Clustering and Dark Matter Halos of Ultra-Massive Passively Evolving Galaxies and Passive Galaxy Groups at $z \sim 1.6$

by

Gurpreet Kaur Cheema

A Thesis Submitted to Saint Mary's University, Halifax, Nova Scotia in Partial Fulfillment of the Requirements for the Degree of Master of Science in Astronomy (Department of Astronomy and Physics)

August 2017, Halifax, Nova Scotia

Approved: Dr. Marcin Sawicki
Supervisor

Approved: Dr. Rob Thacker
Examiner

Approved: Dr. Ivana Damjanov
Examiner

Date: August 9th, 2017.

Acknowledgements

First, I would like to express my gratitude to my supervisor, Dr. Marcin Sawicki for providing me the opportunity to join his group and for his guidance, immense knowledge, encouragement, and most of all for his enthusiasm, which was a key element to motivate me.

Thanks to Liz for the long useful discussions and supporting me throughout the progress of my thesis, and many insightful conversations, about science and otherwise. Thanks to Anneya and Thibaud, for the constructive comments, and guidance. I want to thank professors and staff for their support along my journey. Thanks to the members of my committee, for insightful comments, and encouragement.

Most importantly, I would like to thank my family. Very special thanks to my parents and my brother Pawanpreet for their unwavering love, everlasting support and genuine interest in seeing me succeed at something I love to do. You have always been a constant source of support and love throughout all of my pursuits in life. I would like to express my sincere gratitude to my friends, Pratishtha and Simran for the encouragement and infinite patience.

Thanks a lot to all the people in SMU who have made these two years an unforgettable time of my life. I would like to thank Sofia, Kirsten, Adam, Preet, Nathalie, and Florence. I couldn't wish for the greater company.

Contents

1 Introduction **1**

1.1 UMPEGs 2

1.2 Clustering Studies 4

1.3 Thesis Outline 6

2 Data **8**

2.1 The CFHTLS Deep and Wide Fields 8

2.1.1 Infrared Data 9

2.1.2 Optical Data 10

2.2 Catalog Based on K_s-Selection 11

2.3 Selection of Passive *gzK* Galaxies 11

3 The Angular Correlation Function **19**

3.1 Correlation Function Estimators 21

3.2 Creating the Random Catalog 22

3.3 Measuring $\omega(\theta)$ 23

3.3.1 Pair Computation 23

3.3.2 K_s-selected Sampling 24

3.3.3 Integral Constraint 25

3.3.4 Error Estimation 28

3.4 Clustering Dependence on K_s-magnitude 32

4 The Spatial Correlation Function 35

4.1 Limber Inversion 36

4.2 Redshift Distribution 38

4.3 Estimating the Correlation Length 39

5 Mass Estimation of Dark Matter Halos of UMPEGs 44

5.1 Brief Review of Millennium XXL Simulation 44

5.2 Clustering of Dark Matter Halos at $z \sim 1.6$ 46

5.3 Determination of the Dark Matter Halo Masses 50

6 Groups of Passive Galaxies 53

7 Interpretation 62

7.1 Comparing UMPEGs with other Galaxy Populations 62

7.2 Stellar Mass - Halo Mass Relation 64

7.3 Evolution of UMPEGs to $z \sim 0$ 67

8 Conclusions **70**

Bibliography **72**

List of Figures

2.1 Two color $(z-K_s)$ vs $(g-z)$ diagram representing gzK selection in Wide W1 field. 13

2.2 $(z-H)$ vs $(H-K_s)$ color-color plot for D3 field 15

2.3 Distribution of the PE-gzK in the Deep fields 16

2.4 Distribution of the PE-gzK in the Wide fields 18

3.1 Angular correlation function of PE gzK galaxies as a function of K_s-band magnitude. 31

3.2 Clustering amplitude for gzK-selected passively evolving galaxies as a function of K_s magnitude. 34

4.1 Redshift distribution function for the PE gzK galaxies 40

4.2 Comparison of the correlation lengths of our UMPEGs with previous studies of Passive galaxies. 42

5.1 Comparison of spatial correlation lengths of PE *gzK* galaxies to clustering measurements from the *MXXL* simulation. 49

5.2 Halo mass as a function of halo correlation function from the *MXXL* simulation 52

6.1 Groups of passive *gzK* galaxies in the Wide fields 55

6.2 Angular correlation measurements of groups of the PE *gzK* galaxies 56

6.3 Comparison of spatial correlation lengths of groups of PE *gzK* galaxies to halo clustering measurements from the *MXXL* simulation. 58

6.4 Halo mass as a function of correlation function of the halos from the *MXXL* simulation. 60

7.1 Comoving correlation length r_0 of the UMPEGs in contrast to other galaxy populations. 63

7.2 Stellar mass-halo mass ratio (SHMR) for different PE *gzK* sub-samples . . 65

7.3 Cumulative number density of the PE *gzK* galaxies and DM halos for the *MXXL* simulation. 68

List of Tables

3.1 The clustering amplitude A_ω and masses for the PE *gzK* galaxies as a function of K_s-magnitude bin 34

4.1 r_0 and halo mass measurements for PE *gzK* galaxies 41

5.1 Summary of DM halos from Millennium XXL simulation 48

6.1 Correlation length and halo mass for the PE *gzK* group 57

Abstract

Clustering and Dark Matter Halos of Ultra-Massive Passively Evolving Galaxies and Passive Galaxy Groups at $z \sim 1.6$

by Gurpreet Kaur Cheema

Using a sample of *gzK* selected passive galaxies at $z \sim 1.6$ covering an effective area of 27 deg^2, we used the clustering measurements of Ultra Massive Passively Evolving Galaxies (UMPEGs, $M_{stellar} > 10^{11.4} M_{\odot}$) to determine the masses of their host dark matter halos. We measured the angular and spatial correlation function of UMPEGs and found that UMPEGs cluster more strongly than any other known galaxy population at high redshift. Comparison to the Millennium XXL simulation suggested that their halos are of mass $\sim 10^{14.1} h^{-1} M_{\odot}$. We found that the passive galaxy groups also reside in massive halos, perhaps even more massive than those hosting the UMPEGs. Finally extrapolating the growth of halos hosting the UMPEGs and groups to $z = 0$, we showed that their halos at $z \sim 1.6$ may evolve into massive clusters such as Virgo and Coma in the local Universe.

August 9th, 2017

Chapter 1

Introduction

According to the standard ΛCDM cosmological model, the mass in the Universe is dominated by cold dark matter (CDM). The growth of the first gravitational instabilities which are caused due to quantum fluctuations being inflated to large sizes, leads to the collapse of regions of dark matter (DM) and the formation of first systems, the DM halos. The baryonic matter follows the gravitational well of the DM and galaxies are born as concentrated luminous cores within the DM halos via cooling and condensation of baryons (White & Rees 1978). As the efficiency of gas cooling is related to the gas density, $\Lambda_{cool} \propto \rho^2$ and thus the gravitational potential of the DM halo, the halo mass is a crucial quantity to understand galaxy formation. Following their formation, the galaxies grow their mass by merging of dark matter halos and the associated baryons, progressively assembling more massive systems. The structure formation is hierarchical as the small structures form first and larger

structures assemble later. The evolution of the galaxies in the DM halo involves various internal and external processes such as gas cooling, hydrodynamical effects, star formation, mergers and feedback mechanisms. These processes are also linked to the host DM halos of the galaxies (Behroozi et al. 2010; Contreras et al. 2015). Thus the properties of galaxies are directly coupled to those of the DM halos in which they live. As a result, the properties of galaxies change as they evolve through cosmic time and therefore high redshift galaxies are likely to be different from present day galaxies.

1.1 UMPEGs

The Ultra Massive Passively Evolving Galaxies (UMPEGs) are extreme galaxies at $z \sim 1.6$ with $M_{stellar} > 10^{11.4} M_{\odot}$ that are no longer forming stars. The combination of their very high stellar masses and their low star formation rates makes them rare at this redshift. These galaxies are already quiescent at $z \sim 1.6$ and since the Universe is only $\sim$4 Gyr old at the epoch we observe them, so their massive stellar populations must have assembled very early and rapidly. Assuming that they were on the star-forming main sequence just before becoming quenched, they must have exhibited star formation rates of $\sim 1000 M_{\odot}$/yr or more. While massive and reasonably bright, UMPEGs are exceedingly rare, populating the very massive exponential tail end of the $z \sim 1.6$ galaxy mass function. The number densities of UMPEGs are $\sim 10^{-6}$ Mpc^{-3}dex^{-1} in stellar mass which is a factor of $\sim$300 times lower that of the most common quiescent galaxies with typical mass $\sim 10^{10.5} M_{\odot}$ at $z \sim 1.6$. The

UMPEGs need to be understood better to test the extremes of the hierarchical models of galaxy formation.

The extreme nature of UMPEGs raises the question: what environments do they live in? This issue can be addressed by determining various properties of these galaxies such as their number density and mass function, and examining their companions to understand their merging histories. Along with these studies, it is important to study the DM halo mass of these galaxies. Previous studies have found that galaxy properties such as stellar mass, luminosity, morphology and star formation rate are correlated to host DM halo mass (e.g., Li et al. 2006; Zehavi et al. 2011). Therefore, the DM halo environment plays an important role in shaping the galaxy properties and thus galaxy evolution. The DM halo mass of galaxies is also an important parameter to trace the mass assembly history of these extreme galaxies because dark matter halos grow monotonically with time by merging, independent of the baryonic processes inside the halos.

The UMPEGs are expected to be the most massive, central galaxies of their dark matter halos and these halos, as the UMPEGs themselves, are also expected to be very massive. Studying the behavior of the host dark matter halos of the galaxies and linking the observed properties of the galaxies to the halo masses will be useful in elucidating how the galaxies were built up and how they evolved over time in the context of the standard ΛCDM cosmological model. It is also useful for the purpose of identifying these galaxies' likely $z = 0$ descendants. Ultimately, after constraining the masses of the DM halos hosting these

galaxies, we need to interpret the galaxies in the context of hierarchical growth of structure in the universe.

1.2 Clustering Studies

One of the tools for understanding galaxy formation and evolution is galaxy clustering. Galaxy clustering describes the three-dimensional distribution of galaxies and involves analyzing the statistics of the distribution of galaxies. Qualitatively, we expect that the spatial distribution and hence, clustering of the UMPEGs, reflecting the underlying clustering of their halos, will be stronger than the clustering of "normal" quiescent galaxies at the same epoch. We aim at testing this hypothesis and determining the masses of the halos of UMPEGs by quantitatively comparing their clustering with predictions of dark matter clustering from N-body simulations (Springel et al. 2005; Boylan-Kolchin et al. 2009). In the clustering analysis, the clustering strength of galaxies is evaluated with the correlation function, and is compared to the predictions of the ΛCDM structure formation models.

During the last decade, large observational surveys of galaxies at both low and high redshifts have tremendously improved our knowledge of galaxy evolution, and helped to connect galaxy properties to those of the dark matter distribution using galaxy clustering studies. Several alternative methods, such as halo occupation distribution modelling (Leauthaud et al. 2011; Yang et al. 2003; Berlind & Weinberg 2002), stellar mass-halo mass relationship, halo abundance matching (Kravtsov & Klypin 1999; Conroy et al. 2006; Moster

et al. 2010), and weak gravitational lensing (Brainerd et al. 1996; Hoekstra et al. 2004) provide statistical measures to connect the population of galaxies to their host dark matter halos. Using galaxy clustering measurements, there have been a number of studies of passive galaxies at low redshifts (Coil et al. 2004; Zehavi et al. 2005; Madgwick et al. 2003) and high redshifts (Brown et al. 2008, 2003; Coil et al. 2008; McCracken et al. 2008).

Galaxy clustering is a powerful way to investigate the UMPEGs, since the amplitude of clustering on large scales can provide a measure of the mass of the host DM halos (Mo & White 1996; Sheth & Tormen 1999). Clustering of galaxies is related to the clustering of halos and it is directly related to the formation of halos. In the ΛCDM model, the clustering of DM halos is well understood (Mo & White 2002) where the clustering amplitude is a monotonically increasing function of the halo mass. The halos cluster in such a way that the most massive DM halos have larger clustering strength as measured by the correlation function. The DM halos (and galaxies inside them) form from small perturbations in the early universe which grow with time. The high mass halos are formed in regions with strong, positive perturbations on even larger scales. The large scale collapse accelerates the collapse of the smaller halos, causing an excess of these halos in the general neighborhood and hence, massive systems are strongly clustered. However, large scale perturbations are not needed to form the low mass halos and hence, low mass halos have lower clustering. The galaxy populations that are more clustered are found to be hosted by more massive DM halos (Norberg et al. 2002; Zehavi et al. 2005).

In addition to studying UMPEGs, we also extend our clustering analysis to groups of the passively evolving galaxies (PEGs) at $z \sim 1.6$. These groups can be supposed to be the markers of high redshift protoclusters and as such let us test the extreme range of galaxy environment at $z \sim 1.6$. Here we use group-group clustering to constrain the masses of the DM halos they reside in.

1.3 Thesis Outline

In this work, we used *gzK* selected passive galaxies from the catalogs of Arcila-Osejo (2017) to measure the clustering of the UMPEGs using the correlation function and determine their host DM halo masses. As a consequence, we can relate the UMPEGs from the distant past to the present day by tracing the evolution of these halos predicted within the framework of evolution of dark matter structures.

In Chapter 2, we describe the details of the data as well as the method employed by Arcila-Osejo (2017) for selecting passive high-redshift galaxies: the *gzK* selection technique, the adjusted version of *BzK* selection criteria and *zHK* selection. In Chapter 3, we present the technique to measure the angular correlation function for the passive high redshift galaxies in the four Deep fields and the Wide fields. We further divide the samples into subsamples according to their K_s-band magnitudes to investigate the clustering properties. Chapter 4 covers the spatial correlation function derived from the angular correlation function using the redshift distribution for the passive galaxies. In Chapter 5, we

compare the clustering results of UMPEGs with the clustering measurements of the dark matter halos obtained from the *Millennium XXL* simulation in order to obtain their host DM halo masses. In Chapter 6, we describe the clustering measurements for the passive galaxy groups and determine the halo masses for these groups. A discussion and interpretation of our measurements is presented in Chapter 7. We summarize the main conclusions in Chapter 8.

Throughout this work, we assume flat lambda cosmology ($\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$), the Hubble constant is $h = H_0/100km\ s^{-1}Mpc^{-1} = 0.7$, and the normalization of the matter power spectrum is $\sigma_8 = 0.8$. Unless stated, we use the AB magnitude system (Oke 1974). Stellar masses of galaxies assume the Chabrier (2003) stellar initial mass function (IMF).

Chapter 2

Data

2.1 The CFHTLS Deep and Wide Fields

The data is composed of optical as well as infra-red (IR) data obtained from the 3.6m Canada-France-Hawaii Telescope (CFHT)[1]. The data is within the CFHT Legacy Survey (CFHTLS T0006)-Deep distributed over four fields with an effective area of 2.5 square degrees and within the CFHTLS Wide fields W1 and W4 covering an effective area of about 27 square degrees (Arcila-Osejo 2017).

[1] Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Science de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii.

2.1.1 Infrared Data

For the Wide survey, K_s data is obtained from the Visible Multi-Object Spectrograph (VIMOS) Public Extragalactic K_s Survey Multi-Lambda Survey (VIPERS-MLS) (Moutard et al. 2016) regions within the W1 and W4 fields of the CFHTLS. VIPERS was a follow up of the CFHTLS in the Wide W1 and W4 fields using WIRCam (Puget et al. 2004) corresponding to 100% completeness down to the limiting magnitude $K_s \leq 22.0$ AB mags. Analysis by Arcila-Osejo (2017) shows that these data are 100% complete for PE-gzK galaxies (defined in Section 2.3) to 20.25. The WIRCam is a wide-field near-infrared imager at CFHT consisting of four 2048×2048 pixel cryogenically cooled HgCdTe arrays in a 2×2 array format. It covers a 20 arcmin×20 arcmin field of view with a pixel scale of 0.3 arcsec $pixel^{-1}$. The data reduction was done at CFHT and TERAPIX[2].

In the Deep fields, the IR data are taken from the T0002 release of the WIRCam Deep Survey (WIRDS) (Bielby et al. 2012) that provided deep imaging in these 4 fields. The WIRCam instrument was utilised to obtain all the data except the J-band image for WIRDS-D2 field which was obtained by the WFCAM instrument on the United Kingdom Infrared Telescope(UKIRT). The infra-red data consists of high quality near infrared imaging comprised of deep J, H, and K_s imaging for these fields. The WFCAM instrument is composed of four 2048×2048 $pixel^2$ detectors with each detector covering 13.65 arcmin×13.65 arcmin field of view of the sky with a pixel scale of 0.187 arcsec $pixel^{-1}$. The processing

[2] `http://terapix.iap.fr/`

of WIRDS data was done at TERAPIX in collaboration with CFHT. In case of the Deep fields, the imaging reaches a 50% completeness limit at 24.5 AB mags in J, H, and K_s with an exception of D2 field corresponding to 50% completeness at a limiting magnitude of 24.0 AB mags. Analysis by Arcila-Osejo (2017) shows that these data are 100% complete for PE-gzK galaxies in the Deep fields to 23.25 except in D2 field corresponding to 100% completeness at a limiting magnitude of 22.0 AB mags.

2.1.2 Optical Data

The optical data for both Deep and Wide fields was obtained from the CFHTLS that was carried out from May 2003 to February 2009. This large project was conducted using the wide-field optical imager, MegaCam at MegaPrime. The CFHT MegaCam covers 57.6 arcmin$\times$57.6 arcmin of the sky at a pixel scale of 0.187 arcsec $pixel^{-1}$. The four 1deg$\times$1 deg fields - D1, D2, D3, and D4 were observed in five broad-band filters namely u^*, g', r', i', and z'. The data contain two types of stacks for each image; one with 85 percent best-seeing images and other comprising 25 percent best-seeing image stacks. In this case, Arcila-Osejo (2017) used the one with 25 percent, as it is useful in the study of morphologies and finding close companions. All Wide data was observed in five filters (u^*, g, r', z', and i' or y[3]). The stacks in the Wide survey cover 1 square degrees (19354$\times$19354 pixels) and have a pixel scale of 0.1860". The Wide fields consist of individual pointings that slightly overlap by a few arcmins in both RA and DEC.

2.2 Catalog Based on K_s-Selection

Arcila-Osejo (2017) performed source detection and photometry using SExtractor (Source Extractor,Bertin & Arnouts 1996). Masks were created in order to remove bad regions such as areas near bright stars, cosmic ray trails or dead pixels. As the Spectral Energy Distributions (SEDs) of the passive galaxies are dominated by optically faint long-lived stars, the selection was based on K_s-band. After detecting the objects in K_s images, photometry was done at the K_s band positions in g, z, H (in the Deep fields) and K_s bands. All the pointings in the Wide fields were merged taking care of the duplicate sources due to partial overlapping of adjacent tiles. A full catalog of the detected objects in the Deep and Wide fields was created by Arcila-Osejo (2017) and further used to distinguish between passive and star-forming high-redshift galaxies using the gzK selection technique described in the next section.

2.3 Selection of Passive gzK Galaxies

An important step to study these galaxies at high redshifts is to develop an efficient tool to easily categorize the galaxies into star-forming and passive. Different methods have been used to select passive galaxies at high redshifts based on the spectral signatures of passive old stellar population: Extremely Red Objects (EROs; Messias et al. 2010; McCarthy 2004; Roche et al. 2002), Distant Red Galaxies (DRG; Franx et al. 2003) and color-color

techniques such as *BzK* selection (Daddi et al. 2004). The first two selection techniques are based on red optical to near-infrared colors (e.g., $(R-K) > 5$). Spectroscopic studies showed that there are drawbacks to these methods as these methods select both old passive populations and dusty star-forming galaxies (van Dokkum et al. 2004; Cimatti et al. 2002). On the other hand, *BzK* selection has the ability to identify the passive galaxies at $z \sim 2$. This technique is based on B, z, and $K-$band photometry and provides an efficient selection criterion that is not biased against passive galaxies and highly reddened star-forming galaxies. The criterion used to select galaxies is as follows:

$(z-K) \geq (B-z) - 0.2$, and

$(z-K) < (B-z) - 0.2 \cap (z-K) > 2.5$

for the star forming (*sBzK*s) and the passive galaxies (*pBzK*s) respectively. Various studies have been carried out on passive and star-forming galaxies selected using this technique in the recent years (e.g. Kong et al. 2006; McCracken et al. 2010; Bielby et al. 2014; Arcila-Osejo & Sawicki 2013; Sato et al. 2014; Ishikawa et al. 2015).

We are interested in selecting passive *gzK* galaxies and follow the procedure outlined in Arcila-Osejo & Sawicki (2013). This selection method modifies the *BzK* selection to the available CFHTLS+WIRDS filters to devise a new *gzK* technique shown in Figure 2.1. The cuts designed by comparing locations of galaxy models in the *BzK* color-color diagram (Daddi et al. 2004) to that in the *gzK* color-color space are given by:

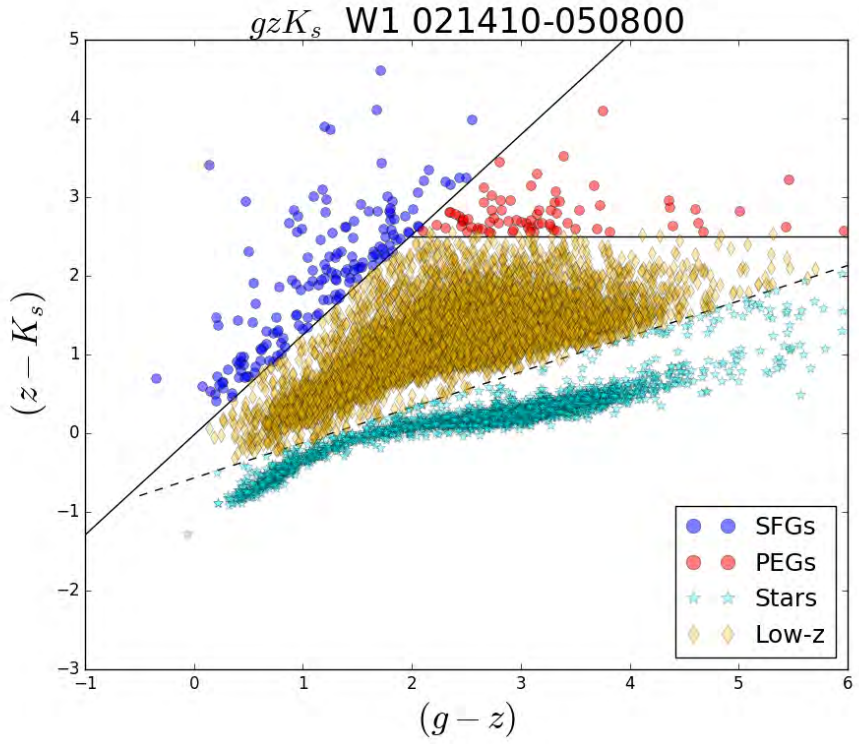

Figure 2.1: Two color $(z-K_s)$ vs $(g-z)$ diagram for the Wide fields representing gzK selection for a single patch in Wide W1 field. The diagonal solid line separates the passive and star-forming galaxies where the red symbols on the right represent passive galaxies and the blue symbols on the left represent the star-forming galaxies. The dashed diagonal line defines the stars that are shown in cyan symbols. Low redshift galaxies are the ones shown as orange symbols. Figure taken from Arcila-Osejo (2017).

$$(z-K_s)-1.27(g-z)\geq -0.022 \tag{2.3.1}$$

representing Star-Forming (SF-gzK) galaxies, and

$$(z-K_s)-1.27(g-z)<-0.022\ \cap\ (z-K_s)\geq 2.55 \tag{2.3.2}$$

representing Passively Evolving (PE-gzK) galaxies. The above Equations (2.3.1) and (2.3.2) are used in the present work to identify star forming and passive galaxies in the Wide fields.

The completeness of passive galaxies in the Wide survey goes up to $K_s < 20.5$ and does not require additional classification. For the Deep fields, the data is not deep enough in the g-band to allow star-forming or passive classification in the gzK plane. In order to solve this problem, a second set of criteria was incorporated to further classify the galaxies in the Deep fields (Arcila-Osejo & Sawicki 2013). This involves zHK color-color selection technique based on $(z-H)$ versus $(H-K_s)$ for galaxies above $(z-K_s) > 2.55$. This is useful in differentiating the old passive galaxies from the dusty SF galaxies when the g-band photometry is too shallow to do so directly in the gzK diagram. At the redshift of interest, the 4000$\mathring{A}$ break is redshifted to higher wavelengths and will lie between the $z'$and K_s bandpasses. H-band is useful in selecting the passive systems by identifying this feature. Therefore zHK color-color selection as shown in Figure 2.2 is used by Arcila-Osejo (2017), and thus in this work, to further classify the galaxies into star-forming and passive. Above

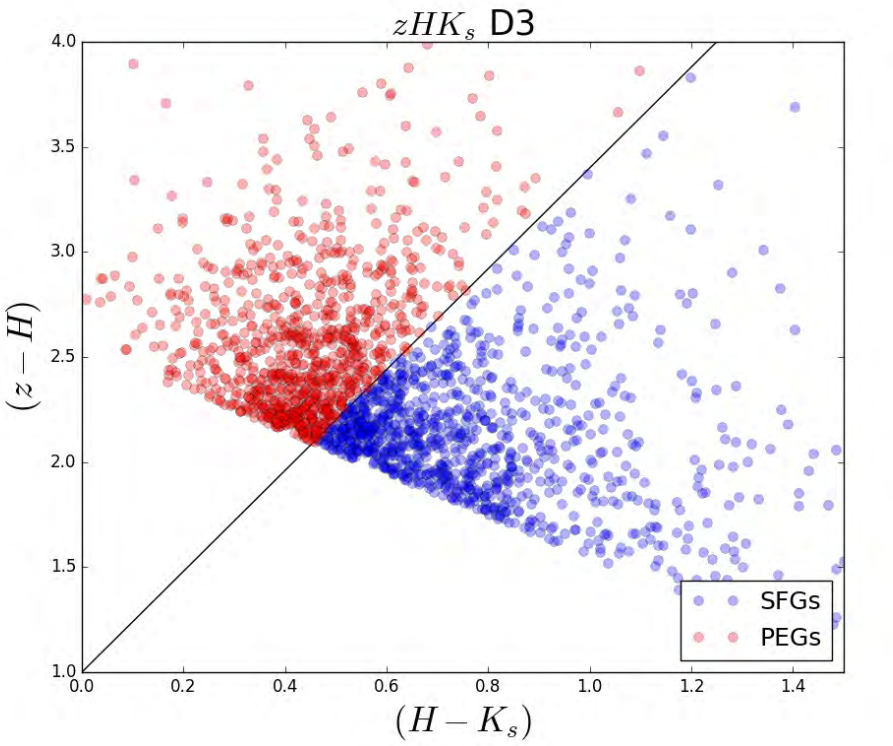

Figure 2.2: $(z-H)$ versus $(H-K_s)$ color-color plot for selecting objects in one of the Deep fields. A galaxy is classified as a passively evolving galaxy based on the position of the galaxy in the color-color plot as defined in Equation 2.3.3. Figure taken from Arcila-Osejo (2017).

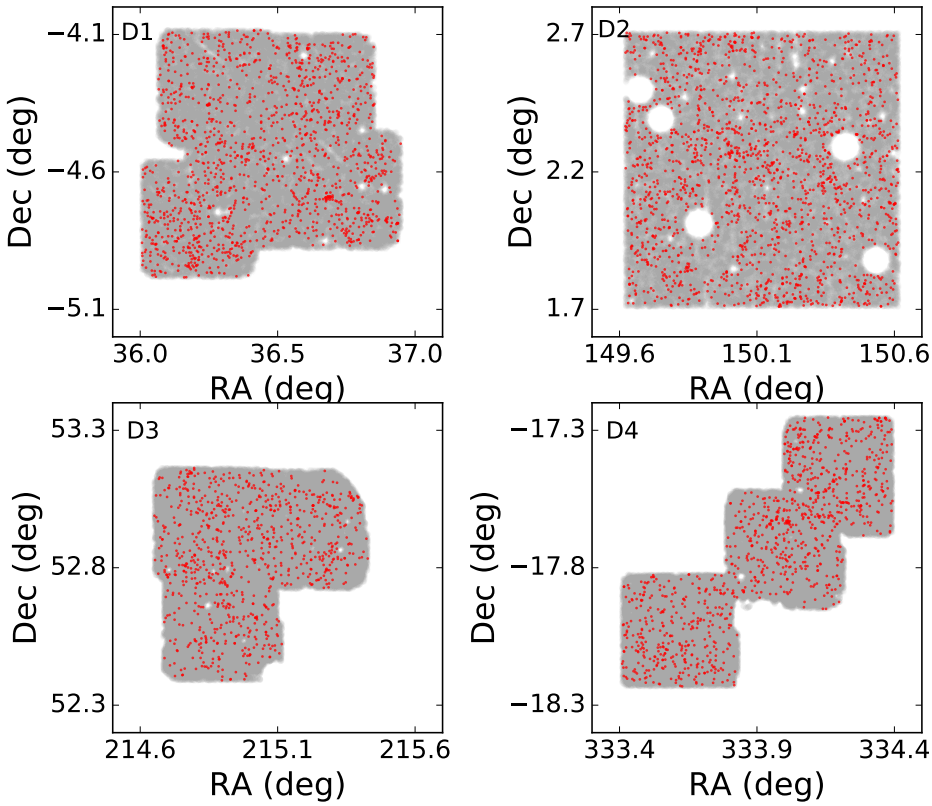

Figure 2.3: The distribution of the PE-*gzK* in the Deep fields D1, D2, D3 and D4. PE *gzK* galaxies with $K_s < 23$ are represented as red points. The gray area shows the layout of the Deep fields along with the boundaries of the survey area. The white empty spaces are the areas with no K_s observation or due to masking of bright stars.

$(z-K) = 2.55$, a galaxy is considered to be a passive system if

$$(z-H) > 2.4(H-K_s) + 1. \tag{2.3.3}$$

Figure 2.3 shows the PE *gzK* galaxies in the Deep fields with $23 < K_s < 20$. In Figure 2.4, we show all the PE *gzK* galaxies to $K_s < 20.25$ along with the layout of our g, z, and

K_s coverage in W1 an W4 fields where the clustering is clearly visible on small scales. The UMPEGs (Ultra Massive Passively Evolving Galaxies) were selected as the passive galaxies brighter in K_s than 19.75, in other words with $log(M_{stellar}/M_{\odot}) \gtrsim 11.4$.

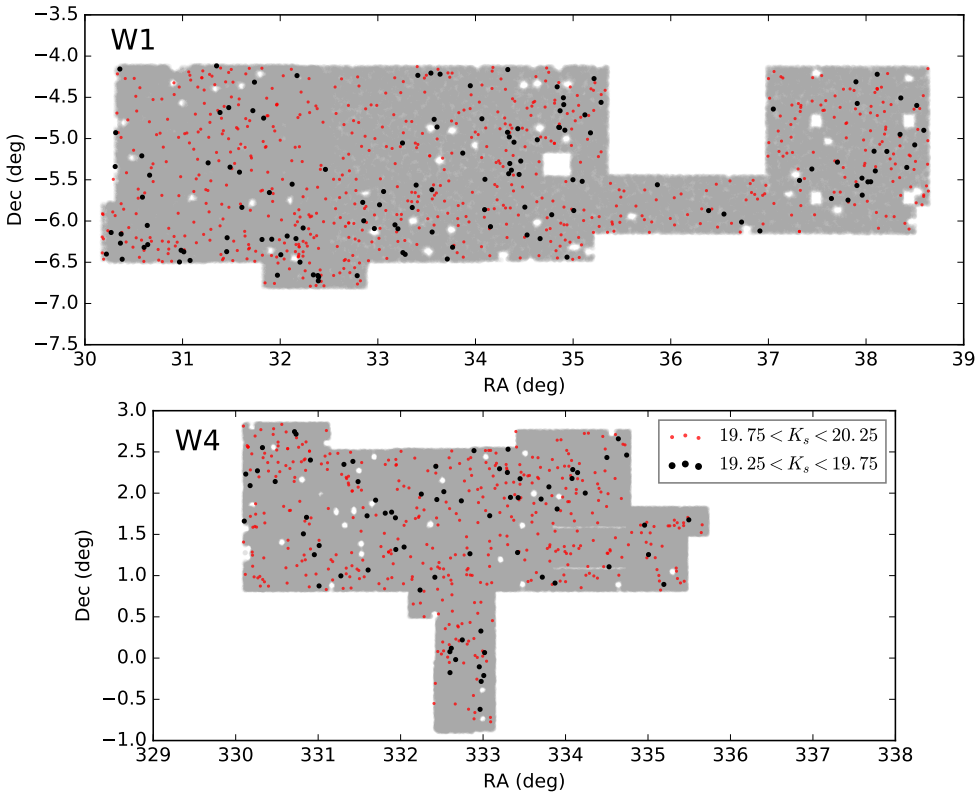

Figure 2.4: The distribution of the PE-*gzK* in the Wide fields W1 (upper sub-panel) and W4 (lower sub-panel). PE *gzK* galaxies in the Wide field W1 and W4 are represented as red points. The gray area shows the layout of the Wide fields along with the boundaries of the survey area. The white empty spaces are the areas with no K_s observation or due to masking of bright stars.

Chapter 3

The Angular Correlation Function

Knowing the positions of galaxies in the survey, the first step is to understand their clustering properties. One of the most widely used statistical quantities to measure clustering in the Universe are the two-point correlation functions, which can be two-dimensional projections (angular correlation function) or three-dimensional measures (spatial correlation function). The correlation function as a clustering measure of galaxies was suggested by Totsuji & Kihara (1969) and then was continued for statistical characterization of the galaxy clustering (Peebles 1980; Maddox et al. 1990; York et al. 2000).

The most useful tool is the spatial correlation function which is not directly measurable. The angular correlation function has an advantage over the spatial correlation function because it only requires the information about the angular positions of the galaxies. This allows the angular correlation to be measured for the wide surveys covering large volumes

to understand the structure in the Universe statistically. The disadvantage is that the angular correlation function is the projection of the spatial correlation function over the redshift distribution of the sample. De-projecting the angular function in order to compute the clustering length is easier for the galaxies with narrow redshift distribution. In contrast, estimation of spatial clustering becomes more difficult at fainter magnitudes, as the redshift distribution gets broader. Another problem is that the redshift distribution is not always known well.

The commonly used and simplest quantitative measure of the degree of clustering is the angular two-point correlation function $\omega(\theta)$ (Peebles 1980). It is defined as measure of the joint probability $dP(\theta)$ of finding a pair of objects in the solid angles $d\Omega_1$ and $d\Omega_2$ separated by an angle θ, with respect to an unclustered random Poisson distribution and is written as

$$dP(\theta) = n[1 + \omega(\theta)]d\Omega_1 d\Omega_2, \tag{3.0.1}$$

where n is the average surface density of galaxies and $\omega(\theta)$ is the two-point correlation function. Thus, $\omega(\theta)$ describes, as a function of angular space, the excess net projected pair clustering of galaxies over the random distribution. The correlation function is defined to be in the range $[-1 \leq \omega(\theta) \leq \infty]$, such that

1. $\omega(\theta) > 0$, the objects are considered to be clumped on the relevant scale,
2. $\omega(\theta) = 0$ implies random distribution of galaxies, and

3. $\omega(\theta) < 0$, implies objects have negative correlation or in other words, are anti-correlated.

3.1 Correlation Function Estimators

There have been a number of different estimators used in the literature to measure the angular correlation function. In the estimators below, *DD* is the data-data pair for fundamental observed distribution of galaxies, but it requires two important corrections: one for the geometry of the field and the other for the relationship between observed galaxies and the edges of the field. These corrections are taken care of through random-random (*RR*) and data-random (*DR*) correlations that are measured using the random catalog discussed in section 3.2.

Various estimators that have been used to compute the angular correlation function are:

$\omega(\theta) = \frac{DD(\theta)}{RR(\theta)} - 1$, Peebles & Hauser (1974),

$\omega(\theta) = \frac{DD(\theta)}{DR(\theta)} - 1$, Davis & Peebles (1983),

$\omega(\theta) = \frac{DD(\theta) - DR(\theta)}{RR(\theta)}$, Hewett (1982),

$\omega(\theta) = \frac{(DD(\theta))(RR(\theta))}{(DR(\theta))^2} - 1$, Hamilton (1993), and

$\omega(\theta) = \frac{DD(\theta) - 2\times DR(\theta) + RR(\theta)}{RR(\theta)}$, Landy & Szalay (LS, 1993). In all these estimators, $DD(\theta)$ refers to unique number of galaxy-galaxy pairs with angular separations between $\theta - \Delta\theta/2 < \theta < \theta - \Delta\theta/2$. $DR(\theta)$ is the number of pairs with the same angular separations between the galaxy catalog and the random catalog of galaxies in the same survey

area. $RR(\theta)$ refers to number of random-random pairs with the same angular separations. In order to reduce Poisson noise from the random points, we want to have $N_R \gg N_D$, but if we do this, we need to do the weighing so as not to give undue weights to the random points. Therefore the counts are weighted by multiplying DR by a factor of $\left(\frac{N_D-1}{N_R}\right)$ and RR by $\left(\frac{N_D(N_D-1)}{N_R(N_R-1)}\right)$ (Adelberger et al. 2005) where N_D and N_R are the total number of data and random points in the survey area respectively. The Hamilton estimator has a subtle advantage in three-dimensional analysis of flux-limited redshift surveys. However, the LS estimator is advantageous as it is less sensitive to the size of the random catalog and also handles edge corrections well (Kerscher et al. 2000), although it requires more computational time. Szapudi & Szalay (1998) demonstrated that the LS estimator has superior shot noise behavior compared to the other estimators, and Labatie et al. (2012) showed that the LS estimator is nearly of minimal variance for a random distribution and has a second order variance decay in $(1/n^2)$ where n is the number of data points. Because of these properties, the LS estimator is adopted for measuring the correlation function in this work.

3.2 Creating the Random Catalog

The survey boundary effects must be taken into account along with other biases while estimating the correlation function. The survey covers only a limited part of the sky and as a result, the objects near masked out regions and near the boundaries of the survey will have fewer neighbors at some angular or spatial separations than they would have in a whole sky

survey. In addition to the edge effects, there is another bias that randoms are often used to account for: non-uniformity in the depth of the survey. That is why the generation of a random sample is a crucial step for an accurate measurement of the correlation function. The non-uniform completeness is not important for our study as we limit our analysis to brighter objects only where we have 100% completeness.

The random catalog of galaxies needs to define a uniformly distributed background with many more objects than the data catalog and with the same observational survey biases as the actual galaxy sample. We generated the random positions in x and y pixels and rejected the random objects that fell in the masked regions or were outside the edges of the survey. The masks were obtained by combining g, z, and K_s-band fits images used for galaxy detection (Arcila-Osejo 2017). In the case of the Wide fields, masks were created for all the tiles separately and the tiles were carefully combined taking care of the overlapping regions of adjacent tiles. To remove the noise in the random pair counts, the random catalog contained about 100 times more objects than the galaxy catalog ($N_R/N_D \sim 100$).

3.3 Measuring $\omega(\theta)$

3.3.1 Pair Computation

The catalogs of the PE *gzK* galaxies created by Arcila-Osejo (2017) provide information about the position of galaxies in x and y pixels as well as angular coordinates: Right Ascen-

sion (RA) and Declination (DEC). By definition, the angular correlation function requires a measure of the number of pairs of points in a dataset lying within a given angular annulus/shell. For every n data points, there are $\frac{n(n-1)}{2}$ pairs and this number increases to $\frac{100n(100n-1)}{2}$ for the random catalog. Counting the number of pairs in each bin in coordinate space involves pairwise distance computations and this task is a challenge to computational power and memory. This problem is solved by reducing the number of distance computations using *kd* trees (Friedman et al. 1977) as it provides a quick and efficient way to count the number of neighbors within different angular shells. Knowing the angular distribution of galaxies, the galaxy-galaxy, random-random and galaxy-random pairs are counted using *kd* trees.

3.3.2 K_s-selected Sampling

The K_s-band magnitude gives an approximate measure of stellar masses of UMPEGs (Arcila-Osejo 2017) as the rest-frame $\sim$8500$\mathring{A}$ light from passive galaxies at $z \sim 1.6$ is dominated by the long-lived low-mass stars that contain most of the stellar mass. The sample was divided into subsamples based on K_s-band luminosity to investigate the luminosity and, by implication, mass dependence of the clustering measurements. Taking advantage of the large PE *gzK* galaxy catalog for the Wide fields, we divided the sample into two differential samples of bin size 0.5 mag: $19.25 < K_s < 19.75$, and $19.75 < K_s < 20.25$. Due to the small area of the Deep fields, the passive population is divided into three subsamples with

bin-size of 1.0 mag: $20 < K_s < 21$, $21 < K_s < 22$, $22 < K_s < 23$. The lower value for the angular range θ is selected in such a way that there is no under counting of galaxy pairs due to galaxy isophotes being too close. The upper limit is smaller than the size of the field. There are no galaxy pairs with separations less than 1.14 arcsec in the Deep fields; thus this value is used as the smallest distance in the calculation of number of pairs for the Deep fields. In case of the UMPEGs in the range $19.25 < K_s < 19.75$ in the Wide fields, there are no galaxy pairs at separations closer than 57.06 arcsec. Due to the effective area for the individual field in the Deep fields survey being less than 1 square degree, $\omega(\theta)$ is computed in angular distance bins of constant logarithmic width $\Delta log\theta = 0.2$ with bins ranging from $log(\theta) = -3.5$ to $log(\theta) = -0.5$ where θ is in degrees. The Wide fields are covering a large number of objects over an effective area of 15.53 square degrees for W1 and 9.56 square degrees for W4. For the brighter UMPEGs in the Wide sample, the angular distance bins consist of the same constant logarithmic width $\Delta log\theta = 0.2$, but different range from $log(\theta) = -3.5$ to $log(\theta) = 0.5$ where θ is in degrees. In this work, the logarithmic binning of angular separations is chosen in order to provide adequate sampling at small scales and in order to avoid excessively fine sampling at large scales.

3.3.3 Integral Constraint

Estimation of $\omega(\theta)$ requires an estimate of the background galaxy density. The mean galaxy density estimate is based on the data sample itself. The bias resulting from the

fact that the angular correlation function is restricted to a limited area Ω of the survey is the "integral constraint (IC)".

For a given angular correlation function, the number of pairs within the separations $[\theta - \Delta\theta/2, \theta + \Delta\theta/2)$ is given by

$$N = n\left(\frac{\delta\Omega_1}{\Omega}\frac{\delta\Omega_2}{\Omega}\right)[1+\omega(\theta)]. \tag{3.3.1}$$

Doubly integrating this quantity over the solid angles Ω_1 and Ω_2 for the total survey area gives us the total number of unique data-data pairs. There is an overestimation of the mean density due to positive correlation between galaxies at small separations (Infante 1994), which is balanced by negative correlation at larger separations. The magnitude of the IC depends on both the field size and the clustering strength.

Let $\omega_{measure}$ be the measured correlation function, and it is related to the actual correlation function ω_{true} (e.g, Sato et al. 2014) by

$$1 + \omega_{measure}(\theta) = f(1 + \omega_{true}(\theta)),$$

where ω_{true} is the true correlation function and f is a scaling factor defined later. Using Equation 3.3.1 and the constraint

$$N = \iint n\left(\frac{\delta\Omega_1}{\Omega}\frac{\delta\Omega_2}{\Omega}\right) f[1+\omega(\theta)],$$

we get $f = \frac{1}{1+IC}$ where IC is the integral constraint. The negative offset is given by integration of assumed true $\omega(\theta)$ over the field Ω (Peebles 1980),

$$IC = \frac{1}{\Omega^2} \iint \omega(\theta) d\Omega_1 d\Omega_2,$$

where Ω corresponds to the solid angle of the survey. The above integral is computed using the following expression including the random-random counts (Roche & Eales 1999; Infante 1994)

$$IC = \frac{\sum RR(\theta) A_\omega \theta^{-\beta}}{\sum RR(\theta)}$$

and is added to the measured value $\omega_{measure}(\theta)$ to compute the $\omega_{true}(\theta)$,

$$\omega_{true}(\theta) \approx \omega_{measure}(\theta) + IC, \tag{3.3.2}$$

where IC is the correction for the bias mentioned above.

The two-point angular correlation function is well approximated by the power law (Peebles 1980):

$$\omega(\theta) = A\theta^{1-\gamma}. \tag{3.3.3}$$

Assuming the above power law form in Equation 3.3.3, the data is fit using a non-linear

least-squares fit to estimate the parameters A_ω and γ to quantify the strength of clustering. For $\omega_{true}(\theta) = A_\omega \theta^{1-\gamma}$, the estimated correlation function is given by $\omega_{measure}(\theta) = A_\omega(\theta^{1-\gamma} - C)$, where $C = \frac{IC}{A_\omega}$. The value of IC is found to range from 0.06 to 0.08 for the Deep fields and 0.04 to 0.06 for the Wide fields. Since the Deep fields cover a relatively small area, the integral constraint has a large effect on the measurements. However, this effect is negligible for the Wide fields.

3.3.4 Error Estimation

The error in the estimation of $\omega(\theta)$ is difficult to model as ideally we need to estimate this error from the variance of an ensemble of independent samples equivalent to the galaxy sample in question. Unfortunately this is not possible, as we have only one realization of our Universe and of our data. One may use mock samples from independent realizations of a realistic model of structure formation, but it would be model dependent. In this work, we used the error estimate on the basis that fluctuation in number of independent data pairs in given bin of θ has a Poisson distribution $\delta DD = \sqrt{DD}$ and similarly for δDR and δRR, the error is $\sqrt{DR}$ and $\sqrt{RR}$, respectively. As the random catalog contains many more objects ($\sim$100 times) than the real catalog, $\sqrt{RR} \gg \sqrt{DD}$ and $\sqrt{RR} \gg \sqrt{DR}$.

The LS estimator can be written as

$$1 + \omega(\theta) = \frac{DD(\theta)}{RR(\theta)} - \frac{2 \times DR(\theta)}{RR(\theta)}. \tag{3.3.4}$$

Neglecting the second term in Equation 3.3.4 ($RR \gg DR$), the first term along with uncertainties, $\left(\frac{DD\pm\sqrt{DD}}{RR\pm\sqrt{RR}}\right)$ can be simplified as $\left(\frac{DD(1\pm1/\sqrt{DD})}{RR(1\pm1/\sqrt{RR})}\right)$. The error on the expression 3.3.4 becomes

$$\frac{DD}{RR}\frac{1}{\sqrt{DD}}(\sqrt{RR} \gg \sqrt{DD}),$$

which can be written as

$$\frac{1+\omega(\theta)}{\sqrt{DD}}.$$

The Poisson expression for computing the error in the angular correlation function in each θ- bin (Landy & Szalay 1993) is thus given by

$$\sigma_\omega = \frac{1+\omega(\theta)}{\sqrt{DD(\theta)}}. \tag{3.3.5}$$

For the Deep fields, the measurements from all different independent fields are combined using a weighted mean. Assuming the points were from the same parent populations with the same mean, but different standard deviations, the weighted average of the angular correlation function is given by

$$\bar{\omega} = \frac{\sum_i(\omega_i/\sigma_i^2)}{\sum_i(1/\sigma_i^2)},$$

where each data point ω_i is weighted inversely by its own variance σ_i. With $w = 1/\sigma_\omega^2$ as

the weight, the uncertainty of the mean σ is given by

$$\sigma^2 = \frac{\sum_i (1/\sigma_i^2)}{(\sum_i (1/\sigma_i^2))^2} = \frac{1}{\sum_i (1/\sigma_i^2)}.$$

The variance of the weighted mean is:

$$\sigma^2 = \frac{\sum_i w_i(\omega_i - \omega)^2}{(\sum_i w_i)} \times \frac{1}{N-1},$$

where N=4 is the number of fields.

The angular correlation measurements for the two Wide fields, W1 and W4 were kept separate and thus treated as independent measurements. These combined $\omega(\theta)$ results from two fields were then fit jointly, which then yielded the desired amplitude value A_ω.

Figure 3.1 summarizes the clustering measurements for *gzK*-selected passive galaxies as a function of the median K_s- magnitude of the samples. The fits to the data were performed over angular scales of 0.01° to 0.32° for the Deep fields and $0.013^\circ$to $0.631^\circ$for the Wide fields. The power law index for the fainter passive galaxies in Deep fields with $22 < K_s < 23$ is found to be: $\gamma = 1.92 \pm 0.12$. The rest of the samples were fitted allowing the amplitude to vary while keeping the slope fixed at 1.92. We clearly found a positive correlation function signal for the massive passive galaxies in both Deep and Wide fields with an angular dependence consistent with slope $\gamma = 1.92$ consistent with the results of Sato et al. (2014) who found γ to be 1.92 for the *gzK*-selected passive galaxies. We note

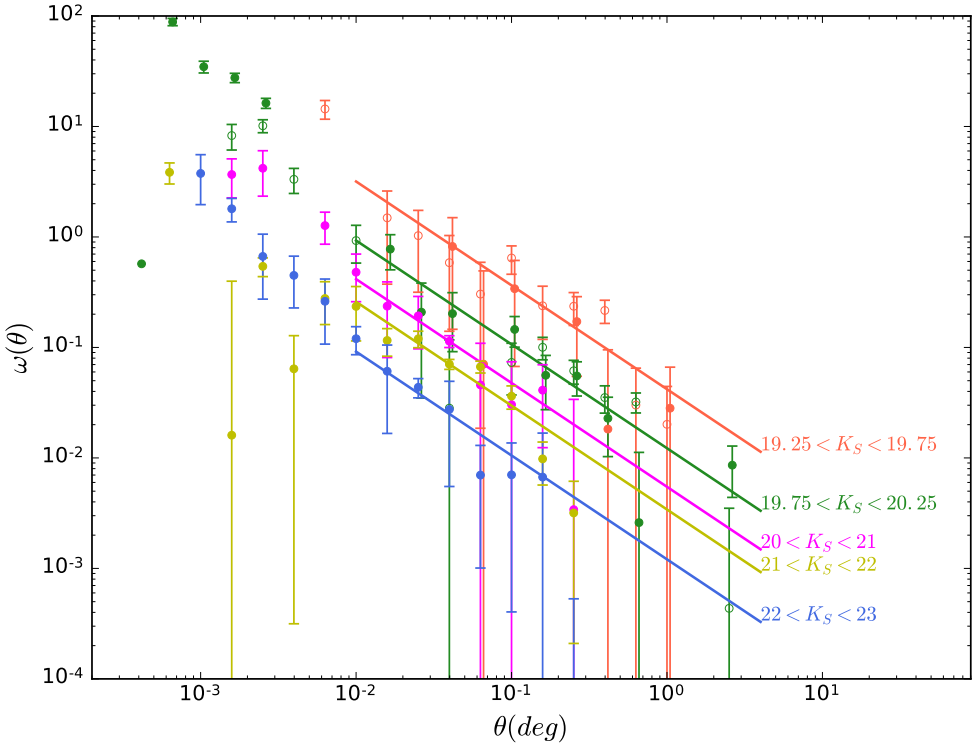

Figure 3.1: The angular correlation function of PE *gzK* galaxies as a function of K_s-band magnitude. The magnitude intervals are $19.25 < K_s < 19.75$, and $19.75 < K_s < 20.25$ for the Wide fields and $20 < K_s < 21$, $21 < K_s < 22$, $22 < K_s < 23$ for the Deep fields. The empty and filled circles for the Wide fields represent W1 and W4 field respectively. Solid lines are the fits to the data with $\gamma = 1.92$ for large angular scales with $\theta > 0.01^\circ$ where one-halo term due to clustering of galaxies within the same halo is negligible.

that McCracken et al. (2010) found the best fitting slope γ for the passive BzK-selected galaxies to be $\gamma \sim 2.3$.

In the case of the lower mass passive galaxies (as seen in Figure 3.1), it is seen that a single power-law approximation does not provide a good fit to the measured correlation results at small scales. The correlation function deviates from the power-law (Zehavi et al. 2004) and there is an excess from the power law at small angular scales. This is due to the 1-halo term that has the contribution from galaxy pairs residing within the same dark halo. This term is determined by clustering of galaxies at small scales which is affected by the dark matter halo substructure (Berlind & Weinberg 2002). Sato et al. (2014) also found that there is a 1-halo term for the PE-gzK galaxies. The power law on large scales comes from the 2-halo term which represents the galaxies that reside in distinct halos and dominates on scales larger than the virial radius of a typical halo.

The fits are done at large θ in order to measure only the clustering of galaxies residing in the distinct halos. The angular correlation function $\omega(\theta)$ is closely approximated by $\theta^{1-1.92}$ in the range $0.01^\circ < \theta < 0.32^\circ$ for Deep fields and $0.013^\circ < \theta < 0.631^\circ$for the Wide fields .

3.4 Clustering Dependence on K_s-magnitude

The K_s-band magnitude range of the UMPEGs, their stellar masses and the clustering amplitudes of the angular correlation function, A_ω, are given in table 3.1. The masses of the

passive *gzK* galaxies are computed using the relation (Arcila-Osejo 2017):

$$log[M_{stellar}] = -0.348K_s + 18.284.$$

It can be clearly seen that the clustering amplitude depends on the K_s-band luminosity and this is shown in Figure 3.2 which shows variation of A_ω with K_s magnitude. The fainter galaxies (with lower K-band luminosity) have lower stellar masses and they have weaker clustering as previously observed in Figure 3.1. On the other hand, the UMPEGs are strongly clustered. Figure 3.2 shows variation of A_ω with K_s-magnitude. Within the uncertainties in the data, the observed correlation function amplitude decreases as a function of magnitude and is consistent with earlier studies (e.g., Sato et al. 2014; Savoy et al. 2011; Ishikawa et al. 2015) that investigated the luminosity dependence of galaxy clustering. Comparing our values with previous authors, Sato et al. (2014) found that the A_ω ranges from 11.32 to 14.49×10^{-3} for the low mass passive galaxies. McCracken et al. (2010) found the amplitude to be 3.2×10^{-2} for K_s =23 and 4.1×10^{-2} for K_s =22 for *BzK*-selected passive galaxies but with a different slope.

During our analysis, it was seen that changing the K_s magnitude binning has an effect on the amplitude and clustering measurement results that resulted in unphysical halo masses. Since the results are highly sensitive to the binning scheme, caution must be exercised in comparing theory and observations in detail.

Sample	K_s[AB mag]	Log($M_*/M_\odot$)	N_g	$A_\omega/10^{-3}$(deg)$^{1-\gamma}$	arcmin (10^{-2})
Wide	19.25-19.75	11.49	132+71	41.88± 6.64	196.54 ± 31.16
	19.75-20.25	11.32	675+434	12.24± 1.51	57.55 ± 7.07
Deep	20-21	11.15	841	5.49± 0.20	25.76± 0.94
	21-22	10.80	2282	3.41 ± 0.37	16.02 ± 1.73
	22-23	10.45	1881	1.21 ± 0.15	5.67 ± 0.69

Table 3.1: The clustering amplitude A_ω and masses for the gzK-selected passive galaxies as a function of K_s-magnitude bin with $\gamma = 1.92$. The estimated parameters are over the angular separation range $0.01^\circ < \theta < 0.32^\circ$ for Deep fields and $0.013^\circ < \theta < 0.631^\circ$for the Wide fields with power law slope fixed at $\gamma = 1.92$.

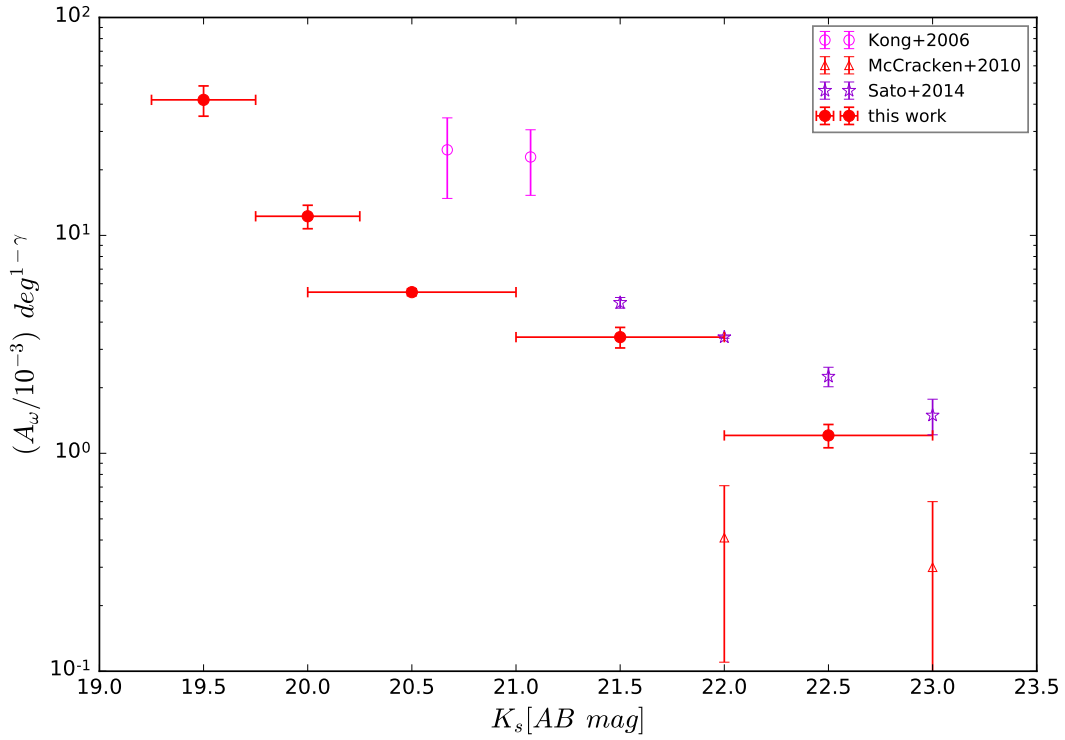

Figure 3.2: The clustering amplitude for gzK-selected passively evolving galaxies as a function of K_s magnitude. The horizontal bars indicate the K_s magnitude intervals defining subsamples (not uncertainties). The measurements are done with γ fixed at 1.92.

Chapter 4

The Spatial Correlation Function

The underlying meaningful physical relation is the full three dimensional spatial correlation function $\xi(r)$. The two-dimensional galaxy clustering, defined by $\omega(\theta)$, as seen in the plane of sky is a projection of the three-dimensional clustering, $\xi(r)$. Similar to the definition (Equation 3.0.1) of $\omega(\theta)$, considering two infinitesimally small spheres centered on two objects, located at r_1 and r_2, the spatial correlation function $\xi(r)$ is defined by the joint probability $dP(r)$ of finding two objects within volume elements $dV_1$and dV_2, at a separation $r = r_1 - r_2$ such that

$$dP(r) = n[1 + \xi(r)]dV_1 dV_2$$

where n is the space density of objects. The spatial correlation function can be described as a power law of the form

$$\xi(r) = \left(\frac{r}{r_o}\right)^{-\gamma},$$

where r is the co-moving distance between the two points, r_o is the characteristic correlation length, and γ is the slope derived from the angular correlation measurements.

We have the observables of angular coordinates and the redshift information of the galaxies. The inverse Limber transformation (Limber 1953) provides a method to connect these two to determine the real space correlation function. Therefore, the clustering properties in terms of co-moving correlation lengths can be determined from the angular correlation function.

4.1 Limber Inversion

The de-projection of the angular correlation function in order to compute the spatial correlation function is done using the Limber inversion. The amplitudes of the power law representation of angular and spatial correlation functions are related by the following equation (Limber 1953; Magliocchetti & Maddox 1999):

$$A_\omega = \frac{H_\gamma r_0^\gamma \int_0^\infty F(z) r_c^{1-\gamma}(z) N^2(z) E(z) dz}{(c/H_0)[\int_0^\infty N(z) dz]^2}, \tag{4.1.1}$$

where A_ω is the amplitude of $\omega(\theta)$, $r_c(z)$ is the radial co-moving distance at redshift z, H_γ is a factor depending on the power-law index slope given by

$$H_\gamma = \Gamma(\frac{1}{2})\frac{[\Gamma(\gamma-1)/2]}{\Gamma(\gamma/2)}, \tag{4.1.2}$$

$E(z)$ is a cosmology-dependent expression given by

$$E(z) \equiv \sqrt{\Omega_m(1+z)^3+\Omega_k(1+z)^2+\Omega_\Lambda}, \tag{4.1.3}$$

where Ω_m is the matter density parameter, and $\Omega_\Lambda$is the cosmological constant and the curvature of space is characterized by $\Omega_k = 1-\Omega_m-\Omega_\Lambda$. $F(z)$ accounts for the redshift evolution of $\xi(r)$ and in this case is assumed to be negligible within the samples considered here and set as $F(z) = 1$. This is the case of "co-moving clustering" where halos expand with the universe. Finally, $N(z)$ corresponds to the redshift distribution of the studied galaxy population which is described in the next section. In order to interpret the 3D clustering measurements of galaxies, redshift information is essential. One of the ingredients required for the formula, that is, the radial co-moving distance between observer and an object at redshift z, is computed using the relation (Hogg 1999)

$$r_c(z) = D_H \int_0^z \frac{dz'}{E(z')},$$

where the function $E(z)$ is defined in Equation 4.1.3, and D_H is the Hubble distance given by $D_H \equiv \frac{c}{H_0}$.

4.2 Redshift Distribution

The redshift distribution of passive galaxies $N(z)$, in the Deep and Wide fields was computed by Arcila-Osejo (2017). The author cross-correlated the PE-gzK samples for the D2 field (subset of COSMOS field) with the catalog of Muzzin et al. (2013) and Wide fields (W1 and W4) with that of Moutard et al. (2016). For the COSMOS field, the photometric redshifts were obtained by Muzzin et al. (2013) by fitting SEDs (Spectral Energy Distributions) of 30 available photometric bands using the photometric redshift code EAZY (Easy and Accurate Redshifts from Yale; Brammer et al. 2008). In the case of the Wide fields, Moutard et al. (2016) determined the redshifts by SED fitting of nine photometric bands with a standard χ^2 template fitting procedure (LE PHARE code; Arnouts et al. 2002).

In order to compute the magnitude dependence of the redshift distribution, Arcila-Osejo (2017) binned the photometric redshifts in magnitude steps of 0.5 width in K_s-band. After that, for each magnitude bin, the author constructed a 2D histogram of the redshift distribution in redshift bins of 0.1 for the passive galaxies and performed Gaussian kernel smoothing on the redshift distribution corresponding to different magnitude bins. The redshift probability distributions for different magnitude ranges as seen in Figure 4.1 were then obtained by normalising the distributions by the area under the curve. It is clearly seen

that the redshift distribution varies with magnitude: the peak of the fainter passive galaxies lies at higher redshifts compared to that of the brighter galaxies. Using these magnitude dependent redshift distributions in Equation 4.1.1, the correlation lengths were obtained for different K_s-mag selected passive gzK samples.

For comparison, we also computed values of r_0 assuming that the PE-gzK redshift distribution could be modelled as Gaussian distribution (Blanc et al. 2008). Their redshift distribution covers the range $1.4 \lesssim z \lesssim 2.5$ and the authors measured $N(z)$ using methods described in Rudnick et al. (2001) using a linear combination of templates followed by the correction for error involved in the estimation of photometric redshifts. Blanc et al. (2008) obtained the best Gaussian fits to the observed distribution with the redshift centered at $\bar{z}$=1.58$\pm$0.04 and width σ_z=0.17$\pm$0.06. The same $\bar{z}$ and σ_z values were assumed while computing r_o for all K_s magnitude subsamples.

4.3 Estimating the Correlation Length

Table 4.1 and Figure 4.2 summarise the values calculated for the correlation length r_0 measured within our K-magnitude selected samples using the Limber inversion. The two different r_0 values are derived from the two different redshift distributions: Arcila-Osejo (2017) and Blanc et al. (2008). The sources of error in the r_o calculations mainly come from two factors:

1. the uncertainty in the measurement of A_ω and

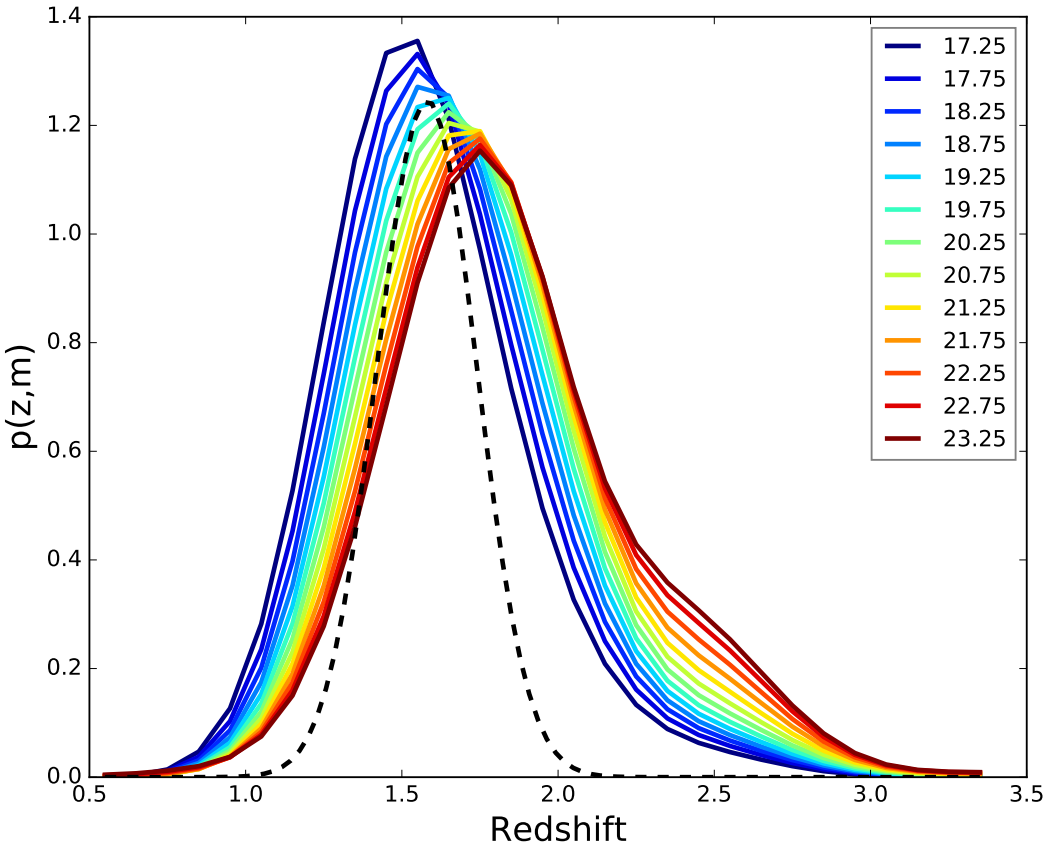

Figure 4.1: Redshift distribution function for the PE *gzK* galaxies by Arcila-Osejo (2017) peaking at $z \sim 1.6$. Figure taken from Arcila-Osejo (2017). The black dashed line represents the redshift distribution of passive galaxies by Blanc et al. (2008). The redshift distribution has been scaled to display it on the same plot.

Sample	K_s[AB mag]	$\log(M_*/M_\odot)$	$r_{0,Blanc}$	$r_{0,Osejo}$	$M_h(\log(h^{-1}M_\odot))$
Wide	19.25-19.75	11.49	21.54±3.73	30.93±5.36	$14.1^{+0.09}_{-0.06}$
	19.75-20.25	11.32	11.43±1.53	16.51±2.19	$13.7^{+0.08}_{-0.10}$
deep	20-21	11.15	7.56±0.14	10.99±0.21	$13.3^{+0.02}_{-0.03}$
	21-22	10.80	5.92±0.33	8.74±0.49	$13.0^{+0.04}_{-0.07}$
	22-23	10.45	3.46±0.22	5.22±0.33	$12.2^{+0.10}_{-0.10}$

Table 4.1: r_0 measurements using Limber inversion for the two different redshift distributions : Arcila-Osejo (2017) and Blanc et al. (2008) and DM halo mass measurements of PE-gzK galaxies.

2. the uncertainty in the redshift distribution $N(z)$.

The first major source of error is the uncertainty in the $\omega(\theta)$ measurements. The uncertainties in the Table 4.1 include uncertainties only due to statistical errors in the measurement of the angular correlation function. To see the effect of redshift distribution on the estimation of r_0, we have used two different redshift distributions to calculate the spatial correlation lengths for each K_s-magnitude selected sample. The systematic errors for the redshift determination are larger than the random errors for $N(z)$. The correlation length is affected by the median redshift and the width of the redshift distribution (McCracken et al. 2010). A larger width in the redshift distribution implies that projection effects are stronger and would result in a larger value of r_0 for a given time or underlying clustering.

It is clearly seen in Table 4.1 that the two different redshift distributions $N(z)$s give different r_0 results. Nevertheless, in both cases, the correlation lengths increase with the increase in K_s luminosity. It has been found that UMPEGs have larger correlation lengths as compared to the fainter passive galaxies indicating stronger clustering.

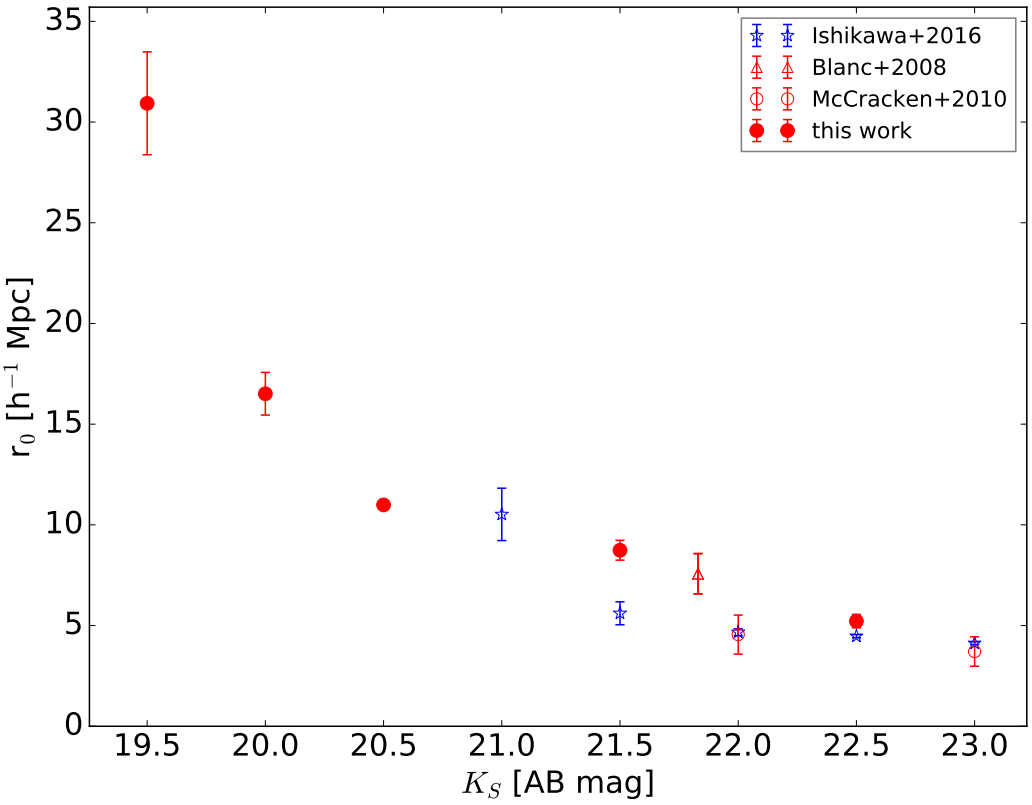

Figure 4.2: Comparison of the correlation lengths of our UMPEGs and PE gzK galaxies (red filled circles) with previous studies of Passive galaxies. Star forming BzK galaxies are shown as blue open stars. All other symbols show passive galaxies. All correlation lengths are in units of $h^{-1}Mpc$ where $h = 0.7$. Also shown are the results from several different authors that used BzK selection criterion to classify K_s selected galaxies as passive galaxies.

In the following analysis and discussion we will use the r_0 values obtained using the redshift distribution of Arcila-Osejo (2017). Figure 4.2 shows a comparison of our r_0 measurements with those of previous studies. For the low mass passive galaxies, the clustering strength is lower and agrees with the results of *BzK* selected passive galaxies by Blanc et al. (2008) at $z \sim 2$ and McCracken et al. (2010) at $z \gtrsim 1.4$. Our results show that the low mass PE *gzK* galaxies have clustering comparable to that of the star-forming galaxies of similar magnitude (Ishikawa et al. 2015). Our UMPEGs are more clustered than the low mass PE *gzK* galaxies. The clustering measurements in our work extend to much brighter passive galaxies than the previous works. The fact that our very luminous passive galaxies cluster more strongly than the fainter *BzK* galaxies (both passive and star-forming) suggests that they reside in much more massive dark matter halos.

Chapter 5

Mass Estimation of Dark Matter Halos of UMPEGs

In this Chapter, we present the clustering measurements of dark matter halos from Millennium *XXL* simulation and the halo correlation function that is an essential ingredient in determining masses of dark matter host halos of UMPEGs.

5.1 Brief Review of Millennium XXL Simulation

We compared the clustering results of UMPEGs with the clustering of dark matter halos in the Millennium-*XXL* Simulation (Angulo et al. 2012). The Millennium-*XXL* Simulation (*MXXL*) is a very large dark-matter high-resolution cosmological N-body simulation, extending the previous Millennium and Millennium-II simulations (Springel et al. 2005;

Boylan-Kolchin et al. 2009) and is the first multi-hundred billion particle run. It adopts a ΛCDM cosmology with WMAP-1 cosmological parameters with the total matter density being $\Omega_m = 0.25$ (in units of critical density); a cosmological constant being $\Omega_\Lambda = 0.75$ (resulting in a flat space geometry); the *rms* linear density fluctuation in 10.96 Mpc spheres, extrapolated to present epoch, $\sigma_8 = 0.9$; and $H_0 = 0.73 km/s/Mpc$. The dark matter in the *MXXL* is distributed through a volume which is equivalent to that of the whole observable Universe up to redshift 0.72 and is 216 times larger than that of the Millennium Simulation (*MS*) (Springel et al. 2005). The *MXXL* simulation follows a non-linear evolution of 6720^3=303,464,448,000 dark-matter particles with mass $6.2 \times 10^9 h^{-1} M_\odot$ within a cubic box of comoving length 3 h^{-1} Gpc. The number of particles in *MXXL* is significantly larger than that used in previous simulations of this type (Springel et al. 2005; Kim et al. 2009; Teyssier et al. 2009). The corresponding mass and force resolutions are $m_p = 8.456 \times 10^9 M_\odot$ (one particle mass) and a force softening of 13.7 kpc. This mass resolution is sufficient to identify host dark matter halos of galaxies with stellar mass greater than $1.5 \times 10^{10} M_\odot$ (De Lucia et al. 2006).

Dark matter simulation in the *MXXL* was done using the cosmological simulation code GADGET-3, an optimised version of GADGET-2 used for the *MS*. The simulation follows the gravitational growth traced by its DM particles and stores it as DM particle positions at 64 discrete time snapshots. The initial conditions are set at a starting redshift of $z = 127$ and the simulation evolves to $z = 0$ with 63 outputs corresponding to various redshifts.

GADGET-3 computes gravitational forces with a Tree-PM method by combining a particle-mesh (PM) scheme with a hierarchical tree method. Halo finding is a two-step procedure: At each snapshot, groups of more than 20 particles are identified as dark matter halos using a Friends-of-Friends (FoF) algorithm (Davis et al. 1985). After that, the SUBFIND algorithm (Springel et al. 2001) finds gravitationally bound subhalos within each FoF halo. The mass of the halo is defined as the conventional virial mass of a halo M_{200}, which is the $M_{200} = M(r < r_{200})$, the mass contained within a sphere of radius that encloses a mean density that is 200 times the critical density. The most massive halo at $z = 0$ has $M_{FoF} = 8.98 \times 10^{15} M_{\odot}$.

Our UMPEGs are expected to reside in massive halos and these objects being rare and unique can only be found in large volumes of the *MXXL*. For this reason The *MXXL* simulation is suited perfectly for studying the host halo masses of UMPEGs.

5.2 Clustering of Dark Matter Halos at $z \sim 1.6$

The mass ranges of the host DM halos of UMPEGs are estimated through comparison of the correlation length measurements with the clustering properties of dark matter halos from the *MXXL* cosmological simulation. The observed clustering properties of galaxies in a K_s-band luminosity range are matched to the clustering of dark matter halos in a certain mass range (Savoy et al. 2011; Kravtsov et al. 2004; Conroy et al. 2006). This method can be used to establish the mass scale of the DM halos hosting the UMPEGs.

We used the halo catalog of *MXXL* simulation at snapshot=36 which corresponds to

$z \sim 1.6$ which is the peak redshift of the UMPEGs being studied in this work. The spatial correlation function of the DM halos is a function of halo mass (Mo & White 1996), and for this reason the halo clustering is studied by selecting all the halos within a given mass range. We obtain large values for the correlation length r_0 for the UMPEGs in Chapter 4 and hence, these galaxies are expected to be hosted by massive dark matter halos. That is why we examine the clustering of massive DM halos with $10^{12} < M_{200} < 7 \times 10^{14} h^{-1} M_{\odot}$. The halo catalog is divided into 11 different samples with halo mass ranges of $12.1 < log(M_{200}) < 14.3$ in steps of $\Delta log(M_{200})$=0.2 where M_{200} is in units of $h^{-1} M_{\odot}$. We used a well-defined sample of halos whose distribution is known within the spatial resolution of the *MXXL* simulation. For every halo sample, we computed the number of pairs as a function of separation in comoving coordinates, r, relative to that of a random distribution to measure the two-point correlation function of the dark matter halos that is given by

$$\xi(r) = \frac{DD(r) - 2DR(r) + RR(r)}{RR(r)},$$

where $DD(r)$ is the unique number of halo pairs in the simulation with separations between $r - \delta r < r < r + \delta r$, $DR(r)$ refers to the number of pairs within the same separations between the halo catalog and the random catalog of halos and $RR(r)$ refers to number of random-random pairs within same range. $\xi(r)$ is computed in bins of constant logarithmic width $\Delta log(r) = 0.13$ with bins ranging from $log(r) = -1.25$ to $log(r) = 1.0$ where r is in $h^{-1}\, Mpc$. The lower value of r is chosen to be $\sim 0.5\ h^{-1}\, Mpc$ in order to avoid the

Halo mass($\log(h^{-1}M_{\odot})$)	N_h	$N[10^{-6}\mathrm{Mpc}^{-3}]$
14.1-14.3	8037	0.10
13.9-14.1	29842	0.37
13.7-13.9	89994	1.14
13.5-13.7	231157	2.94
13.3-13.5	524974	6.67
13.1-13.3	1083218	13.8
12.9-13.1	2076290	26.4
12.7-12.9	3772890	47.9
12.5-12.7	6537824	83.1
12.3-12.5	11003920	139.8
12.1-12.3	18171840	230.8

Table 5.1: Table showing the number of halos at $z = 1.6$ in different halo mass ranges of $12.1 < log(M_{200}) < 14.3$ in steps of $\Delta log(M_{200})$=0.2 where M_{200} is in units of $h^{-1}M_{\odot}$ in *MXXL* simulation. The total number of halos in the halo catalog is 56,406,021.

clustering effects of multiple subhalos within the same halo. Changing the *r* binning for the halos does not affect the results.

As seen in Figure 5.1, halo clustering clearly depends on halo mass. The most massive halos in the range $log(h^{-1}M_{\odot})$=14.1-14.3 cluster more strongly than the low mass ones, which are also more abundant. The halos form from small perturbations in early universe which grow with time. The reason for the mass dependence is that the high mass halos are rare, and only arise in regions with excess material and with strong positive perturbations on even larger scales (Kaiser 1984). These perturbations push the smaller halos and cause them to collapse. However, since they are large scale, this causes an excess of these halos in the general neighborhood and hence, these systems are strongly clustered. Whereas, large scale perturbations are not needed to form the low mass halos, low mass halos have lower

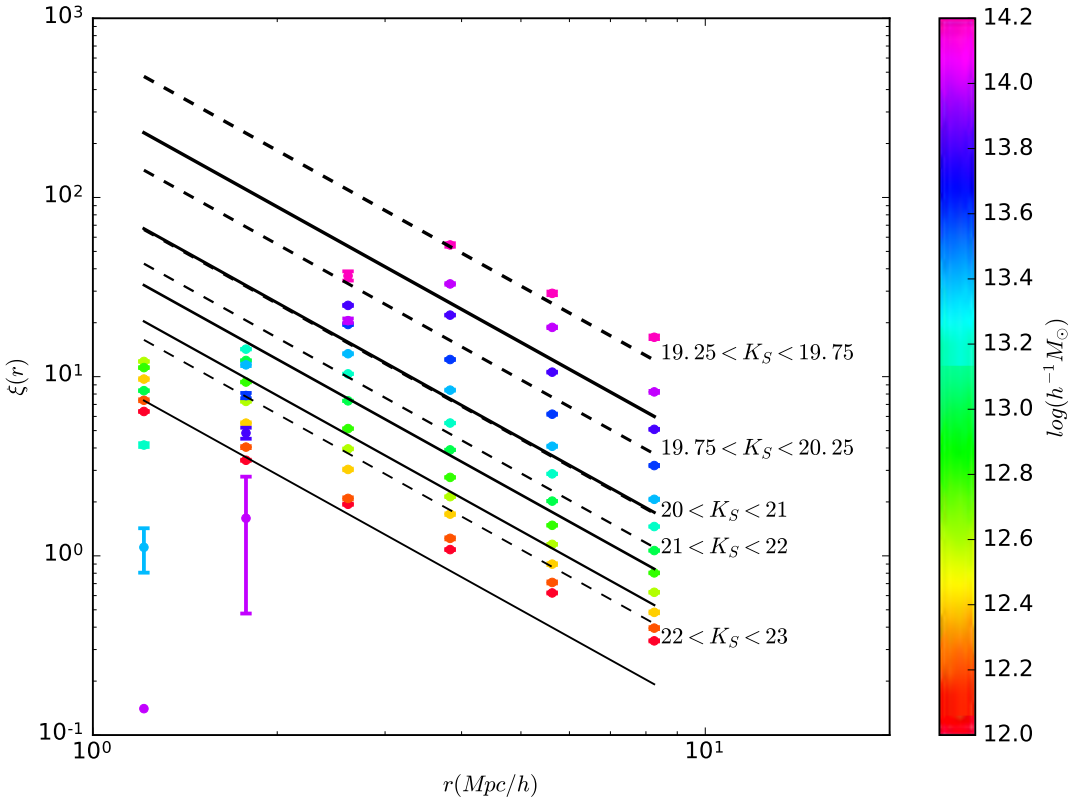

Figure 5.1: Comparison of spatial correlation lengths of K_s-band luminosity ranges (summarised in Table 4.1) to the correlation results from the *MXXL* simulation halo catalog for different mass ranges. The different mass ranges correspond to $log(h^{-1}M_\odot)$=12.1-12.3, 12.3-12.5, 12.5-12.7, 12.7-12.9, 12.9-13.1, 13.1-13.3, 13.3-13.5, 13.5-13.7, 13.7-13.9, 13.9-14.1, 14.1-14.3 with purple points showing the spatial correlation function for the least massive and red for the most massive halos. Dashed lines represent the results of our observed galaxy clustering measurements with redshift distribution computed by Arcila-Osejo (2017). Solid lines represent the clustering measurements done using the Blanc et al. (2008) redshift distribution.

clustering. It is also observed that as the halo mass increases, there are no halo pairs at smaller separations and this trend is more prominent for the most massive halos as seen in Figure 5.1. The abundances of massive halos should decrease strongly at high redshifts because massive halos are rare objects at early times. Simulations encompassing the same volume as our survey suffer from poor statistic because massive halos are rare.

In order to relate the spatial correlation function of halos to our observations of UMPEGs, the observed correlation function of these galaxies (shown as solid and dashed lines in Figure 5.1) are plotted over the *MXXL* simulation results (solid points in Figure 5.1). Clearly, there exists a relationship between halo mass and K_s-magnitude selected passive galaxies where K_s-magnitude is related to stellar mass of these galaxies. The massive halos tend to host the brighter and more massive passive galaxies. The $\xi(r)$ of these galaxies follows a power law with different clustering lengths r_o (listed in Table 4.1) for different K_s magnitude selected samples, with γ fixed at a value of 1.92. This slope provides a proper fit to the spatial correlation function of massive halos. In contrast, the same slope does not fit the simulation results for the lower mass halos (with $log(M_{200}) < 12.5$).

5.3 Determination of the Dark Matter Halo Masses

After measuring the spatial correlation function of the DM halos, we plot the ξ values at a fixed value of r, corresponding to different halo mass ranges in order to infer the dark matter halo masses of UMPEGs directly. Red points and the connecting line in Figure 5.2 show

the relation between clustering strength at $r = 8.25\ h^{-1}$Mpc and DM halo mass in *MXXL* at $z \sim 1.6$. Also plotted (gray lines) are the corresponding $\xi(r)$ values for our observed PE *gzK* samples.

It is seen that the brighter passive galaxies in the range $19.25 < K_s < 19.75$ reside in the most massive halos in the mass range $13.9 < log(M_{200}) < 14.2$ where M_{200} has the units of $h^{-1}M_{\odot}$.

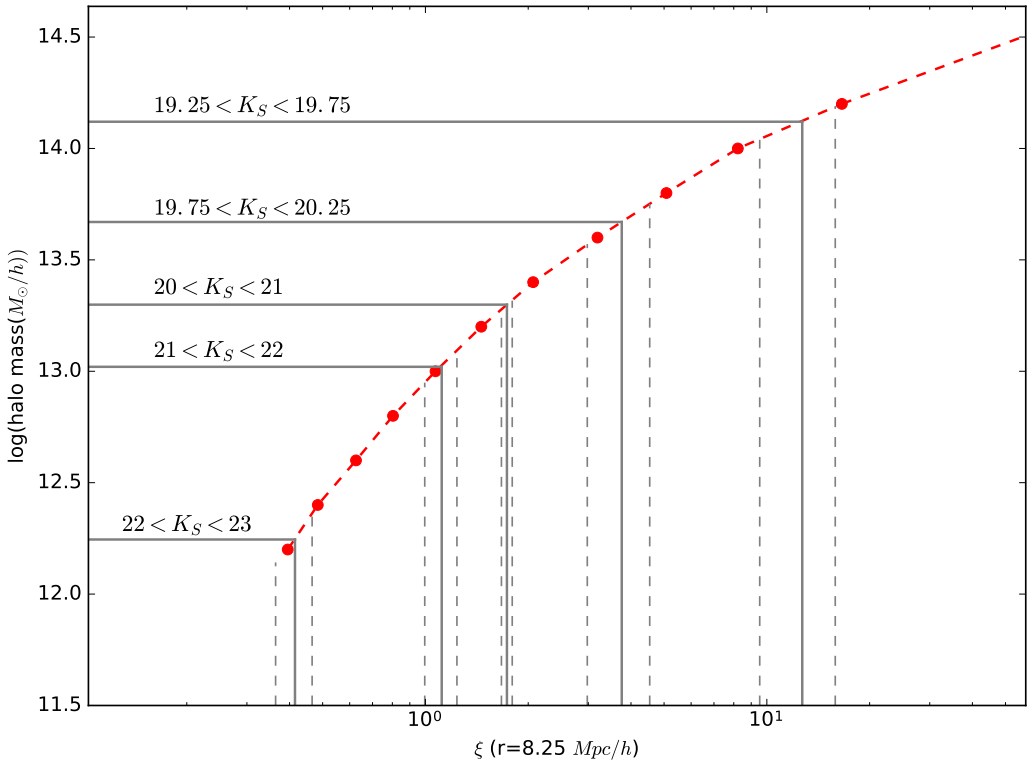

Figure 5.2: Halo mass as a function of correlation function of the dark matter halos from the *MXXL* simulation at $z \sim 1.6$. The points represent the correlation function at a fixed spatial value $r = 8.25\ h^{-1}Mpc$ for the different mass ranges of the halos. The dashed red line is a fit to the data and vertical solid gray lines correspond to the spatial correlation function of the PE-*gzK* galaxies binned according to the K_s-band luminosity with extreme right vertical line representing the correlation function of the most massive passive galaxies and the leftmost being the least massive ones. The dashed lines represent the error bars on the spatial correlation function measurements.

Chapter 6

Groups of Passive Galaxies

In this chapter, we study the clustering of the $z \sim 1.6$ PE-*gzK* galaxy groups to determine their DM halo masses. It is important to analyze clustering of dense environments of the passive galaxies in order to study the evolutionary connections between these groups and highly clustered galaxy populations at high/low redshift such as the UMPEGs at $z \sim 1.6$ and protoclusters at $z \sim 0$.

These groups represent over-densities of the massive PE-*gzK* galaxies and were selected by Arcila-Osejo (2017) from the PE-*gzK* catalog by creating Gaussian density maps where each PE-*gzK* galaxy with $K_s < 20.5$ is modeled by a simple analytic 2D Gaussian profile centered at the position of the PE-*gzK* galaxy and with a FWHM of 1.5 physical Mpc and peak value of 1. These Gaussian profiles are added and every pixel in the resulting density maps is the sum of all the pixels after the superposition of the Gaussians. After

that, Arcila-Osejo (2017) identified the "hot spots" based on pixels whose value is above or equal to 2.7. Then, the author selected objects in a well defined closed contour with 70% level, surrounding these "hot spots" in order to detect the objects belonging to the group. The groups are the ones in which there are at least three objects with $K_s < 20.5$ in close proximity to each other. There was only one group identified in each of the Deep fields D1, D2 and D4. The groups in the Deep fields were not included in the clustering measurements for the passive galaxy groups. There were a total of 31 such groups identified by Arcila-Osejo (2017) as shown in Figure 6.1 in the Wide fields: 15 in W1 and 16 in W4.

In order to compare the clustering of the UMPEGs with the groups, the angular correlation function for these massive groups was computed. Knowing the distribution of these groups in angular space, a similar technique as mentioned in section 3.3 is followed to measure the correlation function. In this case, each group is considered to be a single data-point and the number of data-data, data-random, and random-random pairs are counted with angular separations between $\theta - \Delta\theta/2 < \theta < \theta + \Delta\theta/2$. The angular range chosen for the groups is different as compared to the galaxies as we did not find any pairs of groups at separations closer than 3.37 arcmin. $\omega(\theta)$ is computed in angular distance bins of constant logarithmic width $\Delta log\theta = 0.12$ with bins ranging from $log(\theta) = -1.25$ to $log(\theta) = 1.0$ where θ is in degrees. The θ bin size for these groups were chosen to achieve an optimal S/N ratio. We found that the amplitude for these groups is higher than that of the UMPEGs as seen in Figure 6.2.

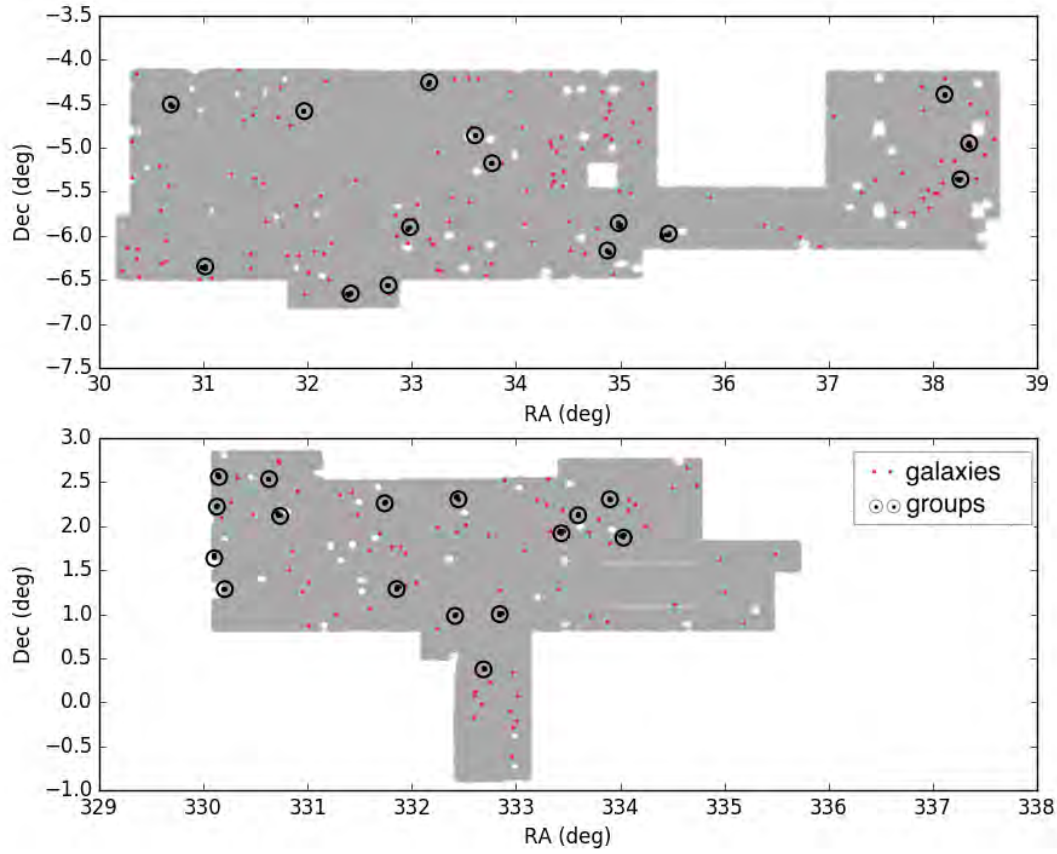

Figure 6.1: Groups of passive *gzK* galaxies with $K_s < 20.5$ in the Wide fields W1 (upper sub panel) and W4 (lower sub panel) are represented as black circles. UMPEGs with $K_s < 19.75$ are shown as red points. The gray region shows the layout of the fields along with the boundaries of the survey area.

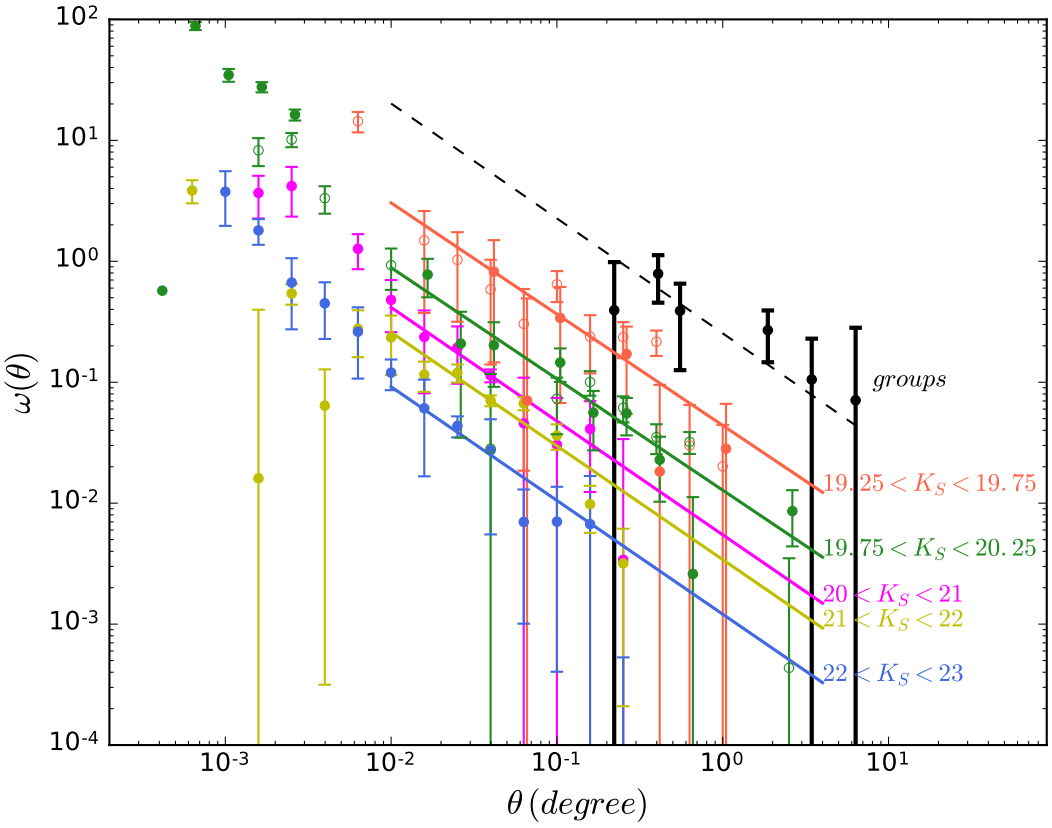

Figure 6.2: The angular correlation measurements (black points) of groups compared with that of the PE *gzK* galaxies. The black dashed line is a fit to the correlation function measurements for the groups selected by Arcila-Osejo (2017).

Sample	K_s[AB]	Log($M_*/M_\odot$)	$A_\omega/10^{-3}\mathrm{deg}^{1-\gamma}$	$r_{0,Osejo}$	$M_h(\log(h^{-1}M_\odot))$
groups	< 20.5		215.45± 30.30	54.85±8.24	$14.6^{+0.05}_{-0.05}$
UMPEGs	< 19.75	11.49	41.88± 6.64	30.93±5.36	$14.1^{+0.09}_{-0.06}$

Table 6.1: Comparison of the clustering amplitude, correlation lengths and DM halo masses for the PE *gzK* groups with UMPEGs. The estimated parameters are with power law slope $\gamma = 1.92$.

After measuring the angular correlation function, the next step is to compute the spatial correlation function for the groups using the Limber inversion (4.1.1). However, the redshift distribution for the groups is likely to be different from that of the passive galaxy population. This is because identifying a group requires identifying not just one galaxy but multiple group members at the same redshift. As the group members have been selected from the passive galaxy catalog with the *gzK* criteria, the probability of finding the group containing the galaxies that already have a low detection probability, is even lower. On the other hand, the probability of detecting groups of galaxies with high individual detection probability is also high. In addition, the selection of group members with $K_s < 20.5$ has confined the group redshift to be in a narrow redshift range, and gives no flexibility for other galaxies of sufficiently different photometric redshifts to join in the narrow redshift distribution.

Using the redshift distribution of the passive galaxy population, we obtain a very high r_0 which results in unphysical halo masses for these massive groups. The difficulty is to have a realistic estimation of the redshift distribution for the group sample. So, we computed r_0 for the modified redshift distributions taking different cases into account. We consider the case where the true number of galaxies in a group is 3 and all the group members need to be

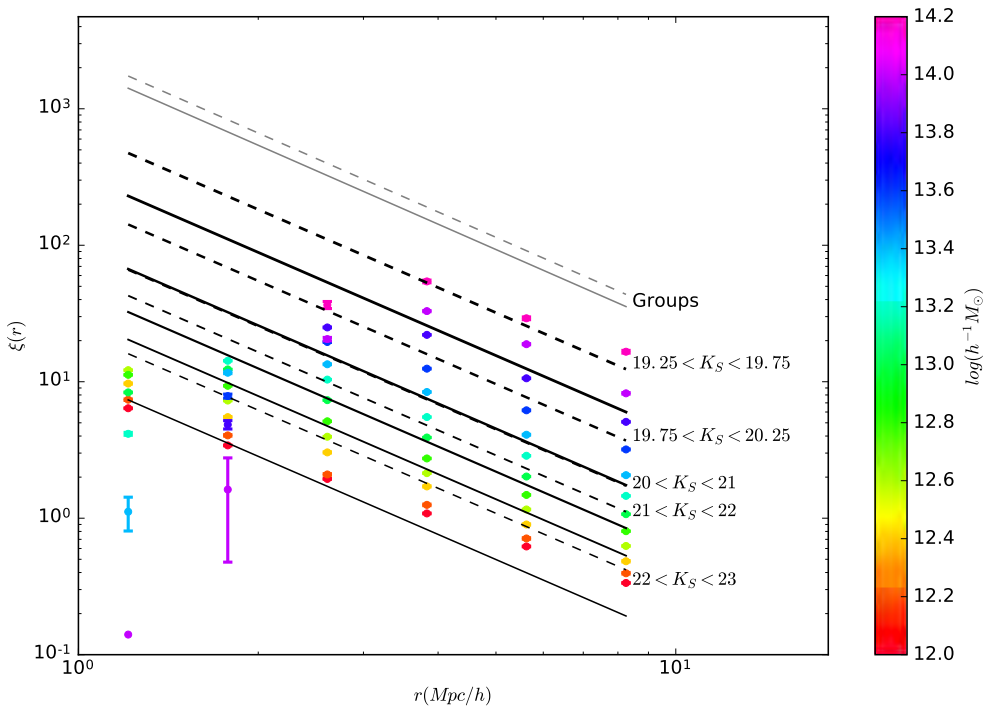

Figure 6.3: Comparison of spatial correlation lengths of groups of PE *gzK* galaxies to the correlation results from the *MXXL* simulation halo catalog for different mass ranges. Same as in Figure 5.1 except the gray dashed and solid line represents the spatial clustering of the groups computed using redshift distribution by Arcila-Osejo (2017) and Blanc et al. (2008) respectively.

detected in order to declare it a group. This results in a narrower redshift distribution and using such a narrow redshift distribution, the resulting correlation length is 54.85 $h^{-1}\,Mpc$. For another case, where true number of galaxies in a group is 5 and 3 group members need to be detected, we obtained $r_0 = 69.02\,h^{-1}\,Mpc$. Increasing the true number of galaxies in a group broadens the redshift distribution and results in larger values for r_0 equivalent to that obtained from using the redshift distribution of the passive population. The correlation length measurements for the groups are listed in Table 6.1. They range from 54.85 to 78.83 $h^{-1}\,Mpc$ corresponding to the two extremes of the redshift distributions. In comparison with the clustering results of the most massive passive galaxies with that of the groups, the groups are more strongly clustered.

The next step was to compare the correlation length measurements of these massive groups to the clustering measurements of the DM halos in the *MXXL* simulation. The halo masses were determined in a similar manner to that described in Chapter 5.

Using the modified narrow redshift distribution, we obtained very massive halos masses of $\log(M_{200}) \sim 14.6$ for these groups at $z \sim 1.6$, where M_{200} is in units of $h^{-1}M_{\odot}$. As mentioned earlier, the correlation length estimation depends on the redshift distribution which affects the measurements of DM halo masses. We cannot estimate the redshift selection function $N(z)$ for these groups with great precision which means that our halo mass measurements are also quite imprecise.

The DM halo masses are estimated considering the dependence of clustering measure-

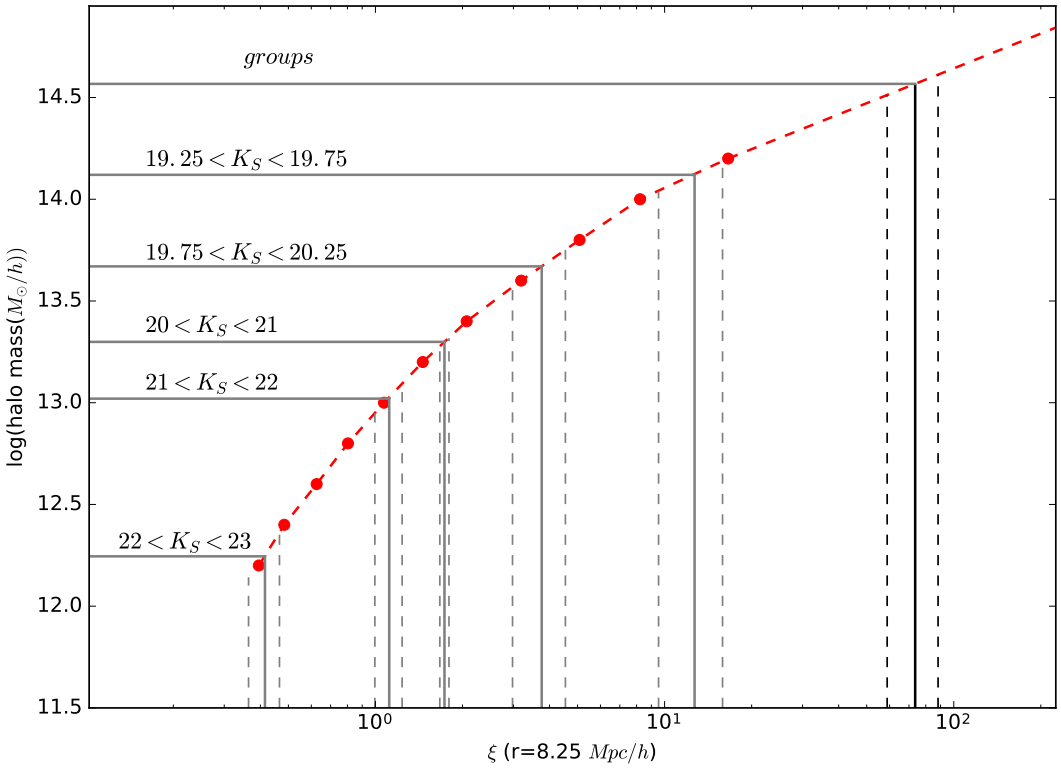

Figure 6.4: Halo mass as a function of correlation function of the dark matter halos from the *MXXL* simulation. Same as in Figure 5.2 except the black vertical line shows the spatial correlation function of the groups at fixed value of $r = 8.25\ h^{-1}$Mpc. The black dashed lines represent the error bars on the spatial correlation function measurements.

ments on the halo mass. An assumption used to interpret the galaxy clustering measurements is that the galaxies populate halos according only to the halo mass. However, this assumption is not precise enough and the halo clustering depends on halo properties other than halo mass; this is referred to as "assembly bias" (Gao & White 2007; Zehavi et al. 2017). As this effect is not taken into account here, the high masses of the halos could be indicative of limitations in our method of estimating dark matter mass from clustering measurements.

In summary, the groups seem to be very highly clustered suggesting that they are associated with very massive halos of mass $10^{14.5}\ h^{-1}M_{\odot}$. However, clustering measurements of the massive groups are sensitive to a number of uncertain effects, so our conclusions remain tentative.

Chapter 7

Interpretation

7.1 Comparing UMPEGs with other Galaxy Populations

After obtaining the correlation length r_0 for passive gzK galaxies at $z \sim 1.6$, we compared the correlation length measurements of these galaxies with clustering measurements of different populations of galaxies from other surveys.

The r_0 for our less massive PE gzK galaxies at $z \sim 1.6$ is comparable to the r_0 measured for the BzK galaxies and EROs at $z \sim 2$. The correlation length r_0 of the UMPEGs is larger than those of other populations of comparable redshift as seen in Figure 5. Comparing with the theoretical predictions for different halo masses, this is consistent with the idea that the UMPEGs are associated with most massive DM halos at high redshifts, halos more massive than the ones hosting any other galaxy type at comparable redshifts. It is seen that based

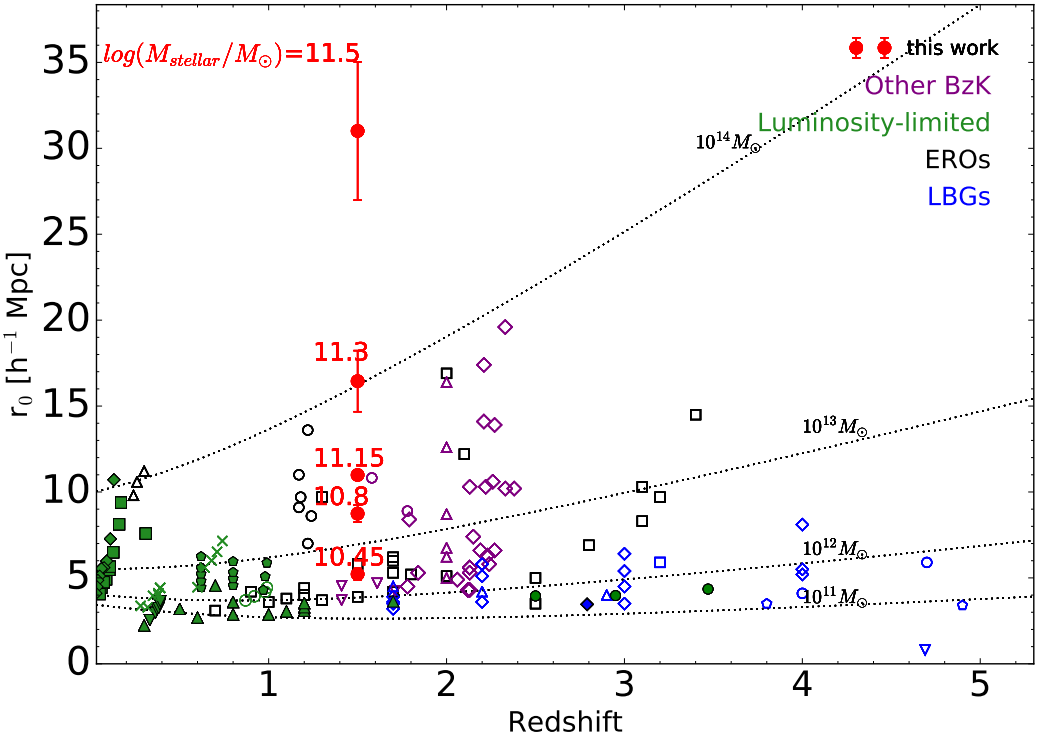

Figure 7.1: Comoving correlation length r_0 of the UMPEGs in contrast to other populations of low and high-redshift galaxies from different surveys. Figure based on the compilation of Durkalec et al. (2015). Black curves show dependence of r_0 for halos at fixed mass as a function of redshift from Press-Schechter analysis of Mo & White (2002). Red solid points represent the gzK-selected passive galaxies from this work in different K_s magnitude bins. Different colors indicate different types of objects selected using different techniques as indicated in the top right corner. Open symbols indicate measurements based on photometric data, while filled symbols are for measurements from spectroscopic data (except for our points). Blue: LBG galaxies (open squares – Foucaud et al. 2003; open circles – Ouchi et al. 2004; open triangles – Adelberger et al. 2005; open reversed triangles – Kashikawa et al. 2006; open diamonds – Savoy et al. 2011; filled diamonds – Bielby et al. 2013; open pentagon – Barone-Nugent et al. 2014). Purple: BzK galaxies (open circles – Blanc et al. 2008; open triangles – Hartley et al. 2010; open reversed triangle – McCracken et al. 2010; open diamonds – Lin et al. 2012). Green: galaxy samples from surveys limited in luminosity (filled circles – Durkalec et al. 2015; filled squares – Norberg et al. 2002; open circles – Coil et al. 2006; filled triangles – Le Févre et al. 2005; filled reversed triangles – Pollo et al. 2006, filled diamonds – Zehavi et al. 2011; filled pentagons – Marulli et al. 2013; crosses – Skibba et al. 2014). Black: EROs or massive red galaxies (open squares – Daddi et al. 2003; filled squares – Zehavi et al. 2011; open circles – Brown et al. 2008).

on the Press-Schechter models (Mo & White 2002) for r_0 of the DM halos, the predicted halo masses of an average mass of $\sim 10^{14} M_{\odot}$ for the passive galaxies have same r_0 as the galaxies at $z \sim 1.6$. This is only a first approximation as this is based on simplified assumptions of of the Press & Schechter theory, which further needs refinement, but is in very good agreement with the halo mass of $10^{13.8} M_{\odot}$ we get from the *MXXL* simulation.

7.2 Stellar Mass - Halo Mass Relation

We investigated the relationship between halo mass and stellar mass of the passive *gzK* galaxies, representing SHMR = $M_{stellar}/M_h$ as a function of halo mass. The mass ratio between stellar content in the galaxy and its host DM halo represents the efficiency with which a galaxy can form and accrete stars and thus relates directly to galaxy formation.

In Figure 7.2, we compared our observed SHMRs as a function of halo mass at $z \sim 1.6$ with the results of numerical simulations by Moster et al. (2013) which are represented as black lines. According to the model proposed by Moster et al. (2013), the SHMR reaches a peak efficiency at halo mass $\sim 10^{12.5}\ M_{\odot}$, and at high halo masses the relation turns over to lower values. The shape of the SHMR results due to different physical mechanisms that prevent star formation in the DM halo. Each process contributes differently at different mass. In the case of the low mass halos, feedback from supernova-driven winds (Larson 1974; Dekel & Silk 1986) is responsible for lowering the star forming efficiency. In contrast, processes such as feedback from active galactic nuclei (AGN; Springel et al. 2005;

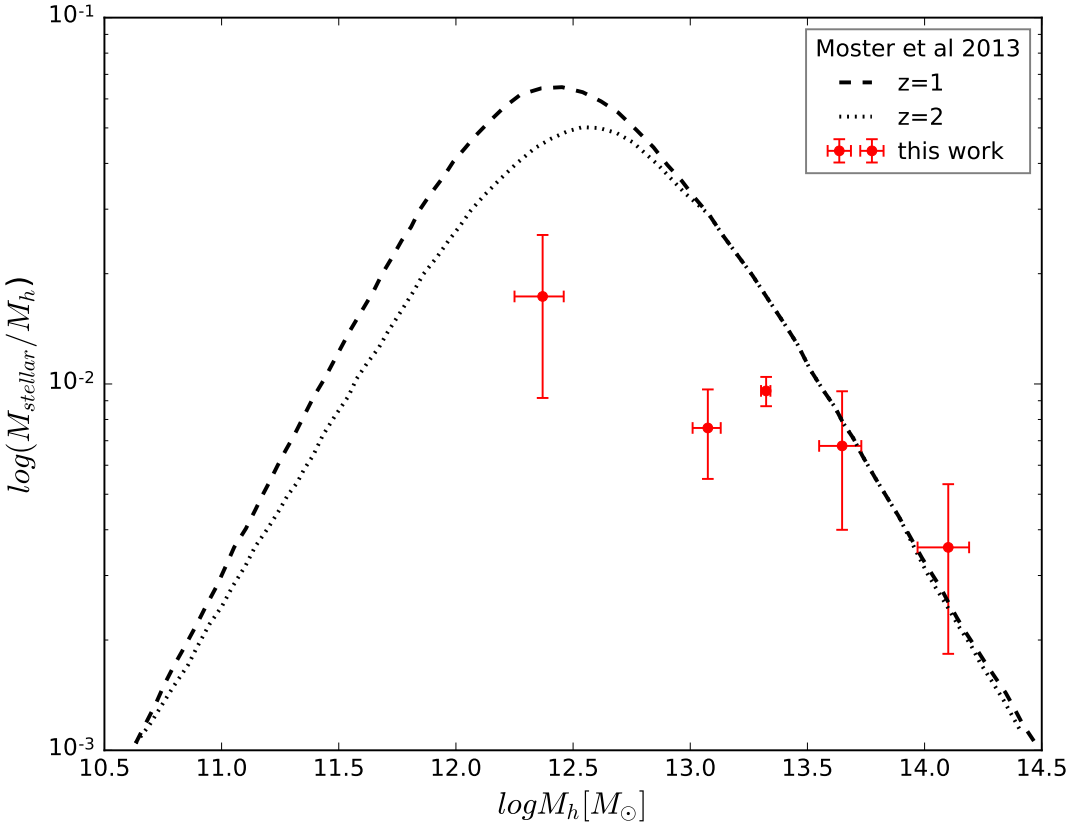

Figure 7.2: Stellar mass-halo mass ratio (SHMR) for different stellar mass selected passive *gzK* sub-samples at $z \sim 1.6$ (filled red circles) as a function of halo mass. The measurements from this work are compared with the model predictions by Moster et al. (2013) at $z = 1$ and $z = 2$ represented by dashed and dotted lines respectively.

Bower et al. 2006; Croton et al. 2006) and gravitational heating dominate in the massive halos. As the UMPEGs are very massive, we are able to probe the high mass end of the stellar-to-halo mass relation. Our data show that the $\log(M_{stellar}/M_h)$ ranges from -2.5 for the high mass sample to -1.78 for the less massive passive galaxies at $z \sim 1.6$.

The SHMR for the most massive passive *gzK* galaxies agrees well with the model predictions by Moster et al. (2013). However, for the lower mass passive galaxies, the model predictions are $\sim 3-4$ times higher than ours in the mass range of $M_{200} < 2 \times 10^{13} M_{\odot}$. This disagreement could be linked to an inefficient AGN-feedback and feedback from supernovae at intermediate masses in the Moster models. Figure 7.2 shows that the UMPEGs contain most of the stellar mass that resides in their DM halos. On the other hand, for the less massive passive *gzK* galaxies, the models with one galaxy per halo may be too simplistic and this is not generally the case. In a more realistic scenario, one halo can be occupied by more than one galaxy. This is supported by detection of the one-halo term in the angular correlation measurements at small separations (see Section 3.3.4). It is likely that this missing mass is attributed to the fact that there is more than one galaxy (could be low-mass passive or star-forming and thus not in our sample) hosted by the same DM halo that contributes to the stellar mass in the halo.

Another explanation of the discrepancy at intermediate masses is that of biasing. If the clustering of the PE-*gzK* galaxies is biased with respect to DM then our measurements could be overestimating halo masses. Halo masses lower by a factor of $\sim 3-4$ would bring

our data into agreement the theoretical predictions.

7.3 Evolution of UMPEGs to $z \sim 0$

We studied the connection between observed galaxies and the simulated DM halos using a variation of a technique called abundance matching (Behroozi et al. 2010; Conroy & Wechsler 2009; Guo et al. 2010). This technique does not require finding the masses of galaxies or halos. Instead, it makes use of galaxy stellar mass function (abundance of galaxies by stellar mass) and halo mass function (abundance of halos by mass). The idea is to solve N(>x) = N($> M_{halo}$), i.e. matching cumulative distributions of the observed galaxy property, x, with the predicted one for halo masses. With these relations in hand, the technique matches the galaxies and halos in a one-to-one manner, assuming that the most massive galaxy is hosted by most massive halo, the second most massive galaxy is hosted by second most massive halo, and so on.

At $z = 1.6$, the halos with mass greater than $1.6 \times 10^{14} M_{\odot}$ (which are the halo masses of our UMPEGs) have a number density of $1.5 \times 10^{-7} Mpc^{-3}$ and the UMPEGs at this redshift have a number density of $1.9 \times 10^{-8} Mpc^{-3}$. There are about 8 times more halos at this redshift than UMPEGs in halos of the same mass. That suggests that every 7 in 8 of these most massive halos are likely to contain something other than an UMPEG, such as a group of PE *gzK* galaxies or massive star-forming galaxies. There is a possibility that these UMPEG-less halos are associated with a galaxy or galaxies in the evolutionary stage

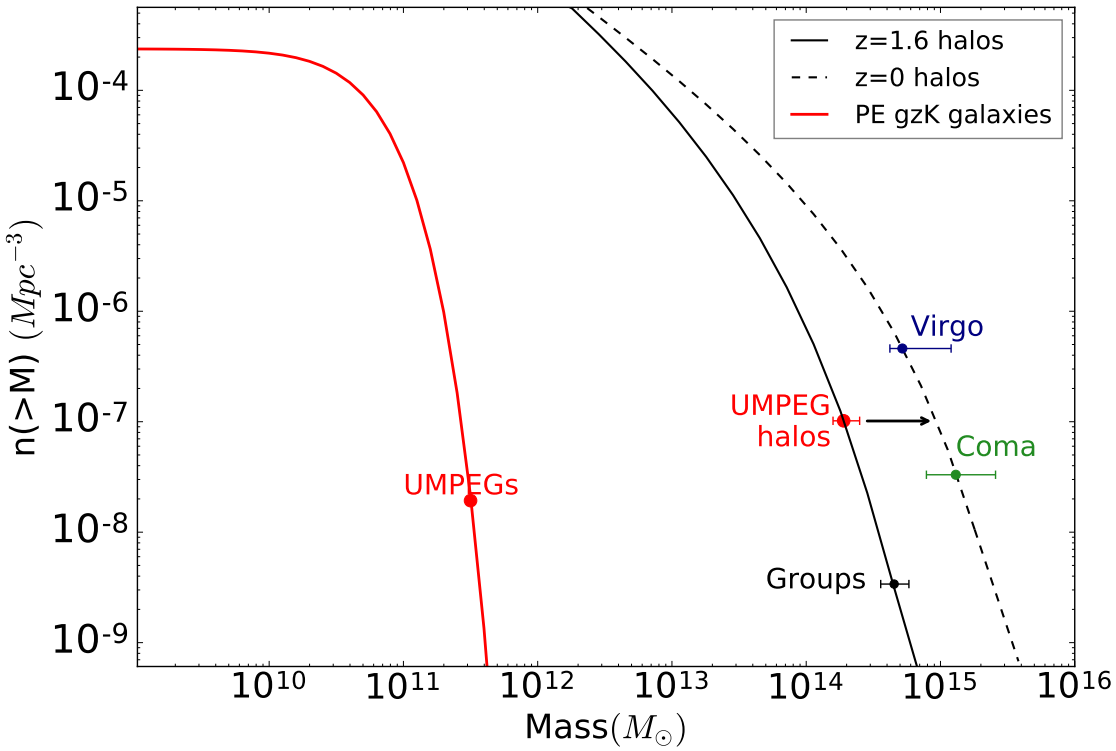

Figure 7.3: Comparison of the cumulative number density of the PE *gzK* galaxies and DM halos for the *MXXL* simulation. The black solid and dashed lines show the cumulative number densities for the halos at $z = 1.6$ and $z = 0$ respectively. The arrow indicates the evolutionary path of the halos of the UMPEGs to $z = 0$. The mass of the Virgo cluster and the Coma cluster are shown on $z = 0$ halo curve. The horizontal error bars represent the range of Virgo mass estimation from different studies (McLaughlin 1999; Ferrarese et al. 2012; Urban et al. 2011) and Coma mass from (Geller & Huchra 1989; Kubo et al. 2007; Gavazzi, R. et al. 2009).

of becoming an UMPEG.

The UMPEGs reside in the most massive halos at $z = 1.6$. Looking at number densities of UMPEG halos in Figure 7.3, we conclude that these halos of the UMPEGs may eventually become the halos of the massive clusters such as Virgo and Coma by the present day. The stellar mass of the central galaxy NGC 4486 (Chabrier (2003) IMF) in the Virgo cluster is $10^{11.57}\ M_{\odot}$ (Forte et al. 2013). For the two main central galaxies NGC 4874 and NGC 4889 (independent of IMF) in the Coma cluster, the stellar masses are $10^{11.98}\ M_{\odot}$ and $10^{12.18}\ M_{\odot}$ respectively (Veale et al. 2017). Since UMPEGs have stellar mass $> 10^{11.5}\ M_{\odot}$, it is plausible that the UMPEGs will become the massive central galaxies of these clusters by growing through different mechanisms (e.g, minor mergers). In this scenario, UMPEGs are the progenitors of some of the massive central galaxies in the clusters in the local Universe.

As discussed in Chapter 6, the PE *gzK* galaxy groups reside in the most massive halos with mass $10^{14.6}h^{-1}M_{\odot}$. These halos are also marked on the $z = 1.6$ halo number density curve in Figure 7.3. These groups could be another channel for the formation of present day massive clusters.

Chapter 8

Conclusions

Using a sample of the massive passively evolving *gzK* galaxies at $z \sim 1.6$ over an unprecedented large area, we used clustering measurements to determine the angular and spatial correlation functions as a means of linking the properties of the galaxies to their dark matter halos. The two-point angular correlation functions for the passive *gzK* galaxies were presented, together with the best power-law fits. Using the observed redshift distributions of these galaxies, we deprojected the spatial correlation function from the angular one and estimated correlation lengths for the UMPEGs as well as for lower mass PE *gzK* galaxies. By comparing our clustering measurements to those of the DM halos from simulations, we estimated the halo masses for the PE *gzK* galaxies.

In this work, our primary results are as follows:

1. We derived the correlation length r_0 for the UMPEGs and found that the UMPEGs

have very strong clustering, larger than any other galaxy population at high redshift. We also confirm previous findings that the correlation length for the clustering of lower mass PE *gzK* galaxies is dependent on the K_s magnitude. In addition to the luminosity dependence, there is a clear enhancement in the clustering of the passive galaxies at small scales as also found by Sato et al. (2014). This is indicative of multiple passive *gzK* galaxies residing in the same dark matter halo.

2. Using the clustering measurements of the DM halos from the Millennium XXL simulation (Angulo et al. 2012), we determined the halo masses and conclude that the UMPEGs inhabit the most massive dark matter halos of mass $\sim 10^{14.1} h^{-1} M_{\odot}$ at their epoch.

3. The UMPEG halos will eventually grow to become halos of mass $\sim 10^{15} M_{\odot}$ by the present day which is comparable to massive clusters such as Virgo and Coma at $z = 0$. The descendants of UMPEGs are likely to reside in massive clusters today, and may be the progenitors of some of the massive central cluster galaxies.

4. We studied the SHMR for the massive passive galaxies. Our measurements for the massive end are in good agreement with the Moster models (Moster et al. 2013). However, there is a discrepancy at lower masses that could be due to the inefficient feedback in the models as compared to PE *gzK* galaxies or multiple galaxies (passive or star-forming) within the same halo.

5. We also investigated the clustering of $z \sim 1.6$ groups of passive galaxies and obtained the correlation length ranging from 54.85 to 78.83 $h^{-1} Mpc$ for different redshift distribu-

tions. These groups are very highly clustered, suggesting that they are associated with very massive halos, $M_h \sim 10^{14.6}\ h^{-1}M_\odot$, and they may also become the centers of $z = 0$ massive clusters.

Bibliography

Adelberger, K. L., Steidel, C. C., Pettini, M., et al. 2005, The Astrophysical Journal, 619, 697

Angulo, R. E., Springel, V., White, S. D. M., et al. 2012, Monthly Notices of the Royal Astronomical Society, 426, 2046

Arcila-Osejo, L. 2017, PhD thesis(in preparation), Saint Mary's University

Arcila-Osejo, L. & Sawicki, M. 2013, Monthly Notices of the Royal Astronomical Society, 435, 845

Arnouts, S., Moscardini, L., Vanzella, E., et al. 2002, Monthly Notices of the Royal Astronomical Society, 329, 355

Barone-Nugent, R. L., Trenti, M., Wyithe, J. S. B., et al. 2014, The Astrophysical Journal, 793, 17

Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, The Astrophysical Journal, 717, 379

Berlind, A. A. & Weinberg, D. H. 2002, The Astrophysical Journal, 575, 587

Bertin, E. & Arnouts, S. 1996, Astron. Astrophys. Suppl. Ser., 117, 393

Bielby, R., Gonzalez-Perez, V., McCracken, H. J., et al. 2014, Astronomy & Astrophysics, 568, A24

Bielby, R., Hill, M. D., Shanks, T., et al. 2013, Monthly Notices of the Royal Astronomical Society, 430, 425

Bielby, R., Hudelot, P., McCracken, H. J., et al. 2012, Astronomy & Astrophysics, 545, A23

Blanc, G. A., Lira, P., Barrientos, L. F., et al. 2008, The Astrophysical Journal, 681, 1099

Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, Monthly Notices of the Royal Astronomical Society, 370, 645

Boylan-Kolchin, M., Springel, V., White, S. D. M., Jenkins, A., & Lemson, G. 2009, Monthly Notices of the Royal Astronomical Society, 398, 1150

Brainerd, T. G., Blandford, R. D., & Smail, I. 1996, The Astrophysical Journal, 466, 623

Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, The Astrophysical Journal, 686, 1503

Brown, M. J. I., Dey, A., Jannuzi, B. T., et al. 2003, The Astrophysical Journal, 597, 225

Brown, M. J. I., Zheng, Z., White, M., et al. 2008, The Astrophysical Journal, 682, 937

Chabrier, G. 2003, Publications of the Astronomical Society of the Pacific, 115, 763

Cimatti, A., Daddi, E., Mignoli, M., et al. 2002, Astronomy & Astrophysics, 381, L68

Coil, A. L., Davis, M., Madgwick, D. S., et al. 2004, The Astrophysical Journal, 609, 525

Coil, A. L., Newman, J. A., Cooper, M. C., et al. 2006, The Astrophysical Journal, 644, 671

Coil, A. L., Newman, J. A., Croton, D., et al. 2008, The Astrophysical Journal, 672, 153

Conroy, C. & Wechsler, R. H. 2009, The Astrophysical Journal, 696, 620

Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2006, The Astrophysical Journal, 647, 201

Contreras, S., Baugh, C. M., Norberg, P., & Padilla, N. 2015, Monthly Notices of the Royal Astronomical Society, 452, 1861

Croton, D. J., Springel, V., White, S. D. M., et al. 2006, Monthly Notices of the Royal Astronomical Society, 365, 11

Daddi, E., Cimatti, A., Renzini, A., et al. 2004, The Astrophysical Journal, 617, 746

Daddi, E., Röttgering, H. J. A., Labbé, I., et al. 2003, The Astrophysical Journal, 588, 50

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, The Astrophysical Journal, 292, 371

Davis, M. & Peebles, P. J. E. 1983, The Astrophysical Journal, 267, 465

De Lucia, G., Springel, V., White, S. D. M., Croton, D., & Kauffmann, G. 2006, Monthly Notices of the Royal Astronomical Society, 366, 499

Dekel, A. & Silk, J. 1986, The Astrophysical Journal, 303, 39

Durkalec, A., Le Févre, O., Pollo, A., et al. 2015, Astronomy & Astrophysics, 583, A128

Ferrarese, L., Côté, P., Cuillandre, J.-C., et al. 2012, The Astrophysical Journal Supplement Series, 200, 4

Forte, J. C., Faifer, F. R., Vega, E. I., et al. 2013, Monthly Notices of the Royal Astronomical Society, 431, 1405

Foucaud, S., McCracken, H. J., Le Févre, O., et al. 2003, Astronomy & Astrophysics, 409, 835

Franx, M., Labbé, I., Rudnick, G., et al. 2003, The Astrophysical Journal Letters, 587, L79

Friedman, J. H., Bentley, J. L., & Finkel, R. A. 1977, ACM Trans. Math. Softw., 3, 209

Gao, L. & White, S. D. M. 2007, Monthly Notices of the Royal Astronomical Society, 377, L5

Gavazzi, R., Adami, C., Durret, F., et al. 2009, Astronomy & Astrophysics, 498, L33

Geller, M. J. & Huchra, J. P. 1989, Science, 246, 897

Guo, Q., White, S., Li, C., & Boylan-Kolchin, M. 2010, Monthly Notices of the Royal Astronomical Society, 404, 1111

Hamilton, A. J. S. 1993, The Astrophysical Journal, 417, 19

Hartley, W. G., Almaini, O., Cirasuolo, M., et al. 2010, Monthly Notices of the Royal Astronomical Society, 407, 1212

Hewett, P. C. 1982, Monthly Notices of the Royal Astronomical Society, 201, 867

Hoekstra, H., Yee, H. K. C., & Gladders, M. D. 2004, The Astrophysical Journal, 606, 67

Hogg, D. W. 1999, ArXiv Astrophysics e-prints

Infante, L. 1994, Astronomy & Astrophysics, 282, 353

Ishikawa, S., Kashikawa, N., Toshikawa, J., & Onoue, M. 2015, Monthly Notices of the Royal Astronomical Society, 454, 205

Kaiser, N. 1984, The Astrophysical Journal Letters, 284, L9

Kashikawa, N., Yoshida, M., Shimasaku, K., et al. 2006, The Astrophysical Journal, 637, 631

Kerscher, M., Szapudi, I., & Szalay, A. S. 2000, The Astrophysical Journal Letters, 535, L13

Kong, X., Daddi, E., Arimoto, N., et al. 2006, The Astrophysical Journal, 638, 72

Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., et al. 2004, The Astrophysical Journal, 609, 35

Kravtsov, A. V. & Klypin, A. A. 1999, The Astrophysical Journal, 520, 437

Kubo, J. M., Stebbins, A., Annis, J., et al. 2007, The Astrophysical Journal, 671, 1466

Labatie, A., Starck, J.-L., Lachiëze-Rey, M., & Arnalte-Mur, P. 2012, Statistical Methodology, 9, 85 , special Issue on Astrostatistics + Special Issue on Spatial Statistics

Landy, S. D. & Szalay, A. S. 1993, The Astrophysical Journal, 412, 64

Larson, R. B. 1974, Monthly Notices of the Royal Astronomical Society, 169, 229

Le Févre, O., Guzzo, L., Meneux, B., et al. 2005, Astronomy & Astrophysics, 439, 877

Leauthaud, A., Tinker, J., Behroozi, P. S., Busha, M. T., & Wechsler, R. H. 2011, The Astrophysical Journal, 738, 45

Li, C., Kauffmann, G., Jing, Y. P., et al. 2006, Monthly Notices of the Royal Astronomical Society, 368, 21

Limber, D. N. 1953, The Astrophysical Journal, 117, 134

Lin, L., Dickinson, M., Jian, H.-Y., et al. 2012, The Astrophysical Journal, 756, 71

Maddox, S. J., Efstathiou, G., Sutherland, W. J., & Loveday, J. 1990, Monthly Notices of the Royal Astronomical Society, 242, 43P

Madgwick, D. S., Hawkins, E., Lahav, O., et al. 2003, Monthly Notices of the Royal Astronomical Society, 344, 847

Magliocchetti, M. & Maddox, S. J. 1999, Monthly Notices of the Royal Astronomical Society, 306, 988

Marulli, F., Bolzonella, M., Branchini, E., et al. 2013, Astronomy & Astrophysics, 557, A17

McCarthy, P. J. 2004, Annual Review of Astronomy and Astrophysics, 42, 477

McCracken, H. J., Capak, P., Salvato, M., et al. 2010, The Astrophysical Journal, 708, 202

McCracken, H. J., Ilbert, O., Mellier, Y., et al. 2008, Astronomy & Astrophysics, 479, 321

McLaughlin, D. E. 1999, The Astrophysical Journal Letters, 512, L9

Messias, H., Afonso, J., Hopkins, A., et al. 2010, The Astrophysical Journal, 719, 790

Mo, H. J. & White, S. D. M. 1996, Monthly Notices of the Royal Astronomical Society, 282, 347

Mo, H. J. & White, S. D. M. 2002, Monthly Notices of the Royal Astronomical Society, 336, 112

Moster, B. P., Naab, T., & White, S. D. M. 2013, Monthly Notices of the Royal Astronomical Society, 428, 3121

Moster, B. P., Somerville, R. S., Maulbetsch, C., et al. 2010, The Astrophysical Journal, 710, 903

Moutard, T., Arnouts, S., Ilbert, O., et al. 2016, Astronomy & Astrophysics, 590, A102

Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, The Astrophysical Journal Supplement Series, 2068

Norberg, P., Baugh, C. M., Hawkins, E., et al. 2002, Monthly Notices of the Royal Astronomical Society, 332, 827

Oke, J. B. 1974, The Astrophysical Journal Supplement Series, 27, 21

Ouchi, M., Shimasaku, K., Okamura, S., et al. 2004, The Astrophysical Journal, 611, 685

Peebles, P. J. E. 1980, The large-scale structure of the universe, Princeton, N.J., Princeton University Press

Peebles, P. J. E. & Hauser, M. G. 1974, Astrophysical Journal Supplement, 28, 19

Pollo, A., Guzzo, L., Le Févre, O., et al. 2006, Astronomy & Astrophysics, 451, 409

Puget, P., Stadler, E., Doyon, R., et al. 2004, WIRCam: the infrared wide-field camera for the Canada-France-Hawaii Telescope

Roche, N. & Eales, S. A. 1999, Monthly Notices of the Royal Astronomical Society, 307, 703

Roche, N. D., Almaini, O., Dunlop, J., Ivison, R. J., & Willott, C. J. 2002, Monthly Notices of the Royal Astronomical Society, 337, 1282

Rudnick, G., Franx, M., Rix, H.-W., et al. 2001, The Astronomical Journal, 122, 2205

Sato, T., Sawicki, M., & Arcila-Osejo, L. 2014, Monthly Notices of the Royal Astronomical Society, 443, 2661

Savoy, J., Sawicki, M., Thompson, D., & Sato, T. 2011, The Astrophysical Journal, 737, 92

Sheth, R. K. & Tormen, G. 1999, Monthly Notices of the Royal Astronomical Society, 308, 119

Skibba, R. A., Smith, M. S. M., Coil, A. L., et al. 2014, The Astrophysical Journal, 784, 128

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629

Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, Monthly Notices of the Royal Astronomical Society, 328, 726

Szapudi, I. & Szalay, A. S. 1998, The Astrophysical Journal Letters, 494, L41

Totsuji, H. & Kihara, T. 1969, Publications of the Astronomical Society of Japan, 21, 221

Urban, O., Werner, N., Simionescu, A., Allen, S. W., & Böhringer, H. 2011, Monthly Notices of the Royal Astronomical Society, 414, 2101

van Dokkum, P. G., Franx, M., Schreiber, N. M. F., et al. 2004, The Astrophysical Journal, 611, 703

Veale, M., Ma, C.-P., Thomas, J., et al. 2017, Monthly Notices of the Royal Astronomical Society, 464, 356

White, S. D. M. & Rees, M. J. 1978, Monthly Notices of the Royal Astronomical Society, 183, 341

Yang, X., Mo, H. J., & van den Bosch, F. C. 2003, Monthly Notices of the Royal Astronomical Society, 339, 1057

York, D. G., Adelman, J., John E. Anderson, J., et al. 2000, The Astronomical Journal, 120, 1579

Zehavi, I., Contreras, S., Padilla, N., et al. 2017, ArXiv e-prints

Zehavi, I., Weinberg, D. H., Zheng, Z., et al. 2004, The Astrophysical Journal, 608, 16

Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2011, The Astrophysical Journal, 736, 59

Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2005, The Astrophysical Journal, 630, 1