

ABSTRACT

Name: Elizabeth A. Holden

Department: Physics

Title: The Accuracy of the Photometric Redshift of Galaxy Clusters

Major: Physics

Degree: Master of Science

Approved by:

Date:

---

Thesis Director

NORTHERN ILLINOIS UNIVERSITY

## ABSTRACT

Galaxy clusters provide the means to measure different astrophysical properties. Accurate measurement of the redshifts of these galaxy clusters is important in these studies. There are two different ways to measure redshift – spectroscopic redshift and photometric redshift. Spectroscopic redshift is more accurate than photometric redshift, but it is also more time-consuming to measure. If one is going to measure the photometric redshift of galaxy clusters, it is important to know how accurate those measurements are.

I examined a catalog of 28,000 galaxy clusters from Sloan Digital Sky Survey (SDSS) data. Their photometric redshift had been calculated by Dr. James Annis, a scientist at Fermi National Laboratory who works on SDSS. Using two different techniques, I extracted the spectroscopic redshift of each cluster from SDSS online databases. I then compared the photometric redshifts and spectroscopic redshifts of the clusters, to measure how accurate the photometric redshifts were.

I examined what factors affect the accuracy of photometric redshift. I compared the accuracy of photometric redshift with the redshift of the cluster and with the number of galaxies in the cluster. I found that at any redshift, all clusters with 20 or more galaxies have a dispersion of less than 0.02, which is considered a small amount of error, small enough to use measurements of photometric redshift.

NORTHERN ILLINOIS UNIVERSITY

THE ACCURACY OF THE PHOTOMETRIC REDSHIFT OF GALAXY CLUSTERS

A THESIS SUBMITTED TO THE GRADUATE SCHOOL  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE  
MASTER OF SCIENCE

DEPARTMENT OF PHYSICS

BY  
ELIZABETH HOLDEN

DEKALB, ILLINOIS

AUGUST 2006

Certification:

In accordance with departmental and Graduate  
School policies, this thesis is accepted in  
partial fulfillment of degree requirements.

---

Thesis Director

---

Date

## TABLE OF CONTENTS

|                      | Page |
|----------------------|------|
| LIST OF FIGURES..... | iv   |
| Chapter              |      |
| 1. INTRODUCTION..... | 1    |
| 2. DATA.....         | 4    |
| 3. RESULTS.....      | 9    |
| 4. CONCLUSION.....   | 23   |
| REFERENCES.....      | 25   |

## LIST OF FIGURES

| Figure  | Page |
|---|------|
| 1. Fraction of objects with spectroscopic data vs. average photometric redshift.....  | 9    |
| 2. Declination vs. right ascension.....   | 10   |
| 3. Histogram of the residuals.....  | 11   |
| 4. Bias of residuals vs. average photometric redshift.....  | 12   |
| 5. Bias vs. average photometric redshift for $10 \leq N_{\text{gals}} < 20$ .....   | 14   |
| 6. Bias vs. average photometric redshift for $N_{\text{gals}} \geq 20$ .....  | 14   |
| 7. Dispersion vs. average photometric redshift for $10 \leq N_{\text{gals}} < 20$ .....   | 16   |
| 8. Dispersion vs. average photometric redshift for $N_{\text{gals}} \geq 20$ .....  | 17   |
| 9. Zoomed-in look at peak near photometric redshift = 0.3.....  | 18   |
| 10. Dispersion vs. $N_{\text{gals}}$ for all photometric redshift.....  | 20   |
| 11. Dispersion vs. $N_{\text{gals}}$ , $0.1 < \text{photometric redshift} < 0.2$ .....  | 20   |
| 12. Dispersion vs. $N_{\text{gals}}$ with best fit lines exponent not constrained to $-0.5$ for all photometric redshift.....                     | 21   |
| 13. Dispersion vs. $N_{\text{gals}}$ with best fit lines exponent not constrained to $-0.5$ , for $0.1 < \text{photometric redshift} < 0.2$ ..... | 22   |

# CHAPTER ONE

## INTRODUCTION

Galaxy clusters are important to science because they can be used as tools to probe many different aspects of our universe. For example, it was by measuring the rotation curves of galaxy clusters that dark matter was discovered. The velocity of galaxies on the outside of a cluster was such that, if there was not more mass that could not be detected visually, those galaxies should not be held gravitationally in the cluster. This was the first indicator that there must be more matter in the universe, which so far can only be detected gravitationally.

Another example of the importance of galaxy clusters is their use in the Dark Energy Survey (DES). Set to begin in 2009, DES will mount a new CCD camera called the DECam to the Blanco telescope at the Cerro Tololo Inter-American Observatory in Chile and will survey 5,000 square degrees of the sky. DES will measure dark energy and dark matter densities using four independent techniques. The first technique relies on the observation of galaxy clusters because their formation is almost solely dependent on the gravitational dynamics of dark matter. By making optical measurements of clusters, including finding their photometric redshift, DES can find the mass of the clusters, which is useful because the mass of clusters is affected by the cosmology of the universe – by both growth density perturbations and the evolution of the volume element. In the

second technique, the weak lensing of galaxies will be measured as a function of photometric redshift, thereby providing another measure of cluster mass. The third technique will measure the angular clustering of galaxies in shells of photometric redshift out to redshifts of approximately 1.1. The fourth technique will examine the light curves (plots of light intensity as a function of time) of supernovae in a range of  $0.3 < \text{redshift} < 0.75$ .

Because three of the four techniques used to probe the nature of dark energy rely on the measurement of photometric redshift, it is clear that photometric redshift is an important quantity in cosmology. The accuracy of the photometric redshift of galaxy clusters is something that must be known if astrophysicists are to feel confident in their measurements. The question is to how best estimate the accuracy of photometric redshift.

To see how accurate the photometric redshift of an object is, it must be compared to the object's spectroscopic redshift. Spectroscopic redshift is a more accurate measurement of redshift than photometric redshift is, but it is also more difficult and time-consuming to measure. Spectroscopic redshifts are obtained by measuring the spectral lines from an object. The peaks at certain frequencies represent different elements that are burning in the object. Because the location of peaks for different elements are known, redshift is found by measuring how far the peaks are shifted from where they are at zero redshift.

Photometric redshifts are obtained by observing the brightness of an object in different color bands, which are created by using filters that only allow a certain range of



frequencies to come through. The brightness of the object in each color band is then compared to what the brightness for that color band is theoretically, at zero redshift.

The reason that spectroscopic redshift is more accurate than photometric redshift is also the reason that it is more time-consuming to measure. In the measuring of spectroscopic redshift, small wavelength bins (of only a couple angstroms) are used, so long integration times are required to get a sufficiently high signal-to-noise ratio for each bin. In the measuring of photometric redshift, larger bins (thousands of angstroms) are used, so the integration times are shorter, but accuracy is sacrificed.

The goal of my research was to examine the accuracy of photometric redshift and the relationship of its accuracy to the number of galaxies in a cluster and to the redshift of the cluster.

## CHAPTER TWO

### DATA

I analyzed all data in a catalog of 28,000 galaxy clusters obtained from imaging data from the Sloan Digital Sky Survey (SDSS). SDSS (York et al. 2000) uses five filters (Fukugita et al. 1996) to image  $\pi$  steradians of the sky. These filters allow light with ranges of frequency in the ultraviolet (u) (3435 angstroms), green (g) (4770 angstroms), red (r) (6231 angstrom), near infrared (i) (7625 angstrom), and infrared (z) (9134 angstrom) sections of the spectrum. The imaging is done using a CCD mosaic in drift-scanning mode (a technique that uses the rotation of the earth to scan the sky) (Gunn et al. 1998) at a 2.5 meter telescope in Apache Point, New Mexico. Two double-spectrographs are mounted on the telescope. After being processed (Lupton et al. 2001, Stoughton et al. 2002, Pier et al. 2003,) and calibrated (Hogg et al. 2001, Smith et al. 2002, Ivezić et al. 2004) the images are selected for spectroscopy (Eisenstein et al. 2001, Richards et al. 2002, Strauss et al. 2002) using an algorithm that ensures highly complete spectroscopic samples (Blanton et al. 2003). The catalog used for this paper included all galaxy clusters with photometric redshifts  $\leq 0.5$  and number of galaxies in a cluster  $\geq 10$ .

The catalog was provided by Dr. James Annis, a scientist at Fermi National Laboratory who works on SDSS, and it included data for each cluster including right ascension (RA) of the brightest cluster galaxy (BCG), declination (dec) of the BCG, number of galaxies in a cluster ( $N_{\text{gals}}$ ), and photometric redshift. The photometric

redshift was found by Dr. Annis using an algorithm based on the assumption that the photometric redshift of the BCG of each cluster is the photometric redshift of the entire cluster. This assumption is reasonable, as shown by Oegerle and Hill (2001). In their examination of 23 galaxy clusters, they found that the redshift of the BCG in each cluster was the same as the redshift of the entire cluster in 19 of the 23 clusters.

To determine how accurate the photometric redshift of each cluster was, I compared each one's photometric redshift with its spectroscopic redshift. This is reasonable, as the error in spectroscopic redshift is negligibly small, on the order of 0.0001 redshift. To obtain the spectroscopic redshift of each cluster, it had to be extracted from SDSS data online. On the SDSS website, there is a portion entitled "My Database" (MyDB), where any user can log in, upload tables of their own, and download SDSS data.

Each object in SDSS has its own unique ObjectID. The ObjectID of an object is used to extract data about the object, such as its spectroscopic redshift (if there is spectroscopic data available for the object). The catalog I was given did not contain the SDSS ObjectID for the BCGs of the clusters. To obtain the ObjectID, a function called "Neighbors Search" was used. "Neighbors Search" is a button available for use on each table of data uploaded to "My Database" which finds objects with a certain radius of a given RA and dec. The objects found by "Neighbors Search" will have a number called a Matched\_ID associated with them, which is the same as that object's ObjectID in SDSS tables.

At this point, the technique diverges into two paths. I used two different methods of determining the spectroscopic redshift of each cluster so as to compare them and see which one of the photometric redshifts deviated from more. For the first method (the BCG method), a Neighbors Search was done in a 5 arc-second radius around the RA and dec of each cluster. This was done so that SDSS would select only the BCG of the cluster, the same object that the RA and dec from the catalog marked. The 5 arc-second spread was there in case the RA and dec measured in the catalog did not match the exact RA and dec found in the latest SDSS data run. In this way, the SDSS ObjectID was determined for the BCG of each galaxy cluster.

To find the spectroscopic redshift of the BCG from “My Database,” a query using Standard Query Language (SQL) was performed. I learned the basics of SQL to do this; the language uses the commands “Select,” “Into,” “From,” and “Join.” In my case, “SELECT n.ra,n.dec,n.pz,n.N<sub>gals</sub>,sp.z,sp.specclass,sp.objID” indicates I want the RA, dec, and photometric redshift from a table called “n” (which is the table created from the Neighbors Search done on the MaxBCG catalog) and spectroscopic redshift, spectral class and ObjectID from a table called “sp” (a table of SDSS objects that have spectroscopic data associated with them). “INTO mydb.spec\_z from mydb.neigh\_z\_grt\_0\_2\_1 as n” shows that the data selected will be put into a new table called spec\_z and renames the table of data from the Neighbors Search “n” (for simplicity in the “Select” command). “Inner join SpecPhoto as sp on n.matched\_ID=sp.objid” shows that the SDSS table “SpecPhoto” will be called “sp” for simplicity, and a join was done between the catalog and the SDSS data. A join looks for matches in two tables of

data from specified headings (in this case, `matched_ID` and `objid`). This is done so that objects found in the Neighbors Search (in this case, the BCG's) will be matched properly with their SDSS data. An inner join means only objects matched between both tables. This way the information I want to extract (given in the "Select" command) only comes from objects that have spectroscopic data associated with them.

In this first method for obtaining spectroscopic redshift, my SQL query returned approximately 6,000 objects. In the second method (the cluster members method), instead of assuming that the spectroscopic redshift of the BCG is that for the entire cluster, the spectroscopic redshifts of individual galaxies in the cluster were found and then averaged. The original catalog was broken into three pieces – one piece for all clusters with photometric redshifts lower than 0.1, one for all clusters with photometric redshifts between 0.1 and 0.2, and one for all clusters with photometric redshifts larger than 0.2. Neighbors Searches with radii of 12, 6, and 5 arc-minutes, respectively, were then done on the data to get the ObjectID of each object within the galaxy cluster. The reason three different groups were created is that, though the size of each galaxy cluster can be assumed to be roughly the same, the distance across a cluster in arc-minutes decreases with increasing distance from us.

Using the ObjectID of each cluster in an SQL query, a join was done on the catalog and the SDSS data to find the spectroscopic redshift, as was done in the BCG method. This method inevitably included objects in the Neighbors Search that were not in the galaxy cluster but happened to lie in the same RA and dec of the cluster. To filter out these noncluster members, several steps were taken. First, all objects with a

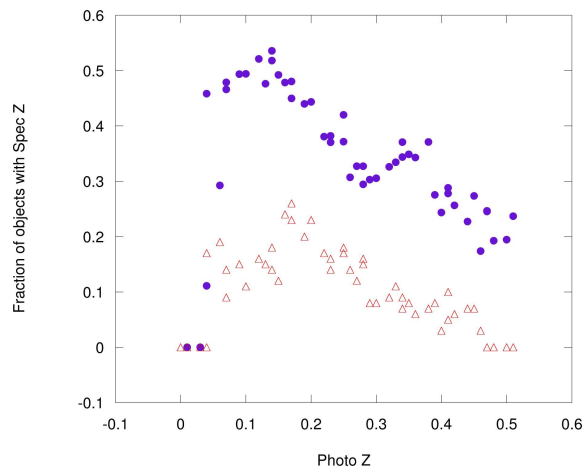
SpecClass of 0 or 1 were omitted. SpecClass is a SDSS notation indicating the spectral class (what type – star, galaxy, etc.) of an object. The classes 0 and 1 are unknown and stars, respectively, and so were eliminated. Second, the residuals – the difference between the photometric redshift and the spectroscopic redshift – of all the objects were found. All objects with residuals greater than 0.1 were eliminated.

After these steps were taken, all objects associated with the same cluster had their spectroscopic redshifts averaged together, so that the spectroscopic redshift for each cluster was an average of the spectroscopic redshift of each galaxy in the cluster. This resulted in approximately 10,000 objects.

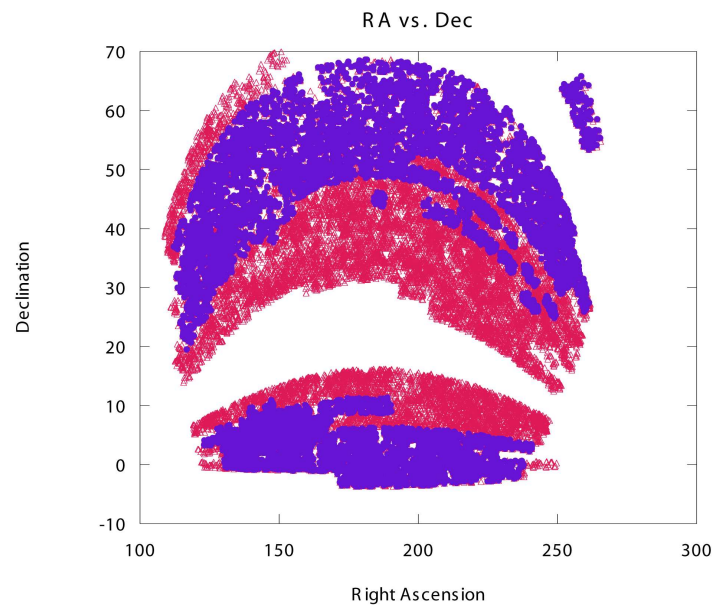
## CHAPTER THREE

### RESULTS

For both the BCG data and the cluster member data, plots were made of the fraction of the original data with spectroscopic redshift in bins of 0.01 photometric redshift. This can be seen in Figure 1. The cluster member data method provided a larger proportion of data with spectroscopic redshift. The reason the amount of data with spectroscopic redshift is never 100% is most likely caused because the original data was gathered from an area of sky where spectroscopic data had not been completely gathered. By plotting RA vs. dec of the original data and RA vs. dec of the data with spectroscopic redshift, one can see that the spectroscopic redshift does not cover the same area. This was done with the cluster member data and can be seen in Figure 2.



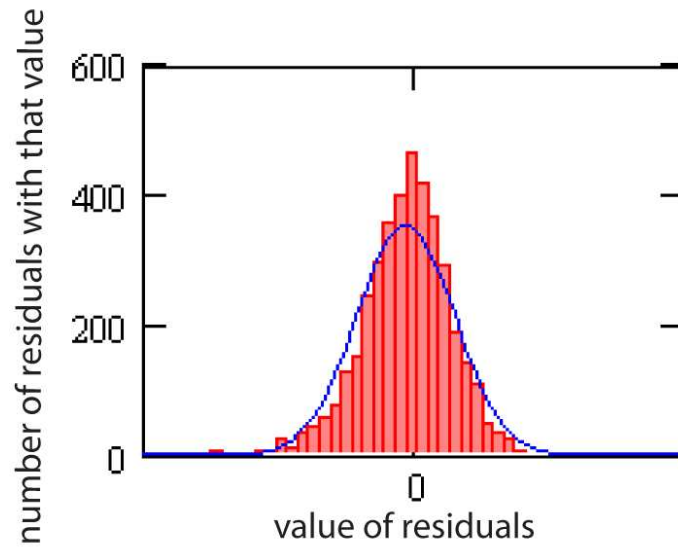
**Figure 1** Fraction of objects with spectroscopic data vs. average photometric redshift. Red triangles are from data obtained using the BCG method, and blue circles are from data obtained using the cluster members method.



**Figure 2** Declination vs. right ascension. This shows the area of the sky covered by the (red dots) MaxBCG data and the area of the sky with (blue dots) spectroscopic data.

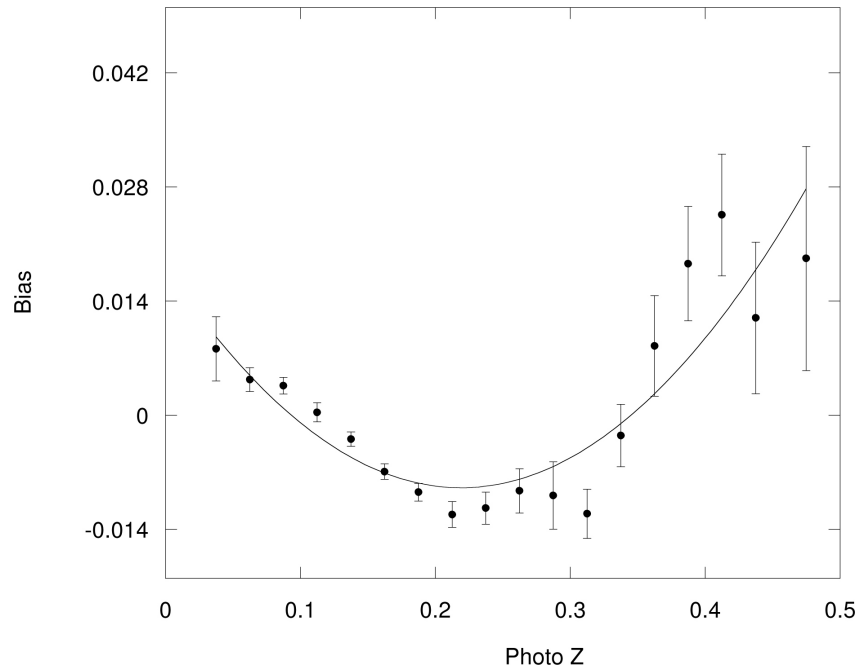
The residuals of the BCG data were then found and grouped in bins of 0.025 photometric redshift. A histogram was done of each bin. The histograms show that the residuals have a Gaussian distribution. The bias and the dispersion of the Gaussian that best fit the data was found. The Gaussian distribution of the residuals was important because then characteristics of the Gaussian curve that fit the histogram, such as bias and dispersion, can be used to analyze the data. An example of one of the histograms can be seen in Figure 3.





**Figure 3** Histogram of the residuals. A bin of  $0.175 < \text{photometric redshift} < 0.2$ , compared to a Gaussian distribution.

By plotting the bias of the bins of residuals as a function of the average photometric redshift of each bin, a relationship between bias and redshift could be found. This plot can be seen in Figure 4. The error bars were found for each measurement of the bias, with a confidence limit of 90%. A second-order polynomial function was fit to the data, with a best fit of  $y = 0.56x^2 - 0.246x + 0.018$ .



**Figure 4** Bias of residuals vs. average photometric redshift. The function that best fits this plot is

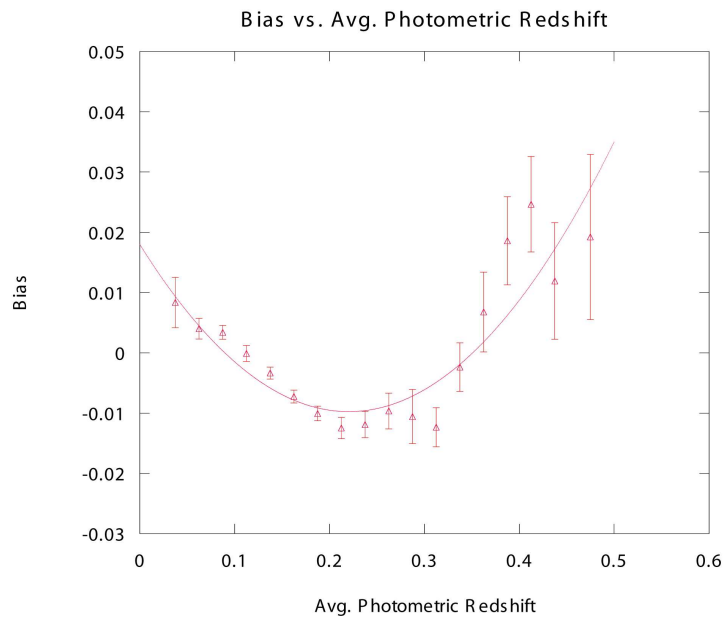
$$y = 0.56x^2 - 0.246x + 0.018.$$

To determine if the equation was a good fit to the data, the  $\chi^2$  of the fit was then found.  $\chi^2$  is the square of the deviation of a sample from its population mean divided by the population variance, where the population has a Gaussian distribution. In this example, the  $\chi^2$  was computed using the formula  $\chi^2 = (1/N)\sum\{[(0.56x_i^2 - 0.246x_i + 0.018) - y_i]/\sigma_i^2\}$ , where  $N$  is the number of data points in the sample,  $x_i$  and  $y_i$  are the  $i$ th points, and  $\sigma_i$  is the dispersion at that point.

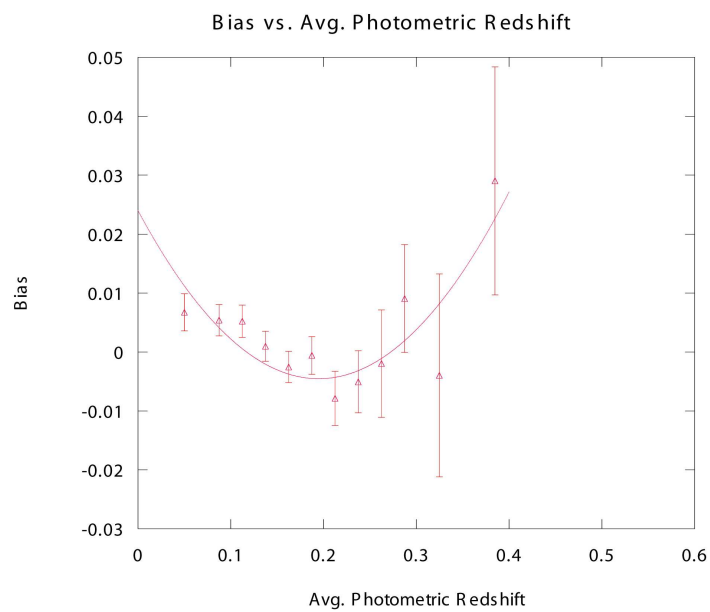
The  $\chi^2$  was used by comparing its value to the corresponding limiting  $\chi^2$  value on a  $\chi^2$  table. The limiting value is based on the number of degrees of freedom of a  $\chi^2$  (which in this case was 17) and is different depending on what significance level is chosen (I chose a significance level of 90%). If a  $\chi^2$  is larger than the limiting value, then that indicates that there is a strong correlation between the data. Because the  $\chi^2$  for the fit was 0.717, and the limiting value is 10.085, the fit was considered a good one.

Using the equation from above, the bias was calculated for every object by entering in each object's photometric redshift. Once the bias was known for each redshift, it could then be subtracted off, moving the value of the photometric redshift closer to the value of the spectroscopic redshift.

Before making this correction, however, it was necessary to make sure this plot will look the same when at both low and high mass. The low mass group was made up of all clusters with  $10 \leq N_{\text{gals}} < 20$ . The high mass group was made up of all clusters with  $N_{\text{gals}} \geq 20$ . When fitting the data from the plot of bias vs. photometric redshift for  $10 \leq N_{\text{gals}} < 20$  (Figure 5) to the equation fitting the data in Figure 2, the  $\chi^2$  is 4.99. When fitting the data from the plot of bias vs. photometric redshift for  $N_{\text{gals}} \geq 20$  (Figure 6) to the equation fitting the data in Figure 4, the  $\chi^2$  is 0.72. These  $\chi^2$  values show that the equation fit to the data for all  $N_{\text{gals}} \geq 10$  fits the data in the separate plots of high and low  $N_{\text{gals}}$ . It can also be seen by eye that the plots are the same as the plot made for all  $N_{\text{gals}}$ . Therefore, this equation can be considered a good fit and can be used.



**Figure 5** Bias vs. average photometric redshift for  $10 \leq N_{\text{gals}} < 20$ . The quadratic equation that best fits this data is  $y = 0.572x^2 - 0.252x + 0.018$ .

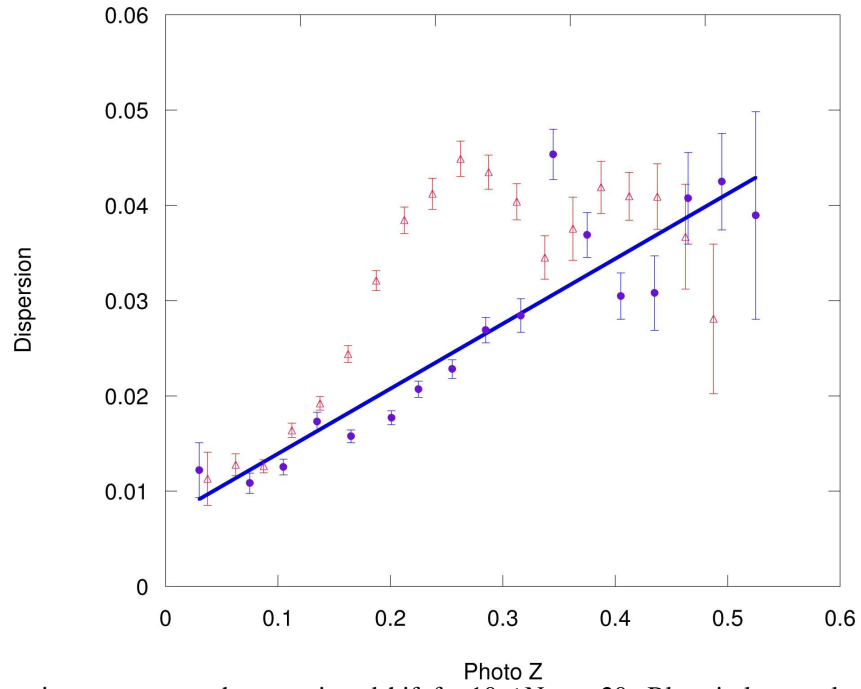


**Figure 6** Bias vs. average photometric redshift for  $N_{\text{gals}} \geq 20$ . The quadratic equation that best fits this data is  $y = 0.752x^2 - 0.293x + 0.024$ .

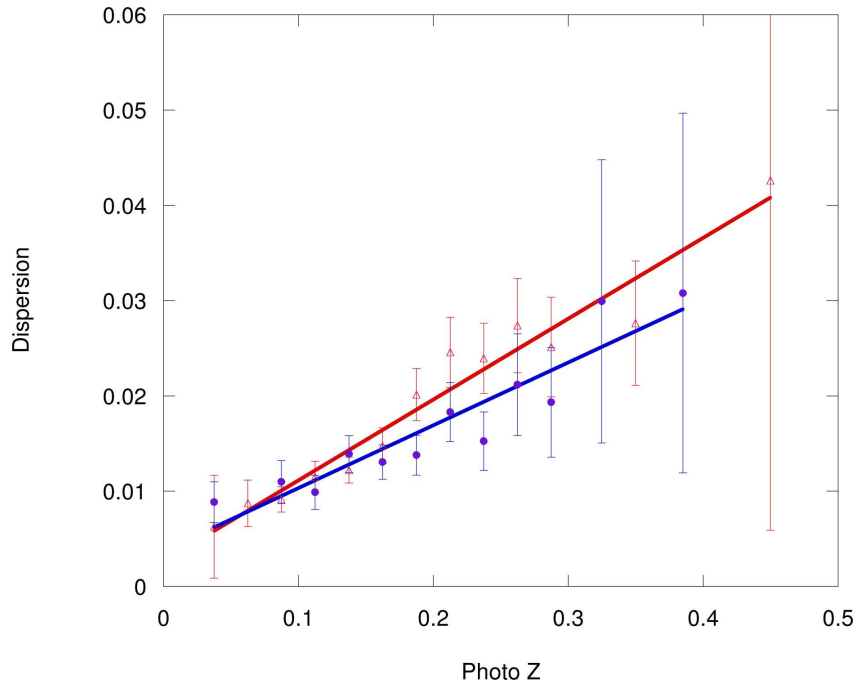
I examined the change produced by correcting the bias of the photometric redshift. The sum of the absolute value of the residuals, divided by the number of residuals, was calculated for both the original and corrected residuals. The average from the original residuals was 0.01753. The average from the corrected residuals was 0.017069. I saw that this was not a big effect, so I did not remove the bias from the data gathered using the cluster members technique.

I then examined the relationship between dispersion and photometric redshift. For both the BCG data and the cluster members data, the data was sorted into two groups: low mass and high mass. The low mass group had  $10 \leq N_{\text{gals}} < 20$ , and the high mass group had  $N_{\text{gals}} \geq 20$ . For each group, the residuals were again put in bins of 0.025 photometric redshift, and the dispersion of each bin was found from the Gaussian best fit to a histogram of the data. For each group, a plot of dispersion vs. average photometric redshift of each bin was done, with error bars with a confidence limit of 90%. The low mass plot can be seen in Figure 7, and the high mass plot can be seen in Figure 8.

A linear relationship between dispersion and photometric redshift can be seen in all plots, except for the cluster members data in the low mass plot (Figure 7). This is perhaps because the radius used in the Neighbors Search was too large for clusters with low  $N_{\text{gals}}$  – since the smaller  $N_{\text{gals}}$ , the smaller the cluster size. Because of this, more things that were not cluster members were included into the data. Because of the isotropy of the universe, all galaxy clusters anywhere in the sky should have the same amount of background objects included. The plateau that is seen in Figure 7 is a result of the inclusion of these noncluster members.



**Figure 7** Dispersion vs. average photometric redshift for  $10 \leq N_{\text{gals}} < 20$ . Blue circles are cluster members technique, red triangles are BCG technique. Best fit line to the BCG technique is  $y = 0.0818x + 0.00713$ .

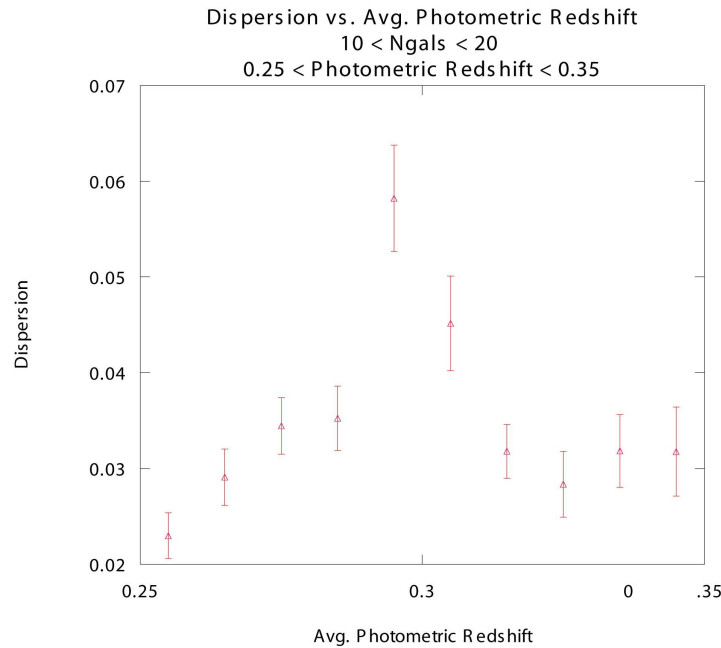


**Figure 8** Dispersion vs. average photometric redshift for  $N_{\text{gals}} \geq 20$ . Red triangles are cluster members data, blue circles are BCG data. Best fit line for BCG data is  $y = 0.0659x + 0.00372$ . Best fit line for cluster members data is  $y = 0.0798x + 0.0034$ .

The linear relationship that occurs overall is somewhat of a surprise, as it was originally thought that the dispersion would remain constant for all redshifts. The increase in dispersion as redshift increases is perhaps caused by the fact that as galaxy clusters get farther away, it is harder to determine what objects are members of the cluster and what objects are not. However, a dispersion of less than 0.1 is considered good for redshifts (Brodwin et al. 1999), and even at a redshift of 0.5, the dispersion is still only 0.04.

In Figure 7, there is a small peak at approximately redshift of 0.3 in the data from the BCG technique. To examine the peak more closely, the data from photometric redshift of 0.25 – 0.35 was broken into smaller bins, and the dispersion of each bin was

plotted against the median photometric redshift of each bin. This can be seen in Figure 9, where it is clear that the peak occurs directly before and after photometric redshift of 0.3.



**Figure 9** Zoomed-in look at peak near photometric redshift = 0.3. Shown in Figure 4.

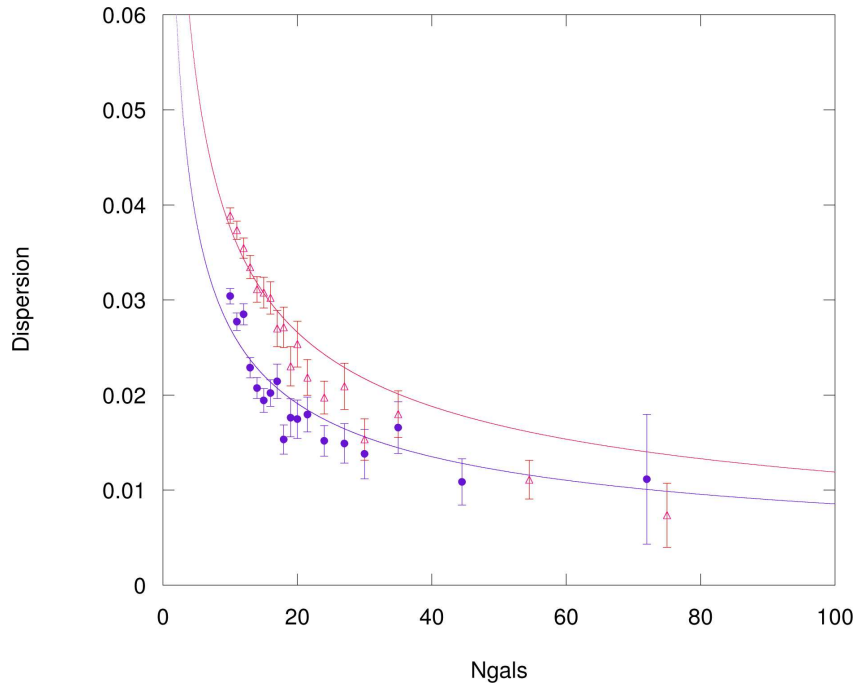
This peak is most likely caused by the accidental inclusion in the data of a couple of objects that are not the BCGs of galaxy clusters but are actually quasars or stars or some other sort of object. For example, at photometric redshift 0.294, there is a quasar with spectroscopic redshift of 0.795, which gives that object a residual much larger than that of any other normal object. There are a couple of objects like this sprinkled throughout the middle of the catalog (0.1-0.2), but they have a larger effect when they are located by 0.3 because there are far fewer objects at that redshift, so there is fewer to balance the mistakes out. The accidental inclusion of points like this is what spurred the data cuts used in the averaging technique described above.



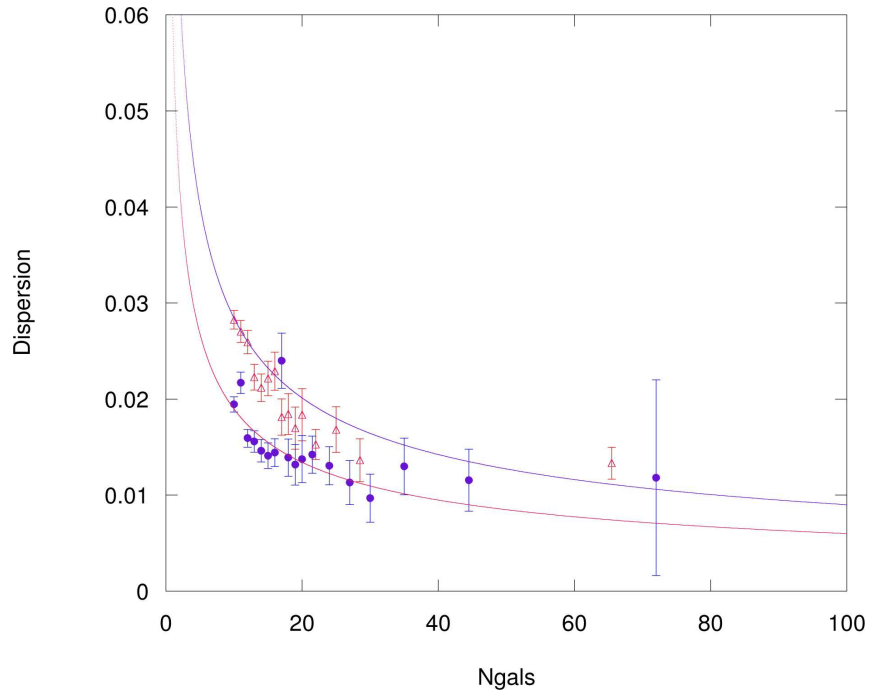
Now I looked at the relationship between the number of galaxies in a cluster and the accuracy of the photometric redshift of that cluster. The residuals for both groups of data were then rearranged into bins of  $N_{\text{gals}}$ . The dispersion was found from the histograms of these bins, and a plot of dispersion vs.  $N_{\text{gals}}$  was made, using the average  $N_{\text{gals}}$  in each bin.

The data showed that dispersion as a function of  $N_{\text{gals}}$  went approximately like  $C/\sqrt{N_{\text{gals}}}$ . Plots of dispersion vs.  $N_{\text{gals}}$  were made for both all photometric redshifts and photometric redshifts in the range 0.1 to 0.2, where I expected the photometric redshift to be very well behaved. Figure 10 shows both sets of data for all photometric redshifts with the exponent of the best fit line constrained to  $-0.5$ . The equation that best fits the BCG data is  $y = 0.0855x^{-0.5}$ , with a  $\chi^2$  of 1.003, and the equation that best fits the cluster member data is  $y = 0.119x^{-0.5}$  with a  $\chi^2$  of 3.51. Figure 11 shows the data for  $0.1 <$  photometric redshift  $< 0.2$ . The equation that best fits the BCG data in Figure 11 is  $y = .06x^{-0.5}$ , with a  $\chi^2$  of 7.62844, and the equation that best fits the cluster member data is  $y = 0.09x^{-.5}$ , with a  $\chi^2$  of 0.497.

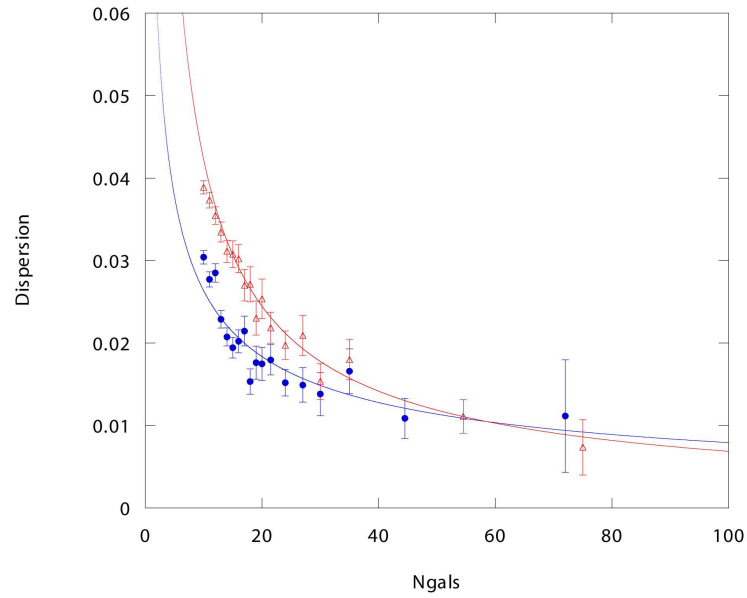
When the exponent is not constrained to  $-0.5$ , the best fit equation for the BCG data is  $y = 0.0876x^{-.5217}$  and the  $\chi^2$  is 0.93721, and the best fit equation for the cluster member data is  $y = 0.2604x^{-.7894}$  and the  $\chi^2$  is 0.26121, shown in Figure 12. For the segment of data in the range  $0.1 <$  photometric redshift  $< 0.2$ , the equation for the best fit line of the BCG data is  $y = 0.0357x^{-0.3026}$  and the  $\chi^2$  is 0.559, and the equation for the best fit line of the cluster member data is  $y = 0.0723x^{-0.4524}$  and the  $\chi^2$  is 0.48669, shown in Figure 13.



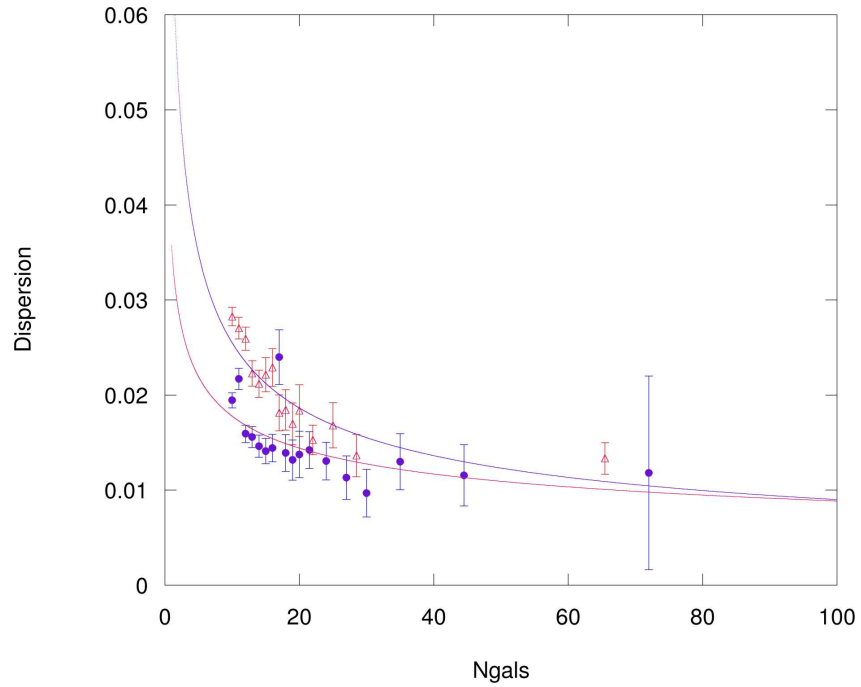
**Figure 10** Dispersion vs.  $N_{\text{gals}}$  for all photometric redshift. Red triangles are cluster members data, blue circles are BCG data. The equation that best fits the BCG data is  $y = 0.0855x^{-0.5}$  and the equation that best fits the cluster members data is  $y = 0.119x^{-0.5}$ .



**Figure 11** Dispersion vs.  $N_{\text{gals}}$ ,  $0.1 < \text{photometric redshift} < 0.2$ . Red triangles are cluster member data, blue circles are BCG data. The equation that best fits the BCG data in Figure 9 is  $y = .06x^{-0.5}$ , and the equation that best fits the cluster member data is  $y = 0.09x^{-0.5}$ .



**Figure 12** Dispersion vs.  $N_{\text{gals}}$  with best fit lines exponent not constrained to  $-0.5$  for all photometric redshift. Blue circles are cluster members data, red triangles are BCG data. The best fit equation for the BCG data is  $y = 0.0876x^{-0.5217}$  and the best fit equation for the cluster members data is  $y = 0.2604x^{-0.7894}$ .



**Figure 13** Dispersion vs.  $N_{\text{gals}}$  with best fit lines exponent not constrained to  $-0.5$ , for  $0.1 < \text{photometric redshift} < 0.2$ . Red triangles are cluster member data, blue circles are BCG data. The equation for the best fit line of the BCG data is  $y = 0.0357x^{-0.3026}$  and the equation for the best fit line of the cluster member data is  $y = 0.0723x^{-0.4524}$ .

## CHAPTER FOUR

### CONCLUSION

I found that no more than 50% of galaxies used in the MaxBCG catalog had spectroscopic data associated with them.

I found that the relationship between the bias of the residuals and the photometric redshift of the clusters is roughly quadratic, but that it does not have a large effect on the average size of the residuals. Therefore correcting for the bias is not necessary.

I found that of the two methods to determine spectroscopic redshift, the BCG technique yields smaller differences between the spectroscopic redshift and the photometric redshift.

The coefficient of the dispersion/ $N_{\text{gals}}$  relationship is small (less than or equal to 0.01). The relationship between dispersion and  $N_{\text{gals}}$  goes approximately like  $C/\sqrt{N_{\text{gals}}}$ .

The increase in errors (dispersion) of photometric redshift goes approximately like  $y = 0.0818x + 0.00713$  for photometric redshifts  $\leq 0.5$  for  $10 \leq N_{\text{gals}} < 20$  and like  $y = 0.0659x + 0.00372$  for  $N_{\text{gals}} \geq 20$  when using the BCG technique to get the spectroscopic redshift. Using that equation, one expects that at a photometric redshift of 0.5, the dispersion is approximately 0.04, which corresponds with the data.

These results show that, if calibration is not a limiting factor, for a Coma mass cluster (a large galaxy cluster, with perhaps  $N_{\text{gals}} \approx 75$ ) the limit of the photometric

redshift of the cluster is about 0.0035, which is remarkably small. A redshift of 0.0035 is about 1000 km/sec, which is on the scale of the redshifts of individual galaxies within a cluster.

## REFERENCES

1. Bahcall, N. A., Hao, L., Bode, P., & Dong, F. 2004, *ApJ*, 603, 1
2. Bahcall, N. A. et al. 2003, *ApjS*, 148, 243
3. Blanton, M.R., Lin, H., Lupton, R. H., Maley, F. M., Young, N., Zehavi, L., & Loveday, J. 2003, *AJ*, 125, 2276
4. Brodwin, M., Lilly, S., Crampton, D. 1999, in *ASP Conf. Ser. 191: Photometric Redshifts and the Detection of High-Redshift Galaxies*, ed. R. Weymann, L. Storrie-Lombardi, M. Sawicki, & R. Brunner, 105 +
5. Eisenstein, D. J. et al. 2001 *AJ*, 122, 2267
6. Fukugita, M. Ichikawa, T. Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, *AJ*, 111, 1748
7. Gunn, J. E. et al. 1998, *AJ*, 116, 3040
8. Haiman, Z., Mohr, J. J., & Holder, G. P. 2001, *ApJ*, 553, 545
9. Hogg, D. W., Finkbeiner, D. P., Schlegel, D. J., & Gunn, J. E. 2001, *AJ*, 122, 2129
10. Ivezić, Z. et al. 2004, *Astronomische Nachrichten*, 325, 583
11. Lupton, R., Gunn, J. E., Ivezić, Z., Knapp, G. R., Kent, S., & Yasuda, N. 2001, in *ASP Conf. Ser. 238: Astronomical Data Analysis Software and Systems X*, ed. F. R. Hardned, F. A. Primiini & H.E. Payne, 269 +
12. Oegerle, W. R., & Hill, J. M. 2001, *AJ*, 122, 2858
13. Pier, J. R., Munn, J. A., Hindsley, R. B. Hennessy, G. S., Kent, S. M., Lupton, R. H., & Ivezić, Z. 2003, *AJ*, 125, 1559
14. Richards, G. T. et al. 2002, *AJ*, 123, 2945
15. Smith, J. A. et al. 2002, *AJ*, 123, 2121
16. Stoughton, C. et al. 2002, *AJ*, 123, 485

17. Strauss, M. A. et al. 2002, AJ, 124, 1810
18. Weller, J., & Albrecht, A. 2001, Physical Review Letters, 86, 1939
19. York, D. G. et al. 2000, AJ 120, 1579