Exploring Galaxy Growth as a Function of Mass and Environment at $z \sim 2$

Jonathan Florez

Second-Year Project

Committee: Shardha Jogee (Research Advisor), Volker Bromm, Caitlin Casey, Neal Evans, Steven Finkelstein

1 Project Overview

The epoch of $z \sim 2 - 4$ is one of the most important and active epochs of galaxy formation. During this time, star formation and black hole accretion peaked, massive protoclusters collapsed into existence, and galaxies underwent significant growth. At low redshifts, shallow surveys have allowed us to efficiently study the growth of galaxies as a function of mass and environment. At higher redshifts, this has remained a challenge because current redshift surveys that are deep enough to observe galaxies in the high-redshift universe have relatively small volume, making it difficult to achieve good enough statistics and a large enough dynamic range in comoving volume to constrain the formation of galaxies at $z > 2$. To advance our understanding of galaxy evolution, we need deep large-area photometric and spectroscopic surveys in order to map out the growth of rare high-mass galaxies at early epochs. In addition, large-area surveys are essential in probing how galaxy properties vary across environment (field, group-like, protoclusters), inform the nature-nurture debate, and characterize the most massive structures in the universe. Such observations will provide important constraints on theoretical models of galaxy evolution, particularly on the growth modes (e.g. mergers, gas accretion, secular processes), star formation, and feedback models.

We plan to make important headway on these fundamental questions by utilizing the powerful synergy of five wide and deep photometric surveys (Dark Energy CAMera *ugriz*, NEWFIRM K-band, *Spitzer*-IRAC 3.6μm and 4.5μm (Papovich et al. 2016), *Herschel*-SPIRE (HerS; Viero et al. 2014) far-IR/sub-millimeter, XMM-Newton and Chandra X-ray Observatory X-ray from the Stripe 82X survey (LaMassa et al. 2013, and private communication)) and the HETDEX blind optical spectroscopic survey (Hill et al. 2008) spanning the redshift range $1.9 < z < 3.5$ in the SHELA/HETDEX legacy field. With an area of 24 deg^2, a sample size of up to ~ 0.6 million galaxies complete in stellar mass above $M_* \sim 2 \times 10^{10} M_\odot$, 0.2 million Ly$\alpha$-based spectroscopic redshifts, and a comoving volume of ~ 0.45 Gpc^3 at $z \sim 2 - 4$, the SHELA/HETDEX field will provide a sample size that is at least an order of magnitude larger than the deepest spectroscopic surveys at these epochs.

The SHELA/HETDEX field surveys will allow us to make significant advancements in understanding the evolution of galaxies and their respective dark matter components due to several factors. First, they will target an interesting epoch ($z \sim 2 - 4$) during which the SFR history and AGN accretion history peaked, massive protoclusters began to collapse, and only 25% of the cosmic stellar mass density was in place. The sample will contain a large dynamic range in stellar masses ($M_* \sim 10^{10} - 10^{12} M_\odot$) and will be complete in stellar

mass down to $M_* = 2 \times 10^{10} M_\odot$, which is a factor of ~ 5 below the characteristic stellar mass at $z \sim 2$. This enables us to study critical processes in the evolution of galaxies using the largest sample to-date of rare high-mass galaxies ($M_* > 5 \times 10^{11} M_\odot$). The volume probed by the survey will host a large range of environments, including voids, fields, and groups, as well as a large number (several tens) of very massive protoclusters. The combination of DECam, NEWFIRM, and *Spitzer*-IRAC surveys (optical-to-mid-IR) is expected to yield photometric redshifts, stellar masses, and extinction-corrected UV-based SFRs for up to 0.6 million galaxies. The addition of HerS (sub-mm) and Stripe 82X (X-ray) data will allow us to study obscured SFRs for the brightest sub-mm galaxies, as well as the connection between quasars and starbursts in our sample. By studying the properties of such massive galaxies at $z \sim 1.9 - 3.5$, we will be able to provide strong constraints on theoretical models of galaxy evolution (Kereš et al. 2005, Guo et al. 2011; Rodriguez-Gomez et al. 2015; Somerville & Davé 2015).

1.1 Questions We Aim to Address Over the Next Four Years

By leveraging the above advantages, we aim to address the following questions over the next four years:

1. How do galaxy properties, such as their number density, merger rates, colors, SFRs, and AGN activity vary as a function of stellar mass, and across different environments, such as field-like, group-like, and massive protoclusters?

2. What is the relationship between star formation and environment at $z \sim 2 - 4$? Is star formation enhanced or suppressed in field, group, and protocluster regions? This is currently a controversial topic where contradictory results (Elbaz et al. 2007, Tran et al. 2010) have been reported at $z > 1.5$.

3. How do stellar mass and environment on small and large scales impact the growth of baryons, black holes, and dark matter in galaxies at $z \sim 2 - 4$?

4. What is the relative importance of galaxy mergers and gas accretion across these different environments in the cosmic web?

5. How efficiently do $z > 2$ galaxies of different halo masses build their stars and central black holes?

1.2 My Second-Year Project

For my second-year project, I am focusing on exploring galaxy growth as function of mass and environment. We are already in a position to explore galaxy properties with mass, but have yet to explore galaxy properties with environment. Once we have HETDEX spectroscopic redshifts (in ~ 2 years), we can get one accurate measure of environment on large scales, however, it is important to know how well we can characterize environment with photometric redshifts alone for two reasons:

1. Measures of environment using HETDEX data will trace Lyman alpha emitters (LAEs). Galaxies with photometric redshifts can provide environment measures on different scales and with different tracers than LAEs.

2. We are in the era of large-area deep photometric surveys (which are less expensive than spectroscopic surveys) and it is important for the scientific community to know how well we can characterize environment and recover trends with environment using photometric redshifts alone.

Over the past year, I have developed numerous measures of environment and applied them to DECam *ugriz* and *Spitzer*-IRAC data in one of the six SHELA fields called SHELA1. I have also applied these measures to light cones from the Illustris simulation at $z \sim 0$ and $z \sim 2$. The scope of my second-year project focuses on addressing the following questions:

1. How well can we measure the galaxy density field with our current photometric redshifts at low redshift ($z < 0.5$), and intermediate redshifts ($z \sim 2 - 4$)?

2. What are the trends between color, specific SFR, and environment at low and intermediate redshifts?

3. Are trends between color and environment predicted in Illustris still detectable when we take into account the uncertainties and systematics of photometric redshifts?

4. Among the most massive galaxies ($M_* > 10^{11}\ M_\odot$) at $z \sim 2 - 4$, what is the nature of extreme star-forming and quiescent galaxies and their relation to AGN activity at $z \sim 2 - 4$?

In the following section I describe my sample selection for the SHELA1 field, the metrics I use to study galaxy environment, and preliminary results. In Section 3 I outline the future short and long-term goals.

2 Measuring the Galaxy Density Field and Galaxy Properties in SHELA1

2.1 Data & Sample Selection

There are six DECam pointings in the SHELA field corresponding to six fields (with some overlap) that cover an area of ~ 24 deg^2. For the purposes of this report, I will only focus on running tests on the west-most field (SHELA1), however, all tests will be extended to the six fields in order to realize the full potential of the SHELA dataset. The SHELA1 field has an area of ~ 4 deg^2, and a total of $\sim 500,000$ sources detected in the optical images with Source Extractor (SExtractor). Matt Stevans of UT Austin and Lalitwadee (Nancy) Kawinwanichakij of Texas A&M currently process the source catalogs such that each object in the DECam images has a counterpart in the *Spitzer*-IRAC mid-IR images. I obtain photometric redshifts and rest-frame colors for all SHELA1 galaxies by running the SED

fitting code EAZY on the DECam and IRAC source catalogs using Sydney Sherman's (UT Austin) EAZY parameter files.

Before I can begin any analysis on the dataset, I first have to ensure that contamination is minimized in the sample that is being used. I do so by applying the following cuts to the data:

1. EAZY needs to fit at least 5 of the 7 available filters in the SED fit

2. The photo-z fit should have a chi-squared value of $\chi^2 < 50$

3. Each source must have signal-to-noise ratio $S/N > 3$ in the r-band

4. Magnitude cut of r-mag < 24.0 at low-z in order to reduce contamination from low luminosity galaxies

5. Only unresolved sources, determined from the DECam r-band half-light radius, are used in the sample ($R_e < 1.2$")

Applying the above cuts and comparing with SDSS DR12 spectroscopic redshifts yields a photometric redshift error of $\sigma_z = 0.041$, determined from equation (7) of Brammer et al. (2008). This yields a total of $\sim 200,000$ sources spanning the redshift range $0 < z < 6$, with $\sim 65,000$ sources falling in the redshift range we are interested in ($z \sim 1 - 4$).

2.2 Measures of Environment

Over the past year, I have studied the environment of SHELA1 galaxies at different redshifts using the following metrics:

i. Galaxy neighbor counts inside projected apertures of fixed radii. This method places an aperture of some fixed radius, which can be varied to study local galaxy density on different scales, on the center of every galaxy and counts the number of neighbors each galaxy has within each aperture (Hogg et al. 2003, 2004).

ii. Projected Nth nearest neighbor distances. This method measures the projected distance between a galaxy and its Nth nearest neighbor, and is particularly useful when studying the density of galaxies in groups and clusters (Dressler 1980; Lewis et al. 2002).

iii. The two-point angular/projected correlation function. The correlation function of galaxies measures the excess probability, relative to a randomly distributed sample of points, of finding two galaxies separated by some distance on the sky. The correlation function is important for testing the clustering of different galaxy populations on both small and large scales, and provides a means of testing cosmological models that predict different things about large-scale structure formation in the universe.

These are the most widely implemented measures of environment in the literature and have yielded a number of useful results relating to the clustering and density field of low-redshift galaxies. At low redshift, it is generally well understood that galaxy colors, SFRs, and

morphologies depend highly on environment with galaxies having earlier-type morphologies, redder colors and lower specific SFRs in higher density environments (Dressler 1980; Kauffmann et al. 2004; Hogg 2004). As a result, redder galaxies are found to exhibit stronger clustering than bluer galaxies at low redshift (Zehavi et al. 2005). In order to test that my environment metrics are working as expected at low redshifts ($z \sim 0.2 - 0.6$) and intermediate redshifts ($z \sim 1 - 3$), I have performed several tests described in the following subsections.

2.3 Tests of Environment at Low Redshift

At low redshift, I use a SDSS DR12 data to check whether it is possible to recover real over-densities in the SHELA1. In Figures 1 and 3 I show the galaxy density field, where density is determined by counting galaxies inside an aperture of fixed radii, and SDSS DR12 galaxies that have high projected densities determined only with spectroscopic redshifts. Figure 2 shows a zoomed in image of the over-density in Figure 3. I show the photometric and spectroscopic redshift distribution around the highest over-densities, as well as the rest-frame color distribution (Figure 4) to see if we detect differences in color from the field.

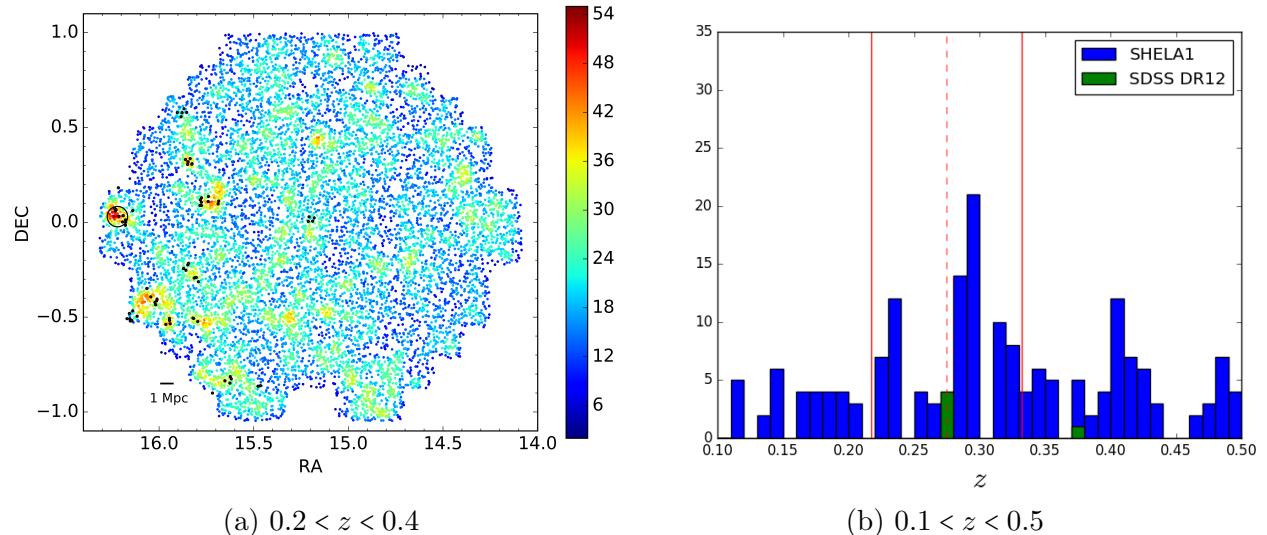(a) $0.2 < z < 0.4$ (b) $0.1 < z < 0.5$

Figure 1: Left: Density field of galaxies with photometric redshift $0.2 < z_p < 0.4$. Color indicates the number of neighbors each galaxy has within a projected distance of 2.5 arcminutes and dark red points are SDSS galaxies that have high (> 70 percentile) density amongst the spectroscopic sample. The large black circle marks the location of one the highest SDSS and SHELA1 over-densities in this redshift range. A 1 Mpc scale is shown at the intermediate redshift. Right: Redshift distribution of SHELA1 galaxies (blue) and for SDSS galaxies (green) inside the circle from the figure on the left for a larger redshift range. The dashed red line marks the spectroscopic redshift of the SDSS galaxy over-density along the line of sight. The solid red lines show the typical deviation from the true redshift that is expected with our photometric redshift error ($\Delta z = 0.041 \times (1 + z_m)$) with $z_m = 0.3$.
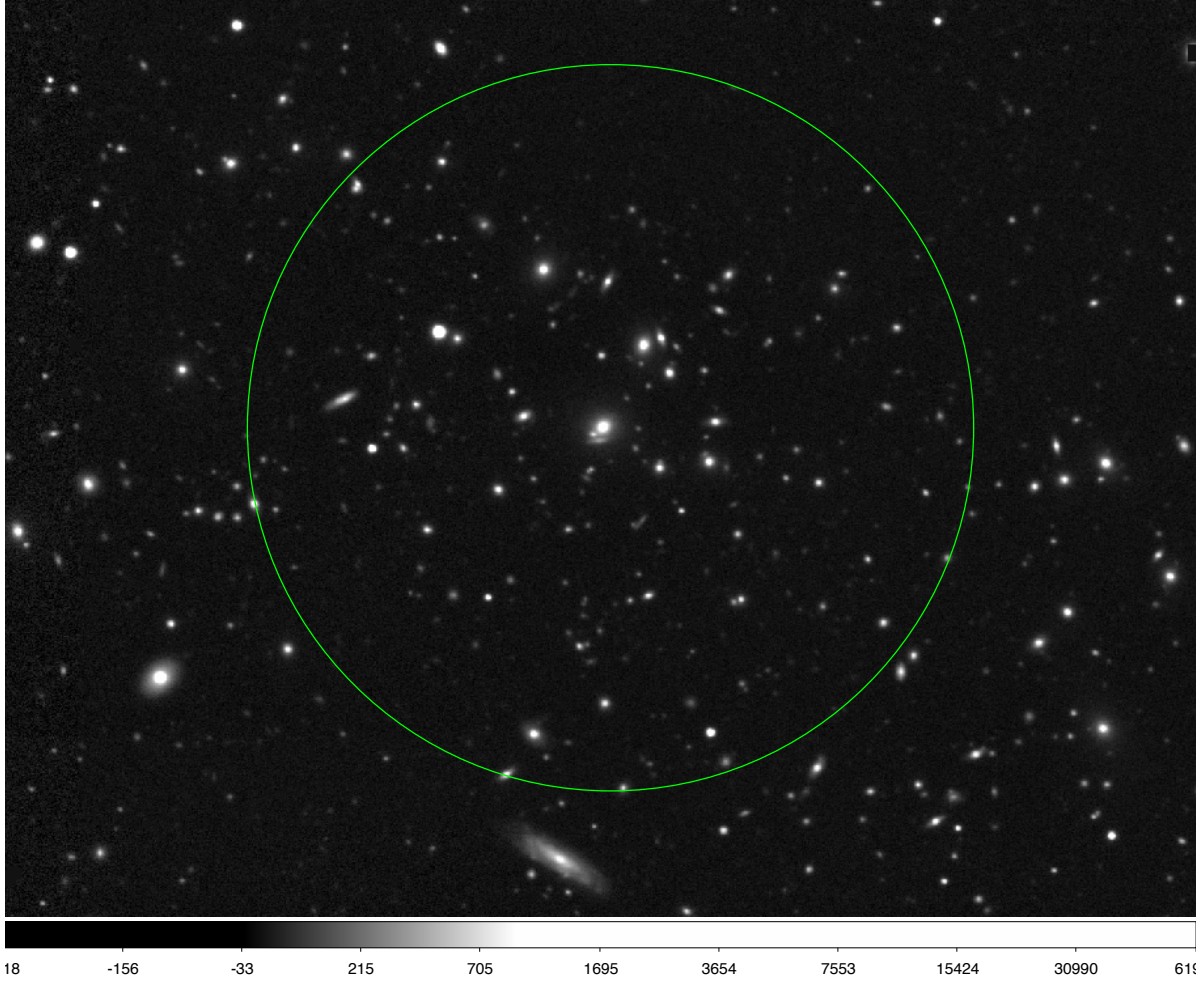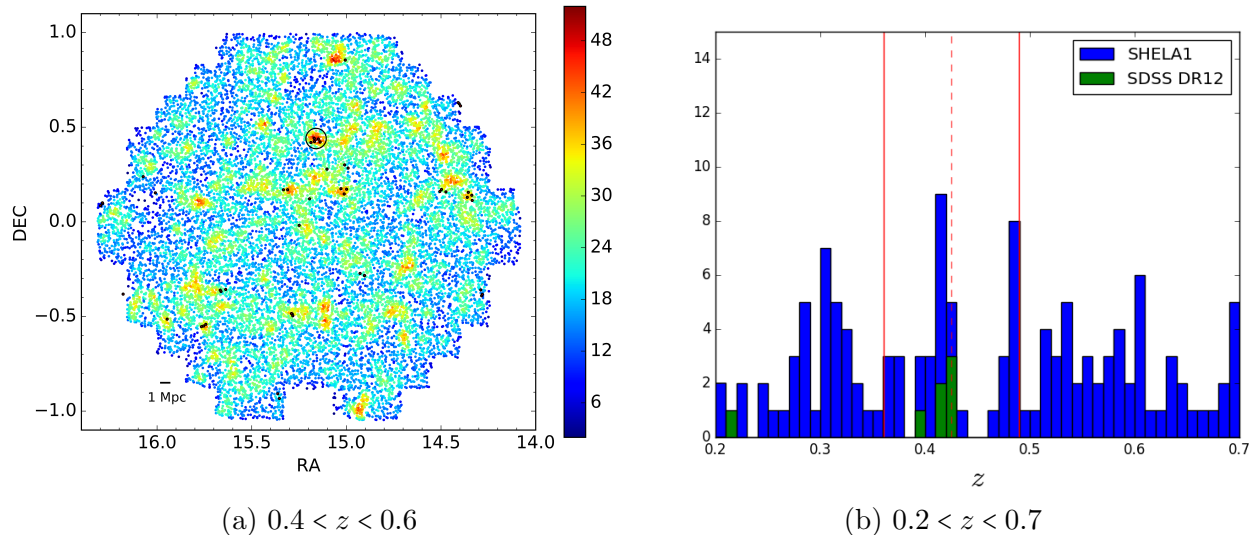
5

Figure 2: Zoomed in r-band image of the over-density shown above (Figure 1) at $z \sim 0.3$ inside a green circle with a diameter of 3.5 arcminutes, which corresponds to a diameter of ~ 1 Mpc at $z \sim 0.3$. The galaxy morphology appears to resemble that of early-types and S0s near the center of the circle.

(a) $0.4 < z < 0.6$ (b) $0.2 < z < 0.7$

Figure 3: Left: Density field of galaxies with photometric redshift $0.4 < z_p < 0.6$. Color indicates the number of neighbors each galaxy has within a projected distance of 2 arcminutes and dark red points are SDSS galaxies that have high (> 70 percentile) density amongst the spectroscopic sample. The large black circle marks the location of one the highest SDSS and SHELA1 over-densities in this redshift range. A 1 Mpc scale is shown at the intermediate redshift. Right: Redshift distribution of SHELA1 galaxies (blue) and for SDSS galaxies (green) inside the circle from the figure on the left for a larger redshift range. The dashed red line marks the spectroscopic redshift of the SDSS galaxy over-density along the line of sight. The solid red lines show the typical deviation from the true redshift that is expected with our photometric redshift error ($\Delta z = 0.041 \times (1 + z_m)$) with $z_m = 0.5$.

At these redshifts, it appears possible to recover real over-densities given our photo-z uncertainty since the SDSS over-densities appear to coincide physically with the projected over-densities we see in SHELA1. If the projected over-densities are real, then the colors of galaxies corresponding to the over-densities should be redder than that of the field galaxies. InFigure 4, I show the rest-frame colors of the galaxies corresponding to the two projected over-densities in the redshift ranges $0.2 < z < 0.4$ and $0.4 < z < 0.6$ (Figures 1 and 3) and also for the field in the same redshift range.

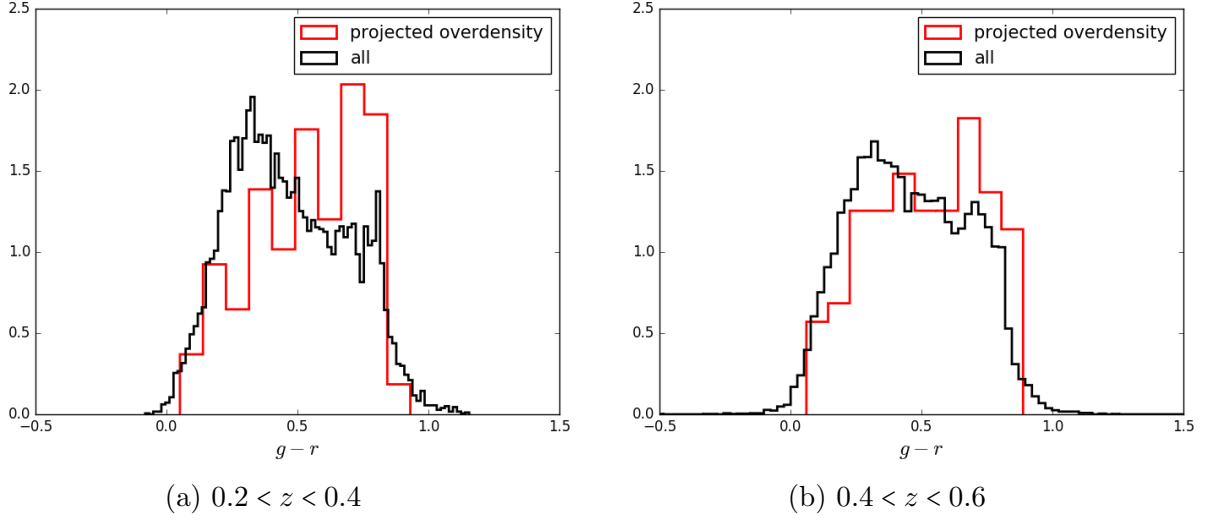(a) $0.2 < z < 0.4$ (b) $0.4 < z < 0.6$

Figure 4: Left: Rest-frame $g - r$ colors for galaxies corresponding to the over-density in $0.2 < z_p < 0.4$ (red) and for all galaxies in the same redshift range (black). Right: Rest-frame colors for galaxies corresponding to the over-density in $0.4 < z_p < 0.6$ (red) and for all galaxies in the same redshift range (black).

In the following figures I measure the projected correlation function of SHELA1 galaxies with $0.2 < z < 0.4$ for different magnitudes. Magnitude cuts are applied using each galaxy's rest-frame r-band absolute magnitude, obtained from EAZY. The correlation function has similar amplitude and slope to that measured from Wang et al. (2013) for SDSS galaxies. I also measure the correlation function for red galaxies and for the field at one magnitude range and find that red galaxies are more clustered at smaller (< 1 Mpc) scales
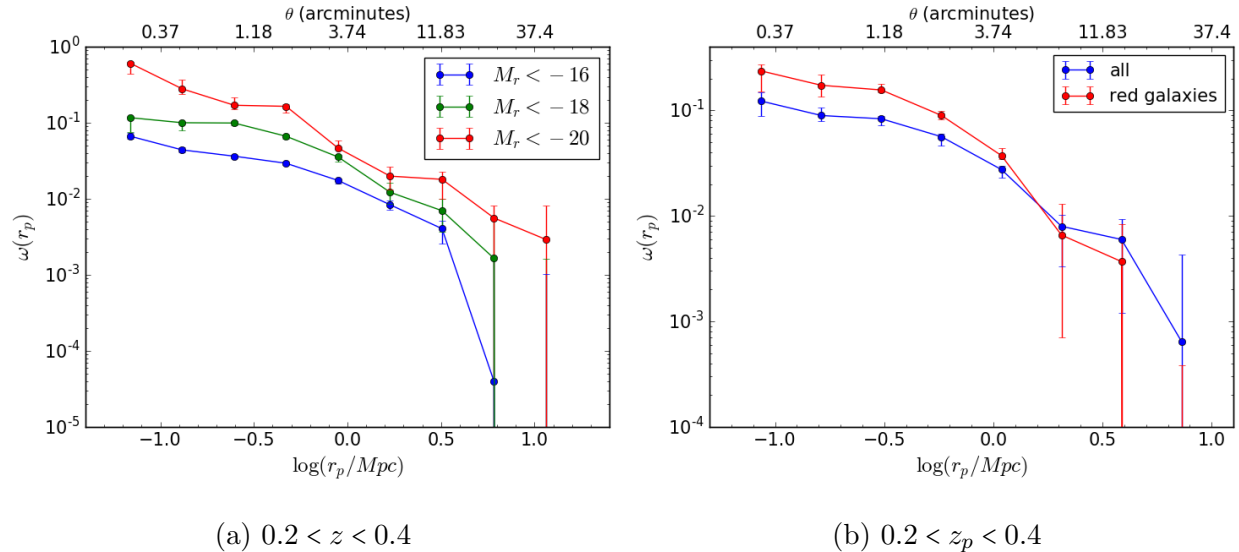
Figure 5: Left: Projected correlation function of SHELA1 galaxies in redshift range (0.2 < z < 0.4) with different absolute magnitudes. Right: Projected correlation function of all SHELA1 galaxies with $M_r < -18$ (blue), and for galaxies in same magnitude range but with color $g - r > 0.5$ (red).

The measures of environment I have shown here demonstrate that it is possible to reproduce known empirical trends in color versus environment and the projected correlation function using photometric redshifts.

2.4 Predictions of Color and sSFR Trends with Environment in Illustris

In the Figures 6 - 13 I show predictions of color and sSFR with environment at $z \sim 0.3, 1, 2$, and 3 using three light cones from the hydrodynamical simulation Illustris (Vogelsberger et al. 2014) with an area of ~ 130 squared arcminutes (see Snyder et al. 2017 for construction of light cones). Given such a small area, it is only possible to probe < 2 Mpc scales with these light cones, however, a member of Lars Hernquist's group, Greg Snyder, has provided us with the code he uses to generate his light cones so it will be possible to generate larger area light cones in the near future. I measure environment mainly by counting galaxies inside projected apertures of fixed radii with an absolute magnitude of $M_r < -20$. At $z \sim 0.3$, I measure environment on galaxies with $M_r < -18$ and inside 2 arcminute apertures I experimented with different radius apertures (1, 1.5, 2 and 3 arcminutes) and found that trends with environment slightly decrease as aperture size increases. For this reason, I only show trends with environment on 1 arcminute scales, corresponding to ~ 1 Mpc scales at high redshift ($z > 1$).
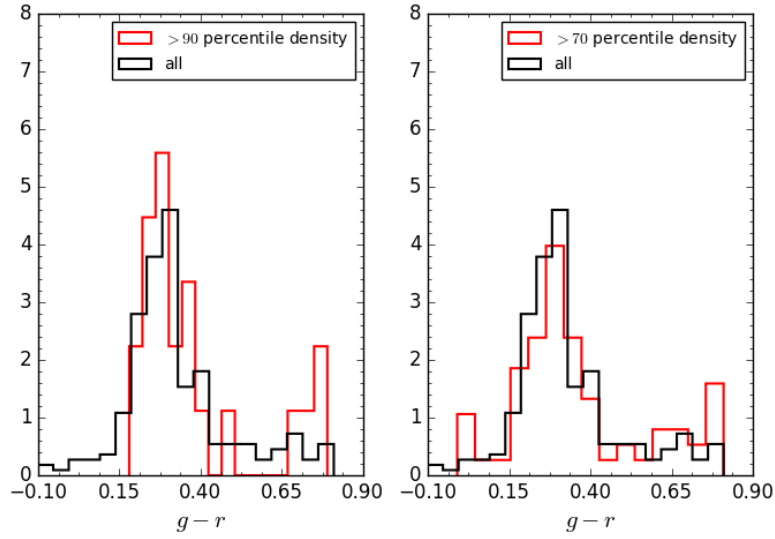
Figure 6: Rest-frame $g-r$ colors for galaxies with $0.2 < z < 0.4$. I show $g-r$ color for galaxies with high densities (red) and for the field (black). Trends appear to decrease as the selection is based on lower percentiles densities (i.e. trends are only apparent in the highest density regions).
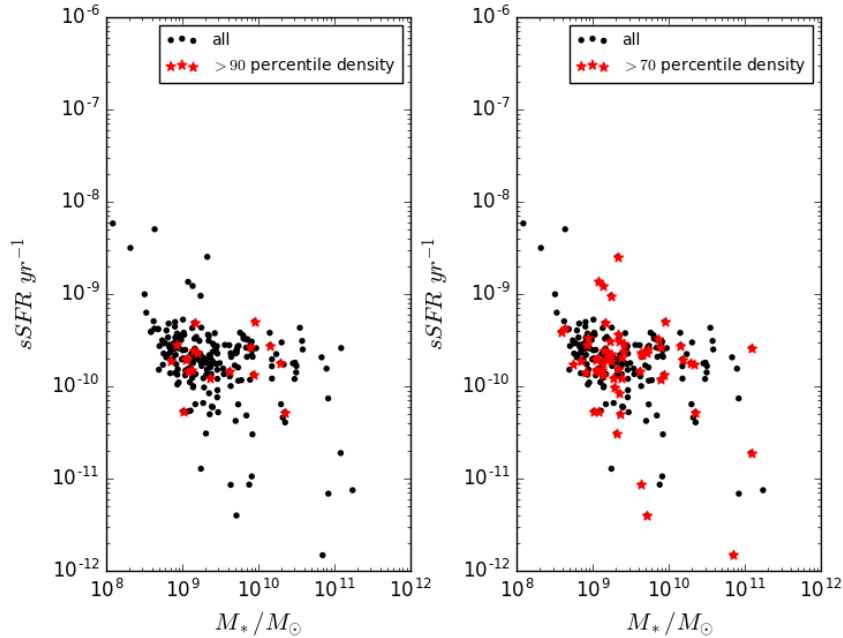

Figure 7: Specific star formation rates vs. stellar mass for galaxies with $0.2 < z < 0.4$ for field galaxies (black) as well galaxies in high-density regions (red).
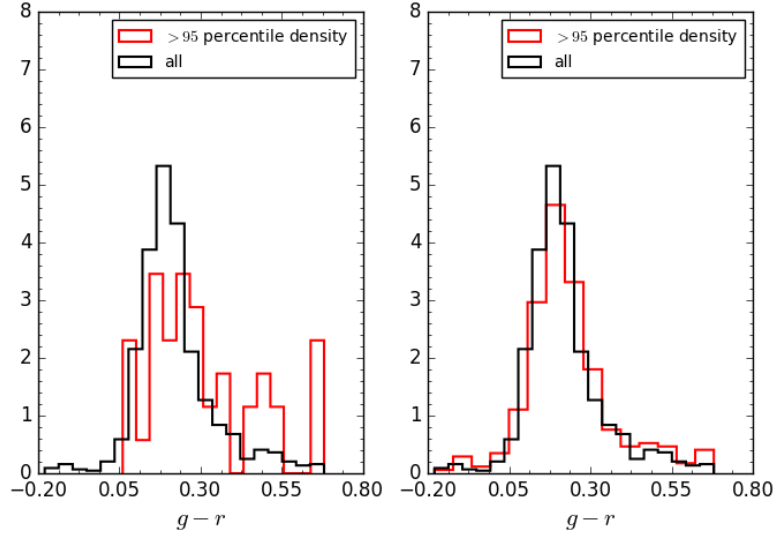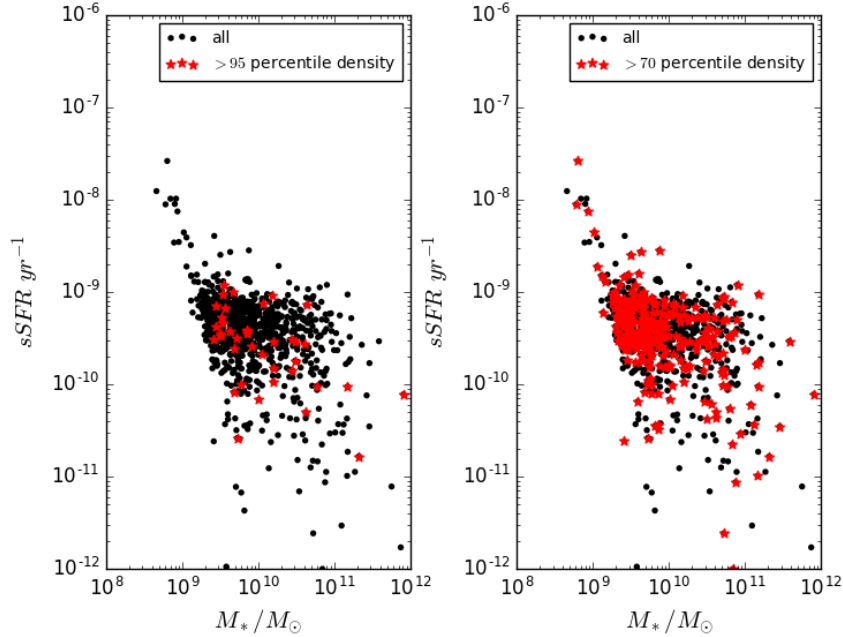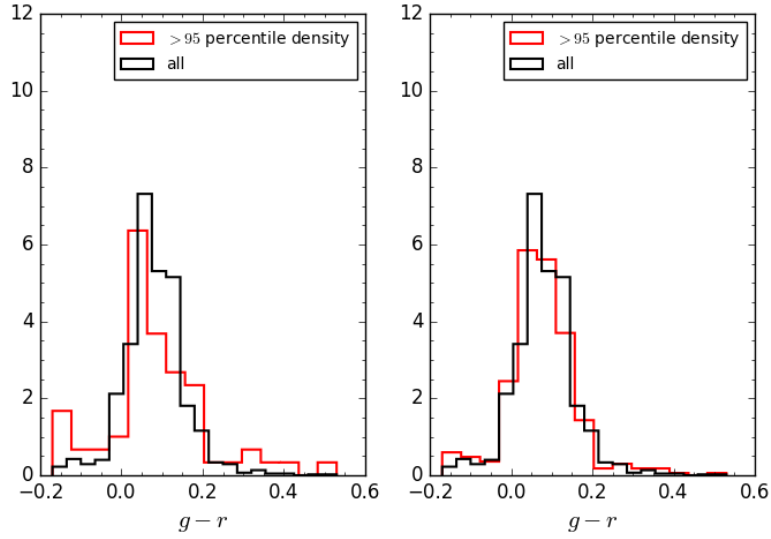
Figure 8: Rest-frame $g-r$ colors for galaxies with $0.75 < z < 1.25$. I show $g-r$ color for field galaxies (black) and for galaxies with high density (red).
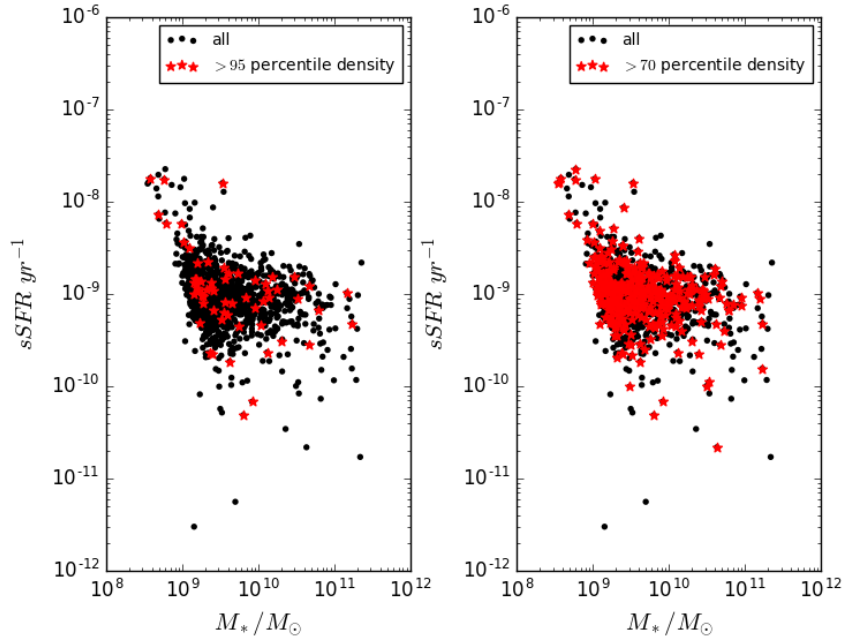

Figure 9: Specific star formation rates vs. stellar mass for galaxies with $0.75 < z < 1.25$ for field galaxies (black) as well galaxies in high-density regions (red).

Figure 10: Rest-frame $g - r$ colors for galaxies with $1.75 < z < 2.25$. I show $g - r$ color for galaxies with high densities (red) and for the field (black).

Figure 11: Specific star formation rates vs. stellar mass for galaxies with $1.75 < z < 2.25$ for field galaxies (black) as well galaxies in high-density regions (red).
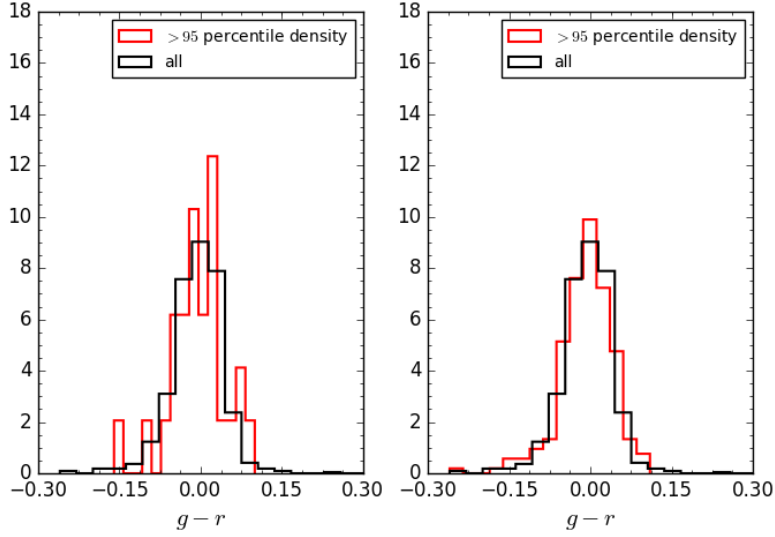
Figure 12: Rest-frame $g - r$ colors for galaxies with $2.75 < z < 3.25$. I show $g - r$ color for galaxies with high densities (red) and for the field (black).
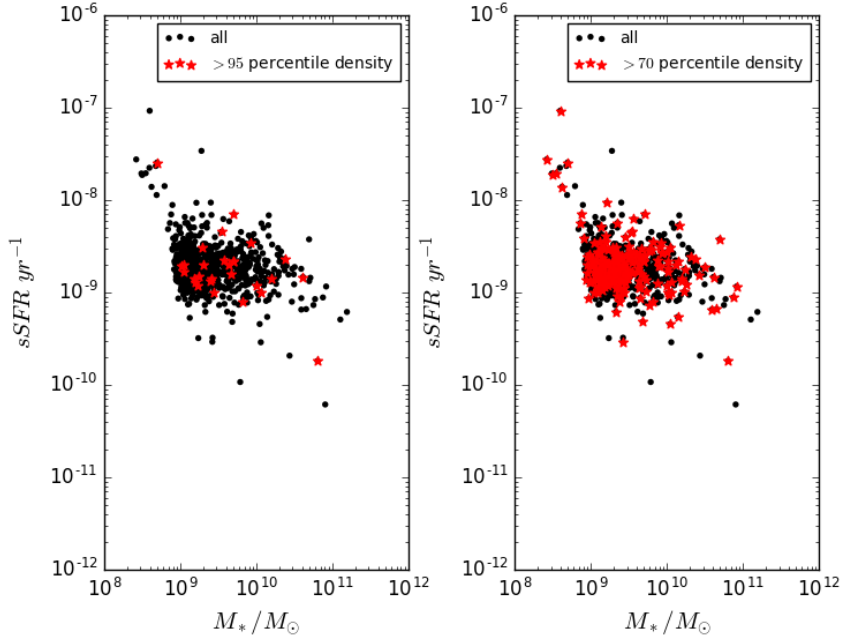

Figure 13: Specific star formation rates vs. stellar mass for galaxies with $2.75 < z < 3.25$ for field galaxies (black) as well galaxies in high-density regions (red).

We find that at $z \sim 0.3$ red galaxies tend to occupy the highest density regions. At $z \sim 1$, high-density regions host massive quenched galaxies as well as star-forming galaxies. At $z \sim 2$, high density regions appear to host both star-forming and quenched galaxies, and at $z \sim 3$ trends with environment are almost nonexistent.

13

As noted, trends with environment appear to decrease with increasing redshift, though it is likely that we are limited by the area of the light cones. Muldrew et al. (2015) reports that > 90% of a protocluster's mass extends over 35 comoving Mpc, corresponding to ~ 30 arcminutes at $z \sim 2$. If density fluctuations occur on such large scales, then it becomes difficult to observe trends with environment using our current Illustris light-cones due to the small area.

2.5   Introducing Redshift Uncertainties to the Light Cones

In this section, I show plots (Figures 14 - 16) that are similar to those shown in the previous section, but now I measure density after introducing uncertainty to the mock galaxy's redshift and plot the color distribution. I shuffle every galaxy's redshift by drawing randomly from a normal distribution centered on the true redshift with standard deviation of $\Delta z \times (1 + z_{true})$ = 0.041. I only do this for redshifts at $z \sim 1 - 2$ because there appears to be no strong trends between environment and color at higher redshifts.
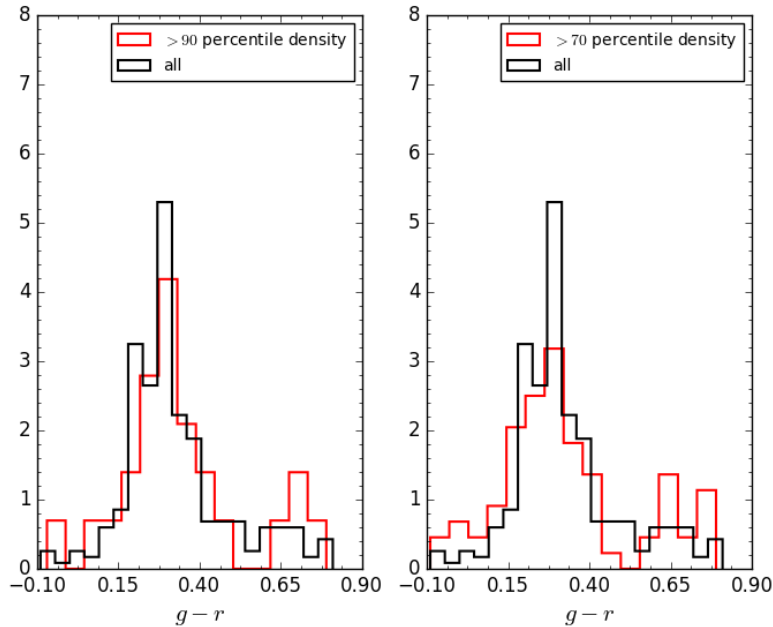

Figure 14: Rest-frame $g - r$ colors for galaxies with $0.2 < z < 0.4$. I show $g - r$ color for galaxies with high densities (red) and for the field (black). Densities are measured after some uncertainty has been introduced to the mock galaxy redshifts. Although trends appear weaker, they are still detectable at the highest densities for this redshift
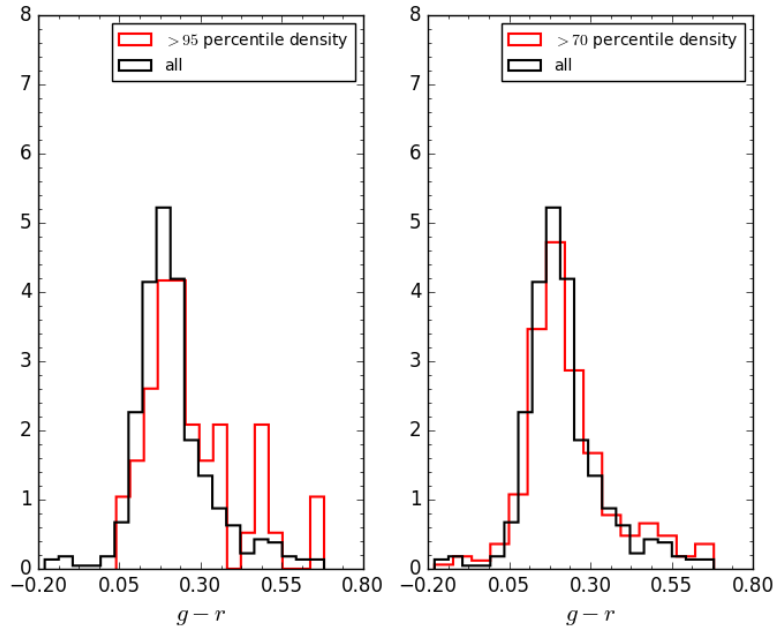
Figure 15: Rest-frame $g-r$ colors for galaxies with $0.75 < z < 1.25$. I show $g-r$ color for galaxies with high densities (red) and for the field (black). Densities are measured after some uncertainty has been introduced to the mock galaxy redshifts. Although trends appear weaker, they are still detectable at the highest densities for this redshift

Figure 16: Rest-frame $g - r$ colors for galaxies with $1.75 < z < 2.25$. I show $g - r$ color for galaxies with high densities (red) and for the field (black). Densities are measured after some uncertainty has been introduced to the mock galaxy redshifts. Although trends appear weaker, they are still detectable at the highest densities for this redshift
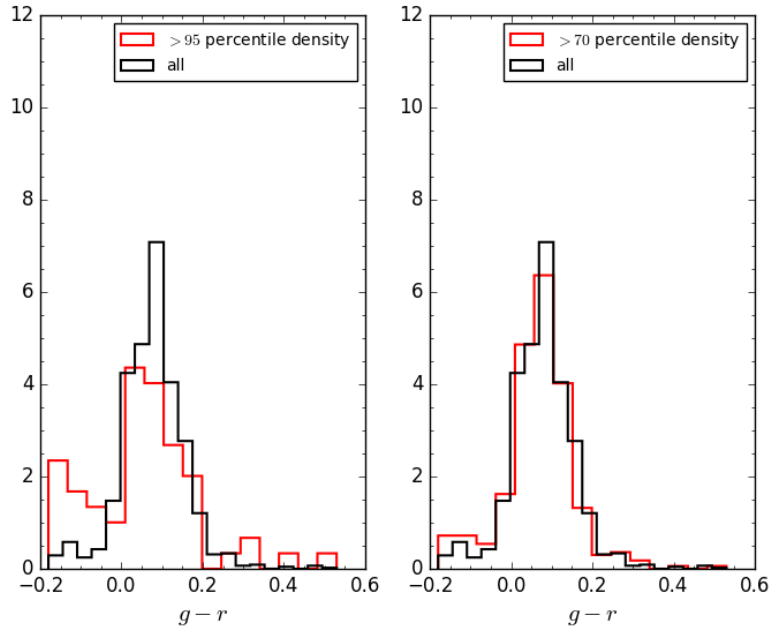
In the Figure 17, I measure the correlation function of Illustris galaxies at $z \sim 2$ on < 1 Mpc scales before and after introducing a redshift uncertainty to the mock galaxies. I also include the correlation function of SHELA1 galaxies at $z \sim 2$. The correlation function is measured for galaxies with $M_r < -20$ in both the data and simulations.
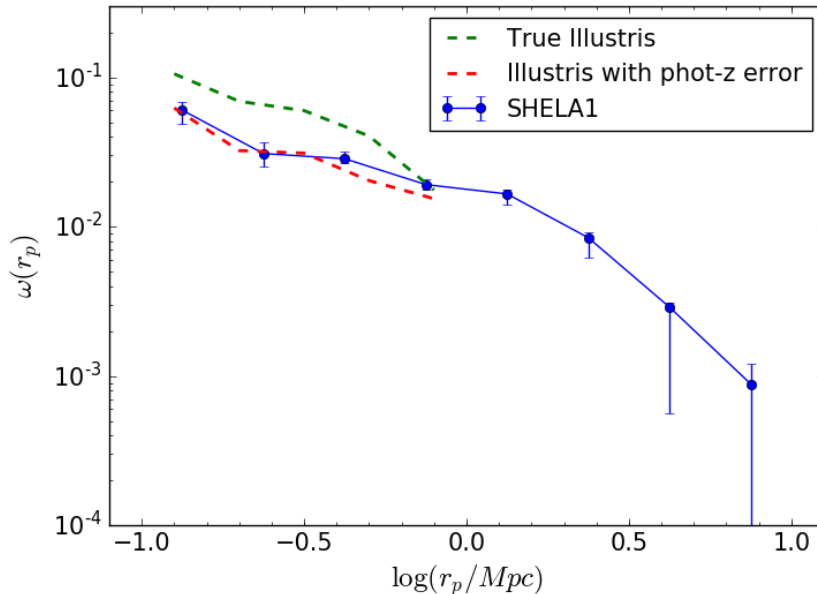
Figure 17: Projected correlation function for SHELA1 galaxies and Illustris galaxies with $M_r < -20$ and $1.75 < z < 2.25$. The effect of introducing a photometric redshift uncertainty to the simulated galaxies on the correlation function is shown as well.

The projected correlation function of Illustris agrees well with the correlation function of SHELA1 at small scales, although the small area of the light cone causes the correlation function to fall off quickly after just a few 100 kpc. Having a larger area light cone will allow us to determine whether the clustering of simulated galaxies on larger scales matches the clustering of SHELA1 galaxies. This analysis can further be extended to determine if the clustering of galaxies with different masses, colors, and SFRs can be measured at high redshift.

2.6 AGN Flux and Color of SHELA Galaxies

We have recently acquired an X-ray catalog from Stephanie La Massa (Stripe-82X, La Massa et al. 2013) that overlaps with the SHELA field. I have begun to examine the AGN flux vs. the color of SHELA galaxies, however, the results are highly contaminated with bad SED fits of SHELA quasars. Fortunately, we have just obtained a photometric redshift catalog for the Stripe-82X quasar sources. I will be able to use this catalog to produce more accurate estimates of the rest-frame color vs. the flux of AGN sources. Furthermore, if the photometric redshifts are accurate enough then I will be able to extend this analysis to include measures of environment.

3 Future Work

With the work I have done so far, and the addition of quasar photometric redshifts, stellar masses and SFRs, it will soon be possible to address the following:

1. What is the relationship between AGN activity, rest-frame colors, SFR, and stellar mass over a large dynamic range in stellar mass ($M_* \sim 10^{10} - 10^{12}$), and particularly for rare high mass galaxies ($> 5 \times 10^{11}$ $M_\odot$) at $z \sim 2 - 4$? These will inform the growth of SMBHs, as well as feedback and quenching of SF activity in galaxies. I will also explore the relationship between AGN activity and growth modes like merger rates if the latter can be accurately derived.

2. What is the relationship between galaxy properties and environment and is it possible to recover these trends given our current photometric redshift uncertainty? Very few studies have analyzed the clustering properties of different galaxy populations at high redshift, thus, I will constrain the formation of large-scale structures and massive systems in the universe by analyzing the angular correlation function and local density of multiple galaxy-types at high redshift.

3. Is it possible to measure galaxy property trends with environment at high-z by using larger area light cones? This can also be tested in simulations that rely on abundance matching and semi-analytic models to populate halos. We will soon receive semi-analytic models from Sadegh Khochfar and Rachel Somerville, as well as abundance matching models from Peter Behroozi

Once HETDEX spectra are available, we will be able to realize the full potential of the SHELA/HETDEX field by performing in-depth studies across a wide range of masses and environments with accurate photometric redshift estimates, spectroscopic redshifts, and halo masses for a large population of galaxies at high redshift. In the coming years I will be able to address the questions stated in Section 1.1 of this report.

4 References

Adams et al. 2011, ApJS, 192, 5
Brammer et al. 2008, ApJ, 686, 1503
Connolly et al. 2002, ApJ, 579, 42
Chiang et al. 2015, ApJ, 808, 37
da Cunha et al. 2008, MNRAS, 388, 1595
Dressler 1980, ApJ, 236, 351
Elbaz et al. 2007, A&A, 468, 33-48
Guo et al. 2011, MNRAS, 413, 101
Hasselfield et al. 2013, JCAP, 7, 008
Heiderman et al. 2009, ApJ, 705, 1433
Hogg et al. 2003, ApJ, 585, L5 Hogg et al. 2004, ApJ, 601, L29 Hill et al. 2008, ASPC, 399, 115
Jogee et al. 2009, ApJ, 697, 1971
Kauffmann et al. 2004, MNRAS, 353, 713

Kriek et al. 2009, ApJ, 700, 221
La Massa et al. 2013, MNRAS, 432, 1351
Lee et al. 2006, ApJ, 642, 63
Lewis et al. 2002, MNRAS, 334, 673
Muldrew et al. 2015, MNRAS, 452, 2528
Papovich et al. 2016, ApJS, 224, 28
Rodriguez-Gomez et al. 2015, MNRAS, 449, 49
Somerville & Davé 2015, ARAA, 53, 51
Tran et al. 2010, ApJ ,719, L126
Viero et al. 2014, ApJS, 210, 22
Vogelsberger et al. 2014, MNRAS, 444, 1518
Zehavi et al. 2005, ApJ, 630, 1