

# Environmental factors affecting Galaxy Morphology

Arun Kannawadi<sup>1\*</sup>, Rachel Mandelbaum<sup>1</sup>, Claire Lackner<sup>2</sup>,

<sup>1</sup>*McWilliams center for Cosmology, Carnegie Mellon University, Pittsburgh, PA 15217, USA*

<sup>2</sup>*Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes for Advanced Study, the University of Tokyo, Kashiwa, Japan*

20 July 2014

## ABSTRACT

According to our current understanding, galaxy shapes and morphologies should depend on various factors such as the local environment. Realistic image simulations for calibration of weak lensing analysis methods that use training samples from the Hubble Space Telescope can therefore be affected by these trends, due to the limited volume of the universe that has been surveyed by Hubble. We will show how redshift slices in a volume-limited subsample of COSMOS can be classified as overdense or underdense (or neither), and how the statistical properties of various morphological parameters such as ellipticity, Sérsic  $n$ , Bulge-to-Total ratio and color differ in these bins. This study requires a careful distinction between environment effects from large-scale structure, which we do not wish to include in simulations, and general trends in the galaxy population with redshift. We conclude with some guidance for how upcoming surveys can use COSMOS data as the basis for weak lensing simulations without having their conclusions overly affected by cosmic variance.

**Key words:**

## 1 INTRODUCTION

***Stress on the importance of WL:*** We are currently in the era of precision cosmology, where systematic errors typically dominate over statistical errors. In almost all branches of observational cosmology, weak lensing introduces systematic error, in addition to itself being a probe of the dark matter and eventually the large scale structure of the universe. As light from the background sources passes through the foreground cluster of galaxies and dark matter profiles, the shapes get distorted. Weak lensing also causes magnification in addition to introducing a shear. Weak lensing effects need to be removed from the Cosmic Microwave Background (CMB) analysis to understand the initial conditions of the universe.

Previous studies have shown that any shape measurement method that uses second moments cannot be independent of the morphology and substructure of the galaxies involved. (**rephrased from GREAT3**). Hence, it is important to check any method that is proposed to measure the shape of the galaxies to check against a simulation that generates images as realistic as possible. With this in mind, the GREAT3 challenge Mandelbaum et al. (2014) was held with the aim to test and facilitate methods for measuring weak gravitational lensing by analyzing astronomical images. Software packages like GalSim<sup>1</sup> can generate galaxy

images fairly accurately, given realistic input parameters. The training dataset used in the challenge comes from the COSMOS survey, which we describe in sufficient detail in Sec. 2. While providing a good quality (high  $S/N$ ) images to be trained over, there could be some bias introduced due to the survey being a narrow one. It is known **cite refs.** that the shape and morphology of galaxies depend very much on their local environment. Hence, local overdensities or underdensities observed in the COSMOS’ field of view may cause the properties of the galaxies to be associated with the redshift, thus unable to distinguish between redshift evolution and large scale structure.

In this paper, we show that these large scale structure effects mentioned are indeed present in the data. The rest of the paper is structured as follows. In Sec. 3.1, we explain how we volume-limit our data and identify overdensities and underdensities and in Sec. 3.2, we explain how we obtain axis-ratios from the images of the galaxies. Finally, we present our results in Sec. 4 and end with conclusions in Sec. 5.

## 2 DATA

COSMOS (Scoville et al. 2007; Koekemoer et al. 2007; Leauthaud et al. 2007) is a flux-limited, narrow deep field survey covering a contiguous area of  $1.64 \text{ deg}^2$  of sky, with images taken using the Advanced Camera for Surveys (ACS) Wide Field Channel (WFC) in the Hubble Space Tele-

\* [akannawa@andrew.cmu.edu](mailto:akannawa@andrew.cmu.edu)

<sup>1</sup> <https://github.com/GalSim-developers/GalSim>

scope (HST). Precise shape measurements, when compared to ground-based surveys, can be made since the full width half-maximum (FWHM) of the point-spread function (PSF) is  $0.12''$ . High resolution images taken through the wide F814W filter (broad  $I$ ), after correcting for the PSF, have led to collection of postage stamp images<sup>2</sup>, originally intended for simulations of ground-based data using SHERA (SHEar Reconvolution Analysis) (Mandelbaum et al. 2011). Lackner and Gunn (2012) fitted parametric models to most of these galaxies including Sérsic  $n$  profile fits, 2 component bulge+disk fits, axis ratios etc. In addition to the ACS/WFC (F814W) imaging, the COSMOS field has also been imaged by Subaru Suprime-Cam ( $B_j, V_j, g^+, r^+, i^+, z^+$ , NB816), the Canada-French Hawaii Telescope (CFHT;  $u^*, i^*$ ) and the KPNO/CTIO ( $K_s$ ).

Photometric redshifts were determined by Mobasher et al. (2007). The accuracy of photometric redshifts for  $F814W \leq 22.5$  is  $\sigma_{\Delta z} = 0.031(1+z_s)$ . The photometric redshift values become noisy beyond  $z$  of 1 for our purposes and the various fits to the galaxies are also not very reliable beyond the apparent magnitude ( $m$ ) value of 23.5. ( $S/N \sim 50$  ? )

We apply the following set of initial cuts to the data, all of which is defined in Leauthaud et al. (2007):

(i) **MU\_CLASS=1**: Unlike SExtractor's stellar index **CLASS\_STAR** which is continuous, the **MU\_MAX** method that compares the peak surface brightness to the background level, provides a clear differentiation of galaxies (**MU\_CLASS=1**) from stars (**MU\_CLASS=2**) and other spurious objects like cosmic rays (**MU\_CLASS=3**)

(ii) **CLEAN=1**: Objects near bright stars or those containing saturated pixels were removed and the rest classify as *clean*.

(iii) **GOOD\_ZPHOT\_SOURCE=1**: This cut requires that there be a good photometric redshift, which typically is equivalent to requiring that the galaxy not be located within the masked regions of the Subaru  $BVIz$  imaging used for photometric redshifts, and that it have a successful match against an object in the Subaru imaging **copied from Mandelbaum 2011**.

We will analyze how these intrinsic shape of the galaxies depends on the environment in which they reside.

### 3 METHODS

#### 3.1 Finding overdensities

Every survey is flux-limited and so is COSMOS and is affected by Malmquist bias. We generate, using the method below, a volume-limited sample that is complete upto  $z = 1$  by applying a cut on luminosity such that only galaxies intrinsically brighter than a certain threshold is considered. This threshold is set on  $K$ -corrected  $i$ -band magnitudes, denoted as  $M_I$ . Since the parent sample contains fainter galaxies, upto  $F814W=25.2$ , we compare the distribution of the  $F814W=23.5$  sample with the samples containing fainter galaxies for high redshift bins, where the incompleteness affects first.

<sup>2</sup> <http://irsa.ipac.caltech.edu/data/COSMOS/images/galaxy-postage-stamps/>

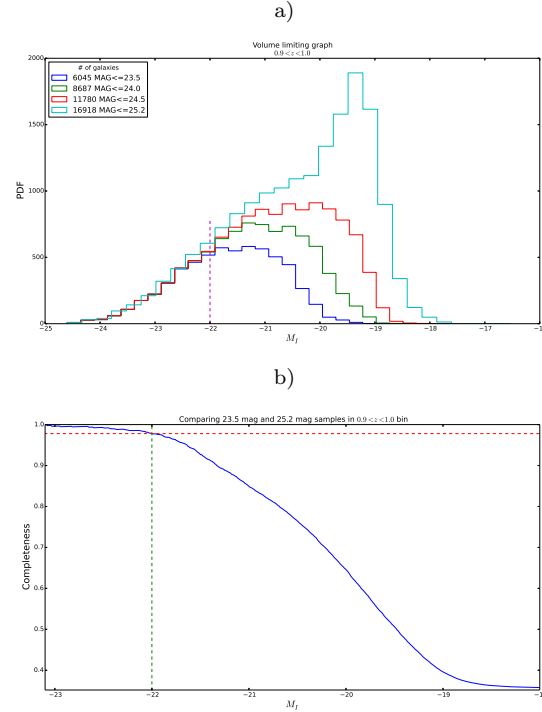


Figure 1. ...

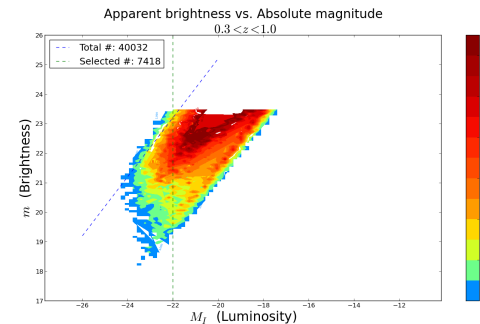


Figure 2. 2-D histogram of galaxies in apparent magnitude ( $m$ ) and absolute magnitude ( $M_I$ ) space.

At  $M_I \sim -22.0$ , we see the sample is beginning to be biased in the  $0.9 < z < 1.0$  bin due to the flux limit. We obtain 97.84% completeness in this bin for  $M_I$ , where completeness is defined as the ratio between the number of galaxies in  $F814W \leq 23.5$  and in  $F814W \leq 25.2$  samples. Fig. 2 shows that at  $M_I < -22$ , we are not affected by flux limit yet. Note that not all galaxies in the sample have postage stamps. Postage stamps may not exist either because the image occurs near the edge of the CCD chips or because the image is too large to fit in a bounding box. **Rachel, please correct my terminology here.** While the former is purely random and doesn't introduce bias, the latter is not. Typically galaxies that are nearby and intrinsically very bright do not have postage stamps associated with them. However, since we used the  $F814W \leq 23.5$  sample, irrespective of the existence of postage stamps, we believe that our conclusions are not affected by this bias.

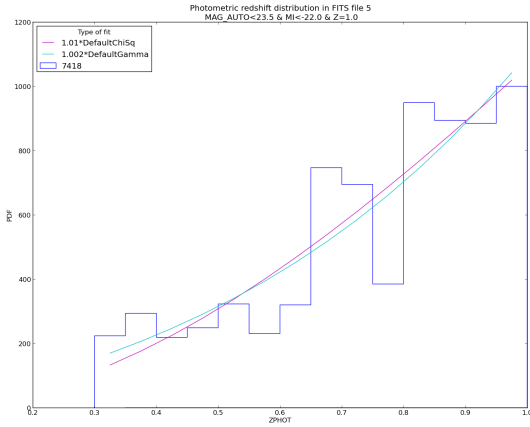


Figure 3.

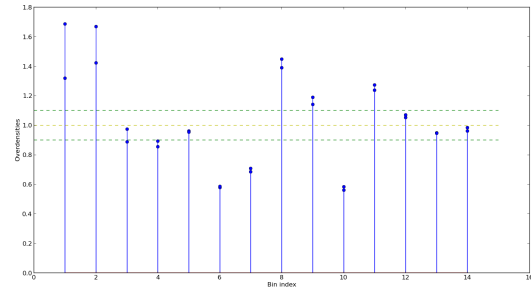
In a wide-field survey, regions of overdensities are identified by computing 3-dimensional comoving densities. But in narrow field surveys like COSMOS, previous work [cite](#) have shown that regions of overdensities can be identified from a 1-dimensional histogram of redshifts, at least for low redshifts.

For our sample of galaxies, volume-limited upto  $z = 1$  by imposing an  $M_I$  cut, we fit parametric models to the histogram of photometric redshifts in order to assign values of overdensity. The figures show chi-squared and gamma functions fitted to the histogram (Baugh and Efstathiou 1993). For low  $z$ , the volume is too small to rely the overdensity values from our model fits and hence the fit is made for  $z \geq 0.3$ . Our binning width,  $\Delta z$  must be small enough so as to be able to identify localized overdensities/underdensities but large enough to not let our conclusions be affected by the errors in photometric redshift. We choose our bins to be 0.05 wide starting from  $z = 0.3$ .

The curves are normalized such that the area under the curves is equal to the number of galaxies considered. **Fig 3 - Histogram with fits**.

Overdensity in a redshift bin is defined as the ratio of the difference between the value given by the histogram and the value predicted by the model to the latter:  $\delta = (N - N_{\text{mod}})/N_{\text{mod}}$ . We leave a 10% margin of safety i.e. iff  $\delta$  from both the fits is greater than 0.1, then we call that redshift bin as an overdense region whereas if  $\delta$  from both the fits is less than  $-0.1$ , then we call that as an underdense region. Redshift bins with  $-0.1 < \delta < 0.1$  are neither too overdense nor too underdense and we call them as ‘unclassified regions’ and discard them from our analysis. **Fig 4 - overdensities vs redshift bins.**

We thus identify the regions  $z = 0.30 - 0.40, 0.65 - 0.75, 0.80 - 0.85$  as overdense,  $z = 0.55 - 0.65, 0.75 - 0.80$  as underdense and  $z = 0.40 - 0.55$  as unclassified. The classification of overdensities and underdensities agrees with the work done by Kovač et al. (2010) on the zCOSMOS field using spectra and 3D density field reconstruction, except for the last two high redshift bins. We believe that this disagreement is due to the errors in our photometric redshifts and



**Figure 4.** Differential comoving densities, integrated within a bin, are plotted. Red line refers to the average comoving density over the entire volume under consideration. The green lines correspond to the 0.9 and 1.1 times the overall average.

the overdensity reported in  $z = 0.875 - 1$  slice is observed by us in  $z = 0.80 - 0.85$  slice.

We see that the region between  $0.85 < z < 1.0$  is neither overdense nor underdense according to our model and hence is not going to be considered for further analysis.

We use therefore relax our luminosity cut so that the sample is volume-limited *not* until  $z = 1$  but until  $z = 0.85$ . This increases the sample size significantly. We choose to impose the cut at  $M_I = -20.8$ , with 95% completeness.

While our method to find overdensities may seem arbitrary, the overdense and underdense regions obtained by the above method have local average number density higher than the global average number density. Refer to Figure ?? . And the unclassified regions fall below or at least lie close to the 10% margin. In computing comoving densities, we have chosen the following values for the  $\Lambda$ CDM parameters:  $\Omega_m = 0.3, \Omega_\Lambda = 0.7$  and  $\Omega_k = 0$  with the Hubble parameter  $h = 0.72$ .

Redshift	# of galaxies	Environment
0.30-0.35	726	Overdense
0.35-0.40	1000	Overdense
0.55-0.60	727	Underdense
0.60-0.65	1070	Underdense
0.65-0.70	2089	Overdense
0.70-0.75	1970	Overdense
0.75-0.80	1159	Underdense
0.80-0.85	2428	Overdense

In the following section, we will compare and analyze the distribution of properties of the galaxies residing in the overdense regions.

*Talk more about environments - Nature and nurture?*

### 3.2 Finding axis ratios ( $b/a$ )

**Should explain in brief Claire’s work.**

If galaxies have elliptical isophotes, its shape and size could be defined by the axis ratio and the area enclosed by a boundary isophote. However, in real galaxy images, the boundary may not be well defined and the shape may not be well approximated by an ellipse. Given the surface brightness (intensity)  $I(\theta)$  of a galaxy image at every angular position  $\theta$ , we can define the tensor of second brightness mo-

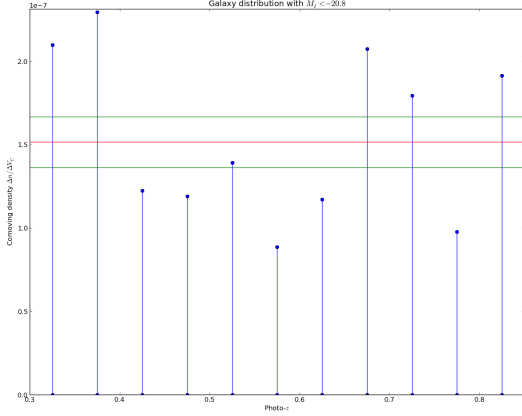


Figure 5.

ments as

$$Q_{ij} = \frac{\int d^2\theta q_I(\theta_i - \bar{\theta}_i)(\theta_j - \bar{\theta}_j)}{\int d^2\theta q_I}, \quad (1)$$

where  $q_I$  is a suitable chosen kernel and  $\bar{\theta}$  is the average angular coordinate, weighted by the same kernel function and  $i, j \in \{1, 2\}$  are the components of  $\bar{\theta}$ . It is common to find in literature two kinds of *complex* ellipticities that quantify the shape of the galaxies -  $\chi$  and  $\epsilon$  defined as

$$\chi = \frac{Q_{11} - Q_{12} + 2iQ_{12}}{Q_{11} + Q_{22}} \quad (2)$$

$$\epsilon = \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22} + 2(Q_{11}Q_{22} - Q_{12})^{1/2}} \quad (3)$$

If the image has elliptical isophotes with axis ratio  $q$ , then

$$|\chi| = \frac{1 - q^2}{1 + q^2}, \quad |\epsilon| = \frac{1 - q}{1 + q}. \quad (4)$$

Conversely, one could compute  $\chi$  or  $\epsilon$  for a galaxy, and obtain an effective, azimuthally averaged (?) axis-ratio  $q$ . However, our method of finding the value of  $q$  is little more complicated. Briefly, we fit two models to each image:

- (i) a Sérsic profile given by the expression

$$S = \Sigma_{1/2} \exp\left(-k(R/R_{\text{eff}})^{1/n} - 1\right), \quad (5)$$

- (ii) two Sérsic component fits: de Vaucouleurs bulge ( $n = 4$ ) + exponential disc profile ( $n = 1$ ),

where

$$R^2 = ((x - x_0) \cos \Phi + (y - y_0) \sin \Phi)^2 + ((y - y_0) \cos \Phi - (x - x_0) \sin \Phi)^2 / q^2,$$

$R_{\text{eff}}$  is the half-light radius of the profile,  $\Sigma_{1/2}$  is the surface brightness at  $R = R_{\text{eff}}$ ,  $(x_0, y_0)$  is the centroid of the image,  $\Phi$  is the profile rotation angle,  $n$  is the Sérsic index,  $k$  is a  $n$ -dependent normalization factor and  $q$  is the axis ratio of the elliptical isophotes, which is the quantity of interest to us. We find the best-fit parameters *simultaneously*, by minimizing the weighted sum of the difference between the image and the PSF-convolved model. More details about the fit can be found in Lackner and Gunn (2012).

## 4 RESULTS

The distribution of the axis ratios that are obtained by the above manner are compared between redshift bins. We use two statistical tests namely the Kolmogorov-Smirnov test and Anderson-Darling test to compare distributions.

We first compare the distribution of the axis ratio in *all* overdense bins and *all* underdense bins. The p-value from both the KS and AD tests are below 0.05, so we believe (95% level) in the ‘alternate hypothesis’ that the overdense and underdense regions have different axis ratio distributions. Next, we will compare distributions between two overdense / underdense regions, where we expect to find similarity, and between an overdense and underdense regions, where we expect the distributions to differ. Figures 7 and 8 show that the distributions are similar when the environments are similar and different when the environments are different. The cumulative distinction functions are also plotted alongside in order to be able to visualize the ‘distance’ between the distributions.

It is useful to consider the root mean square ellipticity since it can characterize the shape of the population/sample of galaxies.

The RMS ellipticities of galaxies in each redshift bin are shown in Figure 9. As one can see, the underdense regions have significantly higher values for RMS ellipticities when compared to the overdense regions. Note that the narrow overdense bin  $0. \leq z < 0.80$  This is consistent with **give physical argument..** From Figs. ?? and ??, the region  $0.4 \leq z < 0.55$  show signs of being marginally underdense but have low RMS ellipticities too that agree with the rest of the overdense regions.

## 5 CONCLUSION

We have shown that

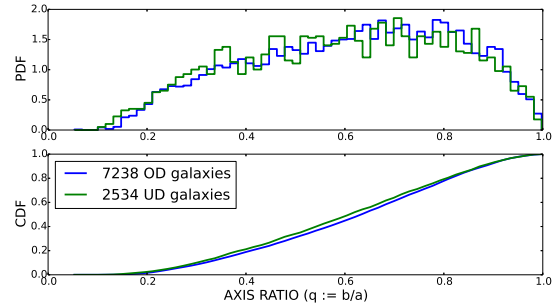
## ACKNOWLEDGMENTS

AK and RM acknowledge the support of NASA ROSES 12-EUCLID12-0004, and program HST-AR-12857.01-A, provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555. AK and RM thank Alexie Leauthaud for many useful discussions.

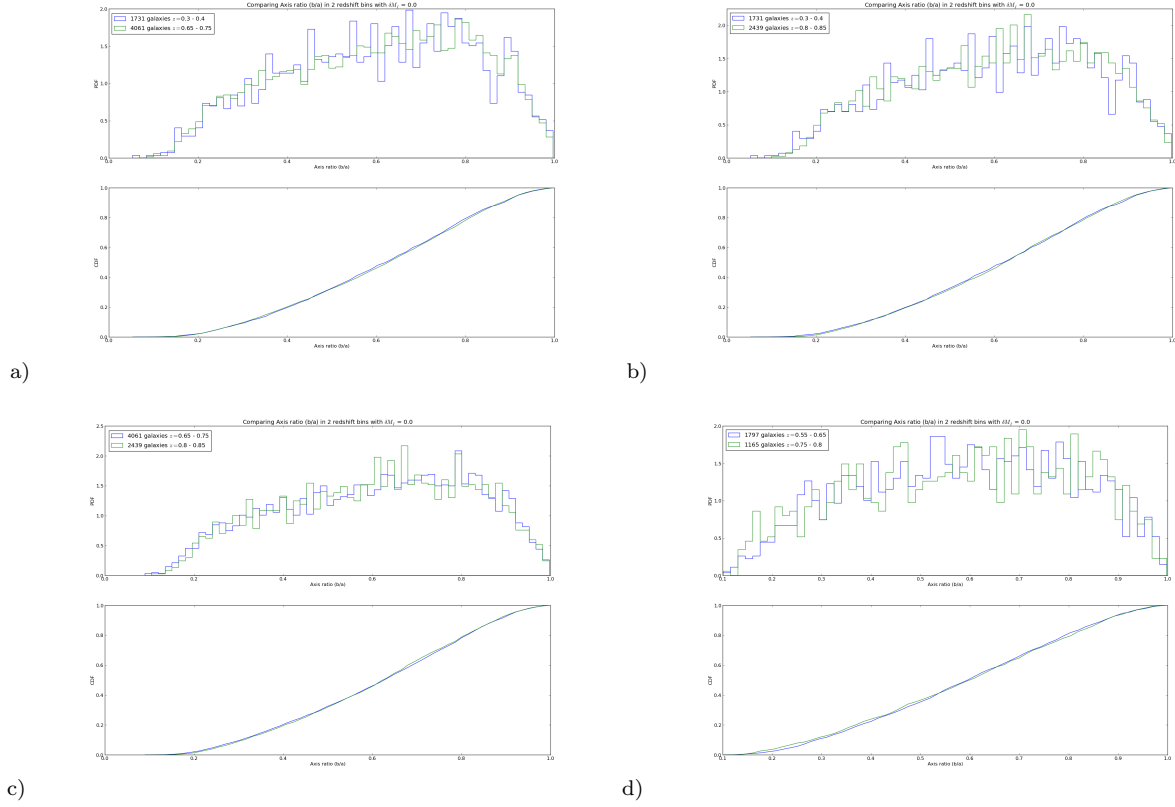
## REFERENCES

- C. M. Baugh and G. Efstathiou. The Three-Dimensional Power Spectrum Measured from the APM Galaxy Survey - Part One - Use of the Angular Correlation Function. *MNRAS*, 265:145, November 1993.
- A. M. Koekemoer, H. Aussel, D. Calzetti, P. Capak, M. Giallisco, J.-P. Kneib, A. Leauthaud, O. Le Fèvre, H. J. McCracken, R. Massey, B. Mobasher, J. Rhodes, N. Scoville, and P. L. Shopbell. The COSMOS Survey: Hubble Space Telescope Advanced Camera for Surveys Observations and Data Processing. *ApJS*, 172:196–202, September 2007. doi: 10.1086/520086.

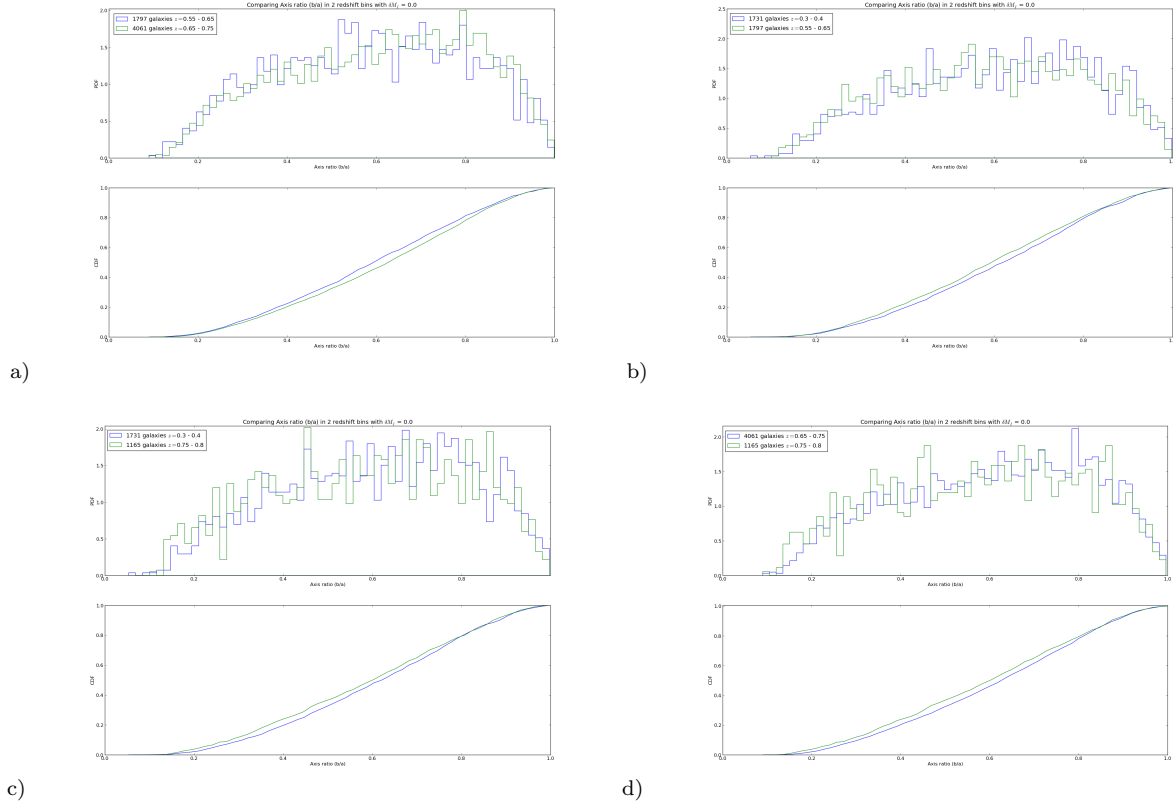
- K. Kovač, S. J. Lilly, O. Cucciati, C. Porciani, A. Iovino, G. Zamorani, P. Oesch, M. Bolzonella, C. Knobel, A. Finoguenov, Y. Peng, C. M. Carollo, L. Pozzetti, K. Caputi, J. D. Silverman, L. A. M. Tasca, M. Scodeggio, D. Vergani, N. Z. Scoville, P. Capak, T. Contini, J.-P. Kneib, O. Le Fèvre, V. Mainieri, A. Renzini, S. Bardelli, A. Bongiorno, G. Coppia, S. de la Torre, L. de Ravel, P. Franzetti, B. Garilli, L. Guzzo, P. Kampczyk, F. Lamareille, J.-F. Le Borgne, V. Le Brun, C. Maier, M. Mignoli, R. Pello, E. Perez Montero, E. Ricciardelli, M. Tanaka, L. Tresse, E. Zucca, U. Abbas, D. Bottini, A. Cappi, P. Cassata, A. Cimatti, M. Fumana, A. M. Koekemoer, D. Maccagni, C. Marinoni, H. J. McCracken, P. Memeo, B. Meneux, and R. Scaramella. The Density Field of the 10k zCOSMOS Galaxies. *ApJ*, 708:505–533, January 2010. doi: 10.1088/0004-637X/708/1/505.
- C. N. Lackner and J. E. Gunn. Astrophysically motivated bulge-disc decompositions of Sloan Digital Sky Survey galaxies. *MNRAS*, 421:2277–2302, April 2012. doi: 10.1111/j.1365-2966.2012.20450.x.
- A. Leauthaud, R. Massey, J.-P. Kneib, J. Rhodes, D. E. Johnston, P. Capak, C. Heymans, R. S. Ellis, A. M. Koekemoer, O. Le Fèvre, Y. Mellier, A. Réfrégier, A. C. Robin, N. Scoville, L. Tasca, J. E. Taylor, and L. Van Waerbeke. Weak Gravitational Lensing with COSMOS: Galaxy Selection and Shape Measurements. *ApJS*, 172: 219–238, September 2007. doi: 10.1086/516598.
- R. Mandelbaum, C. M. Hirata, A. Leauthaud, R. J. Massey, and J. Rhodes. SHERA: SHEar Reconvolution Analysis, August 2011. Astrophysics Source Code Library.
- R. Mandelbaum, B. Rowe, J. Bosch, C. Chang, F. Courbin, M. Gill, M. Jarvis, A. Kannawadi, T. Kacprzak, C. Lackner, A. Leauthaud, H. Miyatake, R. Nakajima, J. Rhodes, M. Simet, J. Zuntz, B. Armstrong, S. Bridle, J. Coupon, J. P. Dietrich, M. Gentile, C. Heymans, A. S. Jurling, S. M. Kent, D. Kirkby, D. Margala, R. Massey, P. Melchior, J. Peterson, A. Roodman, and T. Schrabback. The Third Gravitational Lensing Accuracy Testing (GREAT3) Challenge Handbook. *ApJS*, 212:5, May 2014. doi: 10.1088/0067-0049/212/1/5.
- B. Mobasher, P. Capak, N. Z. Scoville, T. Dahlen, M. Salvato, H. Aussel, D. J. Thompson, R. Feldmann, L. Tasca, O. Le Fèvre, S. Lilly, C. M. Carollo, J. S. Kartaltepe, H. McCracken, J. Mould, A. Renzini, D. B. Sanders, P. L. Shopbell, Y. Taniguchi, M. Ajiki, Y. Shioya, T. Contini, M. Giavalisco, O. Ilbert, A. Iovino, V. Le Brun, V. Mainieri, M. Mignoli, and M. Scodeggio. Photometric Redshifts of Galaxies in COSMOS. *ApJS*, 172:117–131, September 2007. doi: 10.1086/516590.
- N. Scoville, H. Aussel, M. Brusa, P. Capak, C. M. Carollo, M. Elvis, M. Giavalisco, L. Guzzo, G. Hasinger, C. Impey, J.-P. Kneib, O. LeFevre, S. J. Lilly, B. Mobasher, A. Renzini, R. M. Rich, D. B. Sanders, E. Schinnerer, D. Schminovich, P. Shopbell, Y. Taniguchi, and N. D. Tyson. The Cosmic Evolution Survey (COSMOS): Overview. *ApJS*, 172:1–8, September 2007. doi: 10.1086/516585.



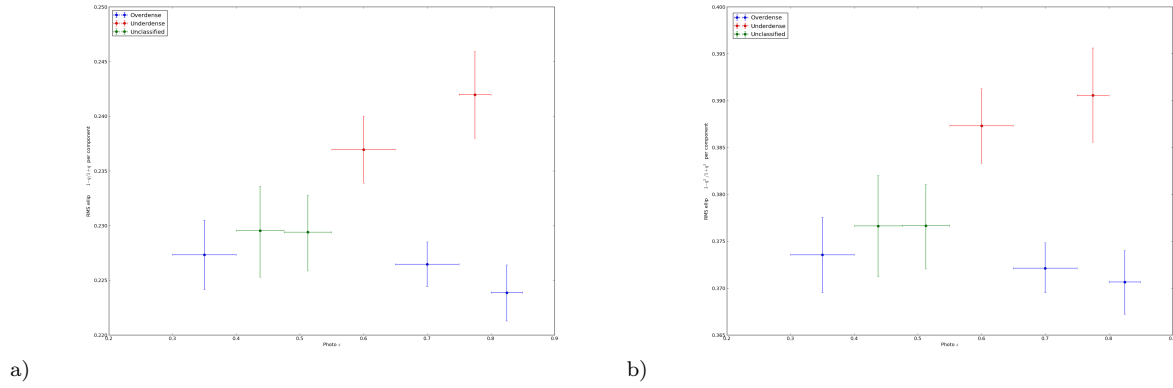
**Figure 6.** KS p-value = 0.004 & AD p-value = 0.001



**Figure 7.** Comparison of axis ratios of galaxies in similar environments.  $p$ -values from the KS and AD test are (will be) given in the plot.



**Figure 8.** Comparison of axis ratios of galaxies in contrasting environments.  $p$ -values from the KS and AD test are (will be) given in the plot.



**Figure 9.** Left: RMS ellipticities with ellipticity defined as  $\frac{1-q}{1+q}$ . Right: RMS ellipticities with ellipticity defined as  $\frac{1-q^2}{1+q^2}$ .