FERMILAB-PUB-09-339-A

arXiv:0907.4383v1 [astro-ph.CO] 24 Jul 2009

Precision Measurements of the Cluster Red Sequence using an Error Corrected Gaussian Mixture Model

Jiangang Hao[1,2], Benjamin P. Koester[3], Timothy A. Mckay [2,4], Eli S. Rykoff[5], Eduardo Rozo[6], August Evrard[2,4], James Annis[1], Matthew Becker[7], Michael Busha[8], David Gerdes[2], David E. Johnston[9], Erin Sheldon[10], Risa H. Wechsler[8]

ABSTRACT

The red sequence is an important feature of galaxy clusters and plays a crucial role in optical cluster detection. Measurement of the slope and scatter of the red sequence are affected both by selection of red sequence galaxies and measurement errors. In this paper, we describe a new error corrected Gaussian Mixture Model for red sequence galaxy identification. Using this technique, we can remove the effects of measurement error and extract unbiased information about the intrinsic properties of the red sequence. We use this method to select red sequence galaxies in each of the 13,823 clusters in the maxBCG catalog, and measure the red sequence ridgeline location and scatter of each. These measurements provide precise constraints on the variation of the average red galaxy populations in the observed frame with redshift. We find that the scatter of the red sequence ridgeline increases mildly with redshift, and that the slope decreases with redshift. We

[1]Center for Particle Astrophysics, Fermi National Accelerator Laboratory, Batavia, IL 60510

[2]Department of Physics, University of Michigan, Ann Arbor, MI 48109

[3]Department of Astronomy and Astrophysics, The University of Chicago, Chicago, IL60637

[4]Department of Astronomy, University of Michigan, Ann Arbor, MI48109

[5]TABASGO Fellow, Physics Department, University of California at Santa Barbara, 2233B Broida Hall, Santa Barbara, CA 93106

[6]Center for Cosmology and Astro-Particle Physics (CCAPP), The Ohio State University, Columbus, OH 43210

[7]Department of Physics, The University of Chicago, Chicago, IL 60637

[8]Kavli Institute for Particle Astrophysics & Cosmology, Physics Department, and Stanford Linear Accelerator Center, Stanford University, Stanford, CA 94305

[9]Department Physics & Astronomy, Northwestern University, Evanston, IL 60208

[10]Center for Cosmology and Particle Physics, Physics Department, New York University, New York, NY 10003

also observe that the slope does not strongly depend on cluster richness. Using similar methods, we show that this behavior is mirrored in a spectroscopic sample of field galaxies, further emphasizing that ridgeline properties are independent of environment. These precise measurements serve as an important observational check on simulations and mock galaxy catalogs. The observed trends in the slope and scatter of the red sequence ridgeline with redshift are clues to possible intrinsic evolution of the cluster red-sequence itself. Most importantly, the methods presented in this work lay the groundwork for further improvements in optically-based cluster cosmology.

Subject headings: Galaxies: clusters - Cosmology: observations - Methods: Data analysis, Gaussian Mixture, Bootstrap

1. Introduction

Galaxy clusters are the largest gravitationally bound systems in our Universe, whose masses, abundance and spatial distribution reflect the growth of structure, composition, and expansion history of the Universe (Evrard 1989; Oukbir & Blanchard 1992). The feasibility of constraining cosmological parameters using galaxy clusters has been demonstrated by many authors (Majumdar & Mohr 2004; Hu 2003; Lima & Hu 2004, 2005) and realistic constraints on cosmological parameters from optically selected galaxy clusters have been implemented recently by Gladders et al. (2007) and Rozo et al. (2009) on the RCS cluster catalog (Gladders & Yee 2005) and maxBCG catalog (Koester et al. 2007a,b) respectively.

The predominantly red, bright, passively evolving red sequence, or "E/S0 ridgeline" galaxies (Visvanathan & Sandage 1977; Annis et al. 1999) found in the cores of clusters of varied richness up to at least $z \sim 1.4$ (Bower et al. 1992; Smail et al. 1998; van Dokkum et al. 1998; Barrientos 1999; Blakeslee et al. 2003; Mullis et al. 2005; Eisenhardt et al. 2005; De Lucia et al. 2007) provide an efficient means for cluster detection, and have been an integral part of modern cluster cosmology. The red sequence itself is ubiquitous in the galaxy population (Renzini 2006, e.g.), and in clusters red sequence galaxies dominate the bright end of the cluster luminosity function (Sandage et al. 1985; Barger et al. 1998). They are extremely tightly clustered in color space, containing old populations of stars whose observed color varies smoothly with redshift (e.g. Gladders & Yee 2000). The pervasiveness of this phenomenon in clusters enables efficient optical cluster detection while a fortuitous color-redshift relation yields accurate photometric redshifts (Gladders & Yee 2000; Koester et al. 2007b). Simple counting of photometrically identified cluster red sequence galaxies (Koester et al. 2007a, e.g) has also been shown to be an effective proxy for cluster mass (Becker et al. 2007; Sheldon et al.

2007; Johnston et al. 2007), with more sophisticated applications yielding improvements in richness as a cluster mass proxy (Rozo et al. 2008). In the era of precision measurements, the extent to which the red sequence can be exploited for cluster cosmology depends on how accurately its characteristics can be measured at a given redshift.

In addition to its relevance to cluster cosmology, the red sequence plays an important role in constraining the complex physical processes that drive galaxy formation and evolution. This includes measurements of the red galaxy luminosity function (Wake et al. 2006; Faber et al. 2007), the clustering of red galaxies in various environments (Zehavi et al. 2005; Coil et al. 2008), and color-magnitude relations of spectroscopically and morphologically identified early-type galaxies (Cool et al. 2006; Mei et al. 2006; Stanford et al. 1998). The high density environments of clusters of galaxies are dominated by red sequence galaxies; the red sequence portion of the galaxy populations in the cores of rich clusters to at least $z \sim 1$ form the basis for various monolithic collapse scenarios (e.g. Bower et al. 1992; Blakeslee et al. 2003; Mei et al. 2009). Faber et al. (2007) summarize some of these results to fill out a picture of galaxy formation that includes a mechanism for the formation of the red sequence.

In color-magnitude space, the red sequence is typically characterized by its slope, zero point, and scatter. Many theoretical modeling have been proposed to explain the red sequence (e.g. Arimoto & Yoshii 1987; Kauffmann & Charlot 1998). Various models posit that in the rest frame, the scatter in the red sequence is driven primarily by age effects, its slope is a manifestation of the mass-metallicity relation, and the zero point is set by combination of age and mass-metallicity differences (e.g. Bernardi et al. 2005; De Lucia et al. 2007; Faber et al. 2007).

Studies of the cluster red sequence have been accomplished by simply measuring the photometric color-magnitude relation (e.g. López-Cruz et al. 2004; De Lucia et al. 2007; Stott et al. 2009), supplemented with HST morphological information (e.g Gladders et al. 1998) and spectroscopy at higher redshift (Mei et al. 2006; Stott et al. 2009). Extra morphological and spectroscopic data allow precise separation of E and S0-types from the rest of the galaxy population, as well as refined identification of cluster members (Blakeslee et al. 2003; Mei et al. 2009). The situation also benefits significantly from precise color measurements afforded by deep, CCD-based imaging (e.g van Dokkum et al. 1998). In the literature, the red sequence has been measured with various levels of scrutiny in dozens of individual clusters.

In the past several years, researchers have turned to the considerable resources of the the Sloan Digital Sky Survey (SDSS) and similar wide field surveys to probe red sequence and elliptical galaxies in various environments (Hogg et al. 2004; Bernardi et al. 2005, 2006;

Cool et al. 2006). These studies include both spectroscopically and morphologically identified red galaxies at $z \sim 0.1$ that aim to constrain galaxy evolution scenarios for the cosmologically-relevant luminous red galaxy (LRG) samples that extend to $z \sim 0.6$ (e.g. Cool et al. 2006).

With the maxBCG cluster catalog, we are positioned to use the SDSS to make one of the most statistically robust photometric measurements of the cluster red sequence, using nearly 14,000 clusters between $0.1 \leq z \leq 0.3$. In this paper we focus on the slope and scatter of the red sequence. We clearly show the systematic effects photometric errors have on the measurement of the underlying slope and scatter of the red sequence, and introduce a method for properly handling these effects. This method, based on an error-corrected Gaussian Mixture Model (ECGMM), reliably recovers the properties of the ridgeline by taking measurement errors into account. After presenting the method, we describe its application to measurement of maxBCG clusters. Of particular relevance to cluster cosmology are the observed mean, scatter, and slope of the E/S0 ridgeline for all maxBCG clusters. These results are presented, along with a discussion of observed trends with redshift.

2. Methods

2.1. Intrinsic properties of red sequence ridgeline

The existence of the red sequence ridgeline is an evidence that cluster galaxies are clustered in color space in addition to real space. The emission from early-type galaxies is dominated by old stellar populations, which gives rise to these remarkably similar galaxy colors. In addition, there is a close mapping between galaxy color and redshift for these galaxies as a result of the restframe 4000 Å break in their spectra. For the SDSS filter sets, the 4000 Å break is within the g band as long as the redshift is below 0.35. Therefore, the most informative ridgeline color for the maxBCG catalog is $g - r$.

In the projected vicinity of a detected cluster, there are both cluster member galaxies and field galaxies. Red sequence galaxies form a part of the member population, whose colors are clustered tightly and can be approximated with a Gaussian distribution with narrow width. On the other hand, the field galaxies' and blue member galaxies' colors are not tightly clustered and can be approximated by a Gaussian distribution with a broader width[1]. The problem of separating the ridgeline from the field can be specified as following:

[1]There are complicated situations where the distribution in color space is not simply unimodal or bimodal, for example when two clusters are seen in projection. For maxBCG catalog, it covers about 7,500 square

What are the two Gaussian components (one for the ridgeline and one for the field) that represent the color distribution in the vicinity of a galaxy cluster? If this double Gaussian is an adequate model for describing this color distribution, the one dimensional Gausian Mixture Model (GMM) is well suited to the problem.

In the traditional applications of GMM (Titterington et al. 1985), measurement errors are not considered. In our case, there are non-negligible measurement errors associated with the galaxy colors. We are interested in measuring the intrinsic color scatter of cluster members, absent contamination by the increasing measurement errors of faint galaxies. Without accounting for the increasing photometric errors, we expect that the color scatter will increase as the measurement errors become larger. While the intrinsic color scatter may increase as redshift increases (because the 4000 Å line break is shifting toward r band and making the $g - r$ color less discriminative.), measurement errors may make us overestimate the increase in intrinsic scatter with redshift. To avoid this problem, we include measurement error into our likelihood function to remove the contamination. We will refer to this as the error-corrected Gaussian Mixture Model and derive the corresponding Expectation Maximization (EM) recursive relation in the following section.

It is clear that we can always improve the fit by adding more Gaussian components, although this is clearly not good in the sense of parsimony. So, we need to somehow decide on the number of Gaussian components by trading off quality of fit against the number of introduced free parameters. To do so, we use the Bayesian Information Criterion (BIC) (Schwarz 1978; Connolly et al. 2000) to determine how many mixtures we should use. The BIC is defined as:

$$BIC = -2 \log \mathcal{L}_{max} + k \log(M) \tag{1}$$

Where k is the number of free parameters. For mixture models with different number of components, we compare their corresponding BIC and select the model with the smallest BIC.

---

degrees with about 14,000 clusters. This leads to about 2 clusters per square degree. Each cluster is about the size of a few arcminutes across, so the chance of two or more overlapped clusters is low. Therefore, a unimodal or bimodal distribution in color space is a good approximation.

2.2. Error-corrected Gaussian Mixture Model

In what follows, we describe how to fit a multi-component Gaussian mixture model to a one dimensional distribution of data with both intrinsic scatter and Gaussian measurement errors. Our method is an extension of the traditional expectation maximization method for GMM (Dempster et al. 1977).

We assume the data are to be modeled by a mixture of N Gaussians fit to the distribution of M data points. The subscript i cycles through N and j cycles through M, and we use μ_i, σ_i and w_i to denote the location, width and weight of each Gaussian component. y_j and δ_j denote the data points and their measurement errors which are assumed to be Gaussian. For brevity, we denote the parameters (μ_i, σ_i and w_i) collectively by θ. Then the likelihood of the parameters given the data and measurement errors is:

$$\mathcal{L}(\theta|y) = \prod_{j=1}^{M} \Big\{ \sum_{i=1}^{N} \frac{w_i}{\sqrt{2\pi(\sigma_i^2+\delta_j^2)}} \exp\Big[-\frac{(y_j-\mu_i)^2}{2(\sigma_i^2+\delta_j^2)} \Big] \Big\} \tag{2}$$

The optimal parameters θ could be estimated by maximizing the above likelihood function. The Expectation Maximization algorithm provides an efficient way to get the maximum likelihood estimators in such a setting. To utilize this, we need to introduce a hidden variable, z_j, which tells which Gaussian component the data point y_j is sampled from. In our case, different from the standard EM prescription, we have non-negligible measurement errors present. After some algebra, we arrive at a set of recursive relations that lead to the maximum of the likelihood (see appendix for details).

2.3. Bootstrapping to increase the robustness of ECGMM

Though the ECGMM is generally stable for estimating the parameters of the Gaussian Components, it can fail occasionally due to very inappropriate choice of initial parameters or some very big measurement errors of certain galaxies. To make our measurement more robust, we introduced a bootstrap-like scheme. Suppose we have M data points. We randomly pick one of the data points and record it. We then repeat this process M times and get M recorded data points. These M points form one resampling set of the original data set. Now, we apply the ECGMM to this new data sample and measure the corresponding parameters. After this, we start a second round, getting another resampling set with M data points in it and measure the parameters using ECGMM again. We repeat this process X times and have X estimates of each parameters. We throw away those outlier estimates (estimates beyond

the upper and lower inner fences[2]) for each parameter and use the mean of good estimates as the value of each parameter. Using this scheme, our resulting parameter estimates are much more robust, at a cost of a tolerable increase in computation time. In this application, we took X to be 50.

2.4. Monte Carlo test of the ECGMM for our application

Before we delve into real data, we first conduct Monte Carlo tests to see whether the ECGMM approach can reliably identify the cluster and background Gaussian components. These tests are used to determine whether this method can reliably recover the true parameters input in the simulation, and to see whether the extracted parameters are generally unbiased with respect to measurement errors.

For this purpose, we generate two Gaussian random data sets, one representing cluster member colors, denoted CL, and the other representing the field galaxies/blue galaxies' colors, denoted as BG. The CL set is generated from $\sim N(0.5, 0.04^2)$ and BG set is generated from $\sim N(0, 0.3^2)$. To represent clusters with different richness, we allow the normalization (also denoted as N_{gals} in the plots) of CL data set to vary as 10, 15, 20, 25, 30, 40, 50, 60 and 70 while keep the normalization of BG set as 30. All the parameters used to generate the mock data are chosen to make the simulation as close to the real data as possible. Then we combine CL and BG to create a mock data set that mimics the colors of both cluster members and background galaxies in a field. It is worth noting that these mock colors are error free so far. Next, we will add some noise to them to mimic the measurements errors. To do this, we first generate random numbers from a uniform distribution in the range of [0, 0.1], which play the role of δ_j in Eq.2. Then, we generate from $N(0, \delta_j^2)$ and add them to the noise free data set to produce a noise added mock color data set. In Fig.1, we plots the results from the ECGMM fitting. The results show that for clusters with $N_{gals} \geq 10$, the method can recover the locations (μ) and widths (σ) of the Gaussian components very well.

Next we test for possible bias in the estimators. For each cluster richness N_{gals}, we regenerate the data as well as errors 200 times and then apply our methods to each to obtain estimates for the parameters. In each case, we calculate the bias of parameters θ (the σ and μ in our case) defined as $E(\hat{\theta}) - \theta$. In Fig.2, we plot the results from both GMM and ECGMM for comparison. Clearly, the introduction of error correction(as shown in the bottom two

[2]In statistics, lower inner fence is defined by $Q_1 - 1.5IQR$ and higher inner fence is $Q_3 + 1.5IQR$, where Q_1 and Q_3 are the first and third quartiles respectively. The IQR is the the interquartile range, defined as $Q_3 - Q_1$.

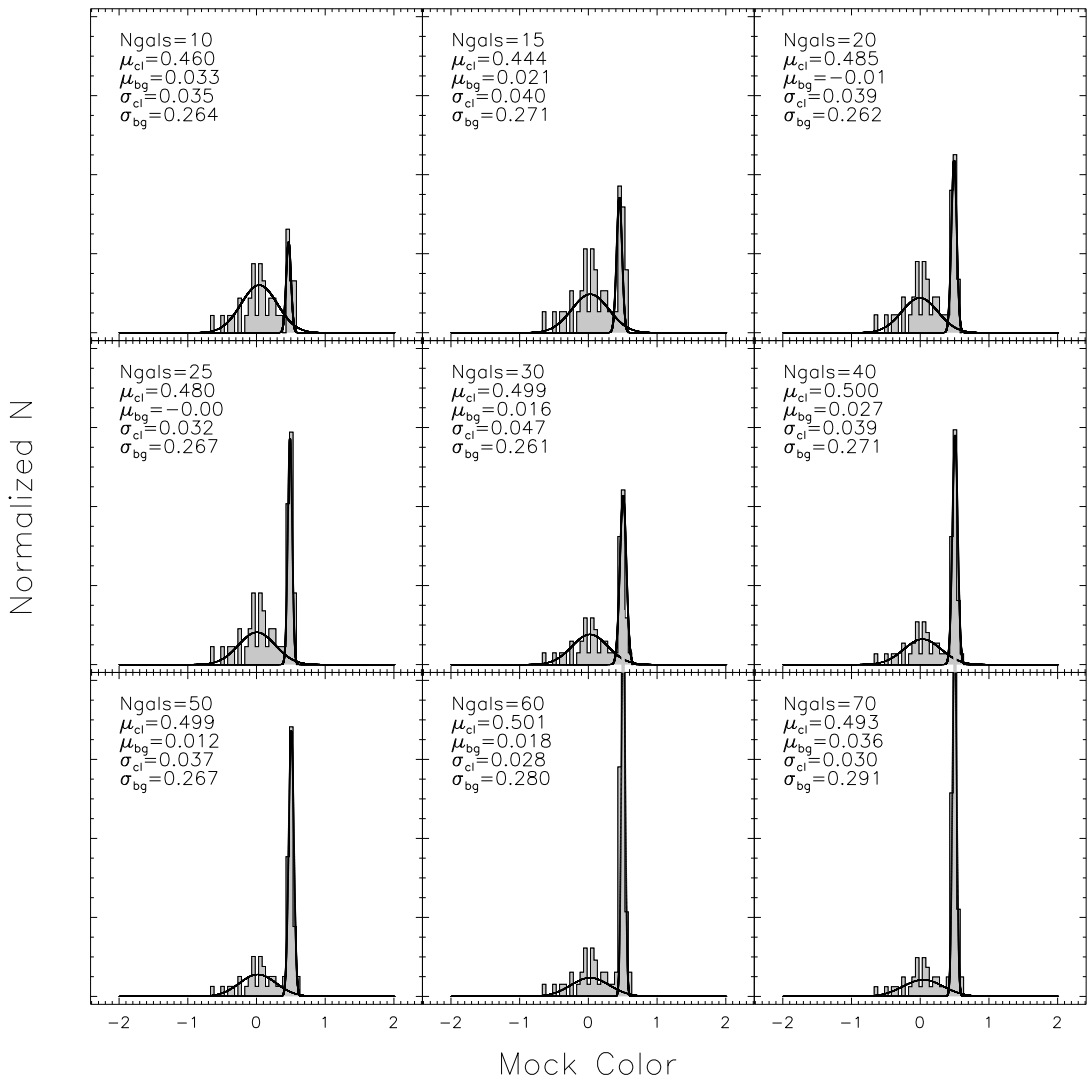

Fig. 1.— Basic Monte Carlo tests of the ECGMM method for fitting mock galaxy $g - r$ color distributions (see text). μ and σ denote the locations and widths of the corresponding Gaussian components, for clusters of increasing richness. The true μ are 0 and 0.5 for BG and CL sets respectively. The true σ are 0.3 and 0.04 for BG and CL sets respectively.

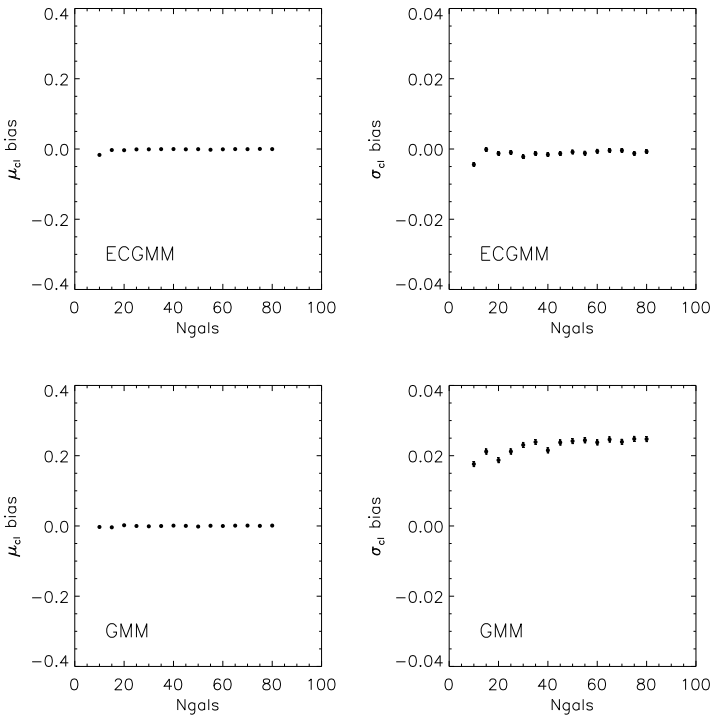

Fig. 2.— Monte Carlo test of the bias, $(E(\hat{\theta}) - \theta)$, of the estimators for the location and width using GMM (bottom two panels) and ECGMM (top two panels) as a function of richness for the cluster component of the mock clusters (see text). The scales of the plots are chosen to be close to the size of the true parameters in the plot to illustrate the fraction precision.

panels) is essential for removal of the bias of the width resulting from measurement error(as shown in the top two panels).

3. Data

3.1. SDSS

Three main resources are included in this work: the SDSS galaxy catalog, the maxBCG catalog, and a value-added SDSS spectroscopic catalog.

The maxBCG cluster sample and the galaxy catalogs used to remeasure cluster richness in this paper are derived from the SDSS (Adelman-McCarthy et al. 2006). The maxBCG cluster sample covers a sky area of about 7500 square degrees. The camera design (Gunn et al.

2006) and drift-scan imaging strategy of the SDSS enable acquisition of nearly simultaneous observations in the u, g, r, i, z filter system (Fukugita et al. 1996). Calibration (Hogg et al. 2001; Smith et al. 2002; Tucker et al. 2006), astrometric (Pier et al. 2003), and photometric (Lupton et al. 2001) pipelines reduce the data into object catalogs containing a host of measured parameters for each object. Galaxies are selected from SDSS object catalogs as described in (Sheldon et al. 2007). In this work we use `CMODEL_COUNTS` as our total magnitudes, and `MODEL_COUNTS` when computing colors. Bright stars, survey edges and regions of poor seeing are masked as previously described (Koester et al. 2007a; Sheldon et al. 2007).

The spectroscopic galaxy catalog is comprised of galaxies from the DR6 of SDSS Value Added Galaxy Catalog. A detailed description about this catalog can be found in Blanton et al. (2005, VAGC).

3.2. Cluster Sample

We obtain sky locations, redshift estimates, and initial richness values from the maxBCG cluster catalog. Details of the selection algorithm and catalog properties are published elsewhere (Koester et al. 2007b,a). In brief, maxBCG selection relies on the observation that the galaxy population of rich clusters is dominated by luminous, red galaxies clustered tightly in color (the E/S0 ridgeline). Since these galaxies have old, passively evolving stellar populations, their $g - r$ color closely reflects their redshift. The brightest such red galaxy, typically located at the peak of the galaxy density, defines the cluster center.

The maxBCG catalog is approximately volume limited in the redshift range $0.1 \leq z \leq 0.3$, with very accurate photometric redshifts ($\delta z \sim 0.01$). Studies of the maxBCG algorithm applied to mock SDSS catalogs indicate that the completeness and purity are very high, above 90% (Koester et al. 2007a). The maxBCG catalog has been used to investigate the scaling of galaxy velocity dispersion with cluster richness (Becker et al. 2007) and to derive constraints on the power spectrum normalization, σ_8, from cluster number counts (Rozo et al. 2009).

4. Measuring the ridgeline location and width of maxBCG clusters

We apply the above prescriptions of ECGMM to the maxBCG cluster catalog and the galaxy catalog (Koester et al. 2007a), measuring the red sequence $g - r$ ridgeline. The procedures are as follows: for each cluster in maxBCG catalog, we choose a scaled aperture R_{200}^{lens} to ensure we are considering equivalent regions of clusters of varied masses and therefore varied richness. R_{200}^{lens} is the critical radius, interior to which the mean mass density

of the cluster is 200 times of the critical energy density. Based on the weak lensing analysis (Johnston et al. 2007; Hansen et al. 2007), the scaling relation between R_{200}^{lens} and the original maxBCG richness N_{200} is given by $R_{200}^{lens} = 0.182(N_{200})^{0.42}$, which ranges from 0.47 Mpc to 1.68 Mpc.

Next, we identify all SDSS galaxies inside this aperture range, fainter than the BCG, and brighter than an i band magnitude corresponding to 0.4 L* at the redshift of the cluster (Koester et al. 2007a). Then, we apply the ECGMM procedure to the $g - r$ colors and corresponding measurement errors of these galaxies. One of the resulting two Gaussian components from the ECGMM will represent the cluster red sequence color distribution while the other represents the background/blue galaxy color distribution. To determine which Gaussian Component belongs to the cluster, we calculate the likelihood of the BCG's $g - r$ color on each Gaussian Component. The component for which the BCG has a higher likelihood is assigned as the cluster component and the other is declared background. In this way, each maxBCG cluster gets a new richness, N_{200}^{lens}. It is worth noting that we apply the above measurements to all maxBCG clusters whose original $N200 \geq 10$. But we will only continue our analysis on a subsample of the clusters whose new measured richness $N_{200}^{lens} \geq 10$ and which have have two identified Gaussian mixture components in order to guarantee the reliability of our measurements. After this selection, we are left with about 7100 clusters and all further analysis are based on them. Clusters falling outside of this selection are not necessarily false clusters, they just fall below the higher richness threshold we imposed for quality control. In Fig.3 and Fig.4, we show the ECGMM fitting of 9 big and 9 small clusters as described above. Their corresponding CMRs are plotted in Fig.5 and Fig.6.

For comparison, we measure the red-sequence location and width using both ordinary GMM and ECGMM. The top panel of Fig. 7 shows the evolution of the average g-r ridgeline location and width measured using ordinary GMM. We observe the well-known trend in the average ridgeline zeropoint, and there is additional apparent strong evolution in the average ridgeline width, which becomes nearly 140% larger by $z = 0.3$. However, from the lower two panels which are measured using ECGMM, one can see very clearly the power of ECGMM in constraining the intrinsic width of the ridgeline without contamination from measurement error. The results show that the mean observed $g-r$ ridgeline location retains the same linear dependence on redshift while the mean scatter of the ridgeline shows a weak dependence on redshift, with the $g - r$ scatter $\sigma(z = 0.1) = 0.051 \pm 0.003$ and $\sigma(z = 0.3) = 0.079 \pm 0.005$ or a broadening by $\sim 55\%$ from $z = 0.1$ to $z = 0.3$. The strong dependence of the scatter on redshift from the GMM is mostly due to the increased measurement errors for cluster members at higher redshift.

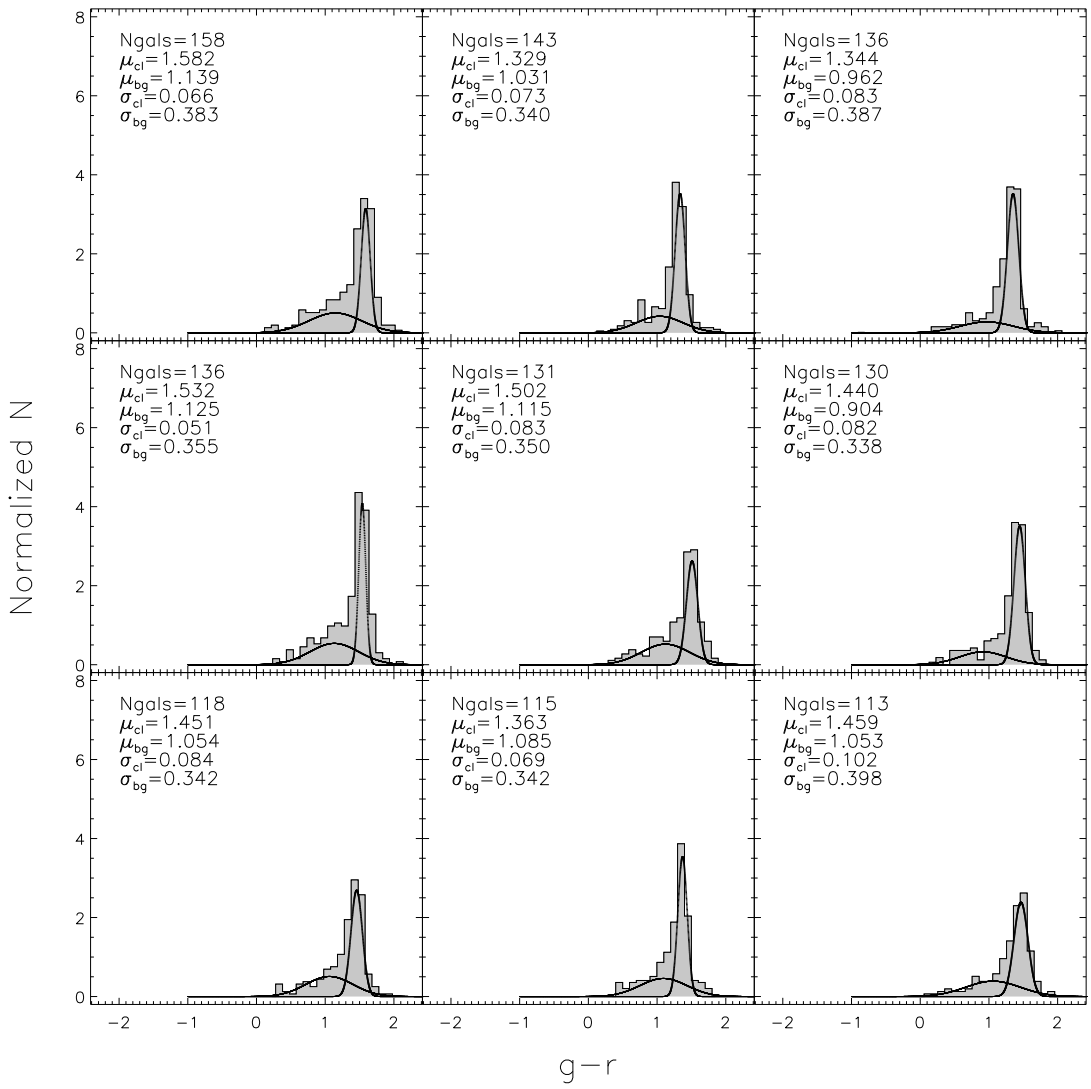

Fig. 3.— The ECGMM fitting to the galaxy color distribution around 9 rich clusters. Note that due to the fact that we corrected for measurement errors, the two Gaussians appear to be narrower than the histogram.

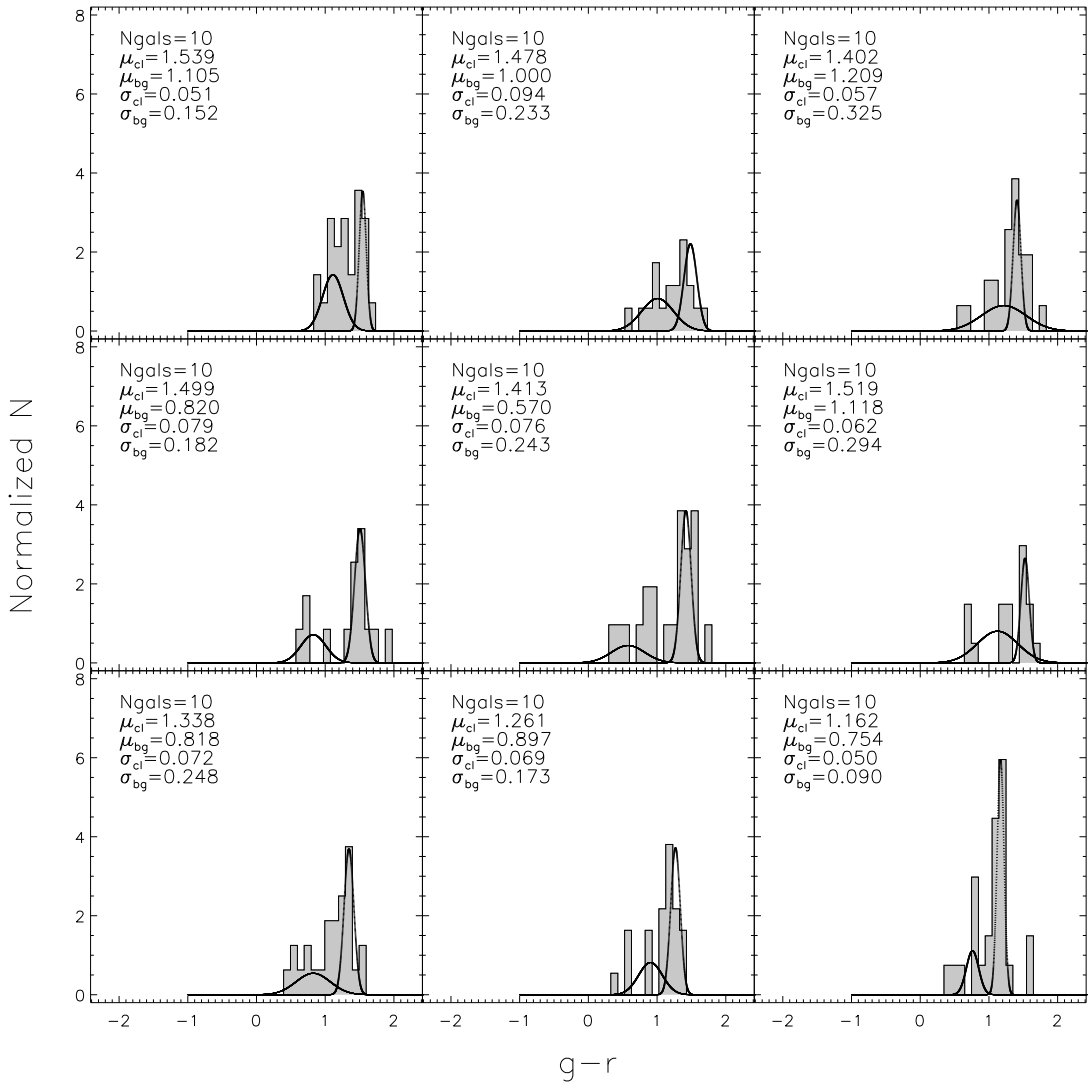

Fig. 4.— The ECGMM fitting to the galaxy color distribution around 9 small clusters. Note that when there are fewer galaxies histogram is no longer a good way to show the distribution.

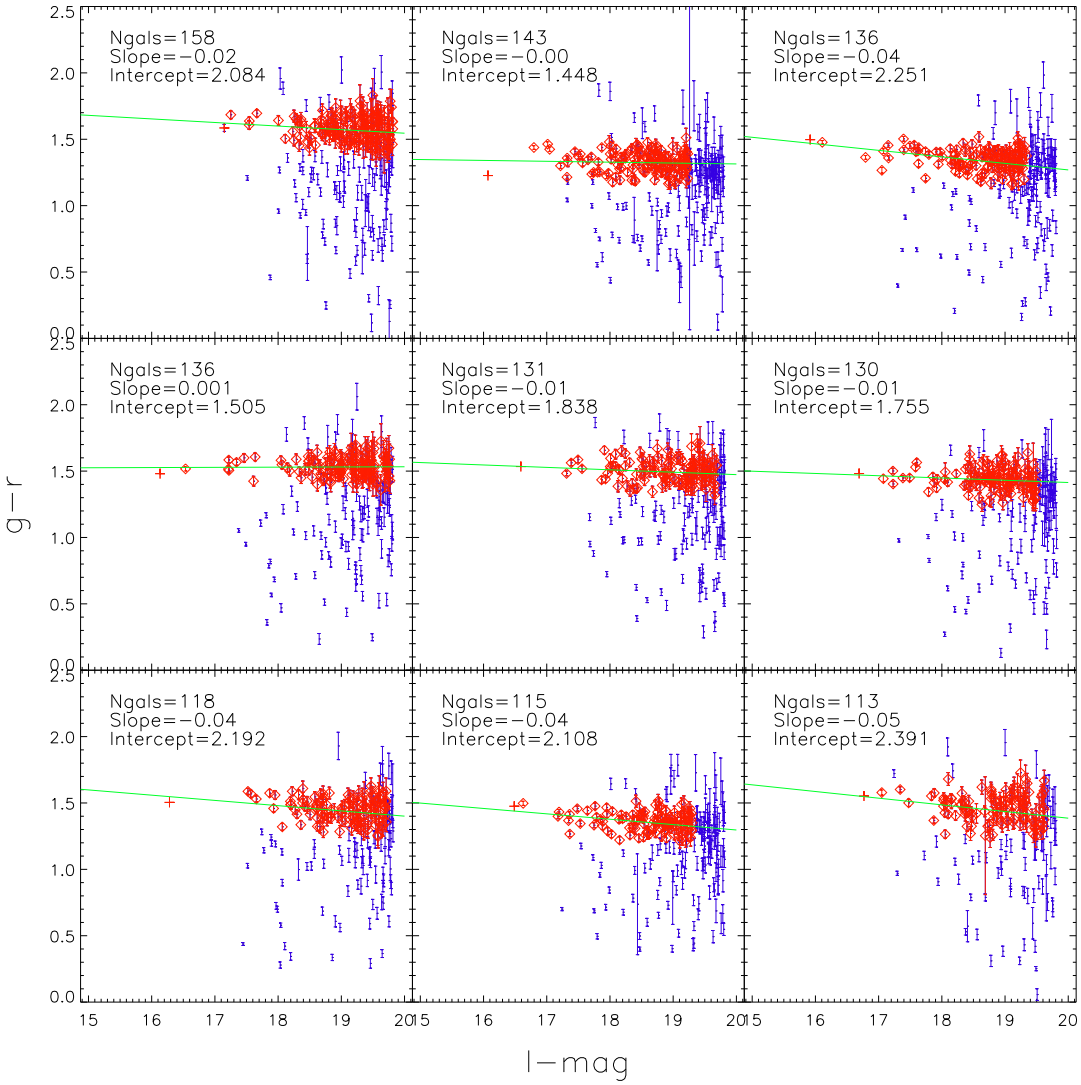

Fig. 5.— The CMR around the 9 rich clusters in Fig.3. The red diamonds and error bars are those from the selected members (red sequence) and the blue dots and error bars are those from field galaxies. The red cross symbol represent the BCG of that cluster. All the galaxies are within R_{200}^{lens} around the BCG, fainter than BCG but brighter than the 0.4L*(see text). The green line is a weighted least square fit (weighted by the inverse square of color errors) to the cluster galaxies.

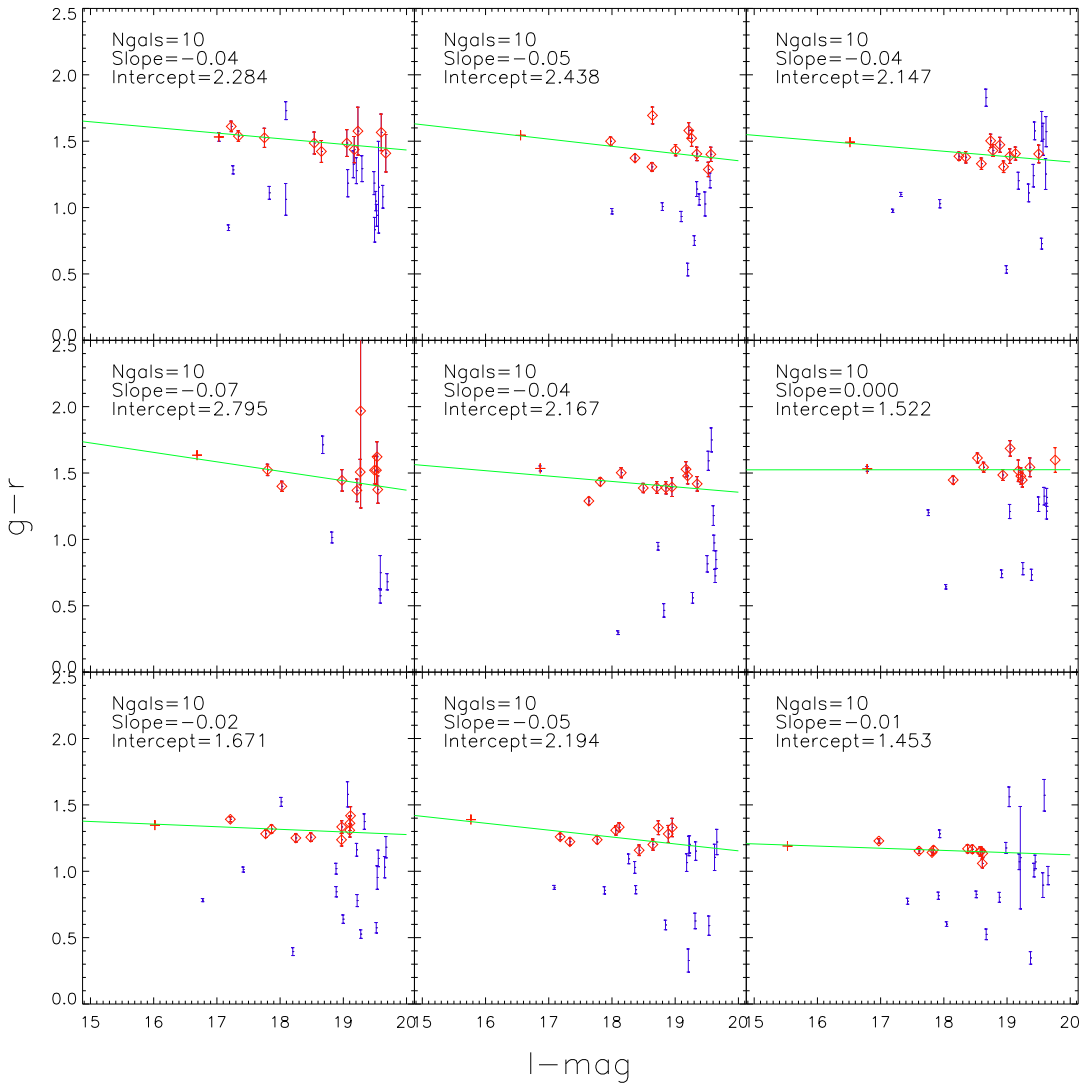

Fig. 6.— The CMR around the 9 small clusters in Fig.4.

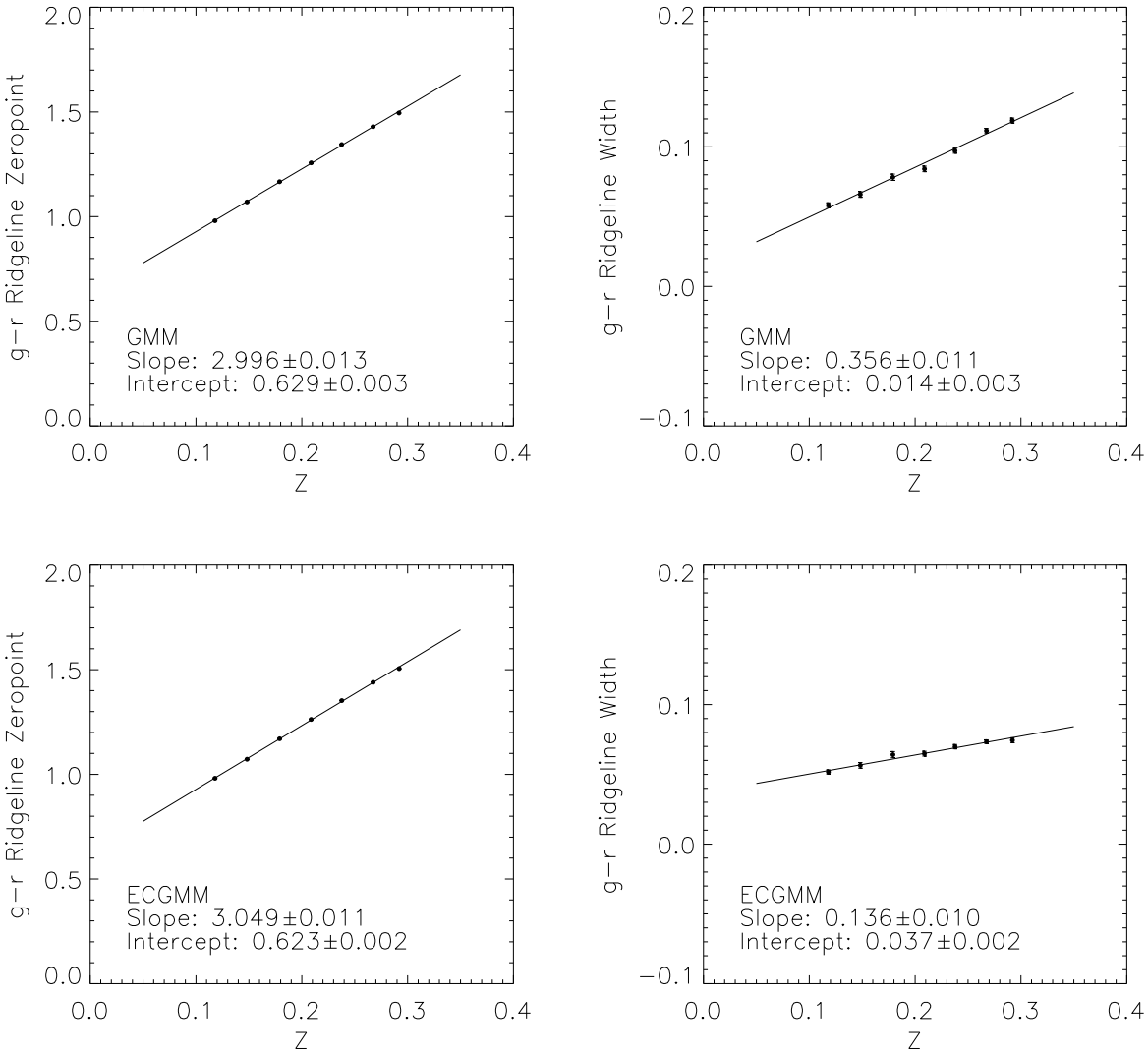

Fig. 7.— Tracking the $(g-r)$ red sequence zeropoint and width as a function of redshift, measured using ordinary GMM (upper panels) and ECGMM (lower panels) respectively. We bin the measured ridgeline color and width into redshift bin of 0.04 and then fit the means with a straight line. After error correction, the broadening of the observed red-sequence width with redshift is greatly suppressed, revealing the effect of photometric errors on the observed broadening.

5. The red sequence ridgeline slope

5.1. Ridgeline slope from galaxy clusters

It has been pointed out that the color-magnitude relation (CMR) of cluster member galaxies has a negative slope (e.g Kodama & Arimoto 1997; Gladders et al. 1998), so that fainter member galaxes are generally bluer. The evolution of these CMR slopes with respect to redshift and richness has been difficult to address, largely due to the lack of a sufficiently large cluster catalog with well measured photometry for all its galaxies. The maxBCG catalog provides about 14,000 galaxy clusters, extending over $0.1 \leq z \leq 0.3$, which enables us to measure the slope of the CMR for clusters with good statistics across a range in both richness and redshift.

Measurement of the slope of the CMR typically proceeds by identification of the cluster red-sequence, followed by some iterative process of outlier removal, and a determination of cluster "member" galaxies which are then used to measure the slope and zeropoint of the CMR. We apply the method described in previous sections to measure the color distribution of individual clusters and to assign the memberships for every cluster by requiring the color difference between the member galaxies and ridgeline within $\pm 2\sigma$ (σ is the convolved ridgeline width, given by the best-fit ECGMM, and the measurement errors of individual member galaxy's color). The richness of the cluster measured by this way is denoted as N_{200}^{lens}. We choose 2σ because this is roughly where the background component's likelihood dominates over the cluster component's likelihood. Based on this identification of membership driven by the ECGMM, we fit for the CMR of clusters galaxies with a straight line using weighted least square fitting. The weights we use are the inverse square of the measurement errors of $g - r$. We call the slope of the fitted line as the ridgeline slope.

The distribution of ridgeline slopes for maxBCG clusters is shown in Fig.8 in bins of $\Delta z = 0.03$. Despite the substantial scatter in slope among individual clusters, we can see from Fig.8 and Fig.9 that the mean slope of the red sequence ridgelines for clusters deviates from zero for $0.1 \leq z \leq 0.3$. For any bin, the error on the mean places the measurement many standard deviations from zero.

In Fig.9, it is apparent that the observed trend of the mean ridgeline slope with redshift is statistically significant: the slope becomes steeper by a factor of 2.5 by $z = 0.3$. In Fig.10, we plot the evolution of the slopes vs richness in each redshift slice, which shows that the dependence of the ridgeline slope on richness is weak, as shown elsewhere (e.g. Hogg et al. 2004). Clearly, the observed slope of the red-sequence is not associated with cluster richness, and is unsurprisingly a strong function of redshift (see Discussion).

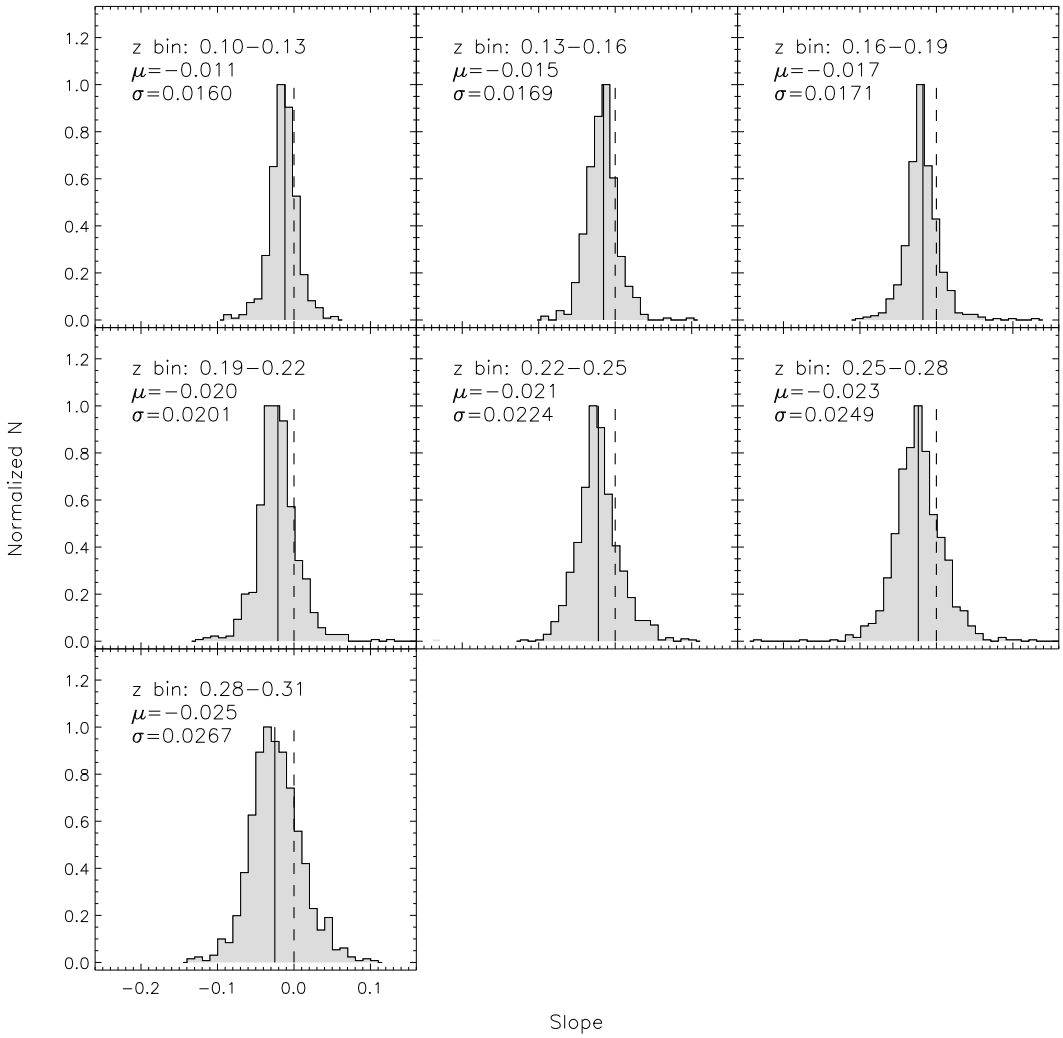

Fig. 8.— The distributions of measured ridgeline slopes for clusters in steps of 0.03 in redshift. μ and σ denote the mean and width of the distribution. The dashed line corresponds to zero.

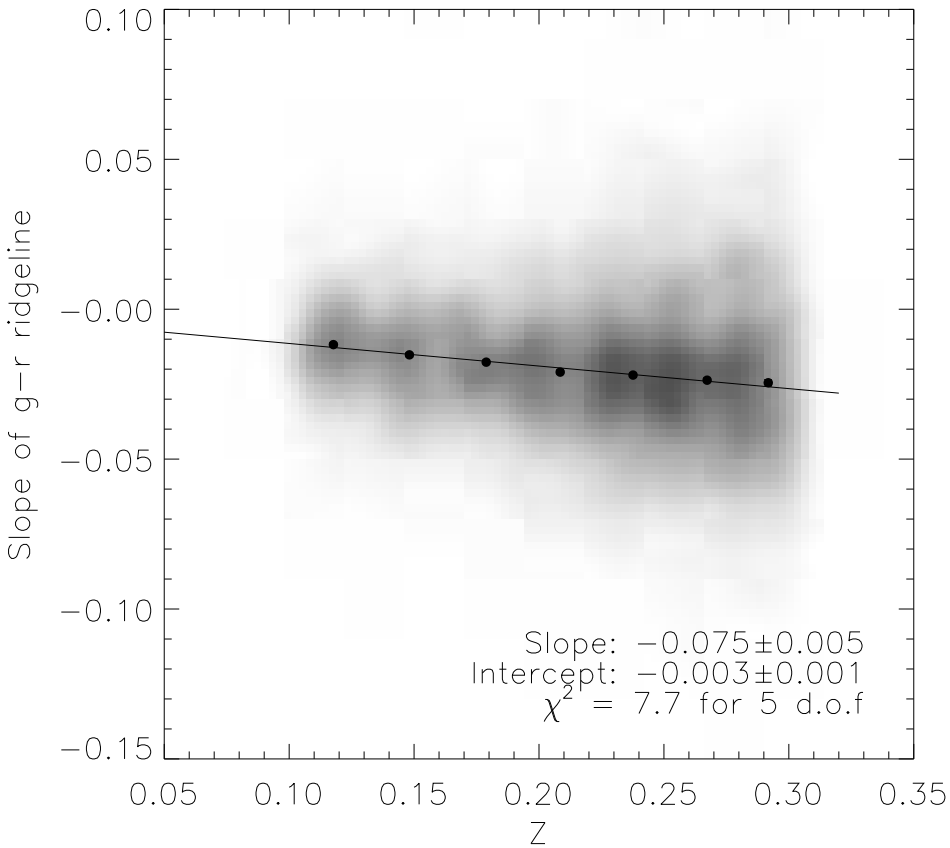

Fig. 9.— Tracking the observed red-sequence slope vs redshift. The gray clouds represent the slope measurements from individual clusters. The black solid circles and error bars are weighted mean and the standard deviation to the weighted mean for each redshift bin ($\Delta = 0.03$). Note that the error bars in the plot are smaller than the symbols.

5.2. Ridgeline slope from spectroscopic data

The above measurement is based only on a photometric determination of red sequence galaxies. The level to which projection plays into this selection is as yet unknown. The true red-sequence galaxy population in some physical volume, either in a cluster or in the field, is contaminated by dusty foreground galaxies which can be rejected via spectroscopy, and by the peculiar velocities of the galaxies themselves.

To address the possibility of foreground contamination, it is interesting to see if the above results are preserved in a spectroscopic sample of galaxies. To achieve this goal, we use galaxies with spectra from DR6 of SDSS Value Added Galaxy Catalog (Blanton et al. 2005, VAGC) for a comparison. Due to the selection effects of the spectroscopic data, we

will choose only the galaxies in redshift from 0.1 to 0.2 and brighter than 0.4 L* magnitude at their respective redshifts. By extension from the photometric sample and from previous work (Hogg et al. 2004), we know that the slope does not vary with environment, so the field sample represented by our spectroscopy should be a fair representation of the expected slope in clusters.

Our procedure is as follows: we first bin the galaxies into bins of size $\Delta z = 0.003$, which corresponds to velocity slices of 900km/sec. The color distribution of the galaxies in each bin shows clear bimodality (top panel of Fig. 11). Then, we separate the red sequence galaxies in each bin using ECGMM. The red sequence galaxies correspond to the Gaussian component with a bigger $g-r$ value and we choose $\pm 2\sigma$ from the peak location as red galaxy samples for each redshift slice, similar to what was done for cluster galaxies. Then, in every bin, we fit the CMR of galaxies' $g-r$ colors and i-band magnitude with a straight line using weighted least square fitting with the weights come from the inverse sequare of the color measurement errors. The weights are the inverse square of the $g-r$ measurement errors. We record the corresponding slopes. In the bottom panel Fig.11, we choose 6 redshift bins ($\Delta z = 0.003$) to illustrate the red/blue galaxy separation and the ridgeline slope fitting in each bin. Finally, we fit the variation of slope with redshift with a line to look for a trend, the results are shown in the left panel of Fig.12. As a comparison to the cluster sample, we also plot the mean variation of the ridgeline slope from clusters in the same redshift range [0.1, 0.2] in the right panel of Fig.12. When the redshift range is changed, the slope of the fitted line for the cluster sample becomes steeper as compared to Fig.9. The reason lies in that the linear fit to the trend is only the first order approximation. But for our purpose here, we just need to require the cluster sample and spectroscopic sample on the same redshift range so as to compare them fairly.

Comparing the two panels in Fig.12 shows that the two slopes from the spectroscopic sample and cluster sample are not different in a statistically significant way. This further confirms our previous observation that the cluster environment will not affect the slope.

6. Discussion

To this point, the measurements have been presented in the observed frame. Photometric cluster detection and the quantities derived (e.g. richness) operate in the frame of the observer, and predictions from galaxy formation models and mock galaxy catalogs can be evaluated in light of these precision measurements. They have particular applicability to calibration of optical cluster detection efforts, especially to those that rely on the properties of the red sequence. The methodologies developed herein allow the "bootstrapping" of op-

tical algorithms: basic cluster finders locate the clusters, and precision measurements (such as these) of said clusters lead to refinements in those algorithms. The extent to which each of these agrees with previous measurements from spectroscopy, other cluster samples, and simulations is left for future work. However, for illustrative purposes, we list the relevant observational considerations to be made in understanding the context of these measurements with respect to previous work in the literature, and then highlight a few of our more interesting results.

In general, there are five places where the comparison to previous work must be treated with caution, which can be summarized as follows: 1) redshifting of the galaxy spectra through the bandpasses under consideration, which imparts trends in the observed colors, 2) selection effects imposed by the color selection (e.g. Franzetti et al. 2007), 3) aperture effects, i.e. the aperture used to measure the color in different banspasses (e.g. Scodeggio 2001; Blakeslee et al. 2006), 4) projection effects. 5) actual evolution in the red sequence.

To some level, any of the aforementioned issues may play into our results: (i) at $z \simeq 0.1$, the CMR of photometrically-selected galaxies is noticeably shallower than previous spectroscopic measurements of the color magnitude relation (Hogg et al. 2004; Cool et al. 2006), (ii) the slopes are almost independent of cluster richness; (iii) the photometric error-corrected scatter of the red-sequence broadens mildly with redshift; (iv) the observed mean slope of the CMR is negative and it becomes more negative as redshift increases.

Naively, we expect that our measurement of the slope of the red-sequence, -0.013 ± 0.0003 mags mag^{-1} at $z = 0.1$, would correspond to the SDSS spectroscopic analysis of Hogg et al. (2004), for which the slope is -0.022 mags mag^{-1} in ${}^{0.1}(g-r)$. In addition to the fact that the Hogg et al. (2004) measurements are k-corrected to the $z = 0.1$ rest-frame, one possible difference comes from our definition of the red sequence: Hogg et al. (2004) use a 2σ clipping algorithm to define the red sequence and to iteratively reject outliers. While they split the sample by Sersic index, sigma-clipping may be more permissive of objects near the "blue cloud" to be included in the red-sequence, while the method presented in this paper automatically accounts for the presence of these objects. Our slope measurements at a given redshift may also be biased shallow, as the initial 2 σ cut derived from the ECGMM fit does not account for the slope in the red-sequence itself, i.e. the cut is applied in the same way regardless of magnitude. Ideally, an iterative procedure would be employed to determine the best-fit line for each cluster and the 2σ cut would be applied as a function of magnitude. Unfortunately, the small number statistics for low richness clusters do not permit this to be implemented in a robust fashion.

Insofar as richness and local density are similar indicators of environment, the second observation (ii) that the slope is almost independent of environment is in basic agreement

with Hogg et al. (2004), who use SDSS spectroscopy at $z \sim 0.1$ to compare galaxies with high ($n \geq 2$) Sersic indicies in different environments characterized by their local density.

After the photometric error correction performed by ECGMM, a trend in the scatter with redshift remains (iii), such that the scatter increases with increasing redshift. At high redshift $z \simeq 1$, the color-magnitude relation has been measured in a handful of clusters (Mei et al. 2009; Koester et al. 2009; Santos et al. 2009, e.g.) with the general conclusion that the restframe scatter in the CMR does not evolve with redshift. More locally, the SDSS Luminous Red Galaxy (LRG) Sample has been used to measure various redshifted frames of bright ($L \gtrsim 2.2L_*$) red galaxies (Cool et al. 2006). Cool et al. (2006) find the intrinsic restframe scatter ${}^{0.16}(g-r) = 35.4 \pm 3.7$ and ${}^{0.37}(g-r) = 43.5 \pm 6.2$ mmags, consistent with no evolution. However, with increased cluster sample, our *observed* frame measurements reveal an increase in the scatter, shown by a statistically significant non-zero slope (the bottom right panel in Fig. 7).

Result (iv) is in qualitative agreement with the results in (Gladders et al. 1998) who find a similar trend in the slope for a sample 44 Abell clusters at $z \leq 0.15$ and 6 clusters at $0.2 \leq z \leq 0.75$, the largest previous study of its kind. In their study of the scatter of the CMR in LRGs, Cool et al. (2006) report no significant trend with redshift in the rest-frame slope of LRGs over $0.16 < z < 0.37$ in either the cluster or the field, but caution that the sample is not-well suited to measuring the slope. The observed factor of 2.5 increase in the magnitude in our measurement of the slope is likely due to a combination of the lack of k-corrections and selection effects (e.g. Franzetti et al. 2007) derived from color cuts that may preferentially include a larger and larger fraction of galaxies with significant star-formation at increasing redshifts.

A further contribution to the inflated slope may come from the choice of the color aperture. van Dokkum et al. (1998) and Scodeggio (2001) note the importance of the use of adaptive apertures, which place the color measurements of large and small galaxies on the same footing. This point motivates our choice of `MODEL_MAGS` from the SDSS, which are derived from the best-fit convolution of the local PSF with a deVaucoleurs model in the r-band. This same best-fit model is then used to compute the flux in both the g and r-bands.

As our results indicated, the intrinsic ridgeline slope may decrease as redshift increase. Though there are many factors that may complicate the interpretation of this results as we discussed in the preceding paragraphs, it is still worthwhile to speculate what may lead to the intrinsic evolution. In general, galaxies are more active at higher redshift (Cowie et al. 1996), and their color distributon spreads wider and toward bluer end. Therefore, clusters at higher redshift are more likely to have member galaxies with wider and bluer color distribution. As a result, when we fit the CMR with a straight line, we tend to have more negative slopes at

higher redshift.

7. Summary

In this paper, we have presented the ECGMM, a new purely photometric method which characterizes the red sequence ridgeline in cluster samples with large statistics. This provides precise measures of the mean variation of the red sequence ridgeline location and scatter (width) with respect to redshift, properly corrected for photometric errors. The measured slopes, scatters, and zeropoints are directly applicable to improved cluster finding efforts and to characterization of known galaxy clusters.

Applying the method to maxBCG clusters approximately recovers known properties of the red sequence, namely its slope and the variation of the slope with redshift, and the insensitivity of the slope to large-scale environment. It also suggests that the scatter of the red-sequence increases mildly with redshift, and that the slope of the red-sequence grows substantially by $z \simeq 0.3$. However, we caution that these observed trends may be attributable to a host of observational effects, including color selection, the lack of k-corrections, and details of the measurement of the individual cluster CMRs. More detailed examination of these issues will be required before their implications for galaxy formation models can be fully understood, and will be the subject of future work. In the meantime, the measurements can serve as an important observational check on models that produce mock galaxy catalogs.

Acknowledgments

JH thank Hyunsook Lee for helpful conversation about the BIC and Mixture models. JH and TM gratefully acknowledge support from NSF grant AST 0807304 and DoE Grant DE-FG02-95ER40899. ESR would like to thank the TABASGO foundation. ER was funded by the Center for Cosmology and Astro-Particle Physics at The Ohio State University and by NSF grant AST 0707985. This project was made possible by workshop support from the Michigan Center for Theoretical Physics.

Funding for the creation and distribution of the SDSS Archive has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the U.S. Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS Web site is http://www.sdss.org/. The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are The Univer-

sity of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, The Johns Hopkins University, the Korean Scientist Group, Los Alamos National Laboratory, the Max-Planck- Institute for Astronomy (MPIA), theMax-Planck-Institute for Astrophysics (MPA), New Mexico State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington

A. The recursive relation for the error corrected Gaussian Mixture Model

In this appendix, we show the derivation of the likelihood function Eq. 2 and the EM recursive relations for the error corrected GMM. To begin with, we introduce the following notations in Table.1. For brevity, we denote the parameters (μ_i, σ_i and w_i) collectively by θ and (t) represents the tth iteration. M represents number of data points and N represents the number of mixtures.

Since we assume the true color distribution can be approximated by mixture of Gaussian distributions, we have the following probability density function for $p(\bar{y}_j|\theta)$:

$$p(\bar{y}_j|\theta) = \sum_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\Big[-\frac{(\bar{y}_j-\mu_i)^2}{2\sigma_i^2}\Big] \tag{A1}$$

Though the true colors are not directly observable, we know that its distribution given the observed colors and measurement errors is approximately Gaussian:

$$p(\bar{y}_j|y_j) = \frac{1}{\sqrt{2\pi\delta_j^2}} \exp\Big[-\frac{(\bar{y}_j-y_j)^2}{2\delta_j^2}\Big] \tag{A2}$$

Notations	Meaning
y_1,..., y_j, ..., y_m:	Observed colors of BCGs and member galaxies.
$\bar{y}_1$,..., $\bar{y}_j$, ..., $\bar{y}_m$:	True colors of BCGs and member galaxies.
z_1,..., z_j, ..., z_m:	Hidden variables that tell which Gaussian component the $\bar{y}_j$ is sampled from.
δ_1,..., δ_j,..., δ_m:	Measurement errors for every y_j.
μ_1, ..., μ_i, ..., μ_n:	Mean of each Gaussian component.
σ_1, ..., σ_i, ..., σ_n:	Width of each Gaussian component.
w_1, ..., w_i, ..., w_n:	Weights of corresponding Gaussian components.

Table 1: The notations used in our derivation of ECGMM algorithm

then, the likelihood function (under the flat priors for θ)

$$L(\theta|y_j) = \int p(\bar{y}_j|\theta)p(\bar{y}_j|y_j)d\bar{y}_j \tag{A3}$$

After integrating over $\bar{y}_j$ and extending to all data points ($\prod_{j=1}^{M}$), we arrive at Eq.2. The optimal parameters could be obtained by maximizing the above likelihood. However, if we introduce hidden variables, z, that tell us which Gaussian component the y_j is sampled from, then the whole maximization process could be significantly simplified. The corresponding pdf of data given z and θ is

$$p(y|z_j = i, \theta^{(t)}) = \prod_{j=1}^{M} p(y_j|z_j = i, \theta_i^{(t)}) = \prod_{j=1}^{M} \frac{1}{\sqrt{2\pi(\sigma_i^{(t)2} + \delta_j^2)}} \exp\Big[-\frac{(y_j - \mu_i^{(t)})^2}{2(\sigma_i^{(t)2} + \delta_j^2)}\Big] \tag{A4}$$

We use w_i denote the weight of each Gaussian Component in the mixture and is given by $w_i = p(z_j = i|\theta)$. The estimation of hidden variable could be related to Eq.A4 by Bayes' Theorem as following:

$$p(z_j = i|y_j, \theta^{(t)}) = \frac{p(z_j = i, y_j|\theta^{(t)})}{p(y_j|\theta^{(t)})} = \frac{p(y_j|z_j = i, \theta^{(t)})p(z_j = i|\theta^{(t)})}{\sum_{i=1}^{N} p(y_j|z_j = i, \theta^{(t)})p(z_j = i|\theta^{(t)})} \tag{A5}$$

The EM algorithm iteratively update the parameters θ by maximizing the expected log likelihood

$$Q(\theta) = \sum_{i=1}^{N}\sum_{j=1}^{M} p(z_j = i|y_j, \theta^{(t)})\Big[-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(\sigma_i^2 + \delta_j^2) - \frac{(y_j - \mu_i)^2}{2(\sigma_i^2 + \delta_j^2)} + \ln p(z_j = i|\theta^{(t)})\Big] \tag{A6}$$

under the constraint $\sum_{i=1}^{N} p(z_j = i|\theta^{(t)}) = 1$. Using the Lagrange Multiplier approach, we redefine

$$\tilde{Q}(\theta) = Q(\theta) - \lambda\Big[\sum_{i=1}^{N} p(z_j = i|\theta^{(t)}) - 1\Big] \tag{A7}$$

with λ as the multiplier.

$$\frac{\partial \tilde{Q}(\theta)}{\partial \mu_i} = \sum_{j=1}^{M} \Big[p(z_j = i|y_j, \theta^{(t)}) \Big(\frac{y_j - \mu_i}{\sigma_i^2 + \delta_j^2} \Big) \Big] = 0 \tag{A8}$$

From Eq.A8, we can arrive at the following recursive relation for μ:

$$\mu_i^{(t+1)} = \frac{\sum_{j=1}^{M} y_j p(z_j = i|y_j, \theta_i^{(t)})/(1 + \delta_j^2/\sigma_i^{(t)2})}{\sum_{j=1}^{M} p(z_j = i|y_j, \theta_i^{(t)})/(1 + \delta_j^2/\sigma_i^{(t)2})} \tag{A9}$$

Similarly, we have

$$\frac{\partial \tilde{Q}(\theta)}{\partial \sigma_i} = \sum_{j=1}^{M} p(z_j = i|y_j, \theta^{(t)}) \Big[\frac{\sigma_i^2(1 + \delta_j^2/\sigma_i^2) - (y_j - \mu_i)^2}{\sigma_i^4(1 + \delta_j^2/\sigma_i^2)^2} \Big] = 0 \tag{A10}$$

Note that since σ_i and δ_j are entangled within the summation, there would not be an simple analytic solution for σ_i. However, since the algorithm is iterative in nature and the major contribution for the update of σ_i is from $(y_j - \mu_i)^2$, we could approximate σ_i in δ_j^2/σ_i^2 with its value in t^{th} iteration. Then we can solve for the $(t+1)^{th}$ iteration relation for σ_i as:

$$\sigma_i^{(t+1)} = \Big[\frac{\sum_{j=1}^{M} (y_j - \mu_i)^2 p(z_j = i|y_j, \theta_i^{(t)})/(1 + \delta_j^2/\sigma_i^{(t)2})}{\sum_{j=1}^{M} p(z_j = i|y_j, \theta_i^{(t)})/(1 + \delta_j^2/\sigma_i^{(t)2})} \Big]^{1/2} \tag{A11}$$

Our numerical test shows that such an approximation works fine in practice. For $w_i = p(z_j = i|\theta)$, we have

$$\frac{\partial \tilde{Q}(\theta)}{\partial w_i} = \sum_{j=1}^{M} p(z_j = i|y_j, \theta^{(t)})/w_i - \lambda = 0 \tag{A12}$$

which leads to

$$w_i = p(z_j = i|\theta) = \frac{1}{\lambda} \sum_{j=1}^{M} p(z_j = i|y_j, \theta^{(t)}) \tag{A13}$$

Using the condition $\sum w_i = 1$, we have $\lambda = M$. Substitute λ back to Eq. A13, we arrive at:

$$w_i^{(t+1)} = \frac{1}{M}\sum_{j=1}^{M} p(z_j = i|y_j, \theta_i^{(t)}) \tag{A14}$$

In the above iteration relations Eq.A9, Eq.A11 and Eq.A14, t and $t+1$ denote the round of iterations. When we ignore the measurement errors δ_j, the above recursive relation reduces to the standard EM recursive relation for Gaussian Mixture Model. The above relations could be easily generalized to multiviate case by simply substituting the data with data matrix, mean with mean vector and variance with covariance matrix, which we will not repeat the formula here. A C++ class that implements the above algorithm is available upon request.

REFERENCES

Adelman-McCarthy, J. K., et al. 2006, ApJS, 162, 38

Annis, J., et al. 1999, in Bulletin of the American Astronomical Society, Vol. 31, Bulletin of the American Astronomical Society, 1391–+

Arimoto, N., & Yoshii, Y. 1987, A&A, 173, 23

Barger, A. J., et al. 1998, ApJ, 501, 522

Barrientos, L. F. 1999, PhD thesis, AA(UNIVERSITY OF TORONTO (CANADA))

Becker, M. R., et al. 2007, ApJ, 669, 905

Bernardi, M., Nichol, R. C., Sheth, R. K., Miller, C. J., & Brinkmann, J. 2006, AJ, 131, 1288

Bernardi, M., Sheth, R. K., Nichol, R. C., Schneider, D. P., & Brinkmann, J. 2005, AJ, 129, 61

Blakeslee, J. P., et al. 2003, ApJ, 596, L143

—. 2006, ApJ, 644, 30

Blanton, M. R., et al. 2005, AJ, 129, 2562

Bower, R. G., Lucey, J. R., & Ellis, R. S. 1992, MNRAS, 254, 601

Coil, A. L., et al. 2008, ApJ, 672, 153

Connolly, A. J., Genovese, C., Moore, A. W., Nichol, R. C., Schneider, J., & Wasserman, L. 2000, ArXiv Astrophysics e-prints

Cool, R. J., Eisenstein, D. J., Johnston, D., Scranton, R., Brinkmann, J., Schneider, D. P., & Zehavi, I. 2006, AJ, 131, 736

Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839

De Lucia, G., et al. 2007, MNRAS, 374, 809

Dempster, A., Laird, N., & Rubin, D. 1977, Journal of the Royal Statistical Society, Series B, 39, 1

Eisenhardt, P. R., et al. 2005, in Bulletin of the American Astronomical Society, Vol. 37, Bulletin of the American Astronomical Society, 1344–+

Evrard, A. E. 1989, ApJ, 341, L71

Faber, S. M., et al. 2007, ApJ, 665, 265

Franzetti, P., et al. 2007, A&A, 465, 711

Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, AJ, 111, 1748

Gladders, M. D., Lopez-Cruz, O., Yee, H. K. C., & Kodama, T. 1998, ApJ, 501, 571

Gladders, M. D., & Yee, H. K. C. 2000, AJ, 120, 2148

—. 2005, ApJS, 157, 1

Gladders, M. D., Yee, H. K. C., Majumdar, S., Barrientos, L. F., Hoekstra, H., Hall, P. B., & Infante, L. 2007, ApJ, 655, 128

Gunn, J. E., et al. 2006, AJ, 131, 2332

Hansen, S. M., Sheldon, E. S., Wechsler, R. H., & Koester, B. P. 2007, ArXiv e-prints, 710

Hogg, D. W., et al. 2004, ApJ, 601, L29

Hogg, D. W., Finkbeiner, D. P., Schlegel, D. J., & Gunn, J. E. 2001, AJ, 122, 2129

Hu, W. 2003, Phys. Rev. D, 67, 081304

Johnston, D. E., et al. 2007, ArXiv e-prints, 709

Kauffmann, G., & Charlot, S. 1998, MNRAS, 294, 705

Kodama, T., & Arimoto, N. 1997, A&A, 320, 41

Koester, B. P., et al. 2009, ArXiv e-prints

—. 2007a, ApJ, 660, 239

—. 2007b, ApJ, 660, 221

Lima, M., & Hu, W. 2004, Phys. Rev. D, 70, 043504

—. 2005, Phys. Rev. D, 72, 043006

López-Cruz, O., Barkhouse, W. A., & Yee, H. K. C. 2004, ApJ, 614, 679

Lupton, R., Gunn, J. E., Ivezić, Z., Knapp, G. R., & Kent, S. 2001, in Astronomical Society of the Pacific Conference Series, Vol. 238, Astronomical Data Analysis Software and Systems X, ed. F. R. Harnden, Jr., F. A. Primini, & H. E. Payne, 269–+

Majumdar, S., & Mohr, J. J. 2004, ApJ, 613, 41

Mei, S., et al. 2009, ApJ, 690, 42

—. 2006, ApJ, 644, 759

Mullis, C. R., Rosati, P., Lamer, G., Böhringer, H., Schwope, A., Schuecker, P., & Fassbender, R. 2005, ApJ, 623, L85

Oukbir, J., & Blanchard, A. 1992, A&A, 262, L21

Pier, J. R., Munn, J. A., Hindsley, R. B., Hennessy, G. S., Kent, S. M., Lupton, R. H., & Ivezić, Ž. 2003, AJ, 125, 1559

Renzini, A. 2006, ARA&A, 44, 141

Rozo, E., et al. 2008, ArXiv e-prints

—. 2009, ArXiv e-prints

Sandage, A., Binggeli, B., & Tammann, G. A. 1985, AJ, 90, 1759

Santos, J. S., et al. 2009, ArXiv e-prints

Schwarz, G. 1978, Annals of Statistics, 6, 461

Scodeggio, M. 2001, AJ, 121, 2413

Sheldon, E. S., et al. 2007, ArXiv e-prints, 709

Smail, I., Edge, A. C., Ellis, R. S., & Blandford, R. D. 1998, MNRAS, 293, 124

Smith, J. A., et al. 2002, AJ, 123, 2121

Stanford, S. A., Eisenhardt, P. R., & Dickinson, M. 1998, ApJ, 492, 461

Stott, J. P., Pimbblet, K. A., Edge, A. C., Smith, G. P., & Wardlow, J. L. 2009, MNRAS, 394, 2098

Titterington, D., Smith, A., & Makov, U. 1985, Statistical Analysis of Finite Mixture Distributions (John Wiley & Sons)

Tucker, D. L., et al. 2006, Astronomische Nachrichten, 327, 821

van Dokkum, P. G., Franx, M., Kelson, D. D., Illingworth, G. D., Fisher, D., & Fabricant, D. 1998, ApJ, 500, 714

Visvanathan, N., & Sandage, A. 1977, ApJ, 216, 214

Wake, D. A., et al. 2006, MNRAS, 372, 537

Zehavi, I., et al. 2005, ApJ, 630, 1

This preprint was prepared with the AAS LaTeX macros v5.2.

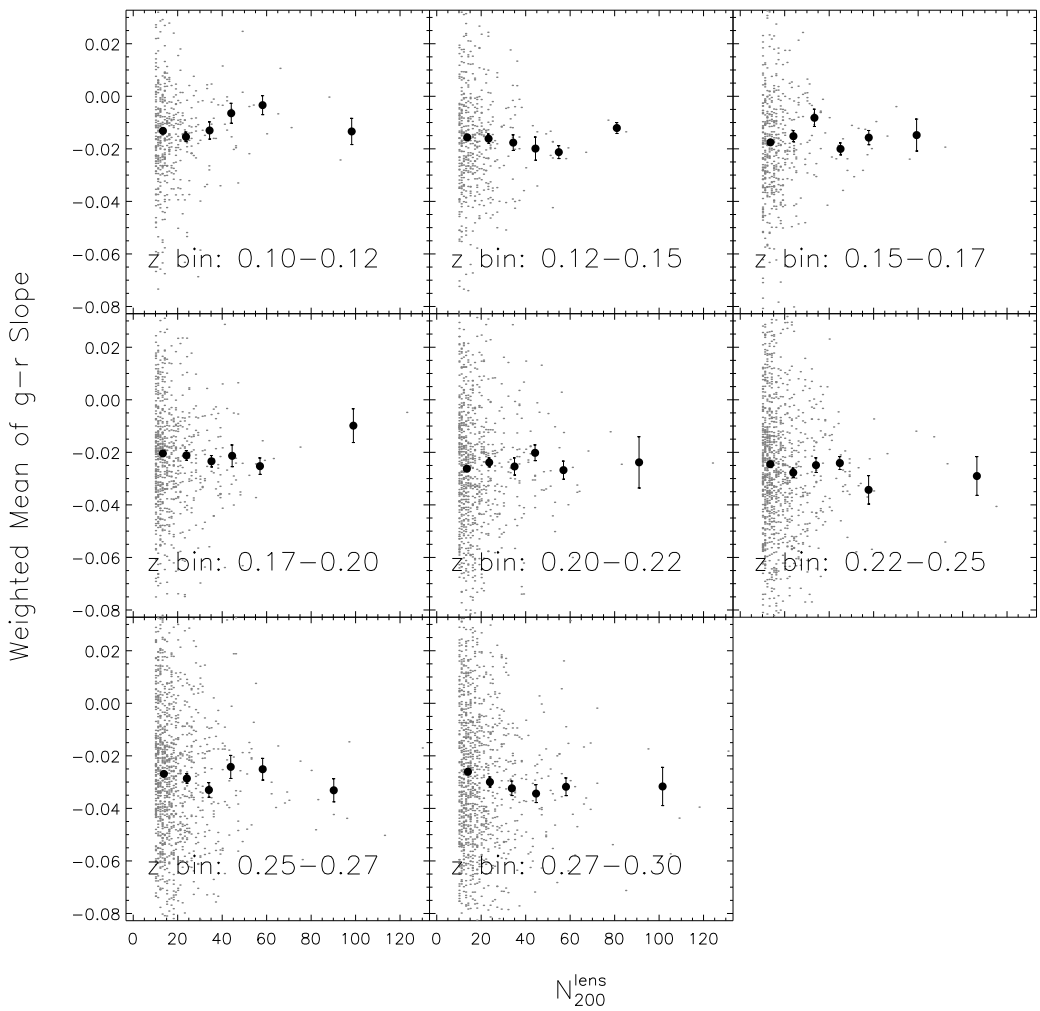

Fig. 10.— The evolution of mean ridgeline slope vs richness at different redshift slices. The richness bins brackets are chosen as N_{200}^{lens} =[10,20,30,40,60,80,161]. The light dark points are from individual clusters. The black solid dots and error bars are weighted mean and standard deviation of the slopes in each N_{gals} bin for every redshift slice. From the plot, we did not see strong trends of the slope evolution w.r.t richness.

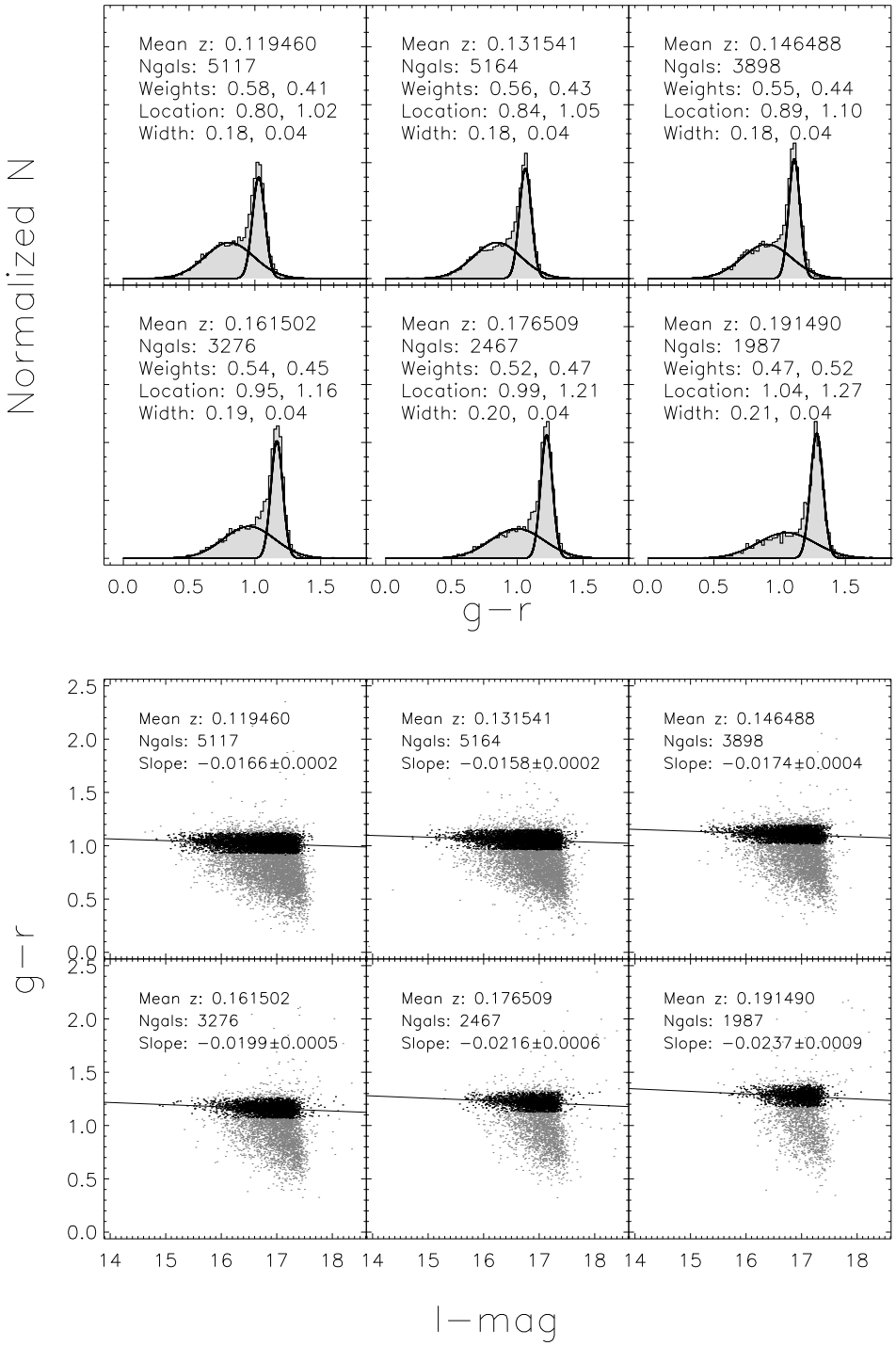

Fig. 11.— Evaluating the ECGMM-derived red sequence slopes in SDSS spectroscopy of field galaxies. The normalized color histograms (top panel) for $\Delta z = 0.003$ slices in spectroscopic redshift clearly show the presence of the red and blue components in the field galaxy distribution. ECGMM is used to separate the two components, the redder of which is to measure the CMR (bottom panel).

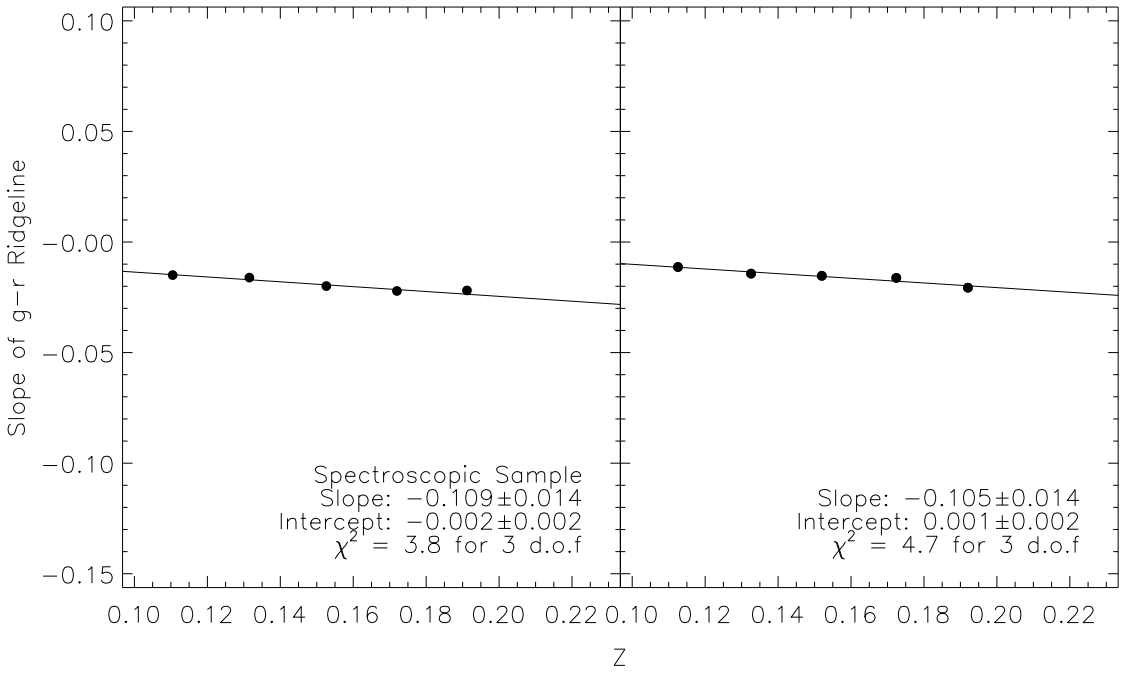

Fig. 12.— The comparison of the evolution of the slopes of CMR for spectroscopic sample and cluster sample. Since the spectroscopic sample is biased due to selection effects at $z \geq 0.2$, we choose both sample in the redshift range from 0.1 to 0.2 and then bin the slopes in redshift bin of 0.02. We then fit straight line to the mean slopes in each bin for spectroscopic sample and cluster sample respectively.