bin. Note that for $N_{VT} = 8$ and $N_{VT} = 9 - 10$ the tangential shear fits are not as good, and there is more scatter in the shear signal. For $N_{VT} = 4$ this is not true.

4.7.3 Mass-Richness Relations After Considering the Effect of the Central BCG

As discussed in Section 4.5.1, we removed any source galaxies within 0.1 h^{-1} Mpc of the BCG in order to try to remove the effect of the mass of the BCG. In Figure 4.24 we plot the mass-richness relations obtained from this analysis and in Table 4.11 we present the numerical results. The low-z bin is affected the least by the change, with the mass coefficient and the slope of the power law almost unchanged. The values for the high-z bin are affected the most, although the error bars are also increased significantly. The values of reduced χ^2 are somewhat better than for the case of no corrections. There is no general trend on the results here, as some numbers are increased and others are decreased. Our conclusion is in agreement with what we found in Section 4.5.1, that the removal of the central region of the halo has a small effect on the final results.

4.7.4 Mass-Richness Relations After Considering the Effect of Halo Miscentering

We next considered the effect of halo miscentering, as discussed in Section 4.5.2. The miscentering correction from Equation 4.9 was applied directly in the `MASS_RICHNESS.pro` routine. The mass-richness plots are shown in Figure 4.25 and the numerical results are in Table 4.12. All of the mass coefficients are increased by about 37% and all of the power law slopes are decreased by about 13% from the

numbers found without any consideration of systematics. The values of reduced χ^2 are almost identical to those found for the case of no corrections.

4.7.5 Mass-Richness Relations After Considering the Effect of Photo-z Bias

Next we considered the effect of photo-z errors, finding and using Σ_{crit} using the sample of spectroscopic redshifts as discussed in Section 4.5.3. We produced a lookup table of values of Σ_{crit} using these spectroscopic redshifts and then had `NFW.pro` refer to this table when finding Σ_{crit} for a particular cluster. The correction factor (Equation 4.24) was then calculated as (following Equation 4.27):

$$correction = \frac{\Sigma_{crit\_actual}}{\Sigma_{crit\_fiducial}} \tag{4.45}$$

$\Sigma_{crit\_fiducial}$ was taken as 7.94×10^{15} kg/m^2, which is the value of Σ_{crit} found for $z_{source} = 0.75$ and $z_{cluster} = 0.55$.

We find that when using the spectroscopic redshifts to correct for photo-z bias, the mass coefficients increase in 3 of 4 cases (except for the high-z case). The power law slopes mostly increase, except for the mid-z case. Most changes in values are quite small; the largest changes in values are for the high-z case where the mass coefficient drops by 9% and the slope increases by 33%.

4.7.6   Mass-Richness Relations After Considering the Effect of Foreground Contamination

We next considered the effect of contamination by foreground galaxies that were misidentified as source galaxies. To do this, we followed the analysis in Section 4.5.4, multiplying the value of average tangential shear by the correction factors described in that section. The results of the fitting and mass-richness analysis after applying this correction are given in Figure 4.27 and Table 4.14. The effect of the foreground galaxy correction is to make the mass coefficient larger in all cases, although it is almost unchanged (6.47 to 6.49) in the case of all-z. The effect on the slope of the power law varies: for low-z it is slightly larger, for mid-z it is smaller, for high-z it is larger (although with larger error bars) and for all-z it is unchanged.

4.7.7   Mass-Richness Relations After Considering All Systematics

We finally found the mass-richness relations for the Stripe 82 coadd including corrections for all of the above systematics. We excluded the central 0.1 h^{-1} Mpc in the fits, using only source galaxies between $0.1 - 3.0$ h^{-1} Mpc from the BCG. We included the correction for halo miscentering in `MASS_RICHNESS.pro`. We also used the spectroscopic redshift lookup table to find values for Σ_{crit} in `NFW.pro` and we applied the foreground galaxy corrections factors when calculating γ_{ave}. Our final plot of mass vs. richness is shown in Figure 4.28 and the final numerical values for the mass-richness relations are shown in Table 4.15.
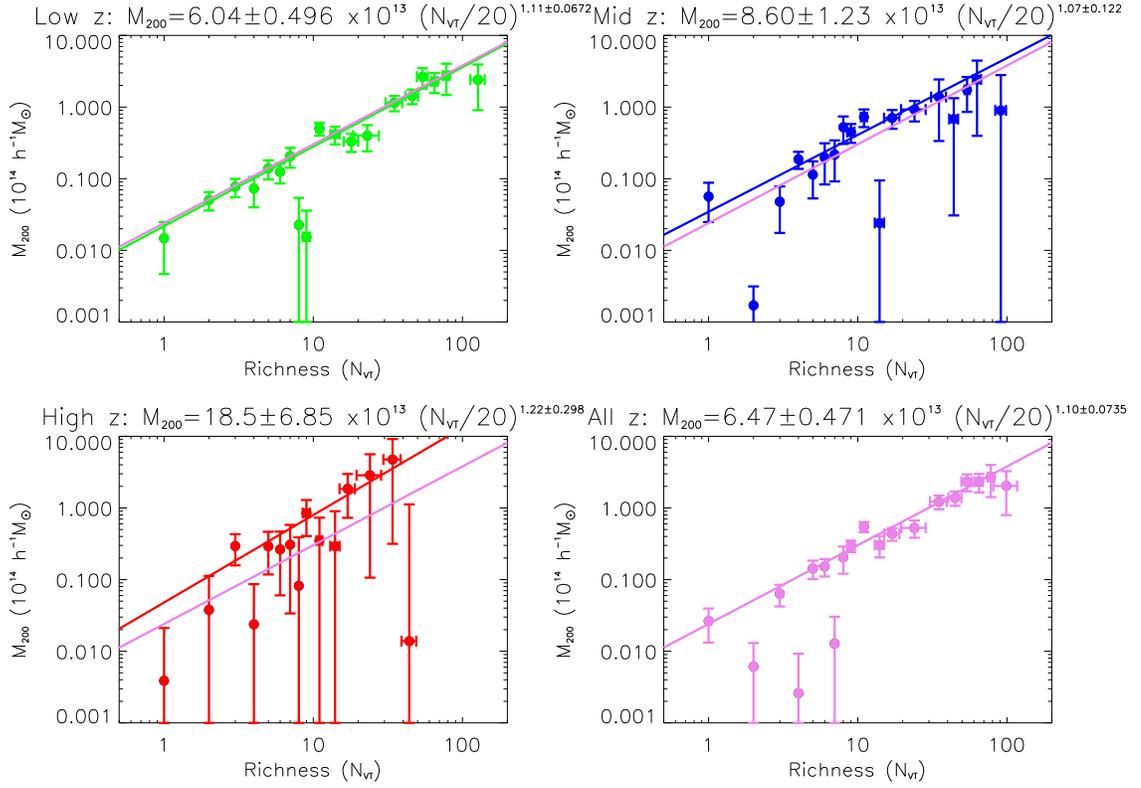
Figure 4.23: The initial mass richness relation (M_{200} vs. N_{VT}) before any corrections were made. The violet line is the fit for all redshift bins.

4.7.8   Comparison of Results

First, we present all mass-richness plots and tables referenced in the preceding sections. We present them all here for easy comparison.

In Figure 4.29 we compare all of the values of the mass coefficient and the power law slope. The situations in these plots are:

1=No corrections

2=BCG correction

3=Halo miscentering correction

Table 4.10: Mass-richness relations for Stripe 82 coadd, not including any corrections. Note that the mass coefficients all have the units $10^{13}\ h^{-1}\ M_{\odot}$.

Redshift Bin	Mass Coefficient	Slope of Power Law	χ^2	Number of dof	Reduced χ^2	Prob
$0.1 \leq z \leq 0.4$	6.04 ± 0.496	1.11 ± 0.0672	20.2	17	1.19	0.264
$0.4 \leq z \leq 0.7$	8.60 ± 1.23	1.07 ± 0.122	32.8	16	2.047	0.00796
$0.7 \leq z \leq 1.0$	18.5 ± 6.85	1.22 ± 0.298	5.15	13	0.396	0.972
$0.1 \leq z \leq 1.0$	6.47 ± 0.471	1.10 ± 0.0735	24.3	17	1.43	0.112
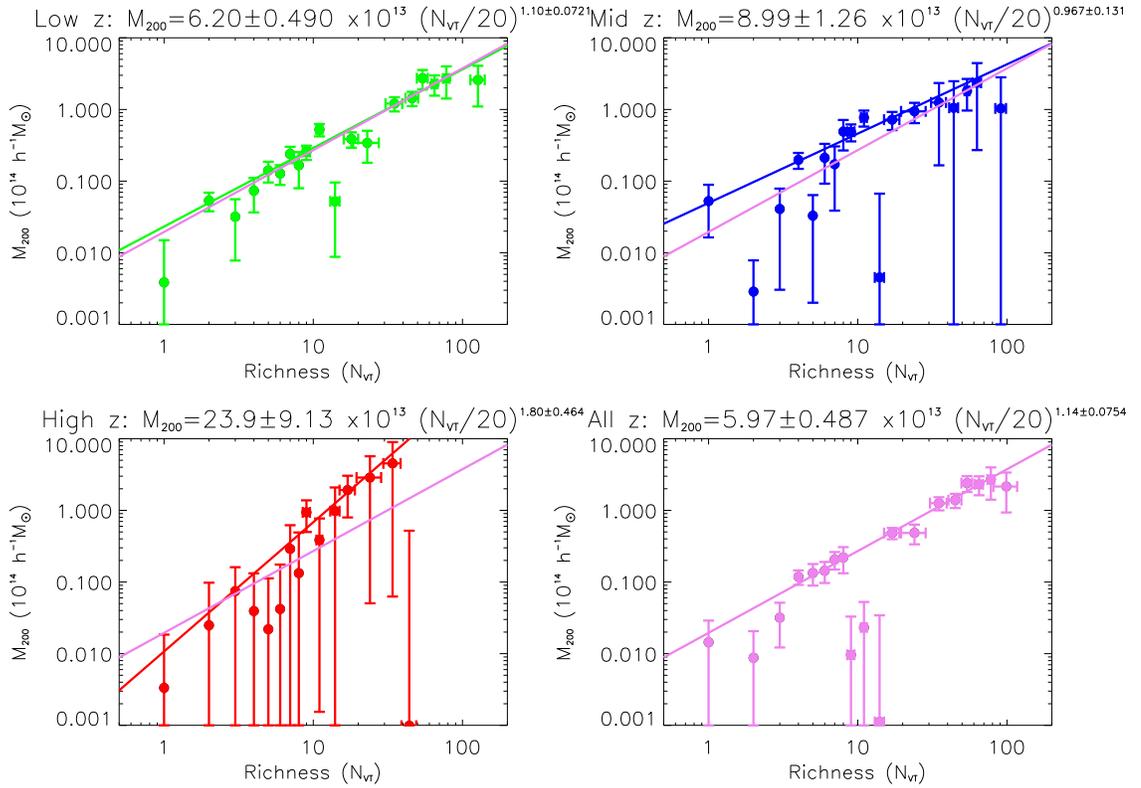
Figure 4.24: The initial mass richness relation (M_{200} vs. N_{VT}) including only a correction for the removal of the central $0.1\ h^{-1}$ Mpc. The violet line is the fit for all redshift bins.

Table 4.11: Mass-richness relations for Stripe 82 coadd, including a correction for the effect of the central BCG. Note that the mass coefficients all have the units $10^{13}\ h^{-1}\ M_\odot$.

Redshift Bin	Mass Coefficient	Slope of Power Law	χ^2	Number of dof	Reduced χ^2	Prob
$0.1 \leq z \leq 0.4$	6.20 ± 0.490	1.10 ± 0.0721	22.6	17	1.33	0.163
$0.4 \leq z \leq 0.7$	8.99 ± 1.26	0.967 ± 0.131	16.3	16	1.020	0.431
$0.7 \leq z \leq 1.0$	23.9 ± 9.13	1.80 ± 0.4648	3.073	13	0.236	0.998
$0.1 \leq z \leq 1.0$	5.97 ± 0.487	1.14 ± 0.0754	13.4	17	0.790	0.707
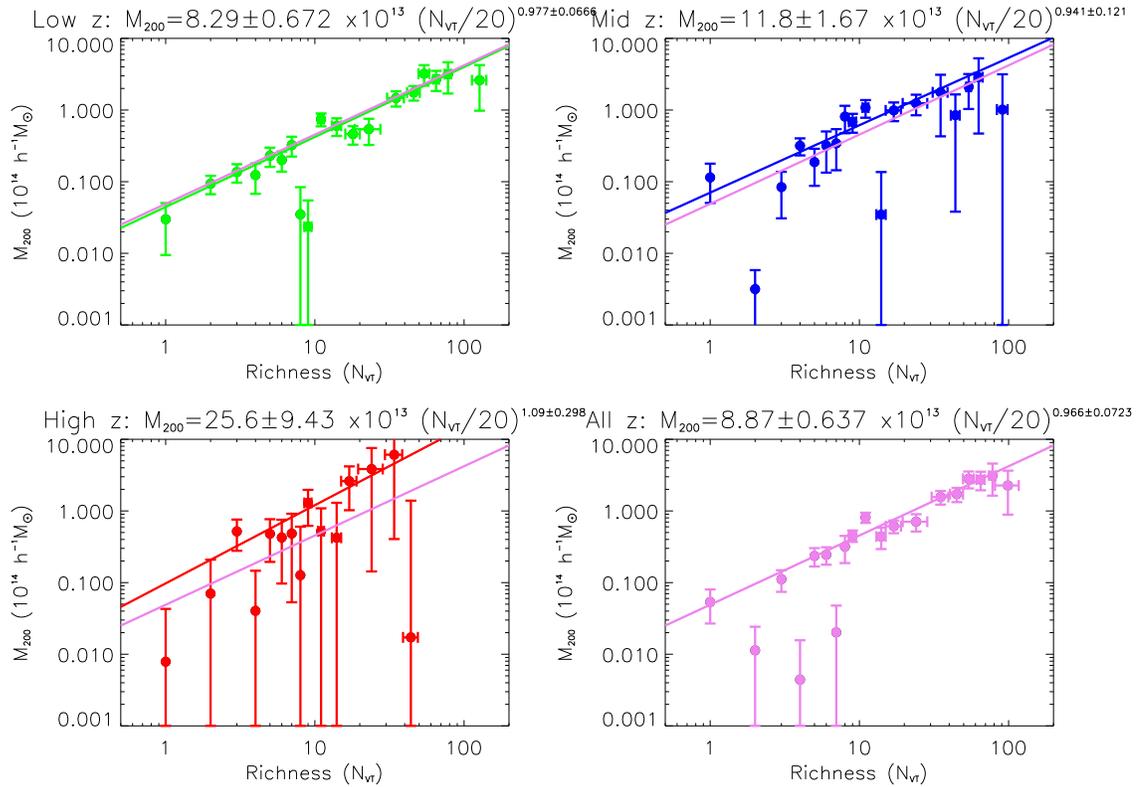
Figure 4.25: The initial mass richness relation ($M_{200} vs. N_{VT}$) including only a correction for the effect of halo miscentering. The violet line is the fit for all redshift bins.

Table 4.12: Mass-richness relations for Stripe 82 coadd, including a correction for halo miscentering. Note that the mass coefficients all have the units 10^{13} h^{-1} $M_\odot$.

Redshift Bin	Mass Coeffi- cient	Slope of Power Law	χ^2	Number of dof	Reduced χ^2	Prob
$0.1 \leq z \leq 0.4$	8.29 ± 0.672	0.977 ± 0.0666	20.5	17	1.20	0.250
$0.4 \leq z \leq 0.7$	11.8 ± 1.67	0.941 ± 0.121	32.8	16	2.049	0.00787
$0.7 \leq z \leq 1.0$	25.6 ± 9.43	1.09 ± 0.298	5.16	13	0.397	0.972
$0.1 \leq z \leq 1.0$	8.87 ± 0.637	0.966 ± 0.0723	24.7	17	1.45	0.102
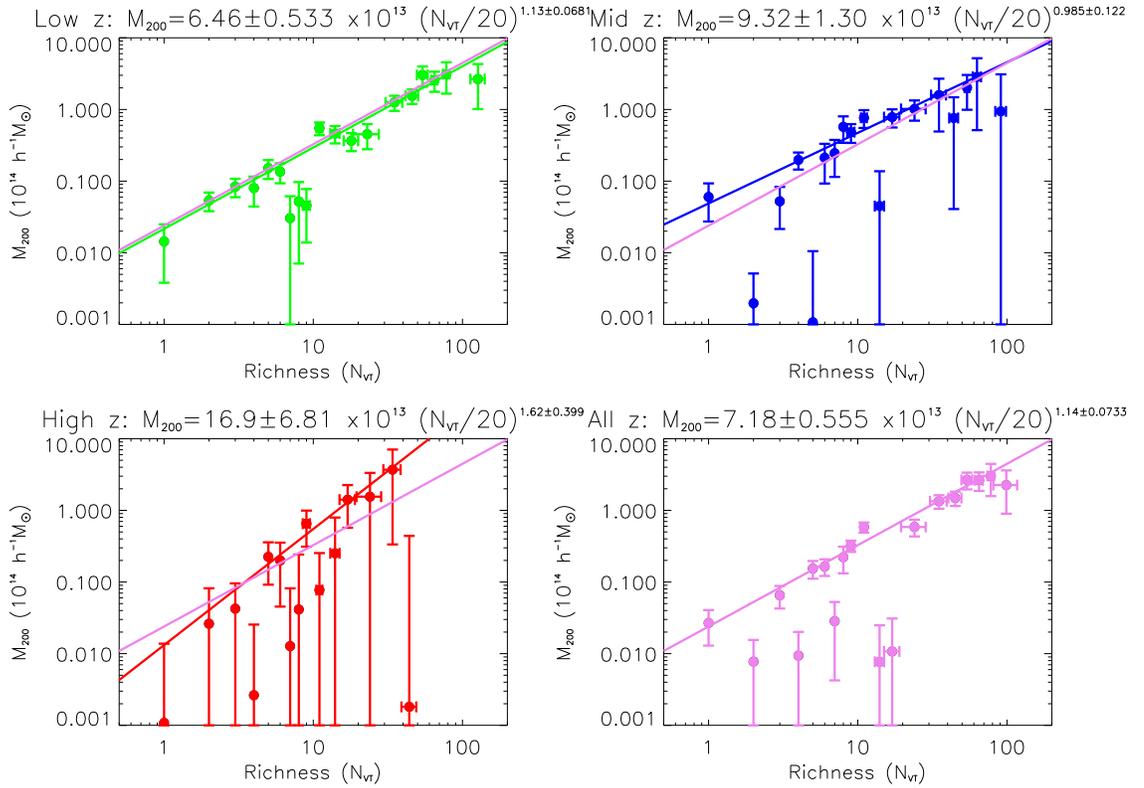
Figure 4.26: The initial mass richness relation (M_{200} vs. N_{VT}) including only a correction for the effect of photo-z bias. The pink line is the fit for all redshift bins.

Table 4.13: Mass-richness relations for Stripe 82 coadd, including a correction for the effect of photo-z bias. Note that the mass coefficients all have the units $10^{13}\ h^{-1}\ M_{\odot}$.

| Redshift Bin	Mass Coefficient	Slope of Power Law	χ^2	Number of dof	Reduced χ^2	Prob
$0.1 \leq z \leq 0.4$	6.46 ± 0.533	1.13 ± 0.0681	26.5	17	1.56	0.0658
$0.4 \leq z \leq 0.7$	9.32 ± 1.30	0.985 ± 0.122	16.6	16	1.036	0.414
$0.7 \leq z \leq 1.0$	16.9 ± 6.81	1.62 ± 0.399	3.068	13	0.236	0.998
$0.1 \leq z \leq 1.0$	7.18 ± 0.555	1.14 ± 0.0733	34.7	17	2.039	0.00689
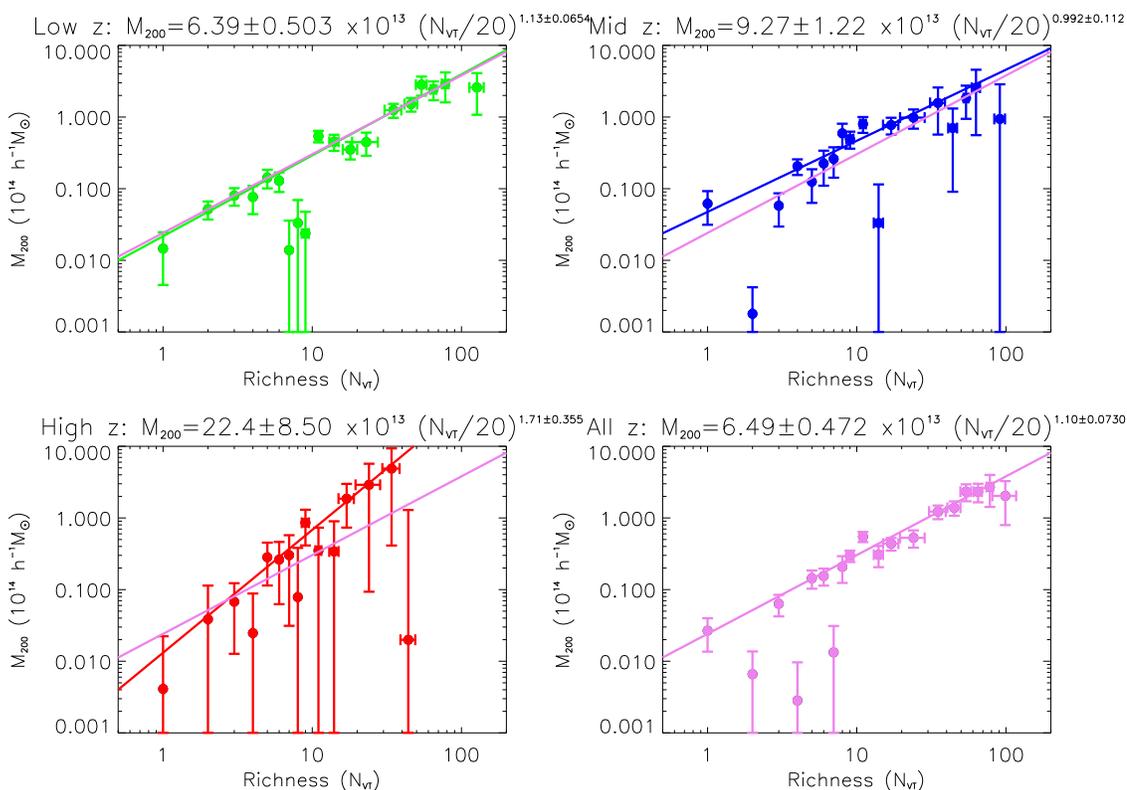
Figure 4.27: The initial mass richness relation (M_{200} vs. N_{VT}) including only a correction for the effect of contamination by foreground galaxies. The violet line is the fit for all redshift bins.

Table 4.14: Mass-richness relation for Stripe 82 coadd, including a correction for the effect of contamination by foreground galaxies. Note that the mass coefficients all have the units $10^{13} \ h^{-1} \ M_{\odot}$.

Redshift Bin	Mass Coefficient	Slope of Power Law	χ^2	Number of dof	Reduced χ^2	Prob
$0.1 \leq z \leq 0.4$	$6.39 \pm 0.0.503$	1.13 ± 0.0654	27.2	17	1.60	0.0548
$0.4 \leq z \leq 0.7$	9.27 ± 1.22	0.992 ± 0.112	23.6	16	1.47	0.0986
$0.7 \leq z \leq 1.0$	22.4 ± 8.50	1.71 ± 0.355	3.00	13	0.231	0.998
$0.1 \leq z \leq 1.0$	6.49 ± 0.472	1.10 ± 0.0730	24.7	17	1.45	0.103
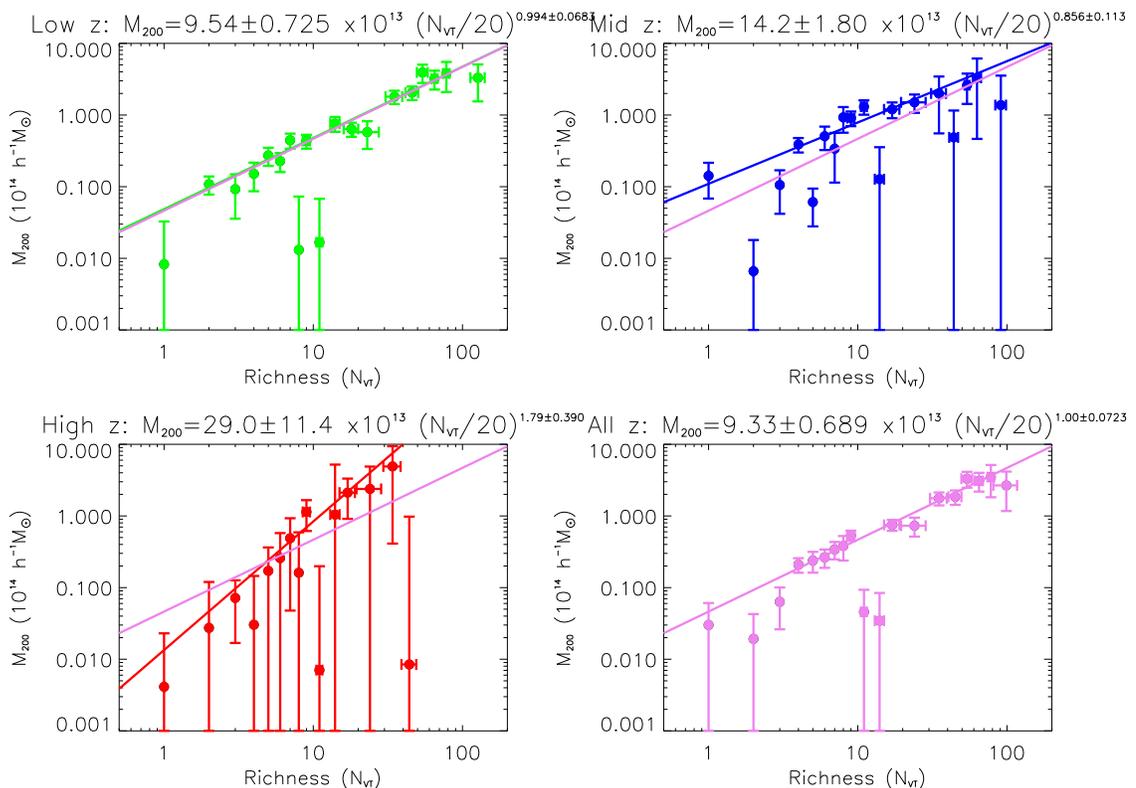
Figure 4.28: The mass richness relation (M_{200} vs. N_{VT}) including corrections for the effect of the central BCG, halo miscentering, photo-z bias and foreground galaxy contamination. The violet line is the fit for all redshift bins.

Table 4.15: Mass-richness relation for Stripe 82 coadd, considering the effect of all systematics. Note that the mass coefficients all have the units $10^{13}\ h^{-1}\ M_\odot$.

Redshift Bin	Mass Coeffi- cient	Slope of Power Law	χ^2	Number of dof	Reduced χ^2	Prob
$0.1 \leq z \leq 0.4$	9.54 ± 0.725	0.994 ± 0.0683	12.4	17	0.732	0.772
$0.4 \leq z \leq 0.7$	14.2 ± 1.80	0.856 ± 0.113	30.0	16	1.88	0.0178
$0.7 \leq z \leq 1.0$	29.0 ± 11.4	1.79 ± 0.390	2.40	13	0.185	0.999
$0.1 \leq z \leq 1.0$	9.33 ± 0.689	1.00 ± 0.0723	19.5	17	1.14	0.303

4=Photo-z correction

5=Foreground galaxy correction

6=all corrections

7=Johnston et al.

8=Simet et al.

Note that the only single correction that seems to have a big effect on the mass coefficient is the correction for halo miscentering. Mass coefficient also changes significantly when the four corrections applied are all applied together (halo miscentering, photo-z, foreground galaxy, central BCG). For the low-z bin, applying these corrections brings the mass coefficient exactly into the range of the two previous results. For the high-z bin, the error bars are always so large that there is no notable change from any of the corrections.

For the power law slope, the corrections have the opposite effect as they did on the mass coefficient. That is, the corrections push the power law slope lower than it was initially. For the low-z bin, the final power law slope of 0.99 ± 0.07 is well
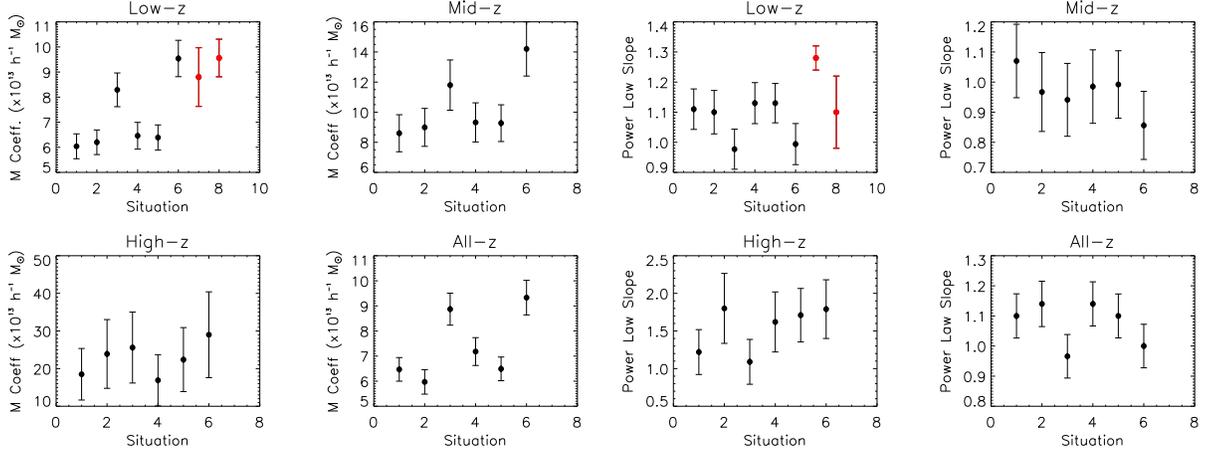
Figure 4.29: Comparison of Stripe 82 mass-richness parameters for all situations. The LEFT FOUR plots compare the mass coefficients and the RIGHT FOUR plots compare the power law slopes. In the low-z plots, we include the data points for Johnston et al. [47] and Simet et al. [77]. The situations are: 1=No corrections, 2=BCG correction, 3=Halo miscentering correction, 4=Photo-z correction, 5=Foreground galaxy correction, 6=all corrections, 7=Johnston et al., 8=Simet et al.

within the range of the Simet et al. measurement ($\alpha = 1.10 \pm 0.12$), but below the measurement of Johnston et al. (1.28 ± 0.04). The final mass coefficient for the low-z bin of 9.54 ± 0.725 is extremely close to that found by Simet et al. (9.56 ± 0.75) and also well within the range found by Johnston et al. (8.8 ± 1.2).

To conclude, our final mass-richness relations are:

$$
\begin{aligned}
Low-z: \quad & M_{200} = (9.54 \pm 0.725) \times 10^{13} h^{-1} M_{\odot} \left(\tfrac{N_{VT}}{20}\right)^{0.994 \pm 0.0683} \\
Mid-z: \quad & M_{200} = (14.2 \pm 1.80) \times 10^{13} h^{-1} M_{\odot} \left(\tfrac{N_{VT}}{20}\right)^{0.856 \pm 0.113} \\
High-z: \quad & M_{200} = (29.0 \pm 11.4) \times 10^{13} h^{-1} M_{\odot} \left(\tfrac{N_{VT}}{20}\right)^{1.79 \pm 0.390} \\
All-z: \quad & M_{200} = (9.33 \pm 0.689) \times 10^{13} h^{-1} M_{\odot} \left(\tfrac{N_{VT}}{20}\right)^{1.00 \pm 0.0723}
\end{aligned}
\tag{4.46}
$$

4.8 Mass-Redshift and Mass-Concentration Relations for Stripe 82

4.8.1 The Mass-Redshift Relation

To examine the mass-redshift relation, we considered the mass coefficient, which is the mass of a galaxy cluster at a richness value of $N_{VT} = 20$ and the power law slope. In Figure 4.30 we plot the mass coefficient and the power law slope vs. redshift for three redshift bins: low-z ($\overline{z} = 0.25$), mid-z ($\overline{z} = 0.55$) and high-z ($\overline{z} = 0.85$). We used FITEXY to find a linear relation between $M_{200}(N_{VT} = 20)$ and z and power law slope and z as:

$$M_{200}(N_{VT} = 20) = (17.2 \pm 1.94)z + (5.21 \pm 6.17)$$
$$slope = (0.0592 \pm 0.142)z + (0.955 \pm 0.374)$$

$$(4.47)$$

Thus it appears that the mass of a halo at $N_{VT} = 20$ does increase with redshift. However the large error bars on the high-z point make the rate of increase with redshift less clear. The power law slope also does appear to increase with redshift, although at redshifts below 0.85 it appears that the power law slope might not have a redshift dependence.

Since the mass coefficient describes the mass of a cluster with $N_{VT} = 20$, this indicates that the mass of a cluster will be higher at higher redshift. This suggests that more dark matter is present in the cluster halos at higher redshift. Since these are likely to be older clusters, they would have had more time to accrete dark matter. Thus it conceptually makes sense that higher redshift clusters would be

more massive. The power law slope describes the rate at which mass increases with richness. The number of galaxies present in a cluster is dependent on the mass of the halo, as more massive halos would be expected to attract more galaxies. The number of galaxies as a function of mass is also connected to the rate of clustering in the universe; when there are more galaxies present in a region of space, there will be more galaxies available to join dark matter halos. Since clustering increases with redshift, we would also expect some dependence of the power law slope on redshift.

It is important to note that there is a difference between real mass evolution with redshift and apparent mass evolution caused by measurement systematics. One particular systematic that might cause the appearance of redshift evolution is the limiting magnitude for cluster finding. The Voronoi tessellation method uses an apparent magnitude cut rather than an absolute magnitude cut. This means that the sample of clusters will be less complete at higher redshift, where galaxies will be biased toward dimmer apparent magnitude simply because of their large distances. Since there are fewer clusters found at higher redshift, there would be poorer statistics at higher redshift, which could bias the ultimate measurement of mass coefficient at higher-z. In order to further test this mass-redshift relation, it would be better to also find this relation using a cluster sample that is limited by absolute magnitude rather than apparent magnitude.

4.8.2   The Mass-Concentration Relation

The mass M_{200} contained in a halo within the radius r_{200} and the concentration c_{200} of the halo are related (recall measurements of this in Chapter 2 for the ten lensing clusters). The relationship between them is called a mass-concentration
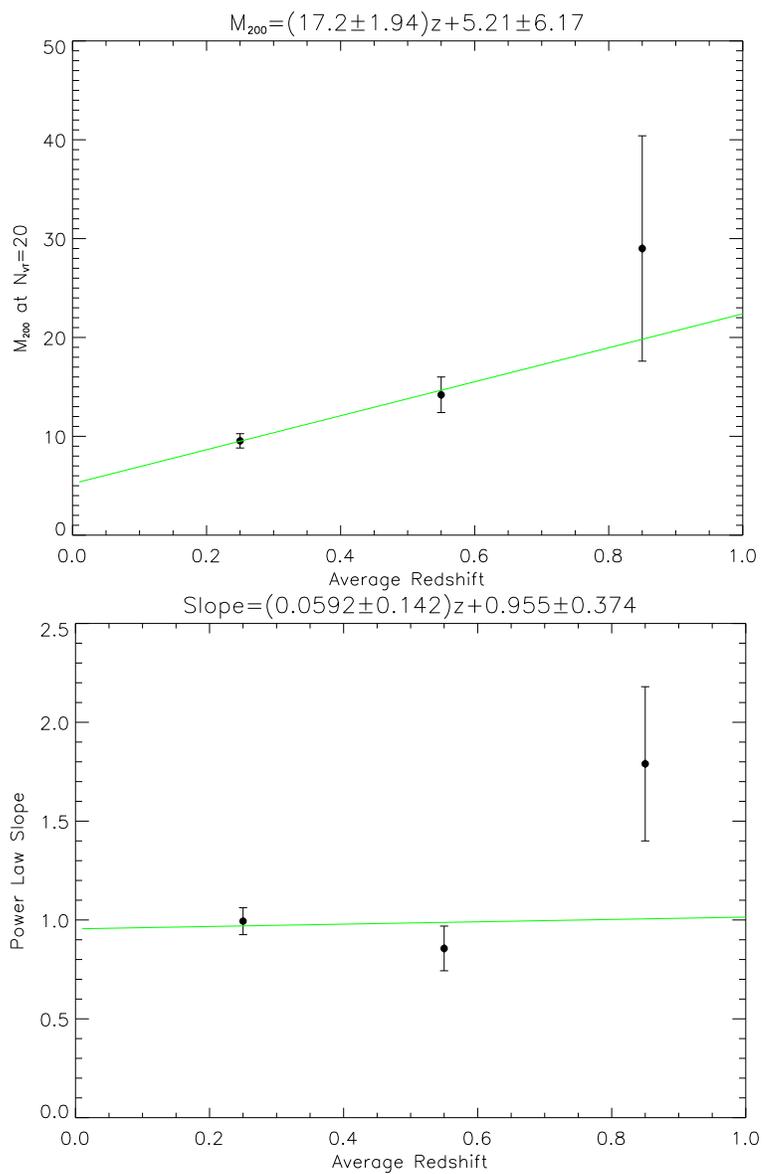
Figure 4.30: The mass at $N_{VT} = 20$ as a function of redshift (upper plot) and the power law slope as a function of redshift (lower plot) for the Stripe 82 coadd.

relation. The mass-concentration relations given in the literature follow the general form [59]:

$$c_{200} = \frac{c_0}{1+z} \left(\frac{M_{200}}{M_0} \right)^{-\beta} \tag{4.48}$$

Mandelbaum et al. (2008) [59] use weak lensing to find a mass-concentration relation of

$$c_{200} = \frac{5.612}{1+z} \left(\frac{M_{200}}{1.0 \times 10^{14} \ h^{-1} \ M_\odot} \right)^{-0.13} \tag{4.49}$$

Duffy et al. (2008) [28] give a mass-concentration relation of

$$c_{vir} = \frac{7.85}{(1+z)^{0.71}} \left(\frac{M_{vir}}{0.0278 \times 10^{14} h^{-1} \ M_\odot} \right)^{-0.081} \tag{4.50}$$

As mentioned in Chapter 2, the process to convert c_{vir} to c_{200} is described at length in the appendix of Johnston et al. We followed this process to find the approximation that

$$M_{200} = 0.93 M_{vir}$$
$$c_{200} = 0.8 c_{vir} \tag{4.51}$$

We then plotted both Equation 4.49 and Equation 4.50 along with the values of c_{200} that we had found from the Stripe 82 coadd. The final values of c_{200} that we found after including all systematics are given in Tables A.1 and A.2 in Appendix A. Our plots of the concentration-mass relation are shown in Figure 4.31. We find that our values for concentration do not closely match the expected relations as given by Mandelbaum and Duffy. Part of the reason is the large error bars on c_{200}, especially in the case of the high-z region. However there does appear to be a systematic offset between our values and the expected values. For the mid-z and high-z bins, the error bars on c_{200} are too large to obtain a sensible concentration-mass relation. For the low-z bin, we obtained a mass-concentration relation using the IDL routine
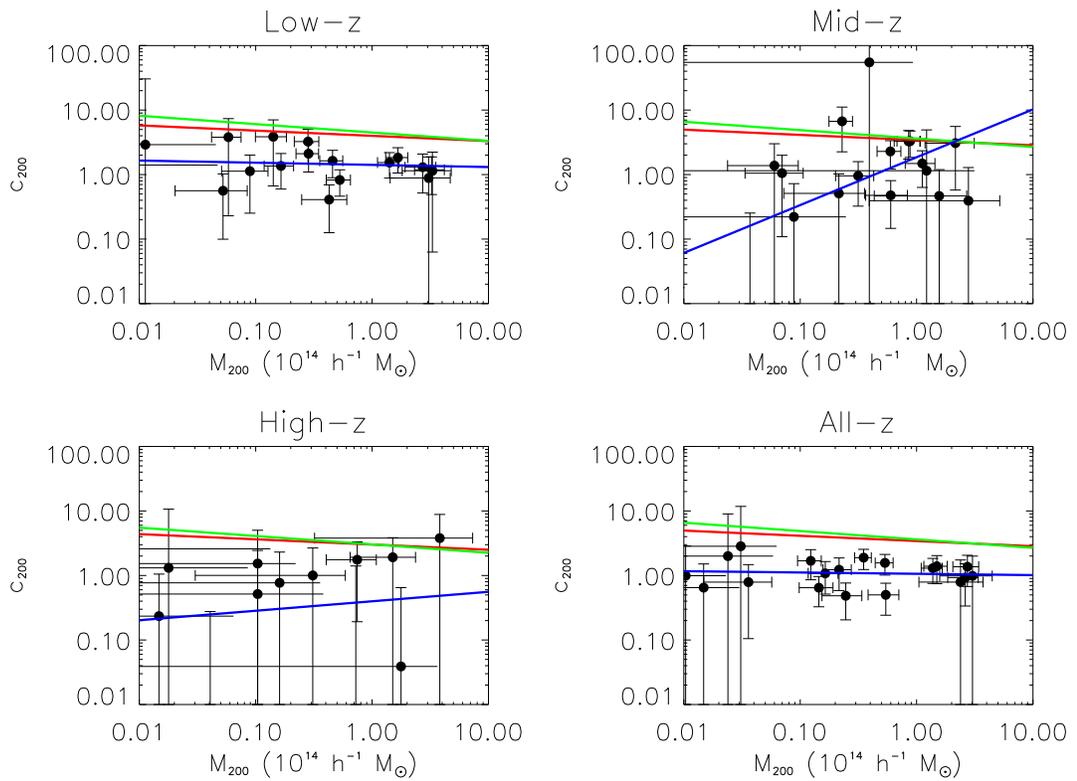
Figure 4.31: A plot of the concentration-mass relation for the Stripe 82 coadd. The red line is the relation from Duffy et al. [28] and the green line is the relation from Mandelbaum et al. [59]. The blue line is our best fit line to the data.

`FITEXY`. For the all-z bin, this routine did not give a sensible result. Thus for the all-z bin we used `LINFIT`, a linear fitting routine that does not include error bars. Our results are

$$
\begin{aligned}
Low-z: \ c_{200} &= \frac{1.76}{1+z} \left(\frac{M_{200}}{1.0 \times 10^{14} \ h^{-1} \ M_\odot} \right)^{-0.033} \\
All-z: \ c_{200} &= \frac{1.64}{1+z} \left(\frac{M_{200}}{1.0 \times 10^{14} \ h^{-1} \ M_\odot} \right)^{-0.020}
\end{aligned}
\tag{4.52}
$$

4.9   Mass-Richness Relations for BCS and DES-SV

4.9.1   BCS Mass-Richness Relations

We ran the `NFW.pro` code on the 763 clusters from Bleem et al. using five bins of richness, using the parameter λ as the richness measure:

I. $10 \leq \lambda \leq 11$

II. $12 \leq \lambda \leq 13$

III. $14 \leq \lambda \leq 15$

IV. $16 \leq \lambda \leq 20$

V. $21 \leq \lambda \leq 65$

We used only one systematics correction, the correction for halo miscentering; we used this one as it had the most significant effect on the Stripe 82 shear profiles. After we found results for M_{200}, we ran them through the `MASS_RICHNESS.pro` code. When we did this we converted λ to N_{VT} using Equation 4.6 so that the final relation would be the same as the others we found. The mass-richness plots are shown in
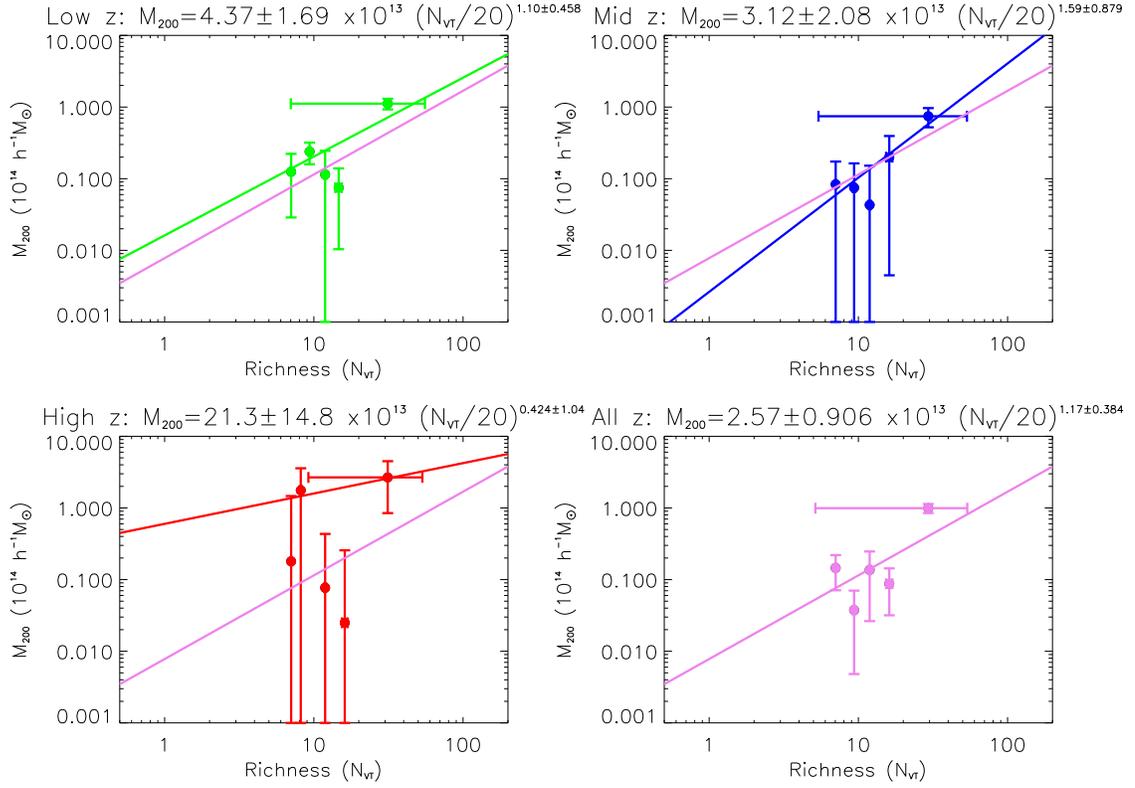
Figure 4.32: The mass richness relation ($M_{200} vs. N_{VT}$) for the BCS data.

Figure 4.32. Note that the error bars are large and the number of points is small; this is because this is a relatively small cluster sample. With a larger sample of clusters in the BCS, we would be able to obtain better statistics. Our final mass-richness relation parameters are shown in Table 4.16.

4.9.2   DES-SV Mass-Richness Relations

We ran the `NFW.pro` code on the preliminary sample of 27,230 VT clusters found in the DES-SV using these bins of N_{VT}:

I. $N_{VT} = 1$

Table 4.16: Mass-richness relation numerical values for the BCS data. Note that the mass coefficients all have the units $10^{13} \ h^{-1} \ M_\odot$.

Redshift Bin	Mass Coefficient	Slope of Power Law	χ^2	Number of dof	Reduced χ^2	Prob
$0.1 \leq z \leq 0.4$	4.37 ± 1.69	1.10 ± 0.458	3.93	3	1.31	0.269
$0.4 \leq z \leq 0.7$	3.12 ± 2.08	1.59 ± 0.879	0.392	3	0.131	0.942
$0.7 \leq z \leq 1.0$	21.3 ± 14.8	0.424 ± 1.04	0.786	3	0.262	0.853
$0.1 \leq z \leq 1.0$	2.57 ± 0.906	1.17 ± 0.384	5.54	3	1.85	0.136

    II. $N_{VT} = 2$

   III. $N_{VT} = 3$

   IV. $4 \leq N_{VT} \leq 6$

    V. $7 \leq N_{VT} \leq 10$

   VI. $11 \leq N_{VT} \leq 15$

  VII. $16 \leq N_{VT} \leq 25$

 VIII. $26 \leq N_{VT} \leq 40$

   IX. $41 \leq N_{VT} \leq 60$

    X. $61 \leq N_{VT} \leq 80$

We used only one systematics correction, the correction for halo miscentering. We then ran the `MASS_RICHNESS.pro` code on the results. Our final mass-richness plots are shown in Figure 4.33 and the numerical results for the mass-richness relations are shown in Table 4.17. Finally in Figure 4.34 we compare the mass-richness
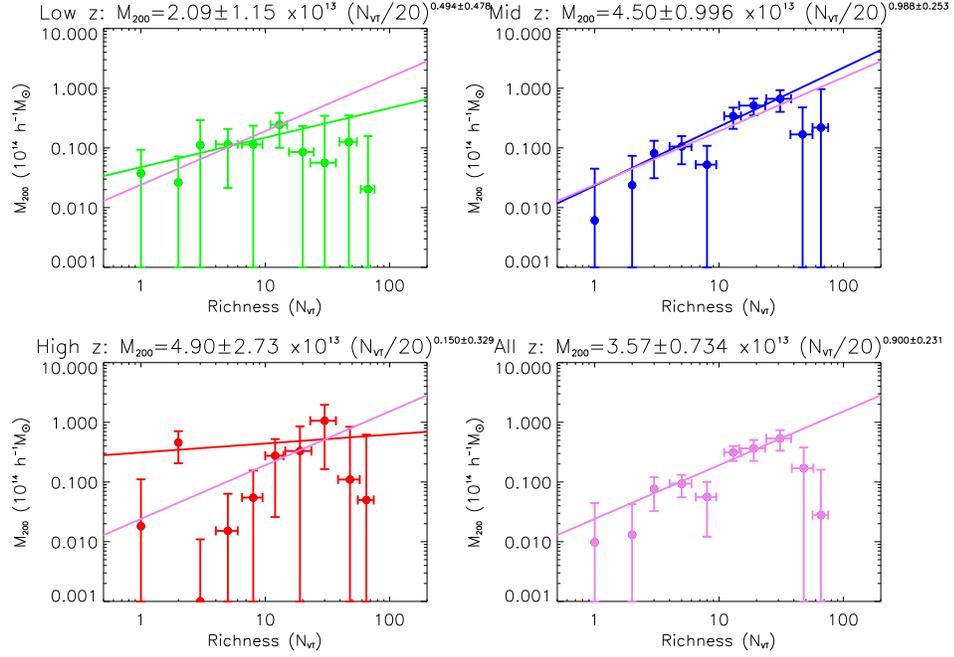
Figure 4.33: The mass richness relation ($M_{200} vs. N_{VT}$) for the DES-SV data. The violet line is the fit for all redshift bins.

Table 4.17: Mass-richness relation numerical values for DES-SV data. Note that the mass coefficients all have the units 10^{13} h^{-1} $M_\odot$.

Redshift Bin	Mass Coefficient	Slope of Power Law	χ^2	Number of dof	Reduced χ^2	Prob
$0.1 \leq z \leq 0.4$	2.09 ± 1.15	0.494 ± 0.478	1.58	8	0.197	0.991
$0.4 \leq z \leq 0.7$	4.50 ± 0.996	0.988 ± 0.253	3.19	8	0.399	0.922
$0.7 \leq z \leq 1.0$	4.90 ± 2.73	0.150 ± 0.329	4.39	8	0.549	0.820
$0.1 \leq z \leq 1.0$	3.57 ± 0.734	0.900 ± 0.231	4.92	8	0.615	0.767

parameters measured in all 3 data sets. The mass coefficients are systematically higher for the Stripe 82 results, while the power law slopes are all within the same ranges for all three data sets.
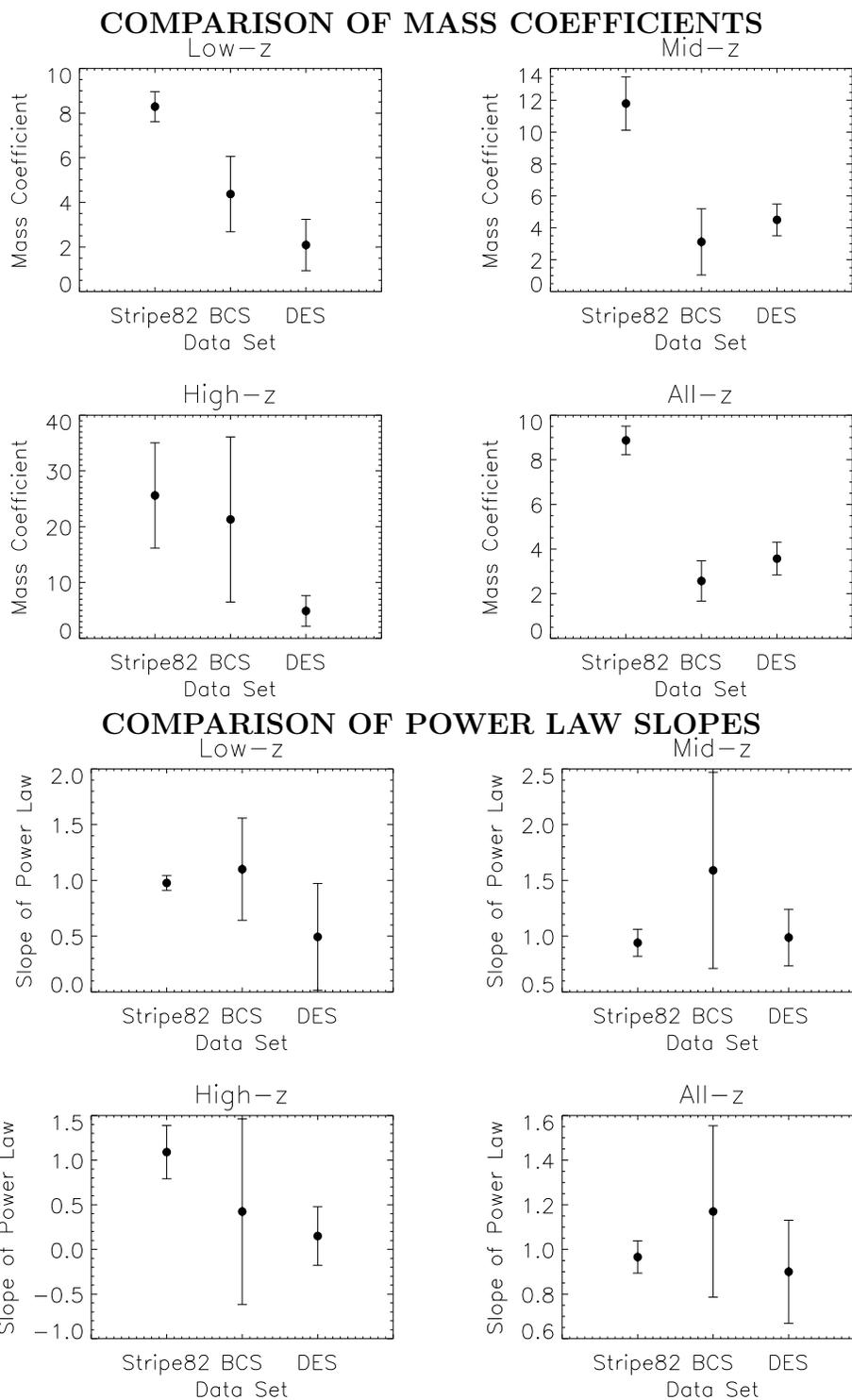
Figure 4.34: Comparison of mass-richness parameters (mass coefficient and slope of power law) for all 3 data sets.

CHAPTER 5

MODELING OF STRONG LENS SDSS J1537+6556

Alice laughed: "There's no use trying," she said; "one can't believe impossible things."

"I daresay you haven't had much practice," said the Queen. "When I was younger, I always did it for half an hour a day. Why, sometimes I've believed as many as six impossible things before breakfast."–Lewis Carroll, from *Alice Through the Looking Glass*

5.1 Introduction

In Chapter 2 we presented estimates of Einstein radii of the ten lensing systems. These estimates were made by fitting circles to the lensing arcs and then measuring the radii of those circles. While this will provide an estimate of the Einstein radius, it is better to model the lensing system in order to obtain better determinations of Einstein radius, lens mass and so on. In this chapter we describe the process of lens modeling and we present initial results for SDSS J1537+6556. This system was chosen as this is one of the four systems found in Chapter 2 to show evidence of overconcentration and it is the only one for which a mass model did not yet exist. In addition we describe use of a matrix-based Bayesian lens modeling routine that has not previously been utilized by the Sloan Bright Arcs Survey group.

In § 5.2 we describe the process of modeling galaxies nearby the arc using `GALFIT` and removing them. In § 5.3 we describe using `LENSVIEW` to fit models to SDSS J1537+6556. In § 5.4 we describe using a new modeling routine developed by H. Lin called `Tri3SrcMatrixFit`. We also present results for SDSS J1537+6556 using `Tri3SrcMatrixFit`. In § 5.5 we compare results from `LENSVIEW` to results from `Tri3SrcMatrixFit`. Finally in § 5.6 we describe an interesting system found during observing for the Dark Energy Survey that initially was thought to be a strong lens candidate.

5.2   Galaxy Modeling

5.2.1   The Process of Galaxy Modeling

Lens modeling involves fitting the lensing arcs in an image to a mass model which would cause the light to be deflected in that way. Before this can be done, the cluster galaxies present in the image need to be subtracted. In order to subtract only the light due to the galaxies, the light profiles observed in each galaxy in the images is fit to a theoretical relation between light intensity and distance that would be expected for such a galaxy.

There are several mathematical relations which are commonly used to predict brightness as a function of radius for elliptical galaxies, including especially the $\acute{S}ersic$ profile and the de Vaucouleurs profile. The $\acute{S}ersic$ profile has the form [21]

$$\log\left(\frac{I(r)}{I_e}\right) = -k\left[\left(\frac{r}{R_e}\right)^{1/n} - 1\right] \tag{5.1}$$
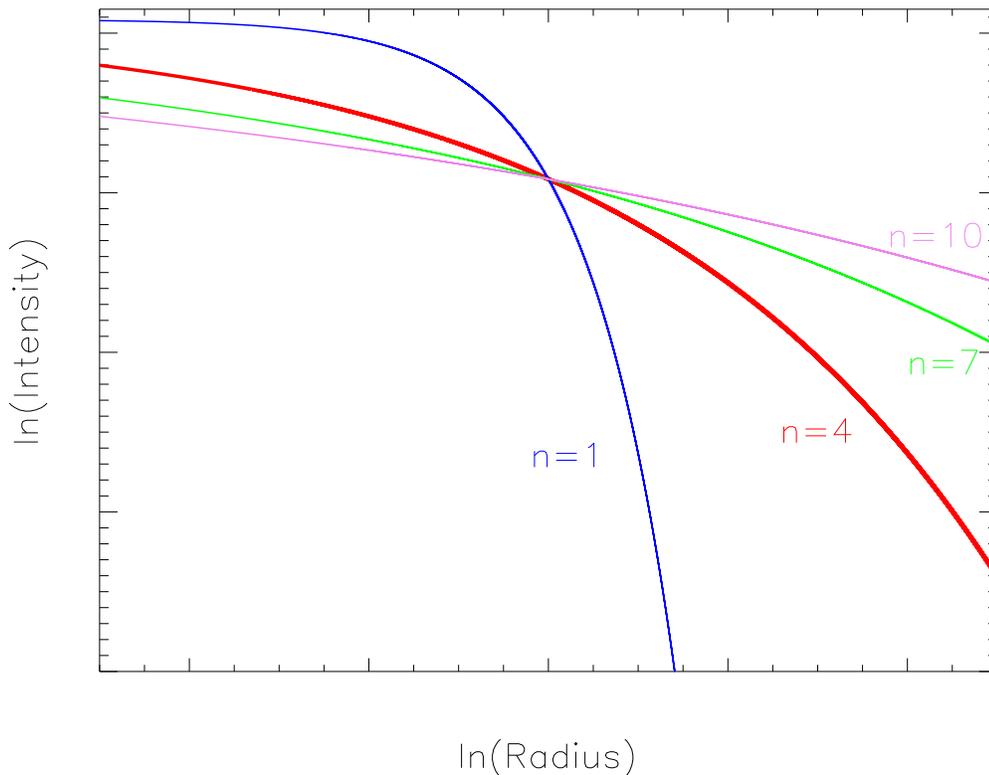
Figure 5.1: A plot of intensity predicted by the Sersic profile as a function of radius. On both x and y-axes we plot the natural logarithm of these quantities.

Here I is the light intensity as a function of radius r, k is a constant, R_e is the radius containing half of the total luminosity, I_e is the light intensity at that distance, and n is the Śersic index. This last parameter controls the curvature of the Śersic profile: as n becomes larger, the profile becomes more centrally concentrated. As can be seen in Figure 5.1, the light falls off with distance much faster for small n than for large n. In the special case of $n = 4$ (the red line in Figure 5.1), we find that

$$I(r) \propto e^{-kr^{1/4}} \tag{5.2}$$

This is called the de Vaucouleurs profile, which is known to be a good fit to many elliptical galaxies. For $n = 1$

$$I(r) \propto e^{-kr} \tag{5.3}$$

This is called the exponential profile. Not all astronomical objects have light profiles well fit by a variation on the Śersic profile; stars are better fit by a profile that is close to a Gaussian, the Moffat profile. The Moffat profile has the form:

$$I(r) = I_0[1 + \left(\frac{r}{\theta}\right)^2]^{-\beta} \tag{5.4}$$

Again $I(r)$ is the intensity as a function of distance from the center of the object, I_0 is the intensity at the center of the object, θ is the full width half max (FWHM) of the profile and β is the atmospheric scattering coefficient [42].

5.2.2   Fitting galaxy models in `GALFIT`

To fit models to the galaxies nearby the BCG, we used the program `GALFIT` [67]. `GALFIT` uses parametric fitting to describe the light distribution of galaxies in the image using a small number of independent parameters. `GALFIT` is run from the command line by being fed a list of parameters. We selected only a small region of the original WIYN image to run in `GALFIT`; we wanted to select a region that contained the entire lensing arc but no larger. The region we selected is shown by the green box in Figure 5.2. In order to run `GALFIT`, several input files are needed along with the image section chosen for galaxy fitting. We also need to produce (1) a sigma image, (2) a mask file and (3) a PSF file. A **sigma image** describes the standard deviation of values in the image section being used. It is made by
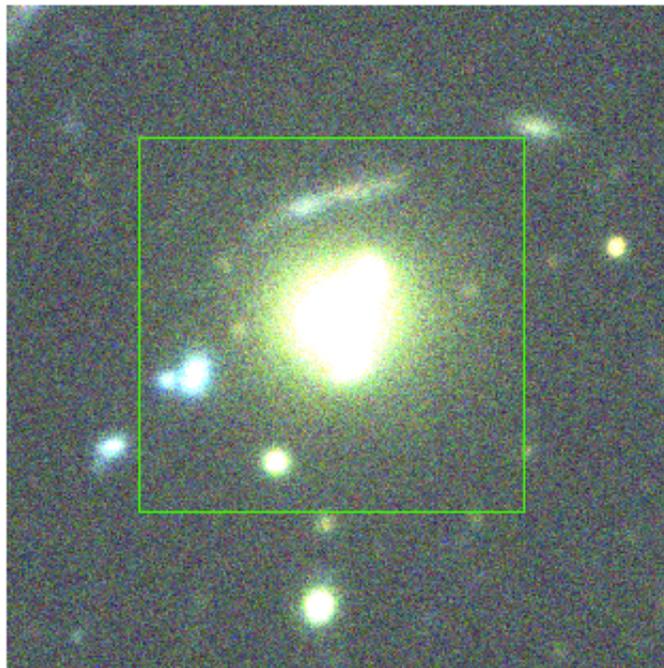
Figure 5.2: The region of the image containing SDSS J1537+6556 that was fit in GALFIT.

taking the original WIYN image and dividing by the GAIN of the image multiplied by the number of exposures (2 in most cases for the WIYN images). To this we add the RDNOISE (readout noise) parameter divided by GAIN times number of exposures, all squared. Finally we take the square root of this image. This can all be done in `IRAF`. Next we made a **mask image**, in which the lensing arcs are masked out. We mask the arcs out because we do not want to include their light profiles with the galaxy profiles. Masks were made by using polygonal regions in DS9, filling them with a Fortran routine called `FILLPOLY.f` (provided by Peng [67]) and using the routine $2MASK$ and then copying this to a .fits file in IRAF. Finally a **point spread function (PSF) image** was made in order to describe how the light was distorted by the atmosphere and the telescope. To do this, a bright but unsaturated star was chosen nearby the lensing system (the star is located at (234.24061, +65.911439)). Then we fit this star with a Moffat fit in `GALFIT`. The resultant model image was used as the PSF image. Note that each of these images must be made for each filter the images were taken in, as `GALFIT` is run only on single-filter images (i.e, g, r or i-band images). Examples of each of these images are shown in Figure 5.3.

To run `GALFIT`, it was necessary to produce a "feedme" file. The list of input and output files and the values of parameters are all put together in a file that traditionally has the suffix ".feedme". Then is invoked from the command line with the name of this input file being declared. A sample "feedme" file is shown in Figure 5.4. Note that at the top we control the input files, the output files, the image region to fit and several other run parameters. Below this is a list of object numbers. For each object in the fitting region we need to declare what type of brightness profile to fit to that object, the coordinates of the object, initial guesses at magnitude, half-light radius, Śersic index, axis ratio and position angle. `GALFIT`
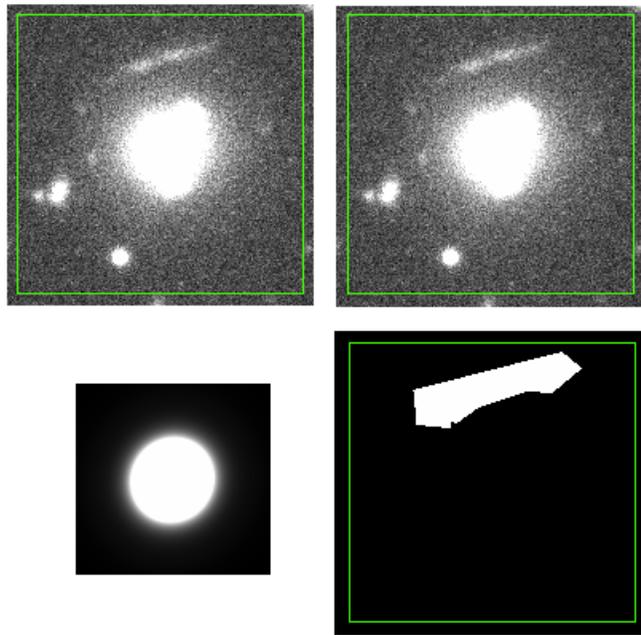
Figure 5.3: Images input to `GALFIT` for SDSS J1537+6556. Going clockwise from top left they are: (1) Original image, (2) sigma image, (3) Mask image and (4) PSF image. The PSF image is 123 pixels square while the other images are 179 x 175 pixels. These are all i-band images.

```
# IMAGE and GALFIT CONTROL PARAMETERS
A) stack_5185_i_2012.fits     # Input data image (FITS file)
B) 5185_gal_i.fits        # Output data image block
C) SIGMA_5185_i.fits  # Sigma image name (made from data if blank or "none").
D) PSF_5185_i.fits   # Input PSF image and (optional) diffusion kernel
E) 1                    # PSF fine sampling factor relative to data.
F) MASK_5185.fits               # Bad pixel mask (FITS image or ASCII coord list)
G) none              # File with parameter constraints (ASCII file).
H) 1280 1494 1998 2159  # Image region to fit (xmin xmax ymin ymax)
I) 200 200              # Size of the convolution box (x y)
J) 25                   # Magnitude photometric zeropoint.
K) 0.14  0.14           # Plate scale (dx dy)   [arcsec per pixel]
O) regular              # Display type (regular, curses, both)
P) 0                 # Create ouput only? (1=yes; 0=optimize).
S)                    # Modify/create objects interactively?

# INITIAL FITTING PARAMETERS
#
#   For object type, the allowed functions are:.
#       nuker, sersic, expdisk, devauc, king, psf, gaussian, moffat,.
#       ferrer, powsersic, sky, and isophote..
#.
#   Hidden parameters will only appear when they're specified:
#       C0 (diskyness/boxyness),.
#       Fn (n=integer, Azimuthal Fourier Modes),
#       R0-R10 (PA rotation, for creating spiral structures).
#.
# -------------------------------------------------------------------
#   par)   par value(s)   fit toggle(s)   # parameter description.
# -------------------------------------------------------------------

# Object number: 0
 0) sersic                # object type
 1)  1377 2086  1  1    # position x, y
 3) 15.9      1          # Integrated magnitude.
 4) 27   1        # R_e (half-light radius)   [pix]
 5)   3.74  1  # Sersic index n (de Vaucouleurs n=4).
 6) 0.0000     0        #     -----.
 7) 0.0000     0        #     -----.
 8) 0.0000     0        #     -----.
 9) 0.88         1          # axis ratio (b/a)..
10) -84.83            1        # position angle (PA) [deg: Up=0, Left=90]
 Z) 0                    # output option (0 = resid., 1 = Don't subtract).

# Object number: 1
 0) sersic                # object type
 1)   1393 2107    1  1   # position x, y
 3) 18.15     1          # Integrated magnitude.
 4) 8.4   1        # R_e (half-light radius)   [pix]
 5) 6.53 1   # Sersic index n (de Vaucouleurs n=4).
 6) 0.0000      0        #     -----
```

Figure 5.4: A section of a sample "feedme" input file for GALFIT.
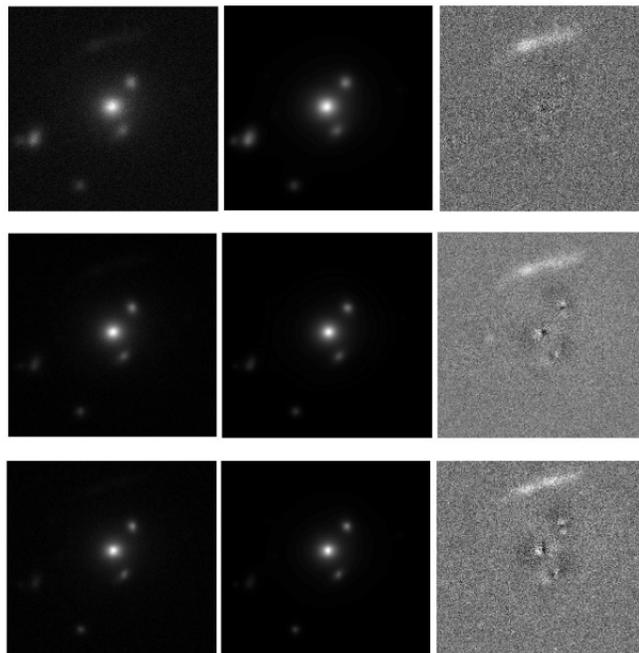
Figure 5.5: `GALFIT` models for SDSS J1537+6556. From top to bottom these are g-band, r-band and i-band images. From left to right they are galaxy images, galaxy models and residual images.

takes these input parameters and tries different values while trying to minimize the value of χ^2. Once we find a set of models that minimize χ^2 and provide reasonable models for all objects in the image, then we take that as our final result. The final results are shown in Figure 5.5. Note that the `GALFIT` produces an image block, which is a FITS image containing 4 different frames. The first frame is empty, and the remaining three contain the image section being fit, the galaxy models produced by the fit and the residual image, which is the image minus the models.

The final parameters found for each of the galaxies in the fitting region for SDSS J1537+6556 are listed in Tables B.1, B.2 and B.3 in Appendix **??**. The final values for reduced χ^2 were 1.152 for g-band, 1.151 for r-band and 1.029 for i-band. Finally we took the fitting regions of the original images in each band and

subtracted the `GALFIT` model images. This produced images in each band containing only the lensing arcs, along with a minimal amount of residual noise not subtracted by `GALFIT`. These images were then used for lens modeling.

5.3 Lens Modeling in *LENSVIEW*

5.3.1 About `LENSVIEW`

`LENSVIEW` contains an implementation of an algorithm called *LENSMEM*. This algorithm proceeds in two cycles, an inner loop and an outer loop. The inner loop finds the best source model that fits the data and the outer loop adjusts the model parameters for the lens and restarts the inner loop [88]. The goal of the inner loop is to minimize the function J where

$$J = C - \lambda S \tag{5.5}$$

In this equation,

$$C = \sum_{ij} (D_{ij} - M_{ij})^2 / \sigma_{ij}^2 \tag{5.6}$$

and

$$S = -\sum_{kl} f_{kl}[ln(f_{kl}/A) - 1] \tag{5.7}$$

In Equation 5.6 C is the χ^2 merit function describing the difference between the data D_{ij} and the model M_{ij}, divided by the variance σ_{ij}^2. In Equation 5.7 S is the entropy in the source model, f_{kl} is the set of source pixels and A is the sky

background. Finally λ in Equation 5.5 is a parameter which can be adjusted by the outer loop of the algorithm in order to change the lens model parameters.

`LENSVIEW` must be installed and then it is run from the command line with a list of parameters like this:

lensview -logfilepath . -tracelevel 2 -pixelres 0.14 -datafile 5185_DATA_r.fits -sourcefile empty30by30.fits -psffile PSF_5185_LV_r.fits -nice 2 -pixelratio 2 -maxiter 100 -paramfile comps_r.txt -srcxoffset 0 -mask MASK_allones.fits -srcyoffset 0 -srcdefaultval 1 -noisefile 5185_VAR_r.fits -targetchisqu 2692 -useminfinder $\$$*

All parameters that can be turned on or off in `LENSVIEW` are described in Wayth (2012) [89]. Note especially that there are several images that are referred to in this command, including (1) a data image, (2) a source image, (3) a PSF image, (4) a mask image and (5) a variance image. In addition there is a parameter file which we called *comps_r.txt*. The data image is the final image produced after using `GALFIT`, containing only the lensing arcs. The source image is a blank image that defines the size of the source plane. The PSF image describes the PSF, the same as used for `GALFIT`. The mask image is used to denote what regions of the image you want to fit, as only those regions with pixel value of 1 will be included in the calculation of χ^2. The variance image is calculated similarly to the Sigma image used in `GALFIT`, but it is only the size of the `LENSVIEW` data image and we do not take the square root of the image at the end of its creation. The parameter file contains the definition of the lens model you wish to fit as well as initial values and step sizes. Options for lens models include the pseudo-isothermal potential, the singular isothermal ellipsoid, the point mass and the NFW profile.

Table 5.1: Mass model parameters found for SDSS J1537+6556 using `LENSVIEW`.

Filter	x-offset	y-offset	Einstein Radius	Axis Ratio	Position Angle	χ^2 (Lens Model)	Entropy
g	0	0	10	0.9	180	87937	-259353
r	0	0	8.36	0.86	145	37339	-33158
i	0	0	8.36	0.84	50.4	33273	-17701

5.3.2 Lens Modeling Results from `LENSVIEW` for SDSS J1537+6556

We used a singular isothermal ellipsoid fit for lens modeling of SDSS J1537+6556. The fit parameters we found are listed in Table 5.1. The lens and source model images we obtained are shown in Figure 5.6. These images are for g (top), r (middle) and i-band (bottom) images. From left to right the images are the input image, the model image, the residual image and the source image. The residual image is the input data image minus the model. Note that the fit was not good for g-band as there is much lower signal to noise in g-band images.

Figure 5.6: Lens model results from LENSVIEW for SDSS J1537+6556. The images are for g (top), r (middle) and i-band (bottom) images. From left to right the images are the input image, the model image, the residual image and the source image.

5.4   Lens Modeling with a Bayesian Lens Fitting Routine: `Tri3SrcMatrixFit`

5.4.1   About `Tri3SrcMatrixFit`

To begin, let us define Bayes' Theorem. Bayes' Theorem originated with the work of Thomas Bayes, an English statistician, philosopher and Presbyterian minister who lived from 1701-1761. Bayes' Theorem can be stated as [29]

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} \tag{5.8}$$

P(A) is called the **prior** and describes the initial probability of event *A*. *P(B)* is the probability of event *B*. $P(B|A)$ is called the **likelihood** and it describes the probability of *B* given event *A*. Finally $P(A|B)$ is called the **posterior probability**; it describes the probability of event *A* given event *B*. In other words, we can get a more precise measure of the probability of event A occuring if we also have some knowledge about an event B that has some effect on event A. This theorem has applications across all fields of knowledge. But in terms of gravitational lensing, we apply Bayes' Theorem by applying it to the model fitting. In all model fitting routines, the goal is to test the goodness of fit by minimizing a function that measures the difference between the fit and the data. This is usually done with a χ^2. In the Bayesian method, a χ^2 is still used to measure goodness of fit, but there is another step in which efforts are made to make the fit as simple as possible. This is done with Occam's razor in mind [81], the idea that the simplest solution is usually the

most realistic. This simplicity is the prior, the extra information we have which we can use to constrain the model fitting.

Suyu et al. (2006) [81] describe a way to apply Bayes' Theorem to gravitational lens modeling. They describe lens modeling in two steps: (1) Model fitting and (2) Model comparison. Model fitting is done by obtaining the best source model parameters by minimizing the function:

$$M(s) = E_D(s) + \lambda E_s(s) \tag{5.9}$$

The quantity s represents the source model parameters. Here $E_D(s)$ is essentially the χ^2 function, $E_s(s)$ is the regularization function and λ is the regularization constant. Model comparison is done by maximizing the simplicity of the model parameters by maximizing the Bayesian evidence (described further below). To do this, it is necessary to find the value of λ that produces the maximum of the evidence and then use that value of λ when applying the regularization. Regularization penalizes any function that is not sufficiently physical, whether by trying to make the source as small as possible (zeroth order regularization), minimzing gradients in the model (first order) or minimizing curvature (second order).

5.4.1.1 Model Fitting

Suyu et al. use $\mathbf{d}$ (the bold indicates a matrix quantity) to represent the vector of data points, s to represent model parameters, λ to denote the regularization

constant, **g** to mean the output of the regularization function and **f** to be the response function, a matrix that relates the data to the model parameters where

$$\mathbf{d} = \mathbf{f}s + \mathbf{n} \tag{5.10}$$

and **n** is the matrix representing noise in the data. Then the posterior probability is

$$P(s|\mathbf{d}, \lambda, \mathbf{f}, \mathbf{g}) = \frac{P(s|\mathbf{g}, \lambda)P(\mathbf{d}|s, \mathbf{f})}{P(\mathbf{d}|\lambda, \mathbf{f}, \mathbf{g})} \tag{5.11}$$

But also

$$P(s|\mathbf{d}, \lambda, \mathbf{f}, \mathbf{g}) = \frac{exp[-M(s)]}{Z_M(\lambda)} \tag{5.12}$$

with

$$Z_M(\lambda) = \int d^{N_s}s \; exp[-M_s(s)] \tag{5.13}$$

Recall that $M_s(s)$ is the function we seek to minimize in order to find the best model parameters. The prior, $P(s|\mathbf{g}, \lambda)$ applies Occam's razor, penalizing any solutions that are not smooth (this is the $P(A)$ term in Equation 5.8). The probability $P(\mathbf{d}|s, \mathbf{f})$ is the likelihood (the $P(B|A)$ term). The quantity $P(\mathbf{d}|\lambda, \mathbf{f}, \mathbf{g})$ is called the **evidence** for the model (the $P(B)$ term).

For the prior, Suyu et al. give the expression

$$P(s|\mathbf{g}, \lambda) = \frac{exp[-\lambda E_S(s|\mathbf{g})]}{Z_s(\lambda)} \tag{5.14}$$

with $E_s(s|\mathbf{g})$ being the regularization (discussed further below) and $Z_s(\lambda)$ being the normalization of the prior probability distribution [81], given as

$$Z_s(\lambda) = \int d^{N_s}s \; exp(-\lambda E_s) \tag{5.15}$$

N_s is the number of model parameters. The likelihood implements the χ^2 test. It is given as

$$P(\mathbf{d}|s, \mathbf{f}) = \frac{exp[-E_D(\mathbf{d}|s, \mathbf{f})]}{Z_D} \tag{5.16}$$

where

$$E_D(\mathbf{d}|s, \mathbf{f}) = \frac{1}{2}(\mathbf{f}_s - \mathbf{d})^T \mathbf{C}_D^{-1}(\mathbf{f}_s - \mathbf{d}) = \frac{1}{2}\chi^2 \tag{5.17}$$

where $\mathbf{C}_D$ is the covariance matrix for the data and Z_D is the normalization for the probability, given by

$$Z_D = (2\pi)^{N_d/2}(\det \mathbf{C}_D)^{1/2} \tag{5.18}$$

Finally the evidence is given in terms of a logarithm in order to simplify some of the terms. It is

$$\log P(\mathbf{d}|\lambda, \mathbf{f}, \mathbf{g}) = -\lambda E_s(s_{MP}) - E_D(s_{MP}) - \tfrac{1}{2}\log(\det \mathbf{A}) + \tfrac{N_s}{2}\log\lambda + \lambda E_s(\mathbf{0}) + \tfrac{1}{2}\log(\det \mathbf{C})$$
$$- \tfrac{N_d}{2}\log(2\pi) + \tfrac{1}{2}\log\left(\det \mathbf{C}_D^{-1}\right) \tag{5.19}$$

The quantity s_{MP} indicates the model parameters of the most probable solution. As previously mentioned, N_s is the number of source model parameters, N_d is the total number of data points and $\mathbf{C} = \nabla\nabla E_S(s)$. E_D is the regularizing function for the data and $E_D = \frac{1}{2}\chi^2$. Finally

$$\mathbf{A} = \nabla\nabla\mathbf{M}(\mathbf{s}) \tag{5.20}$$

5.4.1.2 Model Comparison

The regularizing function E_s penalizes a solution that is overly complicated. The regularizing function can be of several different forms, including *zeroth order*, *gradi-*

ent and *curvature* regularization [81]. Zeroth order regularization tries to minimize intensity at every source pixel, while gradient regularization tries to minimize the difference in intensity between adjacent pixels. Finally curvature regularization tries to minimize the second derivatives in the intensity in x and y-directions. All regularizations try to minimize rates of change of intensity, again trying to keep the source light distributions as simple as possible. Zeroth order regularization can be written as [81]

$$E_S(s) = \frac{1}{2} \sum_{i=1}^{N_s} s_i^2 \qquad (5.21)$$

where the s_i are the source pixels. Gradient and curvature regularization are described by Equations A2 and A3 in Suyu et al.

In order to accomplish model comparison, we want to maximize the posterior for the regularization constant (λ) given the data, the response function for the model parameters and the choice of regularization. We need to find a value of λ that will allow us to maximize this posterior. Suyu et al. give the posterior as

$$P(\lambda|\mathbf{d}, \mathbf{f}, \mathbf{g}) = \frac{P(\mathbf{d}|\lambda, \mathbf{f}, \mathbf{g})P(\lambda)}{P(\mathbf{d}|\mathbf{f}, \mathbf{g})} \qquad (5.22)$$

Suyu et al. assume a flat prior in $\log \lambda$, meaning any value is possible so $P(lambda)$ cannot be constrained. The quantity to consider is then is $P(\mathbf{d}|\lambda, \mathbf{f}, \mathbf{g})$ which was the evidence in Equation 5.19.

In Figure 5.7 we see Figure 3 from Suyu et al. [81]. This plot gives the expected relation between λ and $\log P(\mathbf{d}|\lambda, \mathbf{f}, \mathbf{g})$ (they work with the log(P) in order to simplify some of the terms). The solid line in this plot represents the result of Equation 5.19 as a function of λ. Note that $\log(P)$ reaches a maximum value for a particular value of λ. Thus what we are looking for when using `Tri3SrcMatrixFit` is the value of

Figure 5.7: A plot of regularization constant (λ) vs. log of Bayesian evidence (Log(P)) from Suyu et al. (2006) [81] (solid line). The dotted and dashed lines are plots of Eq. 20 from Suyu et al. The evidence, rescaled to fit graph, is shown as the dot-dashed line.

λ that maximizes the evidence. Then from there we seek to find source and lens model parameters by minimizing the χ^2 function.

5.4.1.3   Using `Tri3SrcMatrixFit`

`Tri3SrcMatrixFit` is run using parameter files in a manner similar to `GALFIT` and `LENSVIEW`. The source code is written in Fortran. The program is invoked by simply stating *tri3srcmatrixfit* on the command line once the Makefile has been run and the code is compiled. `Tri3SrcMatrixFit` operates (like `GALFIT`) using

two major loops: an inner loop called *tri3subroutinematrixfit.f* and an outer loop called *tri3srcmatrixfit.f*. The inner loop searches for the best source plane model parameters given a value or range of values of λ. The outer loop takes these source model parameters and searches for the best lens plane model parameters, finding model parameters for a singular isothermal ellipsoid (SIE) profile. Both inner and outer loops have their own parameter files by which starting values and ranges can be controlled. Figure 5.8 shows the parameter files for the outer (upper image) and inner (lower image) loops. The upper (lens plane) file lists the system name (note that 5185 is an alternative informal name for SDSS J1537+6556 assigned during lens searches using the interacting galaxies method) and then the image names. Note that the input images are the same types as used by LENSVIEW, including a data image, a source image, a variance image, a mask image and a PSF image. Parameters that are then read in include the x and y-centers of the image (where the BCG was before subtraction), the Einstein radius in pixels, the axis ratio and the position angle given counterclockwise from the x-axis. A core radius and truncation radius can be optionally be given to use a pseudo-Jaffe rather than an SIE profile. The lower (source plane) file takes input on the starting and ending values for λ as well as a λ step size and a parameter determining which type of regularization to use. The second line in this file control the center and the size of the source plane image.

Initially we kept the lens plane input parameters constant and varied λ to find which values of λ maximize the evidence. The resultant plots are shown in Figure 5.9. For zeroth order regularization we found the evidence is maximized at $\lambda = 0.00012$ while for first order regularization we found the maximum evidence at $\lambda = 0.00008$; plots are shown in Figure 5.9 (in both cases we used the r-band data image).

```
1

5185
5185_DATA_r.fits
model_src_test.fits
5185_VAR_r.fits
mask_new.fits
PSF_5185_LV_r.fits
7  0.1  1
91. 1 1.
89. 1 1.
59.    1 1
0.87  1 0.1
10.    1 10.
0. 0 0.
0. 0 0.
0

WHAT-STUFF-IS
number of parameters; f_tolerance parameter (Start bigger, like
0.1...negative means don't fit with amoeba, accept input parameters);
1 is do fit, 0 is project it
x_Center; 0 fix parameter, 1 let float; step value  (where BCG was)
y_center; 0 fix parameter, 1 let float
Einstein radius; 0 fix parameter, 1 let float  (in pixels)
q (axis ratio, b/a);.
Position angle  ;CCW from x axis?
r_core
r_trunc
```

```
0.00012d0 0.00012d0 0.1d0  0  0.1d0
96. 91. 0.5 0.5 1.e-10    16. 16. 16. 16.  8    0 0 0
64. 64. 4. 4.  30 30 -0.1

;THIS FILE CORRESPONDS TO unit 11 in fortran file

first row is lambda start, lambda end, lambda step size (Delta
lambda), ireg (tells which regularization to use), lambdapot (not
implemented)
2nd row controls size of output image
Row 2:  x-coord of center of source plane
.       y-coord of center of source plane
.       size of source plane relative to lens plane (x)  0.5 means 1/2
        size
.       size of source plane relative to lens plane (y)
.       Distance from center (+) in x (Within this region, fit,
        outside ignored...these are in source plane)
.       Distance from center (-) in x
.       Distance from center (+) in y
.       Distance from center (-) in y.
.       Distances could be smaller or bigger than FITS image,
        model_src_test.fits but difference will be blank),
.       2 means every other pixel gets mapped back exactly.

Ignore line 3, it is not implemented in code


This is the read command for this file.

read(11,*) xs0, ys0, dxs, dys, small,
     $              dxs1, dxs2, dys1, dys2, nmodtri,
     $              ncuthd, radhd, nmodhd
```

Figure 5.8: Parameter files for `Tri3SrcMatrixFit`. The upper file is for the outer loop (lens plane fitting). The lower file is for the inner loop (source plane fitting).

Figure 5.9: The relation between log(evidence) and λ for SDSS J1537+6556. The plot on the left is for zeroth order regularization while the plot on the right is for first order regularization.

Table 5.2: Mass model results for SDSS J1537+6556 found from `Tri3SrcMatrixFit`.

Filter	x-center	y-center	Einstein Radius	Axis Ratio	Position Angle	χ^2 (Lens Model)	χ^2 (Source Model)	log (Evidence)
g	91	89	8.4	0.87	10	2369514	63051	-45443
r	91	90	8.3	0.87	10	35999	14323	-19662
i	91	90	8.3	0.87	10	32313	12650	-18786

We then fixed λ at these values as we then allowed lens plane parameters to vary, seeking to find the best lens and source plane model fits.

5.4.2   Lens Modeling Results from `Tri3SrcMatrixFit` for SDSS J1537+6556

After running `Tri3SrcMatrixFit` on SDSS J1537+6556 we obtained the lens model parameters listed in Table 5.2. Lens and source models for SDSS J1537+6556 are shown in Figure 5.10. In this figure, the three rows are for g (top), r (middle)

and i-band (lower) images. From left to right they are: (1) input image, (2) model image, (3) residual image, (4) source image, (5) second source image and (6) finer source image. The input image is the image with only the lensing arc that was output by `GALFIT`. The model image is the model of the arc that was produced by `Tri3SrcMatrixFit`, considering the χ^2 only in the highlighted region. The residual is the input image minus the model image. The first source image is that produced directly by the `Tri3SrcMatrixFit` program. The second and third source images were produced by a subsidiary program also written by H. Lin that takes position and magnitude information output by `Tri3SrcMatrixFit` and uses the IDL routine `Triangulate` to construct a Delaunay triangulation of points above background in the source plane. It then uses the IDL routine `Trigrid` to display the points. The second source image uses a coarse grid and the last one uses a finer grid to display the points.

Figure 5.10: Lens model results from `Tri3SrcMatrixFit` for SDSS J1537+6556. In this figure, the three rows are for g (top), r (middle) and i-band (bottom) images. From left to right they are: (1) input image, (2) model image, (3) residual image, (4) source image, (5) second source image and (6) finer source image.

The subsidiary program also produces a plot of points above background. In this plot, the red points are points with brightness greater than 0 and the green points are points with brightness greater than the defined background (we used 100). The red points, being produced by a model of a singular isothermal ellipsoid, trace out a caustic. The green points represent the source model. It appears from Figure 5.11 that the source object falls almost along one boundary of the caustic.

Figure 5.11: Points above background in the source plane for SDSS J1537+6556. The plots are for g (left), r (middle) and i-band (right). The red points are points with brightness greater than 0 and the green points are those with brightness greater than 100.

5.5 Comparison of Results from `LENSVIEW` and
`Tri3SrcMatrixFit`

Both modeling programs produce very similar source models, appearing as a long filament. In both cases, the g-band fit has the largest errors, because of lower signal to noise. The χ^2 values for the lens model fits from `LENSVIEW` are 87937, 37339 and 33273 while those from `Tri3SrcMatrixFit` are 63051, 14323 and 12650. Thus the fits are significantly better for `Tri3SrcMatrixFit` than for `LENSVIEW`. The results for Einstein radius are similar for both programs, with $\theta_E = 8.4''$ for `LENSVIEW` and $\theta_E = 8.3''$ for `Tri3SrcMatrixFit`. We found $\theta_E = 8.5 \pm 2.8''$ in Chapter 2. Both programs found axis ratio around 0.86, but there was significant disagreement about position angle.

So our final conclusion about the source is that it is a filamentary object falling along a caustic. Recalling Figure 1.11 in Section 1.7.2 it we see that a source galaxy falling along the astroid caustic would produce a single large arc, along with several much smaller counter images. It is not surprising that the counter images are not visible, as they would likely be significantly dimmer than the main arc. There also are no visible secondary arclets visible in the model images produced by `Tri3SrcMatrixFit`.

5.6 Butterfly Collecting: The Vizcacha Galaxy

During observing for the Dark Energy Survey on the night of November 14, 2013, I noticed the presence of an object in a z-band image from Chip S-16 that had the potential to be a gravitational lens. This object has the coordinates (347.028508,

Figure 5.12: A closeup image of the Vizcacha Galaxy from the raw z-band image taken at the 4-m Blanco telescope on November 11, 2013.

-54.0222056); I give it the informal designation of "Vizcacha Galaxy" due to the sighting of one of these chinchilla-like rodents before sunset on this night. A closeup image of the z-band data is shown in Figure 5.12 while a wider-field image showing surrounding galaxies in shown in Figure 5.13. This object stood out because the center of the object appeared like a large elliptical galaxy and it was surrounded by a faint, diffuse circle that could have potentially been an Einstein ring. I notified H. Lin and T. Diehl of this finding, and we planned to further investigate this object when data was available in multiple wavelength bands to see what colors the objects in the image had.

After further investigation it was found that this object was in DES-SV data (from 2012-2013 observing season) and so the multi-band observations did exist. The images with color information and stacked observations gave evidence that this was not a gravitational lensing system, but was rather a spiral galaxy with a ring-component. This was evident because the ring surrounding the galactic core was flocculent and not well defined as lensing arcs typically are. Although this was determined not to be a lensing system, it is nontheless interesting as a well-defined

Figure 5.13: A larger image of the Vizcacha Galaxy from the raw z-band image taken at the 4-m Blanco telescope on November 11, 2013.

ring galaxy. A ring galaxy is believed to be formed when a disk galaxy experiences a direct impact be another system [12]. The disk galaxy is gravitationally distorted to form a ring of star-forming material while the impactor forms the new ring galaxy's nucleus. In Figure 5.14 we display a wide section of this image, while in Figure 5.15 we display a large image of the galaxy.

Figure 5.14: A wide-field color image of the Vizcacha Galaxy.

Figure 5.15: A color image of the Vizcacha Galaxy.

CHAPTER 6

CONCLUSION

If you would attain to what you are not yet, you must always be displeased by what you are. For where you are pleased with yourself there you have remained. Keep adding, keep walking, keep advancing.–St. Augustine of Hippo

6.1 Conclusions: Evidence of Overconcentration Among Low-Mass Strong Lensing Clusters

In Chapter 2 we describe 10 strong lensing galaxy clusters of redshift $0.26 \leq z \leq 0.56$ that were found in the Sloan Digital Sky Survey. We present measurements of richness (N_{200}), mass (M_{200}), and velocity dispersion for the clusters.

We find that in order to use the massrichness relation from Johnston et al. [47], which was established at mean redshift of 0.25, it is necessary to scale measured richness values up by 1.47. Using this scaling, we find richness values for these clusters to be in the range of $22 \leq N_{200} \leq 317$ and mass values to be in the range of $1 - 30 \times 10^{14} h^1 \ M_\odot$. We also present measurements of Einstein radius, mass, and velocity dispersion for the lensing systems. The Einstein radii (θ_E) are all relatively small, with $5.4 \leq \theta_E 13$.

Finally, we consider if there is evidence that our clusters are more concentrated than CDM would predict. We find that six of our clusters do not show evidence of overconcentration, while four of our clusters do. We note a correlation between

overconcentration and mass, as the four clusters showing evidence of overconcentration are all lower-mass clusters. For the four lowest mass clusters the average value of the concentration parameter c_{200} is 11.6, while for the six higher-mass clusters the average value of c_{200} is 4.4. ΛCDM would place c_{200} between 3.4 and 5.7.

6.2 Conclusions: The Accuracy of the Richness Extrapolation

In Chapter 3 we describe follow-up observations at Apache Point Observatory (APO) of six of the ten strong lensing galaxy clusters described in Chapter 2. Follow-up observations were taken in 2012 as some sections of the galaxy clusters being studied were not present in the original images taken at the WIYN telescope in 2009. In order to complete the analysis described in Chapter 2, it was necessary to extrapolate the total cluster richness. Our goal in Chapter 3 was to confirm that this extrapolation relation was reliable.

We describe the process of data reduction and magnitude calibration using IRAF, LACosmic, SCAMP, SWARP and SExtractor. We then measure the total number of cluster galaxies present in the APO images and compare this to the number of cluster galaxies found in the previous WIYN images. We finally combine APO and WIYN results to give total measured richnesses for these six clusters. We compare the total measured richnesses to the total extrapolated richnesses and find there is a good match. Thus we conclude that the extrapolation relation previously used is highly reliable. Finally we present new imaging of another galaxy cluster serendipitously noticed in the APO imaging.

6.3   Conclusions: A Higher Redshift Mass-Richness Relation

In Chapter 4 we present mass-richness relations for three data sets: the Sloan Digital Sky Survey Stripe 82 Coadd (Stripe 82), the Blanco Cosmology Survey (BCS) and the Dark Energy Survey Science Verification data (DES-SV). These mass-richness relations are presented for four redshift bins, $0.1 \leq z \leq 0.4$, $0.4 \leq z \leq 0.7$, $0.7 \leq z \leq 1.0$ and $0.1 \leq z \leq 1.0$.

We present samples of galaxy clusters for each of the 3 data sets and describe how these clusters were found using a Voronoi tessellation cluster finder for Stripe 82 and DES and using the red sequence method for BCS. Richness for Stripe 82 and DES-SV was given in terms of N_{VT} and for BCS was given in terms of richness measure λ. Our sample includes $19,376$ clusters in Stripe between $0.1 \leq z \leq 0.98$ and $1 \leq N_{VT} \leq 127$. It also includes 763 clusters in the BCS between $0.13 \leq z \leq 0.75$ and $10 \leq \lambda \leq 64.3$ and $27,230$ clusters in the DES-SV between $0.12 \leq z \leq 1.68$ and $1 \leq N_{VT} \leq 538$. We present measurements of richness for clusters in each of these data sets. We then describe how we measured cluster masses with measurements of weak lensing shear.

We describe the effects of systematics on weak lensing shear and thus cluster mass (M_{200}) results. These systematics include the effect of the central BCG, halo mis-centering, foreground galaxy contamination and photometric redshift uncertainty. We present mass-richness relations with each of these effects considered separately as well as considered altogether. We also conducted two null tests and one final systematics test; these included measurements of orthotangential shear for all clus-

ter bins, measurements of weak lensing shear around random points and analysis of dependence of shear on sky position.

We present values for the mass coefficient and the power law slope for mass-richness relations found using Stripe 82 data for each of the four redshift bins. Our final results for mass richness relations for the Stripe 82 coadd including corrections for all systematic effects are:

$$
\begin{aligned}
Low-z: \quad & M_{200} = (9.54 \pm 0.725) \times 10^{13} h^{-1} M_{\odot} \left(\tfrac{N_{VT}}{20}\right)^{0.994 \pm 0.0683} \\
Mid-z: \quad & M_{200} = (14.2 \pm 1.80) \times 10^{13} h^{-1} M_{\odot} \left(\tfrac{N_{VT}}{20}\right)^{0.856 \pm 0.113} \\
High-z: \quad & M_{200} = (29.0 \pm 11.4) \times 10^{13} h^{-1} M_{\odot} \left(\tfrac{N_{VT}}{20}\right)^{1.79 \pm 0.390} \\
All-z: \quad & M_{200} = (9.33 \pm 0.689) \times 10^{13} h^{-1} M_{\odot} \left(\tfrac{N_{VT}}{20}\right)^{1.00 \pm 0.0723}
\end{aligned}
\tag{6.1}
$$

Our result for the mass coefficient is very consistent with the results of other groups at low redshift. Our result for the power law slope at low redshift is consistent with one of the other groups but outside the range of the other.

We also considered the mass-redshift and mass-concentration relations using the Stripe 82 data. We find that the mass coefficient shows evidence of increasing with redshift, while the power law slope may not. Our results for the concentration parameter have large error bars which contributes to a systematic offset between our final mass concentration relations and those from other groups.

Finally we present mass-richness relations for all four redshift bins for clusters found in the BCS and for clusters found in the DES-SV. We applied only the halo miscentering correction to these data sets.

6.4 Conclusions: A Mass Model of SDSS J1537+6556

We present initial lens models for SDSS J1537+6556 of the mass distribution in the cluster and of the lensed source galaxy. We describe how the image was prepared for lens modeling by using `GALFIT`. We then describe lens model results from two modeling routines: First, from `LENSVIEW` and then from `Tri3SrcMatrixFit`, a new Bayesian lens-modeling routine.

Both `GALFIT` and `Tri3SrcMatrixFit` give similar results, with $\theta_E \approx 8.3''$ and axis ratio about 0.86. Position angle estimates cover a wide range. The final source image describes a long filamentary source object that is very close to the astroid caustic. Its position along the caustic in the source plane leads to the formation of a single distinct arc in the image plane, along with other small counter images that were not detected in our images.

Finally we briefly describe a system found during observing for the Dark Energy Survey that was initially thought to be a strong lens. Later it was found to be a ring galaxy. We present initial images taken at the Blanco Telescope along with color images of the system produced after observing.

REFERENCES

[1] S. Allam, et al.. Merging Galaxies in the Sloan Digital Sky Survey Early Data Release. *The Astronomical Journal*, 127:1883-1899, 2004.

[2] S. Allam, et al.. The 8 O'Clock Arc: A Serendipitous Discovery of a Strongly Lensed Lyman Break Galaxy in the SDSS DR4 Imaging Data. *The Astrophysical Journal*, 662, L51, 2007.

[3] S. Allen, A. Evrard, and A. Mantz.. Cosmological Parameters from Observations of Galaxy Clusters. *Annual Review of Astronomy and Astrophysics*, 49, 1.

[4] J. Annis et al.. The SDSS Coadd: 275 Deg^2 of Deep SDSS Imaging on Stripe 82. *The Astrophysical Journal*, 2011.

[5] J. Annis and J. Kubo. private communication, 2010.

[6] R. J. Barlow. *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences.* New York: John Wiley and Sons, 1999.

[7] M. B. Bayliss, et al.. GeminiGMOS Spectroscopy of 26 Strong Lensing Selected Galaxy Cluster Cores. *The Astrophysical Journal Supplement*, 193, 8, 2011.

[8] M. Becker, T. McKay, B. Koester, et al.. The Mean and Scatter of the Velocity-Dispersion-Optical Richness Relation for MaxBCG Galaxy Clusters. *The Astrophysical Journal*, 669, 905, 2007.

[9] A. Berlind, et al.. Percolation Galaxy Groups and Clusters in the SDSS Redshift Survey Identification, Catalogs, and the Multiplicity Function. *The Astrophysical Journal*, 167, 1, 2006.

[10] E. Bertin et al.. Astromatic programs. Retrieved from http://www.astromatic.net/.

[11] E. Bertin and S. Arnouts. SExtractor: Software for source extraction. *Astronomy and Astrophysics*, 317, 393, 1996.

[12] J. Binney and S. Tremaine. *Galactic Dynamics*. Princeton, N.J.: Princeton University Press, 1987.

[13] L. E. Bleem et al.. The Blanco Cosmology Survey: An Optically-Selected Galaxy Cluster Catalog and a Public Release of Optical Data Products. Submitted to *The Astrophysical Journal Supplement*, March 27, 2014.

[14] H. Bohringer. Lecture Notes in Observational Cosmology. LMU, Spring 2010.

[15] T. Broadhurst, et al.. The Surprisingly Steep Mass Profile of A1689, from a Lensing Analysis of Subaru Images. *The Astrophysical Journal*, 619, L143, 2005.

[16] T. Broadhurst and R. Barkana. Large Einstein Radii: A Problem for ΛCDM. *Monthly Notices of the Royal Astronomical Society*, 390, 1647, 2008.

[17] E. Buckley-Geer, et al.. The Serendipitous Observation of a Gravitationally Lensed Galaxy at z=0.9052 from the Blanco Cosmology Survey: The Elliot Arc. *The Astrophysical Journal*, 742, 48, 2011.

[18] S. Carroll. *Spacetime and Geometry*. San Francisco: Addison Wesley, 2004.

[19] B. Carroll & D. Ostlie. *An Introduction to Modern Astrophysics*. Boston: Pearson Addison Wesley, 2007.

[20] D. Clowe et al.. A Direct Empirical Proof of the Existence of Dark Matter. *The Astrophysical Journal*, 648, L109-L113, 2006 September 10.

[21] F. Combes, P. Boisse, A. Mazure & A. Blanchard. *Galaxies and Cosmology*. Berlin: Springer, 2002.

[22] C. Cunha. Nearest-Neighbor Codes for Photometric Estimates of Redshift Statistics. Retrieved from http://www.stanford.edu/ ccunha/nearest/.

[23] M. Davis et al. and the DEEP2 Collaboration. Data retrieved from http://deep.ps.uci.edu/DR4/home.html.

[24] S. Desai et al.. The Blanco Cosmology Survey: Data Acquisition, Processing, Calibration, Quality Diagnosics and Data Release. *The Astrophysical Journal*, 757, 83, 2012.

[25] H. Diehl, S. Allam, J. Annis, et al.. The Sloan Bright Arcs Survey: Four Strongly Lensed Galaxies with Redshift > 2. *The Astrophysical Journal*, 707, 686, 2009.

[26] S. Dodelson. *Modern Cosmology*. Boston: Academic Press, 2003.

[27] E. Drabek, H. Lin, et al.. Spectroscopy of Two Strong Lensing Systems. In preparation, 2012.

[28] A. Duffy, J. Schaye, S. Kay, & C. Dalla-Vecchia. Dark matter halo concentrations in the Wilkinson Microwave Anisotropy Probe year 5 cosmology. *Monthly Notices of the Royal Astronomical Society*, 390, L64, 2008.

[29] J. Eldred. *Solomonoff Induction*, talk given at meeting of Fermilab Society of Philosophy, April 10, 2014. Available from http://www.youtube.com/watch?v=yXkEpzexYIU.

[30] D. M. Elmegreen. *Galaxies and Galactic Structure*. New Jersey: Prentice Hall, 1998.

[31] A. Evrard, et al.. Virial Scaling of Massive Dark Matter Halos: Why Clusters Prefer a High Normalization Cosmology. *The Astrophysical Journal*, 672, 122, 2008.

[32] C. Fedeli. The Effects of Baryonic Cooling on the Concentration-Mass Relation. *Monthly Notices of the Royal Astronomical Society*, 424, 1244, 2008.

[33] M. Gladders & H. Yee. A New Method for Galaxy Cluster Detection. I. The Algorithm. *The Astronomical Journal*, 120, 2148, 2000.

[34] M. Gralla, et al.. Sunyaev Zel'dovich Effect Observations of Strong Lensing Galaxy Clusters: Probing the Over-Concentration Problem. *The Astrophysical Journal*, 737, 74, 2011.

[35] Gravitational Lensing Wiki. Retrieved from http://gravitationallensing.pbworks.com/w/page/15553257/Strong%20Lensing.

[36] D. Gross. Keynote Talk at 2012 Fermilab Users Meeting.

[37] S. Hansen, et al.. Measurement of Galaxy Cluster Sizes, Radial Profiles and Luminosity Functions from SDSS Photometric Data. *The Astrophysical Journal*, 633, 122, 2005.

[38] J. Hao. *Optical Galaxy Cluster Detection Across a Wide Redshift Range*. PhD Dissertation, University of Michigan.

[39] J. Hao et al.. A GMBCG Galaxy Cluster Catalog of 55,424 Rich Clusters from SDSS DR7. *Astrophysical Journal Supplement*, 191, 254H, 2010.

[40] M. Hasselfield et al.. The Atacama Cosmology Telescope: Sunyaev-Zel'dovich Selected Galaxy Clusters at 148 GHz from Three Seasons of Data. *Journal of Cosmology and Astroparticle Physics*, Issue 07, article id 008, 2013.

[41] C. Heymans. Weak Gravitational Lensing: Lecture Notes. IPM School on Gravitational Lensing Techniques, March 2008, Tehran, Iran.

[42] G. Hill. An Image Quality Figure of Merit Based on Moffat Functions? *HET Technical Reports*, http://het.as.utexas.edu/HET/TechReports/moffat/moffat.html, 2010.

[43] C. Hirata & U. Seljak. Shear Calibration Biases in Weak Lensing Surveys. *Monthly Notices of the Royal Astronomical Society*, 343, 459, 2003.

[44] W. Hu and A. Kravtsov. Sample Variance Considerations for Cluster Surveys. *The Astrophysical Journal*, 584, 702, 2003.

[45] W. Jaffe. A Simple Model for the Distribution of Light in Spherical Galaxies. *Monthly Notices of the Royal Astronomical Society*, 202, 995, 1983.

[46] D. Johnston. Routine *CLOSE_MATCH_RA_DEC*. Retrieved from http://www.astro.washington.edu/docs/idl/cgi-bin/getpro/library36.html.

[47] D. Johnston, E. Sheldon, R. Wechsler, et al.. Cross Correlation Weak Lensing of SDSS Galaxy Clusters II: Cluster Density Profiles and the Mass-Richness Relation. ArXiv: Astro-ph 0709.1159v1, 2007.

[48] N. Kaiser, G. Squires and T. Broadhurst. A Method for Weak Lensing Observations. *The Astrophysical Journal*, 449, 460, 1995.

[49] C. R. Keeton. A Catalog of Mass Models for Gravitational Lensing. Retrieved from http://redfive.rutgers.edu/ keeton/gravlens/catalog.pdf.

[50] C. Kochanek, P. Schneider and J. Wambsganss. *Gravitational Lensing: Strong, Weak and Micro*. Berlin: Springer-Verlag, 2003.

[51] B. Koester, et al.. MaxBCG: A Red-Sequence Galaxy Cluster Finder. *The Astrophysical Journal*, 660, 221, 2007a.

[52] B. Koester, et al.. A MaxBCG Catalog of 13,823 Galaxy Clusters from the Sloan Digital Sky Survey. *The Astrophysical Journal*, 660, 239, 2007b.

[53] R. Kormann, P. Schneider M. Bartelmann. Isothermal elliptical gravitational lens models. *Astronomy and Astrophysics*, 284, 285, 1994.

[54] D. Kubik. *Strong Gravitational Lensing Systems Found in the Sloan Digital Sky Survey*. Masters Thesis, Northern Illinois University, 2007.

[55] J. Kubo, et al.. Dark Matter Structures in the Deep Lens Survey. *The Astrophysical Journal*, 702, 980, 2009.

[56] J. Kubo, S. Allam, et al.. The Sloan Bright Arcs Survey: Discovery of Seven New Strongly Lensed Galaxies from z = 0.66-2.94. *The Astrophysical Journal*, 724, L137, 2010.

[57] M. Lima, C. Cunha et al.. Estimating the redshift distribution of photometric galaxy samples. *Monthly Notices of the Royal Astronomical Society*, 390, 118, 2008.

[58] H. Lin, et al.. Discovery of a Very Bright, Strongly Lensed z=2 Galaxy in the SDSS DR5. *The Astrophysical Journal*, 699, 1242, 2009.

[59] R. Mandelbaum et al.. A Halo Mass-Concentration Relation from Weak Lensing. *Journal of Cosmology and Astroparticle Physics*, Issue 08, id. 006, 2008.

[60] S. Mollerach and E. Roulet. *Gravitational Lensing and Microlensing*. Hackensack, N.J.: World Scientific, 2002.

[61] NASA, N. Benitez, T. Broadhurst, H. Ford, M. Clampin, G. Hartig, G. Illingworth, the ACS Science Team and ESA. Image published online at ¡http://hubblesite.org/newscenter/newsdesk/archive/releases/2003/01/image/a¿.

[62] R. Narayan & M. Bartelmann. Lectures on Gravitational Lensing. ArXiv: Astro-ph 9606001v2, 1997.

[63] J. Navarro, C. Frenk, and S. White. A Universal Density Profile from Hierarchical Clustering. *The Astrophysical Journal*, 490, 493, 1997.

[64] M. Oguri and R. Blandford. What is the Largest Einstein Radius in the Universe? *Monthly Notices of the Royal Astronomical Society*, 392, 930, 2009.

[65] M. Oguri, et al.. Subaru Weak Lensing Measurements of Four Strong Lensing Clusters: Are Lensing Clusters Overconcentrated? *The Astrophysical Journal*, 699, 1038, 2009.

[66] M. Oguri, et al.. Combined Strong and Weak Lensing Analysis of 28 Clusters from the Sloan Giant Arcs Survey. *Monthly Notices of the Royal Astronomical Society*, 420, 3213, 2012.

[67] C. Peng. Galfit. Retrieved from http://users.obs.carnegiescience.edu/peng/work/galfit/galfit.h

[68] F. Prada, et al.. Halo Concentrations in the Standard ΛCDM Cold Dark Matter Cosmology. *Monthly Notices of the Royal Astronomical Society*, 423, 3018, 2012.

[69] C. L. Reichardt et al.. Galaxy Clusters Discovered Via the Sunyaev-Zeldovich Effect in the First 720 Square Degrees of the South Pole Telescope Survey. *The Astrophysical Journal*, 763, 127, 2013.

[70] E. Rozo, et al.. Constraining the Scatter in the Mass-Richness Relation of maxBCG Clusters with Weak Lensing and X-Ray Data. *The Astrophysical Journal*, 699, 768, 2009.

[71] E. Rozo, et al.. Cosmological Constraints from the SDSS MaxBCG Cluster Catalog. *The Astrophysical Journal*, 708, 645, 2010.

[72] E. S. Rykoff et al.. Robust Optical Richness Estimation with Reduced Scatter. *The Astrophysical Journal*, 746, 178, 2012.

[73] E. S. Rykoff et al.. redMaPPer. I. Algorithm and SDSS DR8 Catalog. *The Astrophysical Journal*, 785, 104R, 2014.

[74] P. Schechter. On the Significance of the Luminosities of First-Ranker Members of Sparse Groups of Galaxies. *The Astrophysical Journal*, 209, 670, 1976.

[75] York et al.. Data retrieved from http://skyserver.sdss3.org/CasJobs/.

[76] E. Sheldon et al.. Cross-correlation Weak Lensing of SDSS Galaxy Clusters. III. Mass-to-Light Ratios. *The Astrophysical Journal*, 703, 2232, 2009.

[77] M. Simet et al.. The Sloan Digital Sky Survey Co-add: Cross-correlation Weak Lensing and Tomography of Galaxy Clusters. *The Astrophysical Journal*, 748, 128, 2012.

[78] The SDSS Collaboration. SDSS Data Release 7: Sky Coverage. Image retrieved from http://www.sdss.org/dr7/coverage/.

[79] M. Soares-Santos, et al.. The Voronoi Tessellation Cluster Finder in 2+1 Dimensions. *The Astrophysical Journal*, 727, 45, 2010.

[80] D. Sobel. *A More Perfect Heaven.* Walker and Company, 2011.

[81] S. Suyu, et al.. A Bayesian analysis of regularized source inversions in gravitational lensing. *Monthly Notices of the Royal Astronomical Society*, 371, 983, 2006.

[82] J.L. Tinker, D.H. Weinberg & Z. Zheng. Redshift-Space Distortions with the Halo Occupation Distribution I: Numerical Simulations. *Monthly Notices of the Royal Astronomical Society*, 368, 85, 2006.

[83] D. Tody. IRAF in the Nineties. Astronomical Data Analysis Software and Systems II, A.S.P. Conference Series, Vol 52, eds. R.J. Hanisch, R.J.V. Brissenden, & J. Barnes, 173., 1993.

[84] P. Van Dokkum. Cosmic-Ray Rejection by Laplacian Edge Detection. *The Publications of the Astronomical Society of the Pacific*, Volume 113 1420, 2001.

[85] A. Vikhlinin et al.. Chandra Cluster Cosmology Project III: Cosmological Parameter Constraints. *The Astrophysical Journal*, 692, 1060, 2009.

[86] The VIPERS Collaboration. Data retrieved from http://vipers.inaf.it/rel-pdr1.html.

[87] The VVDS Collaboration. Data retrieved from http://cesam.oamp.fr/vvds/vvds_criteria.php.

[88] R. B. Wayth and R. L. Webster. Lensview: Software for Modelling Resolved Gravitational Lens Images. *Monthly Notices of the Royal Astronomical Society*, 372, 1187, 2006.

[89] R. Wayth. Lensview. Retrieved from http://cira.ivec.org/dokuwiki/doku.php/staff/rwayth/le

[90] S. Weinberg. *Cosmology*. Oxford University Press, 2008.

[91] A. West, et al.. The Sloan Bright Arcs Survey: Modeling of Four Strong Gravitational Lens Systems with z ¿ 2. *The Astrophysical Journal* submitted, 2012.

[92] Z. L. Wen, J. L. Han & F. S. Liu. Erratum: Galaxy Clusters Identified from the Sloan Digital Sky Survey DR6 and their Properties. *The Astrophysical Journal Supplement*, 187, 272W, 2010.

[93] M. Wiesner. *On the Properties of Ten Strong-Lensing Systems Found in the Sloan Digital Sky Survey*. Masters Thesis, Northern Illinois University, 2010.

[94] R. Williamson et al.. A Sunyaev-Zel'dovich-selected Sample of the Most Massive Galaxy Clusters in the 2500 deg^2 South Pole Telescope Survey. *The Astrophysical Journal*, 738, 139, 2011.

[95] C. Wright and T. Brainerd. Gravitational Lensing by NFW Halos. *The Astrophysical Journal*, 534, 34, 2000.

[96] D. York, et al.. The Sloan Digital Sky Survey: Technical Summary. *The Astronomical Journal*, 120, 1579, 2000.

[97] A. Zitrin, T. Broadhurst, Y. Rephaeli, & S. Sadeh. The Largest Gravitational Lens: MACS J0717.5+3745 (z=0.546). *The Astrophysical Journal*, 707, L102, 2009.

[98] A. Zitrin, et al.. Strong-Lensing Analysis of a Complete Sample of 12 MACS Clusters at z ¿ 0:5: Mass Models and Einstein Radii. *Monthly Notices of the Royal Astronomical Society*, 410, 1939, 2011.

[99] Y. Zu et al.. Cosmological Constraints from the Large-Scale Weak Lensing of SDSS MaxBCG Clusters. *Monthly Notices of the Royal Astronomical Society*, 439, 1628, 2014.

[100] F. Zwicky. On the Masses of Nebulae and of Clusters of Nebulae. *The Astro-physical Journal*, 86, 217, 1937.

APPENDIX A

APPENDIX
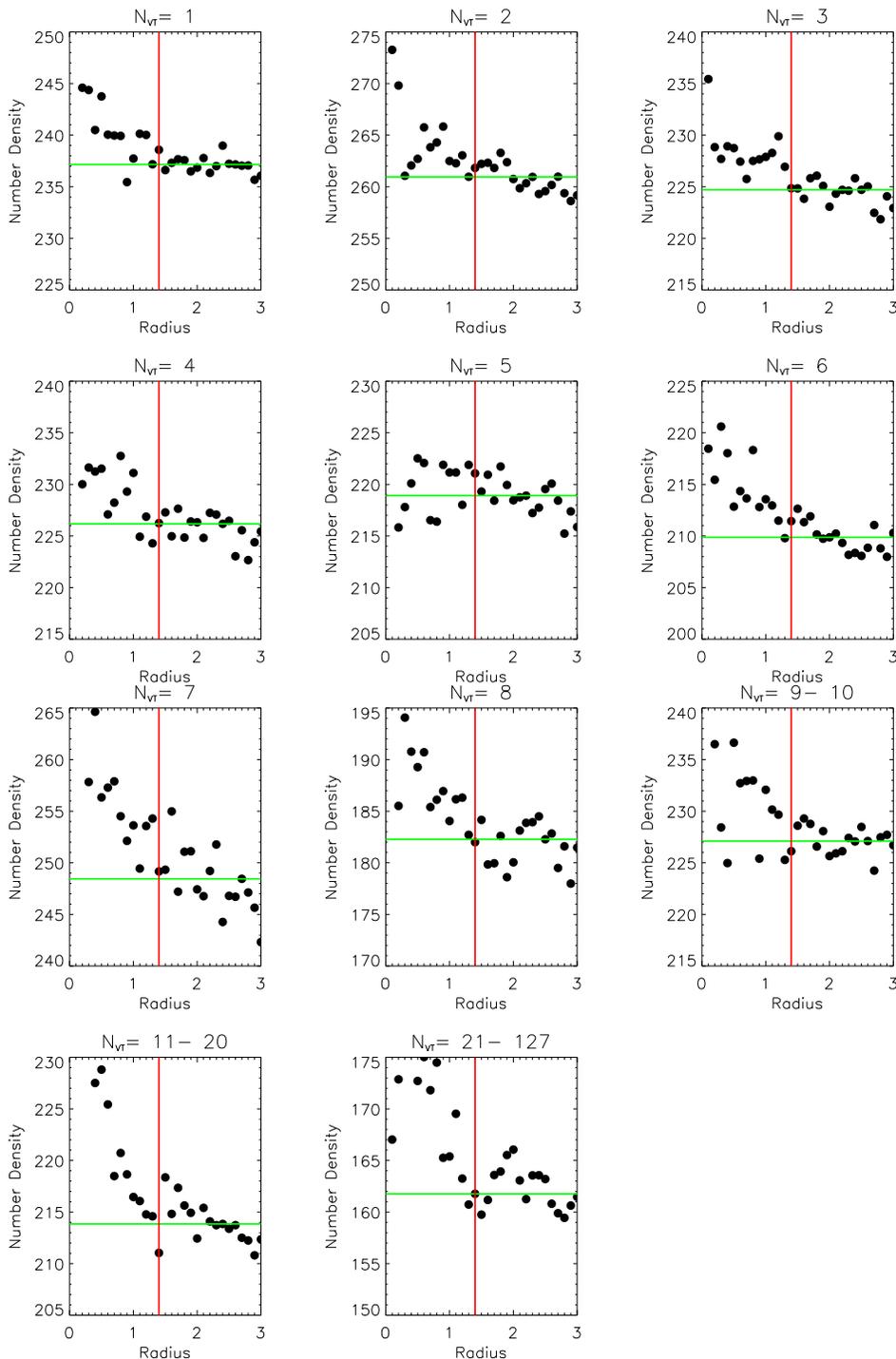
Figure A.1: Plots of number density nearby clusters as a function of radius for all richness bins and $0.1 \leq 0.4$. The green line is the median value of the correction between $1.4 - 3.0h^{-1}$ Mpc and the red line indicates the $1.4h^{-1}$ Mpc point where calculation of the median begins.
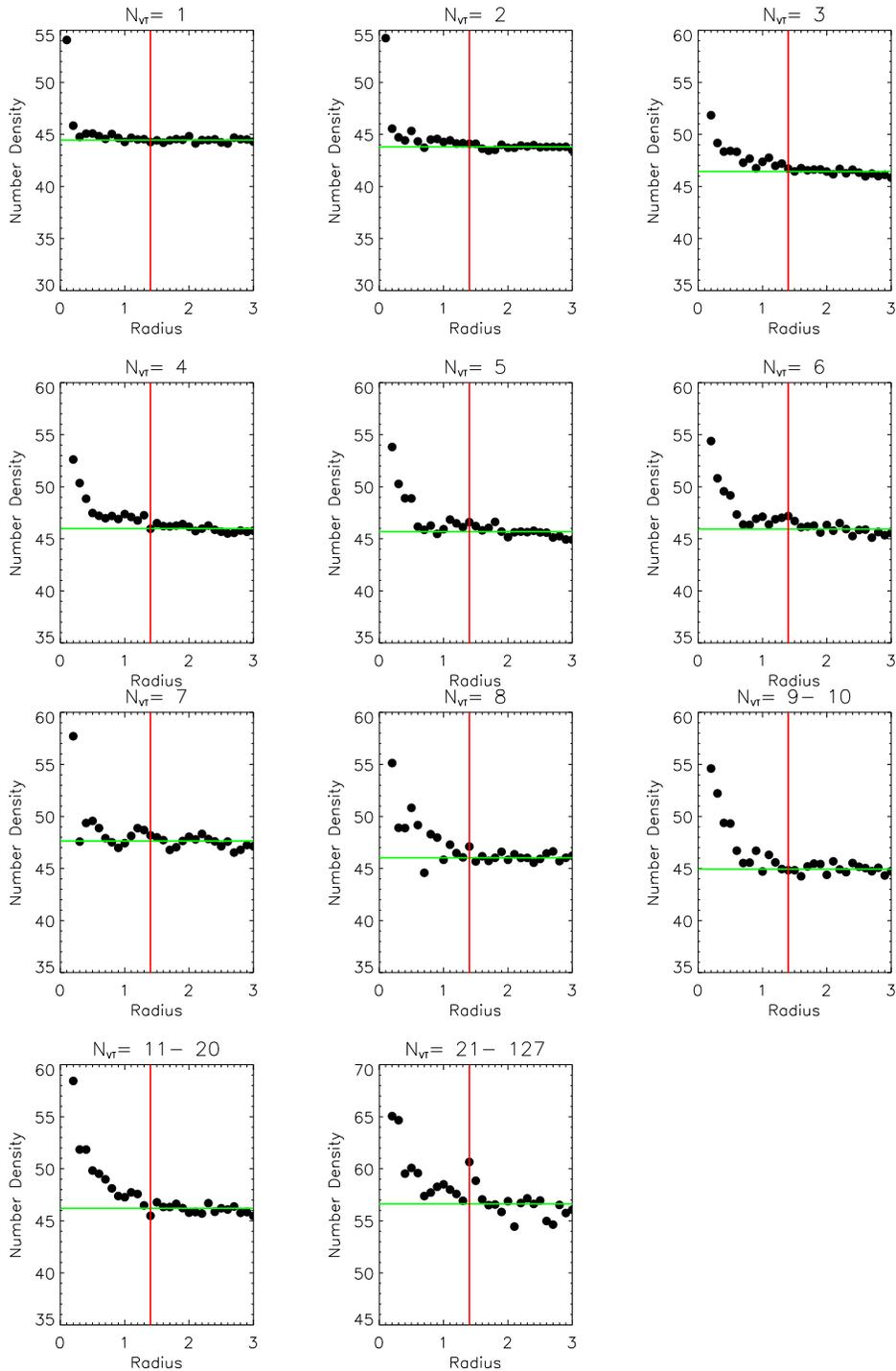
Figure A.2: Plots of number density nearby clusters as a function of radius for all richness bins and $0.4 \leq 0.7$. The green line is the median value of the correction between $1.4 - 3.0h^{-1}$ Mpc and the red line indicates the $1.4h^{-1}$ Mpc point where calculation of the median begins.
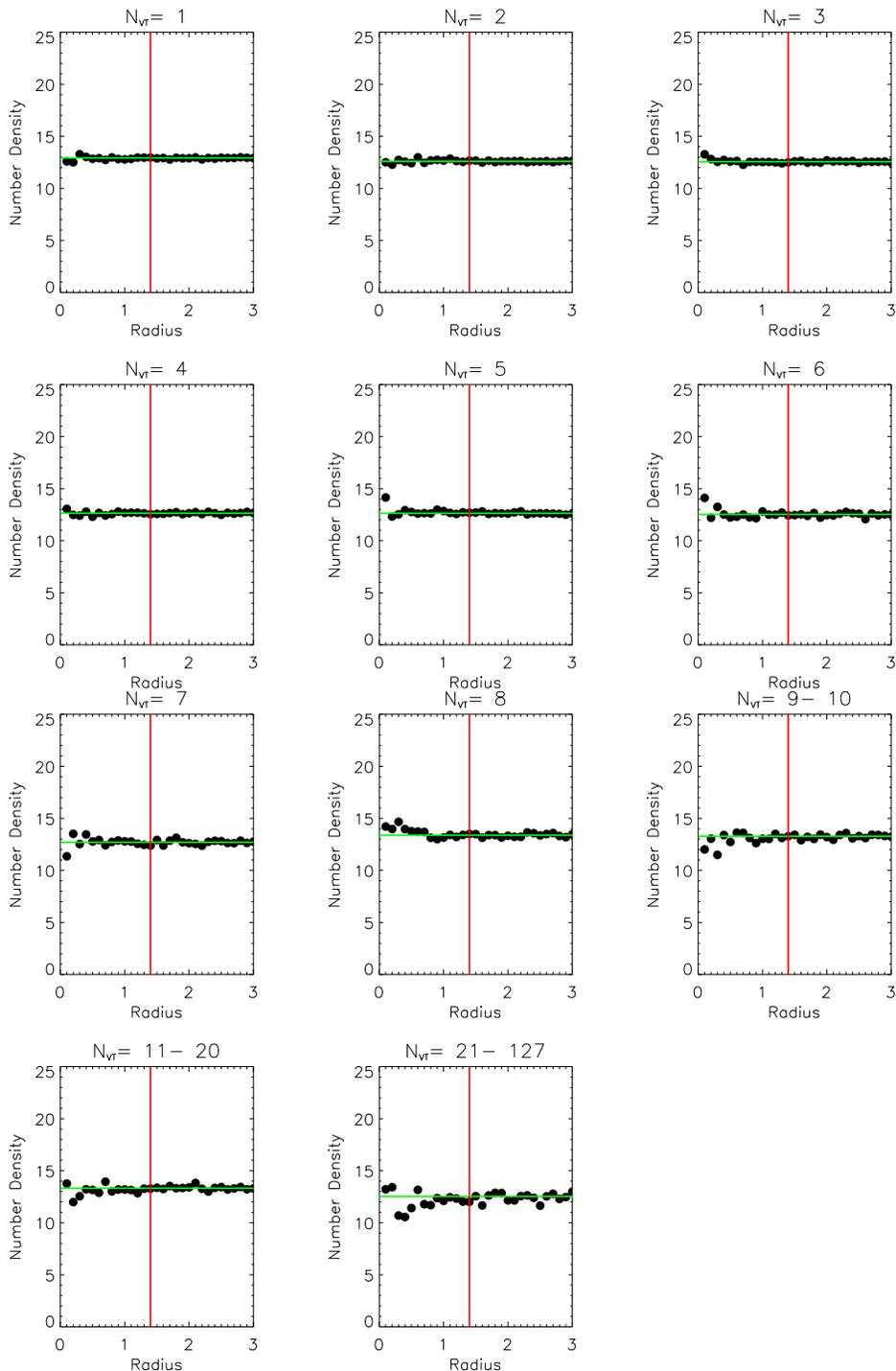
Figure A.3: Plots of number density nearby clusters as a function of radius for all richness bins and $0.7 \leq 1.0$. The green line is the median value of the correction between $1.4 - 3.0h^{-1}$ Mpc and the red line indicates the $1.4h^{-1}$ Mpc point where calculation of the median begins.
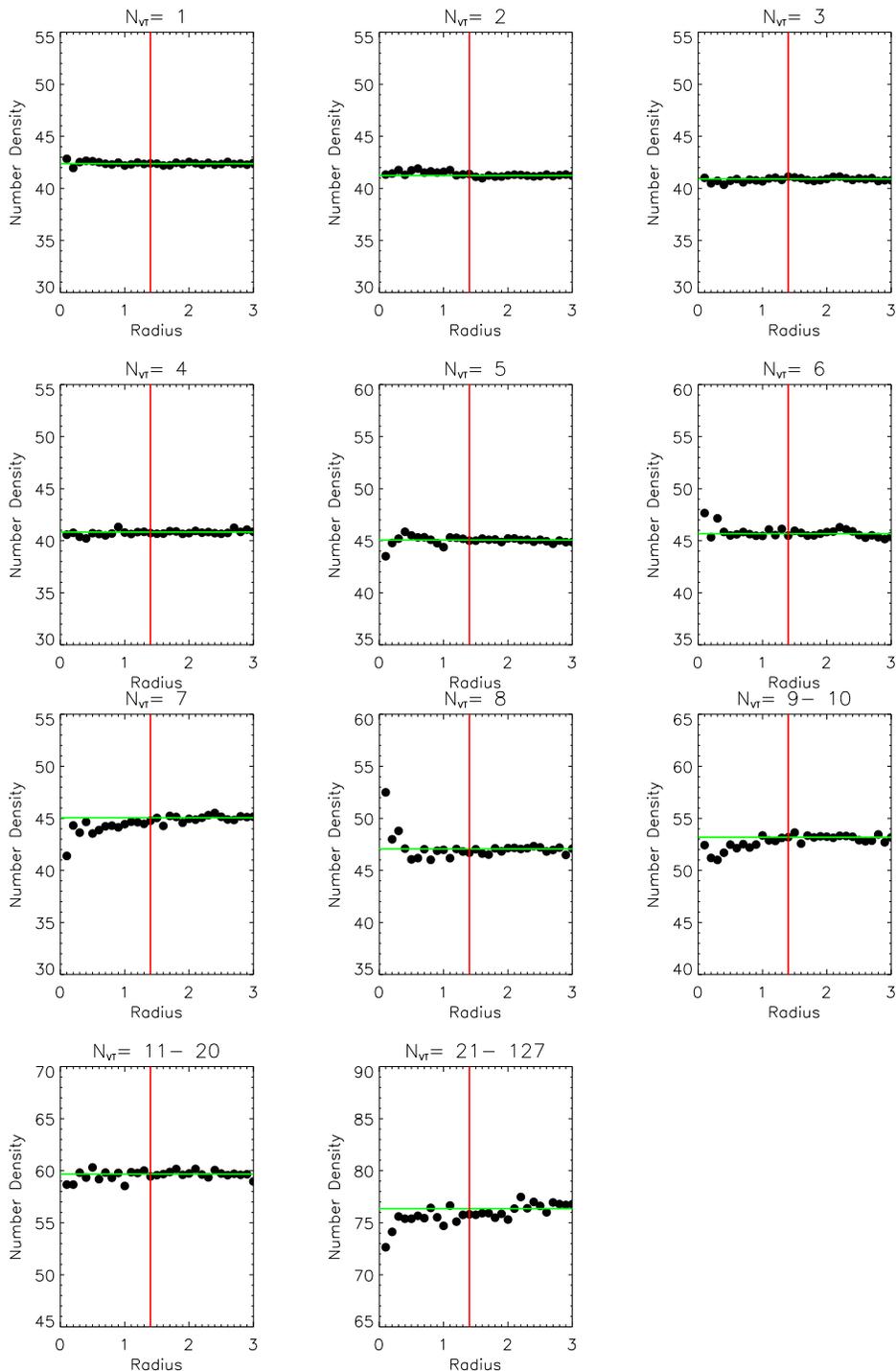
Figure A.4: Plots of number density nearby clusters as a function of radius for all richness bins and $0.1 \leq 1.0$. The green line is the median value of the correction between $1.4 - 3.0h^{-1}$ Mpc and the red line indicates the $1.4h^{-1}$ Mpc point where calculation of the median begins.

Table A.1: Values of N_{VT}, M_{200}, reduced χ^2 for fit to 0 tangential shear, reduced χ^2 probability for this fit and c_{200} for low-z and mid-z bins in the Stripe 82 coadd, including all systematics for low-z and mid-z. Note that all values of M_{200} have the units $10^{14}\ h^{-1}\ M_\odot$.

N_{VT}	M_{200}	χ^2_{red}	**Prob**	c_{200}
LOW-Z				
1±0	0.00406±0.012009	1.416802	0.067667	0.252202±0.70859
2±0	0.058286±0.016521	1.73268	0.008492	3.799757±3.569984
3±0	0.052419±0.032134	1.93739	0.001797	0.561834±0.462221
4±0	0.089228±0.038313	1.806692	0.004925	1.129056±0.876895
5±0	0.164858±0.046633	2.494385	0.000014	1.358829±0.76223
6±0	0.141509±0.042059	1.827822	0.004201	3.856865±3.191844
7±0	0.281613±0.067054	2.259911	0.000121	3.247561±1.781743
8±0	0.008465±0.03857	2.260948	0.000119	1.404976±12.036672
9±0.5	0.284505±0.062433	2.54702	0.000009	2.108155±1.011853
11±0.5	0.011267±0.034215	4.600511	0	2.909726±27.526752
14±1	0.527766±0.12408	4.498949	0	0.824398±0.362104
18±2	0.457485±0.102334	2.712129	0.000002	1.634428±0.740969
23±4.5	0.427442±0.178904	2.590297	0.000006	0.408939±0.283339
35±4.5	1.410066±0.296783	3.017968	0	1.535705±0.653672
46±4.5	1.666487±0.357641	2.80289	0.000001	1.834109±0.775471
54±4.5	3.244358±0.929298	2.282188	0.000099	1.184691±0.694519
65±3	2.72311±0.802419	3.056965	0	1.299974±0.782216
78±4	3.299089±1.485384	1.966073	0.00143	1.146059±1.083152
127±14	3.061078±1.630495	1.420342	0.06628	0.884632±0.991859
MID-Z				
1±0	0.069772±0.03618	1.576941	0.024949	1.051204±0.942445
2±0	0.003561±0.006142	1.588484	0.023113	0.017414±0.056651
3±0	0.059976±0.036219	1.33745	0.10583	1.37736±1.628322
4±0	0.229013±0.052323	2.396917	0.000035	6.694057±4.441216
5±0	0.037±0.020071	2.371531	0.000045	0.001301±0.251882
6±0	0.314665±0.113355	2.374317	0.000043	0.956172±0.631958
7±0	0.214724±0.142441	1.507642	0.038993	0.510683±0.513062
8±0	0.599321±0.2341	2.001758	0.001073	0.476461±0.329924
9±0.5	0.595091±0.137589	2.745069	0.000001	2.287043±1.071839
11±0.5	0.878344±0.199879	2.707946	0.000002	3.290453±1.444648
14±1	0.088256±0.158563	1.290609	0.135599	0.220921±0.502494
17±2	0.853927±0.211785	2.448511	0.000022	3.277201±1.541722
24±4.5	1.118405±0.319318	1.839168	0.003854	1.485424±0.853651
35±4	1.563332±1.130345	1.780024	0.006007	0.466469±0.727339
44±3	0.392595±0.53786	1.834253	0.004001	54.881158±485.101411
54±2	2.155516±0.974095	1.458665	0.05275	3.071444±2.495965
63±4	2.787909±2.396291	1.164216	0.248069	0.392854±0.893397
91±7.5	1.218674±1.920435	0.745867	0.835177	1.148463±3.781818

Table A.2: Values of N_{VT}, M_{200}, reduced χ^2 for fit to 0 tangential shear, reduced χ^2 probability for this fit and c_{200} for high-z and all-z bins in the Stripe 82 coadd, including all systematics for high-z and all-z. Note that all values of M_{200} have the units 10^{14} h^{-1} $M_\odot$.

N_{VT}	M_{200}	χ^2_{red}	**Prob**	c_{200}
1±0	0.002029±0.009283	0.957746	0.529986	0.01483±0.12344
2±0	0.01475±0.049737	0.86649	0.671578	0.234656±0.816177
3±0	0.040665±0.031091	1.545216	0.03069	0.001878±0.274297
4±0	0.017853±0.067707	1.276644	0.145645	1.313342±9.352414
5±0	0.103945±0.116542	1.842378	0.003761	1.530302±3.495593
6±0	0.160461±0.197662	0.698788	0.884504	0.769622±1.537464
7±0	0.309319±0.279084	0.58758	0.961559	1.002459±1.664817
8±0	0.103975±0.277	0.897874	0.623441	0.515606±1.92646
9±0.5	0.745032±0.341688	1.077766	0.353493	1.755921±1.563842
11±0.5	0.004744±0.128912	0.853321	0.691342	2.588025±199.342227
14±1	0.724868±2.88679	1.021296	0.433272	0.002621±1.402359
17±2	1.504825±0.856889	1.009955	0.450107	1.91472±1.935946
24±4.5	1.77267±1.865614	1.336321	0.106479	0.038911±0.609364
34±4.5	3.817381±3.497042	1.239977	0.174752	3.79427±5.040852
44±5	0.006809±0.781065	1.269146	0.151272	0.398401±48.844501
ALL-Z				
1±0	0.014803±0.015163	1.44742	0.056448	0.648094±0.861629
2±0	0.010365±0.012625	1.780743	0.005975	0.991389±1.873042
3±0	0.035904±0.021049	2.718256	0.000002	0.788744±0.683776
4±0	0.123229±0.028352	2.877786	0	1.687073±0.828339
5±0	0.144676±0.046665	3.311249	0	0.645999±0.318774
6±0	0.164156±0.047078	3.059856	0	1.078554±0.564484
7±0	0.215766±0.058829	2.782672	0.000001	1.227187±0.642782
8±0	0.246157±0.092297	2.795151	0.000001	0.48445±0.279183
9±0.5	0.350722±0.057498	3.741433	0	1.889484±0.65386
11±0.5	0.030831±0.031881	5.860152	0	2.839587±8.926458
14±1	0.023933±0.034446	4.452306	0	1.996493±7.008187
17±2	0.534629±0.093642	4.03183	0	1.566853±0.551982
24±4.5	0.543445±0.160251	3.544194	0	0.500548±0.258482
35±4.5	1.378453±0.282283	3.196332	0	1.296939±0.543255
45±4.5	1.492566±0.338548	2.476048	0.000017	1.390023±0.643546
54±4.5	2.732788±0.683912	2.57956	0.000006	1.366189±0.69135
65±4	2.618092±0.76672	3.12242	0	0.928388±0.592224
78±4	3.022221±1.441194	1.758291	0.007049	0.984739±1.013918
99±18	2.385235±1.338059	0.968976	0.512555	0.79108±0.94815

APPENDIX B

APPENDIX2

Table B.1: Fit parameters in the g-band for SDSS J1537+6556. Reduced χ^2 for this fit is 1.152294. Parameters in brackets were held constant.

Object	Fit Type	X-Coord	Y-Coord	Magnitude	Half-Light Radius	Sersic Index	Axis Ratio	Position Angle
1	sersic	1374.3979 ± 0.0237	2084.9402 ± 0.0205	17.4389 ± 0.0186	28.3697 ± 0.7329	3.5575 ± 0.0510	0.8737± 0.0046	88.8789 ± 1.3666
2	sersic	1390.4496 ± 0.0530	2105.5269 ± 0.0556	19.1156 ± 0.3483	43.0002 ± 40.0398	13.0348 ± 3.9879	0.7732± 0.0263	31.8516 ± 4.8259
3	sersic	1383.0238 ± 0.0624	2064.4194 ± 0.0569	20.3739 ± 0.0439	7.2176 ± 0.5294	3.4674 ± 0.3756	0.5045± 0.0231	-51.9336 ± 2.1955
4	sersic	1346.4375 ± 0.1048	2017.4507 ± 0.1069	21.4388 ± 0.0238	1.8511 ± 0.2304	4.0000	0.6272± 0.1143	70.5943 ± 11.0105
5	psf	1294.6576 ± 0.1056	2055.7161 ± 0.0985	22.1252 ± 0.0260	–	–	–	–
6	sersic	1308.4723 ± 0.1581	2060.7151 ± 0.6147	20.9757 ± 0.2128	4.3045 ± 0.3514	1.1978 ± 0.2103	0.5956± 0.0362	-28.8738 ± 6.3180
7	sersic	1306.9716 ± 0.0926	2055.6104 ± 0.2699	20.6643 ± 0.1399	3.2199 ± 0.3708	3.5630 ± 0.8741	0.7511± 0.0578	-75.3914 ± 16.3601
8	psf	1330.2479 ± 0.5937	2079.4106 ± 0.6174	24.3845 ± 0.1997	–	–	–	–
9	psf	1322.7611 ± 0.6326	2110.9307 ± 0.6915	24.3218 ± 0.1881	–	–	–	–
10	psf	1436.5002 ± 0.5514	2098.8923 ± 0.5940	24.1311 ± 0.1577	–	–	–	–
11	sky	–	–	256.6260	–	–	–	–

Table B.2: Fit parameters in the r-band for SDSS J1537+6556. Reduced χ^2 for this fit is 1.151182. Parameters in brackets were held constant.

Object	Fit Type	X-Coord	Y-Coord	Magnitude	Half-Light Radius	Sersic Index	Axis Ratio	Position Angle
1	sersic	1376.3577 ± 0.0089	2086.2585 ± 0.0083	15.9208 ± 0.0105	27.4349 ± 0.4225	3.7070 ± 0.0323	0.8837± 0.0020	-86.1759 ± 0.6301
2	sersic	1392.4614 ± 0.0165	2106.8474 ± 0.0183	18.0033 ± 0.0787	18.3057 ± 3.6837	10.5145 ± 0.9491	0.6833± 0.0111	33.3883 ± 1.3199
3	sersic	1385.2336 ± 0.0228	2065.6501 ± 0.0227	18.9964 ± 0.0088	6.7768 ± 0.1261	4.0000	0.4283± 0.0106	-47.1949 ± 0.7511
4	sersic	1348.3701 ± 0.0352	2018.5625 ± 0.0361	20.0365 ± 0.0084	1.5629 ± 0.0753	4.0000	0.7314± 0.0443	67.5400 ± 6.2341
5	psf	1297.0903 ± 0.0692	2057.2061 ± 0.0665	21.7565 ± 0.0260	–	–	–	–
6	sersic	1310.9547 ± 0.0730	2062.4309 ± 0.1230	20.2751 ± 0.0425	5.5600	4.0000	0.7025± 0.0392	-14.8913 ± 7.3438
7	sersic	1309.0724 ± 0.0411	2057.2107 ± 0.0552	20.0828 ± 0.0467	2.7471 ± 0.2587	4.0000	0.8026± 0.0335	-63.5804 ± 12.4048
8	psf	1324.5046 ± 0.2854	2111.4158 ± 0.3045	23.3713 ± 0.0935	–	–	–	–
9	psf	1439.8379 ± 0.2102	2099.2566 ± 0.2083	23.1049 ± 0.0793	–	–	–	–
10	sky	–	–	848.0933	–	–	–	–

Table B.3: Fit parameters in the i-band for SDSS J1537+6556. Reduced χ^2 for this fit is 1.028762.

Object	Fit Type	X-Coord	Y-Coord	Magnitude	Half-Light Radius	Sersic Index	Axis Ratio	Position Angle
1	sersic	1377.3176 ± 0.0090	2086.4600 ± 0.0083	15.9269 ± 0.0107	27.2935 ± 0.4125	3.7337 ± 0.0286	0.8836± 0.0023	-84.8530 ± 0.6336
2	sersic	1393.4456 ± 0.0156	2107.0481 ± 0.0176	18.1580 ± 0.0627	12.8645 ± 1.8865	8.3298 ± 0.6435	0.6526± 0.0118	29.7365 ± 1.2284
3	sersic	1386.3014 ± 0.0225	2065.8818 ± 0.0225	18.9981 ± 0.0300	6.9839 ± 0.3888	4.4912 ± 0.2397	0.4079± 0.0117	-46.8339 ± 0.7253
4	sersic	1349.4187 ± 0.0293	2018.7106 ± 0.0310	20.0554 ± 0.0116	1.6087 ± 0.0889	4.0000	0.7050± 0.0460	57.6272 ± 5.6987
5	psf	1298.2356 ± 0.0849	2057.3516 ± 0.0774	22.1870 ± 0.0399	–	–	–	–
6	sersic	1311.5455 ± 0.1076	2061.4131 ± 0.2619	20.5614 ± 0.1061	4.9438 ± 0.2603	1.5076 ± 0.2129	0.5282± 0.0259	-16.5862 ± 4.3823
7	sersic	1309.9556 ± 0.1430	2057.1079 ± 0.1427	20.8710 ± 0.1789	3.9186 ± 0.9491	4.0000	0.6778± 0.0789	-77.4156 ± 15.8052
8	psf	1331.7040 ± 0.1791	2081.4048 ± 0.1908	23.0551 ± 0.0816	–	–	–	–
9	psf	1325.3127 ± 0.2944	2111.8164 ± 0.3160	23.3717 ± 0.1076	–	–	–	–
10	psf	1440.1631 ± 0.2262	2099.1506 ± 0.2371	23.1588 ± 0.0998	–	–	–	–
11	sky	–	–	1591.1017	–	–	–	–