UNIVERSITY OF CALIFORNIA
RIVERSIDE

Evolution of Galaxies in Different Environments Over Cosmic Time

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Physics

by

Nima Chartab Soltani

March 2021

Dissertation Committee:
Dr. Bahram Mobasher, Chairperson
Dr. Laura Sales
Dr. Brian Siana

The Dissertation of Nima Chartab Soltani is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgements

I would like to thank everyone who helped me during my Ph.D. years and made this dissertation possible. Words cannot express how grateful I am to my parents and my brother for their tremendous understanding and encouragement throughout my life. I would not have made it this far without them.

I would like to express my special thanks to my advisor, Professor Bahram Mobasher, for his dedicated support and guidance. He has provided me with great opportunities and gave me the freedom to pursue various projects. His insightful discussions and suggestions over these past several years have been invaluable.

I am very grateful to all members of the Physics and Astronomy Department at UCR. My sincere thanks go to Drs. Laura Sales and Brian Siana for serving as my dissertation committee members and spending their time to read my thesis.

A very special thank you to Zahra Sattari for her moral support and great help in writing research papers and completing this thesis. During the last two years of my Ph.D. she has always encouraged me throughout this experience, especially when the workload was heavy.

I would like to thank Dr. Ali Ahmad Khostovan for always being so helpful in numerous ways, especially in the first years of my Ph.D. when I just started my research work. I will never forget the great memories and fun activities we have had together.

I am indebted to all my incredible friends and groupmates: Zahra Sattari, Abtin Shahidi, Marziye Jafariyazani, Dr. Melanie Simet, Stephen Hannon, Amanda Pagul, and also Drs. Shoubaneh Hemmati, Behanm Darvish, Hooshang Nayyeri and Irene Shivaei. We had lots of constructive discussions during our group meetings and my research benefited from their feeback and suggesstions.

During my Ph.D., I was very fortunate to collaborate with different research teams, including CANDELS, MOSDEF and Hawaii Two-0. I want to greatly thank the productive meetings and collaborations, which helped me improve my research. In particular, I would like to thank Professors David Sanders, Sune Toft, Istvan Szapudi, Dr. Conor McPartland, Lukas Zalesky and John Weaver for the great discussions in our weekly meetings and also for the teamwork during observing nights.

I finally wish to thank my best friends Sadra Sadraddini, Babak Bahrami Asl, Ehsan abedi, Ali Fahminiya and Hamidreza Akbari. (Too many to list here, but you know who you are!) Their friendship will be cherished for life.

The text of this dissertation, in part, is a reprint of the material as it appears in journal articles entitled *Large Scale Structures in the CANDELS Fields: The Role of the Environment in Star Formation Activity* (2020ApJ...890....7C), and *The MOSDEF Survey: Environmental dependence of the gas-phase metallicity of galaxies at* $1.4 \leq z \leq 2.6$ (2021ApJ...908..120C).

To my parents for all the support.

ABSTRACT OF THE DISSERTATION

Evolution of Galaxies in Different Environments Over Cosmic Time

by

Nima Chartab Soltani

Doctor of Philosophy, Graduate Program in Physics
University of California, Riverside, March 2021
Dr. Bahram Mobasher, Chairperson

This thesis focuses on the effect of the environment of galaxies on their star formation activity and the metal content of their interstellar medium. I develop a technique to reconstruct the underlying number density field of galaxies in the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS). I then use these measurements to estimate stellar mass and environmental quenching efficiencies out to $z \sim 3.5$. I find that the environmental quenching efficiency increases with increasing stellar mass and decreasing redshift. I show that a dense environment can quench very massive galaxies as early as $z \sim 3$. These observations provide a supporting argument for a scenario where the termination of cool gas accretion happens in a dense environment, and the galaxy starts to consume its remaining gas reservoir in depletion time. The depletion time is shorter for massive galaxies, so they are expected to become quenched faster. I also study the gas-phase metallicity of galaxies in different environments using near-IR spectroscopy of the MOSFIRE Deep Evolution Field (MOSDEF) survey. Based on gas-phase oxygen abundance computed from rest-frame optical emission lines, I find that at

a fixed stellar mass, galaxies in overdense regions have lower metallicity than their field counterparts at $z \sim 2.3$, but they become more metal-rich as they evolve to $z \sim 1.5$. My results suggest that the efficient gas cooling mechanisms at high redshifts result in the prominent accretion of primordial metal-poor gas into the galaxies in overdensities. However, as galaxies evolve to the lower redshifts ($z < 2$), the shock-heated gas in overdensities with massive halos cannot cool down efficiently, ramping up the gas-phase metallicity of galaxies. Future surveys will provide statistically significant samples of high redshift galaxies in diverse environments that can further test results from the present study. I build a machine learning model to transfer the knowledge gained in fields with a wealth of observations to those which lack such extensive observations. The technique provides valuable information to optimize the observing strategy for future surveys and offers complimentary data in the wavebands not accessible by these surveys.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The standard ΛCDM model of cosmology implies that the structures in the universe form hierarchically, meaning that small dark matter halos merge to form the larger ones. Galaxies within these structures undergo different evolutionary phases, all affected by the interplay between galaxies and their respective environments. This co-evolution of galaxies and large scale structures leaves imprints on the observable properties of galaxies that can be used as a proxy to study both the nature of processes involved in the evolution of individual galaxies and their assembly to form large scale structures.

1.1 Galaxy formation and evolution

First galaxies were born from primordial overdensities devoid of any metals. They start to build their stars rapidly as they are fed with more cool gas available in the early epoch of structure formation. According to the galaxy formation theory of White and Rees (1978), when a cloud of cold gas from intergalactic medium (IGM) falls into the potential well of a galaxy halo, it gets shock-heated to the virial temperature and then slowly cools down (e.g., Silk, 1977). This shock-heated gas starts to cool down from

the inside out as it dissipates its internal energy, meaning that the cooling radius (i.e., a radius containing the cooled gas) grows with time and feeds more cold gas into the disk of a galaxy as a fuel to maintain its star formation rate. However, Birnboim and Dekel (2003) found that a virial shock does not get formed in low-mass halos ($\lesssim 10^{11} M_{\odot}$), which are prevalent at high redshifts ($z > 2$). Moreover, Dekel and Birnboim (2006) found that at $z \gtrsim 2$, cold gas accretion can also be maintained in big halos due to funneling of cold dense gas through filaments. They found that galaxies residing in halos with a mass $\gtrsim 10^{12} M_{\odot}$ are surrounded by shock-heated gas almost independent of the redshift; however, cold gas can still penetrate the hot media at $z \gtrsim 2$. As halos grow to lower redshifts ($z < 2$), the cold gas budget of the galaxies decreases, which explains the suppression of the cosmic star formation rate in the late universe.

Additionally, feedback processes regulate the population of massive and low-mass galaxies. Förster Schreiber et al. (2019) observed that active galactic nuclei (AGN) activities and outflows are prevalent in massive galaxies ($M_{*} \gtrsim 10^{10.5} M_{\odot}$). On the other hand, supernovae explosions and stellar winds are more effective for low-mass galaxies, removing part of the interstellar medium gas into the circumgalactic medium (CGM) (Larson, 1974). Although all these galaxy formation and evolution theories were established to explain general observables such as cosmic star formation history or luminosity function of galaxies, more direct observational pieces of evidence for such theories are limited as of today, especially at high redshifts.

1.2 Galaxy environments

The galaxy environment plays a significant role in the star formation activity of galaxies (Peng et al., 2010; Quadri et al., 2012; Scoville et al., 2013; Lin et al., 2014; Lee et al., 2015; Darvish et al., 2016; Balogh et al., 2016; Nantais et al., 2016; Kawinwanichakij

et al., 2017; Guo et al., 2017; Fossati et al., 2017; Ji et al., 2018; Pintos-Castro et al., 2019; Contini et al., 2020; Old et al., 2020; Ando et al., 2020). Massive halos traced by overdensities at low and intermediate redshifts lack any cold gas accretion as predicted by simulations (Kereš et al., 2005). Thus, galaxies within massive halos start to starve and shut down their star formation. However, this star formation shut down, known as quenching, does not happen for all the galaxies simultaneously. Once the parent halo of a galaxy group is surrounded by the shock-heated gas and no more fresh cold gas is provided for the galaxies within the halo, galaxies will consume their remaining gas reservoirs at a specific time scale. McGee et al. (2014) showed that this timescale is shorter for massive galaxies as they consume their gas at a higher rate. In other words, in the absence of cool gas accretion, massive galaxies will enter the quenching phase sooner.

Moreover, galaxy interactions and mergers can also affect the star formation activity of galaxies (e.g., Larson and Tinsley, 1978; Donzelli and Pastoriza, 1997). More interestingly, galaxy interactions and mergers leave morphological imprints on the galaxies. As a pioneer work, Dressler (1980) found that the morphology-density relation exists in the local universe, such that late-type star-forming disky galaxies preferentially reside in lower density regions while early-type passive ellipticals prefer dense regions. This observation can be understood with tidal interaction of galaxies as well as other mechanisms that remove gas from galaxies such as ram pressure stripping (e.g., Boselli and Gavazzi, 2014). Ram pressure stripping (i.e., the removal of a gas from a galaxy due to its interaction with the intra-cluster medium) is found to be effective only for low-mass galaxies at low redshifts (Gunn and Gott, 1972). The most convincing observation of ram pressure stripping conducted in the local universe is the observed tens of kpc tails of gas seen in several cluster galaxies (e.g., Crowl et al., 2005; Machacek et al., 2006). Therefore, it is known that, to some extent, the ram pressure stripping

is responsible for the environmental quenching of galaxies in the local universe. Other mechanisms such as mergers and pure starvation we mentioned above are not very effective at low redshifts. The merger cross-section is small at $z = 0$ due to a high velocity dispersion of massive clusters present in the local universe, making merger interactions less effective. Also, gas depletion due to pure starvation takes > 5 Gyr in the local universe (Balogh et al., 2016) which rules this candidate out from being the main source of environmental quenching at $z \sim 0$.

At the early epoch of cluster formation (high redshifts), ram pressure stripping is less important as halos have not grown yet to be massive enough to provide higher velocities for infalling galaxies. On the other hands, pure starvation gets important at high redshifts as the depletion time is < 1 Gyr at $z \sim 2$ for galaxies with $M_* \gtrsim 10^{10.5} M_\odot$ (McGee et al., 2014). Therefore, the mechanisms that quench galaxies at high redshifts may differ from those playing a significant role in the local universe.

Another piece of the puzzle is the metal content of the interstellar medium of galaxies. This indicator, known as gas-phase metallicity, traces both star formation rate and cold gas accretion rate of galaxies (e.g., Mannucci et al., 2010). It is observed that the gas-phase metallicity of galaxies is strongly correlated with their stellar mass such that massive galaxies have higher gas-phase metallicity out to $z \sim 3.5$ (e.g., Sanders et al., 2020). It is also observed that at a fixed stellar mass, galaxies with higher star formation rates have lower gas-phase metallicities (e.g., Mannucci et al., 2010). At first glance, the inverse correlation between SFR and gas-phase metallicity seems unphysical since as stars form, they produce heavy elements, release those heavy elements to the ISM and eventually increase the gas-phase metallicity of the galaxy. Although this explanation is valid for isolated galaxies with no inflow or outflow (closed system), the infall of cold metal-poor gas is the primary driver of star formation in galaxies. Gas accretion from IGM adds chemically poor gas into the ISM, which lowers the galaxy's

gas-phase metallicity. This makes a complicated interplay between gas accretion rate, star formation rate, and gas-phase metallicity (Lilly et al., 2013).

In addition, simulations and observations suggest that the star formation rate and/or cold gas accretion rate depend on the galaxy's environment. Therefore, it is expected to observe a correlation between gas-phase metallicity and the environment of galaxies. If the termination of cool gas accretion is responsible for the quenching of galaxies, so the gas-phase metallicity of galaxies should rapidly increase during the depletion time when they continue the ongoing star formation rate, but their gas reservoirs do not replenish with fresh gas. Moreover, over cosmic time, stars within the galaxies produce heavier elements and feedback processes expel those heavy elements to the CGM and IGM, so the gas accreted in overdensities at the late universe is pre-processed and has higher metallicity. Using IllustrisTNG simulations, Gupta et al. (2018) predicted that the infalling gas in the denser environment is more metal-enriched compared to low-density regions. In the local universe, it is also observed that at a given stellar mass, galaxies in overdensities have higher gas-phase metallicities (e.g., Cooper et al., 2008; Ellison et al., 2009; Peng and Maiolino, 2014; Wu et al., 2017; Schaefer et al., 2019). However, the environmental dependence of galactic properties at $z > 1$ is still a matter of debate. Thus, future analyses are essential to constrain the theoretical picture of galaxy evolution in the early phase of structure formation.

In Chapter 2 and 3 of this thesis, using the deepest photometry available from Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS)) and wealth of spectroscopic data with a ground-based Keck telescope, we tackle a fundamental question regarding the environmental dependence of both star formation activity (quiescent fraction) and gas-phase metallicity out to $z \sim 3$. These investigations help address questions in the literature about the co-evolution of galaxies and their environments at high redshifts.

1.3 Ongoing and future surveys

Ongoing surveys and future missions will broaden our understanding of the evolution of galaxies within large-scale structures. Ongoing Hawaii Two-0 (H20) survey covers a total area of 20 deg^2 with Subaru Hyper-Suprime-Cam (HSC) in the Spitzer Legacy Survey (SLS) fields. The combination of area, depth, and wavelength coverage will provide a statistically significant number of overdensities (protoclusters) at high redshifts. This will revolutionize studies about the early stage of structure formation as > 100 overdensities (protocluster) are expected to be found in 20 deg^2 of H20 fields. However, accurate and unbiased photometric redshifts are essential to map large-scale structures and identify overdensities. Euclid will provide deep observations in near-IR bands over the same fields once it is launched, and hence these deep fields with 20 deg^2 coverage will have HSC/optical, Euclid/near-IR and Spitzer/IR observations. Therefore, the number of bands will not be as extensive as other existing small-area fields such as COSMOS. In Chapter 4, we investigate the possibility of transferring information from a field with extensive observations such as COSMOS to the field where a limited number of bands are available. We develop a machine learning-based model to predict missing-band fluxes for galaxies and show how this can be useful to measure accurate photometric redshifts and physical parameters such as stellar mass and star formation rate. In the end, Chapter 5 highlights the significant results of the thesis.

Chapter 2

Large Scale Structures in the CANDELS Fields

Abstract

We present a robust method, weighted von Mises kernel density estimation, along with boundary correction to reconstruct the underlying number density field of galaxies. We apply this method to galaxies brighter than HST/F160w $\leq$ 26 AB mag at the redshift range of $0.4 \leq z \leq 5$ in the five CANDELS fields (GOODS-N, GOODS-S, EGS, UDS, and COSMOS). We then use these measurements to explore the environmental dependence of the star formation activity of galaxies. We find strong evidence of environmental quenching for massive galaxies ($M \gtrsim 10^{11} M_{\odot}$) out to $z \sim 3.5$ such that an over-dense environment hosts $\gtrsim$20% more massive quiescent galaxies compared to an under-dense region. We also find that environmental quenching efficiency grows with stellar mass and reaches $\sim$ 60% for massive galaxies at $z \sim 0.5$. The environmental quenching is also more efficient in comparison to the stellar mass quenching for low mass galaxies ($M \lesssim 10^{10} M_{\odot}$) at low and intermediate redshifts ($z \lesssim 1.2$). Our findings concur thoroughly with the "over-consumption" quenching model where the termination of cool

gas accretion (cosmological starvation) happens in an over-dense environment and the galaxy starts to consume its remaining gas reservoir in depletion time. The depletion time depends on the stellar mass and could explain the evolution of environmental quenching efficiency with the stellar mass.

2.1 Introduction

It is now well established that the observed properties of galaxies are correlated with their host environment. In the local Universe, the environmental dependence of galaxy morphology and star formation rate (SFR) confirms that early-type passive galaxies often reside in dense environments, such as galaxy groups and clusters, whereas late-type and star-forming systems are mostly found in less dense environments, so-called field (e.g., Dressler, 1980; Kauffmann et al., 2004; Balogh et al., 2004; Peng et al., 2010; Woo et al., 2013). However, the situation is not entirely settled at intermediate to high redshifts. While there is convincing evidence for a density-morphology relation at intermediate redshifts (e.g., Capak et al., 2007), the exact trend in the density-SFR relation remains controversial. Some studies show a reversal relation so that on average the SFR is higher in dense environment (Elbaz et al., 2007; Cooper et al., 2008), others find no significant correlation (Grützbauch et al., 2011; Scoville et al., 2013; Darvish et al., 2016) and some observe the same relation as in the local Universe (Patel et al., 2009). Recently, an increasing number of studies have found that the locally observed environmental quenching persists at least out to $z \sim 2$ (e.g., Fossati et al., 2017; Guo et al., 2017; Kawinwanichakij et al., 2017; Ji et al., 2018). Therefore, a comprehensive study is needed to ascertain the role of the environment in star formation activity of galaxies at high redshifts.

Accurate measurement of the environment of galaxies is needed before any such study can be performed. One can use a variety of density estimators to quantify the environment in which galaxies are located. Darvish et al. (2015) have reviewed and compared different density estimators, including adaptive weighted kernel smoothing, 10^{th} and 5^{th} nearest neighbors, count-in-cell, weighted Voronoi tessellation and Delaunay triangulation. Comparing with simulations, they found that the weighted kernel smoothing method is more reliable than widely-used nearest neighbor and count-in-cell methods. Although kernel density estimation is a powerful and reliable technique for estimating the density field of galaxies, there are considerable ambiguities in the selection of the appropriate kernel function and optimized kernel window size (bandwidth). The selection of the bandwidth is the most crucial step in kernel density estimation. Small bandwidth results in an under-smoothed estimator, with high variability. On the other hand, large bandwidth causes an over-smoothed (biased) estimator. Boundary problem is another common issue regardless of the density estimator and the net effect is an underestimation of density near the boundaries since galaxies beyond the edge of the survey are missed. In this paper, we develop a new technique, weighted von Mises kernel density estimation considering boundary correction to reconstruct the density field of galaxies.

While measurement of density enhancement is available in contiguous wide-area surveys such as the Cosmic Evolution Survey (COSMOS) (Scoville et al., 2007), studying the influence of environment on the evolution of low mass galaxies ($M \lesssim 10^{10} M_{\odot}$) requires deep surveys that are often performed over much smaller areas because of the trade-off between the area coverage and the depth in galaxy surveys. The Cosmic Assembly Near-IR Deep Extragalactic Legacy Survey (CANDELS; Grogin et al., 2011; Koekemoer et al., 2011) includes extensive data in five fields, ideal for any study of the evolution of galaxies with redshift. The CANDELS provides: (1) Multi-waveband deep

data from the Hubble Space Telescope (HST) and Spitzer Space Telescope for all the five fields; (2) Accurate measurement of the photometric redshifts, stellar mass and SFRs with their probability distributions; (3) Extensive spectroscopic observations for galaxies; (4) Constraints on the cosmic variance using five widely separated fields. These make the CANDELS fields ideal for such studies. The challenge, however, is to perform a reliable estimate of the density measurements for such fields with limited volume.

In this paper, we make a publicly available catalog of density measurements for 86,716 galaxies brighter than F160w $\leq$ 26 AB mag at $0.4 \leq z \leq 5$ in all the five CANDELS fields using weighted von Mises kernel density estimation with taking into account the boundary effect. We use a grid search cross-validation method to optimize the bandwidth of the kernel function. In order to reduce the projection effect, we use full photo-z probability distribution function (PDF) of individual galaxies (Kodra et al. in prep.).

The paper is organized as follows: In Section 2.2, we discuss the data and describe the Spectral Energy Distribution (SED) fitting procedure to measure the physical properties of galaxies. Section 2.3 describes our methodology for measuring the local environment of galaxies and presents the galaxy environment catalog and large scale structure maps. In Section 2.6, we explore the role of environment in the star formation activity of galaxies. We discuss our results in Section 2.7 and summarize them in Section 2.8.

Throughout this work, we assume a flat ΛCDM cosmology with $H_0 = 100h$ kms^{-1}Mpc^{-1}, $\Omega_{m_0} = 0.3$ and $\Omega_{\Lambda_0} = 0.7$. All magnitudes are expressed in the AB system and the physical parameters are measured assuming a Chabrier (2003) IMF.

2.2 Data

We use the HST/F160w (H-band) selected catalogs of the five CANDELS fields covering a total area of $\sim$ 960 arcmin^2: GOODS-S (Guo et al., 2013), GOODS-N (Barro et al., 2019), COSMOS (Nayyeri et al., 2017), EGS (Stefanon et al., 2017), and UDS (Galametz et al., 2013). The comoving coverage area of each field as a function of redshift is shown in Figure 2.1.

The Catalogs are a combination of CANDELS wide, deep, and Hubble Ultra-Deep Field (HUDF) regions. The point source 5σ limiting AB magnitude ranges from $\sim$ 27.4 to $\sim$ 29.7 in the wide and HUDF area, respectively. However, the 5σ limiting magnitude is brighter for the extended objects and depends on the surface brightness profile of sources. The limiting magnitude in the wide field reaches $H_{lim} \sim 26$, which corresponds to the 50% completeness at the median size of sources (Guo et al., 2013).

We utilize the updated CANDELS photometric redshift catalog (Kodra et al. in prep.), which provides accurate photometric redshifts with normalized median absolute deviation ($\sigma_{\rm NMAD}$) of $\sim$ 0.02, combined with the spectroscopic/3D-HST grism redshifts (z_{best}). The catalog also contains redshift probability distribution functions (PDFs) of galaxies determined by the minimum Frechet distance method. The Frechet distance (Alt and Godau, 1995) is a measure of similarity between two curves (e.g., two measurements of photo-z PDFs). The best PDF is obtained based on the minimum of the Frechet distance among six independent z-PDF measurements.

In this work, we measure the local number density for a total of 86,716 galaxies selected based on the following criteria (Table 2.1):

- Removing the stars by requiring SExtractor's stellarity parameter CLASS_STAR $<$ 0.9.

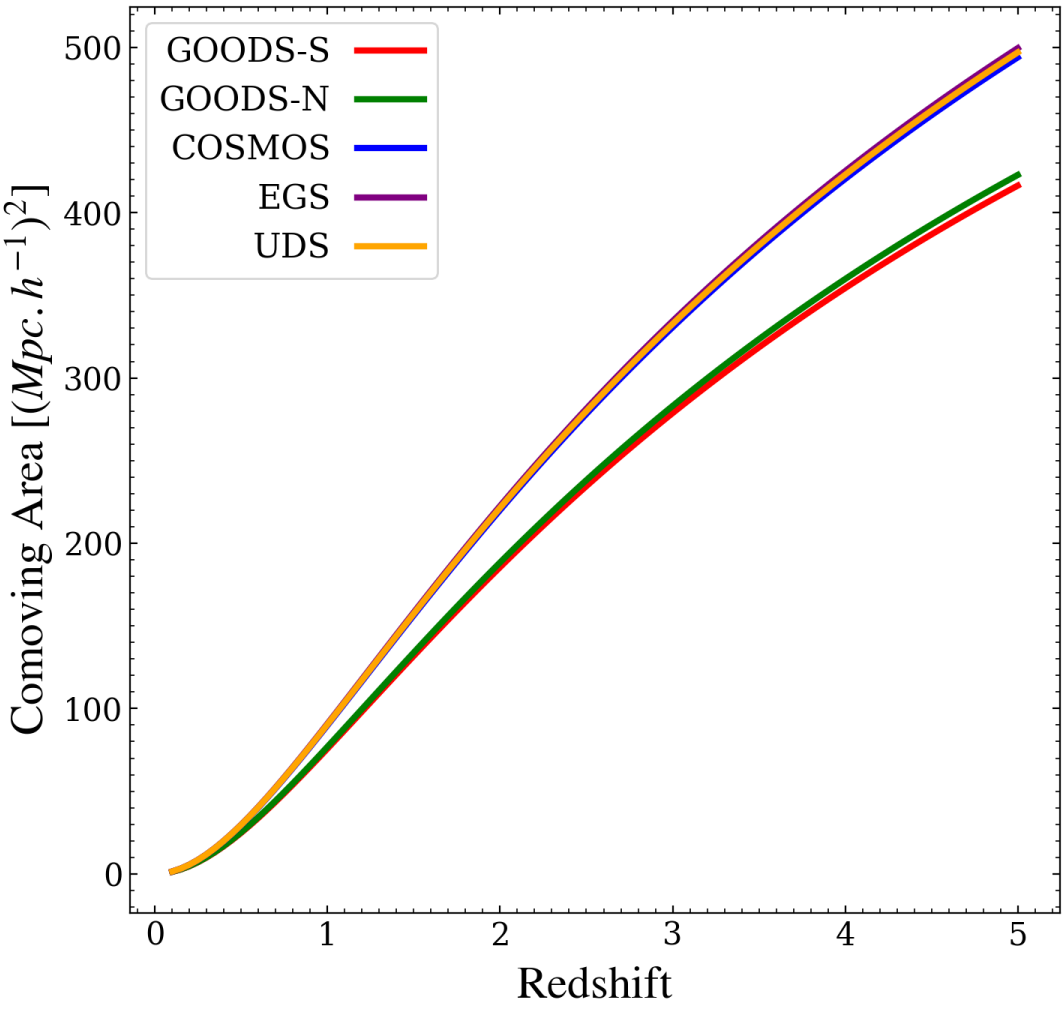


Figure 2.1: The comoving areal cover of the fields as a function of redshift. Each field covers $\sim 25h^{-2}$ Mpc^2 comoving area at $z \sim 0.5$ up to $\sim 450h^{-2}$ Mpc^2 at $z \sim 5$.

- Covering a redshift range of $0.4 \leq z \leq 5$. We select galaxies with greater than 95% probability of being in this redshift range. We limit our analysis to $z \geq 0.4$ due to the small volume of the survey at lower redshifts.

- A cut on H-band magnitude to remove the sources fainter than 26 AB mag. Although the fields have different 5σ limiting magnitudes, we use a similar magnitude cut to have homogeneous and comparable samples.

2.2.1 Stellar Mass & Star Formation Rates

We perform Spectral Energy Distribution (SED) fitting to derive physical parameters of galaxies such as stellar mass and star formation rate (SFR). We use the LePhare code (Arnouts et al., 1999; Ilbert et al., 2006) combined with a library of synthetic spectra generated by the Bruzual and Charlot (2003) population synthesis code. To perform SED fitting, we fix redshifts on z_{best} from the updated version of the CANDELS photometric redshift catalog. We assume an exponentially declining star formation history with nine e-folding times in the range of $0.01 < \tau < 30$ Gyr. We adopt the Chabrier (2003) initial mass function, truncated at 0.1 and 100 $M_{\odot}$, and Calzetti et al. (2000) attenuation law to apply dust extinction ($E(B-V) \leq 1.1$). The code also includes emission lines using Kennicutt (1998) relation between SFR and UV luminosity, as described in Ilbert et al. (2009). Three different stellar metallicities are considered: $Z = 0.02, 0.008$, and 0.004.

The LePhare code computes fluxes in all given bands for each template, then finds the template with minimum χ^2 based on the model and observed fluxes. The best values of the physical parameters come from the template with the minimum χ^2. The code also provides the median of the stellar mass (M), SFR and specific SFR (sSFR=SFR/M) along with uncertainties obtained from the marginalized probability

Table 2.1: Summary of the data used in this work

Field	Reference	Area (arcmin^2)	5σ depth (AB)	All objects	Final sample[a]	spec/grism z[b]
GOODS-S	Guo et al. (2013)	170	27.36	34930	14200	16%
GOODS-N	Barro et al. (2019)	170	27.8	35445	15746	18%
COSMOS	Nayyeri et al. (2017)	216	27.56	38671	18896	7%
EGS	Stefanon et al. (2017)	206	27.6	41457	19670	13%
UDS	Galametz et al. (2013)	202	27.45	35932	18204	7%

a Criteria: (1) CLASS_STAR < 0.9, (2) Probability of being in $0.4 \leq z \leq 5$ greater than 95%, (3) H$\leq$ 26 AB mag.
b The percentage of galaxies in the final sample with confirmed spectroscopic/grism redshifts.

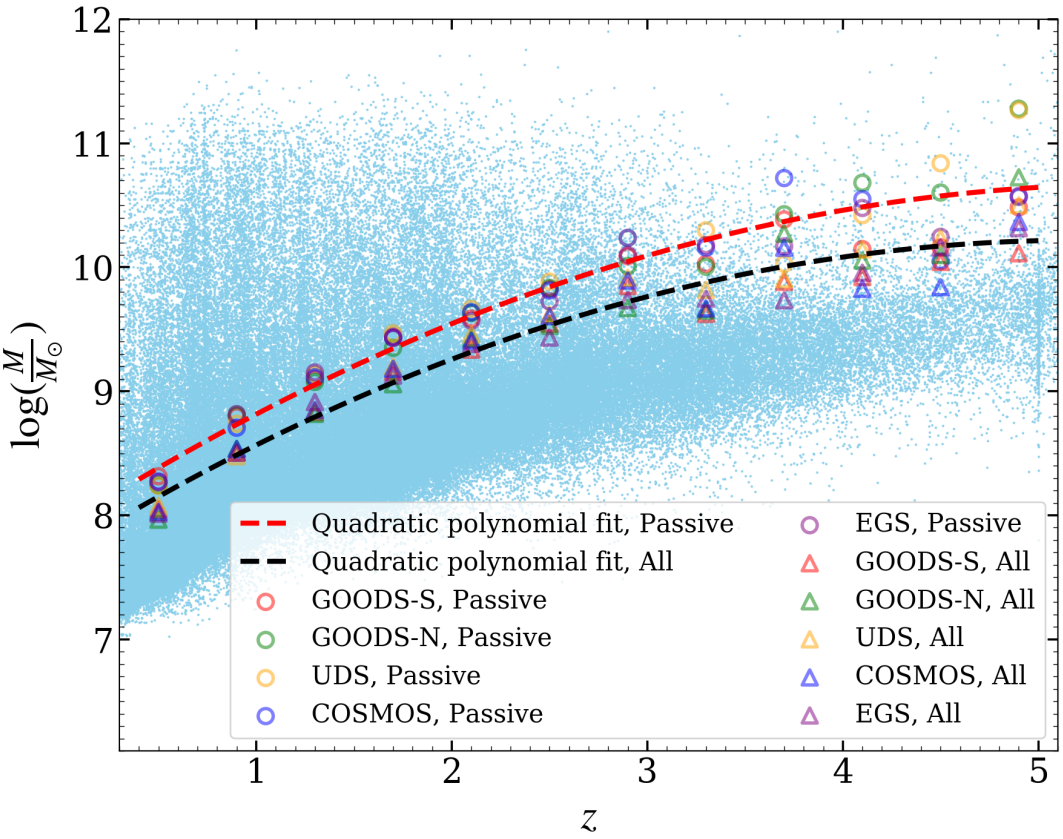


Figure 2.2: The stellar mass of galaxies as a function of their redshifts. Red (black) dashed line represents the 95% stellar mass completeness limit ($H_{lim} = 26$) for the passive (all) galaxies, determined using the method of Pozzetti et al. (2010).

distribution (probability$\propto e^{-\chi^2/2}$) of each parameter. In this work, we use the median values for stellar mass, SFR and sSFR. We also obtain U,V and J rest-frame colors from best-fit SEDs.

Figure 2.2 shows the distribution of stellar mass as a function of redshift for galaxies in the five fields. The stellar mass completeness limit (95%) associated with H_{lim}=26 is determined using the method introduced by Pozzetti et al. (2010). We divide the sample into redshift bins, separately in the case of all and quiescent populations. We utilize rest-frame U,V and J colors along with Muzzin et al. (2012) criteria to select quiescent galaxies at $z < 4$. Beyond this redshift, we use a sSFR cut derived from

the first quartile ($<$ 25% percentile) of the sSFR distribution to build a sub-sample of passive galaxies. We then measure the limiting stellar mass, M_{lim}, for galaxies in the sub-sample, defined as the stellar mass a galaxy would have if it had a magnitude equal to the adopted magnitude limit of the survey (H_{lim}). If we consider constant mass-to-light ratio, then M_{lim} for a galaxy with stellar mass M can be computed as $\log M_{lim} = \log M + 0.4(H - H_{lim})$. The stellar mass completeness limit ($M_{min}(z)$) is the 95th percentile of the M_{lim} distribution for the 20% faintest sources at each redshift bin. Thus, if we take a sample of galaxies with the stellar mass higher than the completeness limit, less than 5% of galaxies will be missed from the sample. As shown in Figure 2.2, the stellar mass completeness limit is higher for passive galaxies with higher mass-to-light ratios. Hence, we adopt the completeness limit derived from passive galaxies, which can be modeled with a quadratic polynomial; $\log(M_{min}(z)/M_{\odot}) = 7.90 + z - 0.09z^2$.

2.3 Measuring Galaxy Environment

The environment of a galaxy is defined as the density field where that galaxy resides. To reconstruct the density field, we consider narrow redshift intervals (z-slices) and treat each z-slice as a two-dimensional structure. Using kernel density estimation, we calculate the density field within each z-slice. The location of a galaxy with photo-z is probabilistic and best identified by its redshift PDF. Thus, a galaxy with photometric redshift contributes to all z-slices. The contribution of each galaxy to different redshift intervals is proportional to the area under the photo-z PDF that lies within that interval. This introduces the weighted approach for density estimation. The density field associated with each galaxy is derived from the weighted sum of surface densities at different z-slices using the full redshift PDF of the galaxy. Therefore, the surface density, σ, of a

galaxy at any given coordinate (RA, DEC) is,

$$\sigma_{(\mathrm{RA,DEC})} = \sum_{\mathrm{j}} \omega_{\mathrm{j}} \sigma^{\mathrm{j}}_{(\mathrm{RA,DEC})} \tag{2.1}$$

where $\sigma^{\mathrm{j}}_{(\mathrm{RA,DEC})}$ is the surface number density field at the position (RA,DEC) in the j^{th} z-slice and ω_j is the probability of the desired galaxy to be in the j^{th} z-slice. Although $\sim$ 12% of our sample have spectroscopic/grism redshift, we do not use them to determine ω_j. This assures that our method is not biased in favor of galaxies with accurate spectroscopic/grism redshifts. Therefore, we rely on uniformly calculated photometric redshifts, with well-calibrated probability distributions (Kodra et al. in prep.) in ω_j estimation. We use the area underneath the photometric redshift PDFs to obtain the likelihood of a galaxy to be in the j^{th} z-slice.

In order to measure $\sigma^{\mathrm{j}}_{(\mathrm{RA,DEC})}$, we use the weighted von Mises kernel density estimation technique corrected for the boundary effect. In the following sections, we describe different steps for estimating σ^j: an estimate of the redshift bin size (Section 2.3.1), weighted von Mises kernel density estimation (Section 2.3.2), bandwidth selection of the kernel function (Section 2.3.3) and boundary correction (Section 2.3.4).

2.3.1 Selection of Redshift Slices

It is important to optimize the width of redshift slices to account for the extended structures. While the redshift of a galaxy can be used to measure its location along the line of sight, the estimate can be affected by Redshift Space Distortion (RSD) due to the peculiar velocity of galaxies. The RSD effect is cosmological model dependent such that a galaxy cluster with internal velocity dispersion of Δv will be extended in comoving

space ($\Delta\chi$) as,

$$\Delta\chi = \frac{\Delta v}{H_0} \frac{(1+z)}{\sqrt{\Omega_{m_0}(1+z)^3 + \Omega_{\Lambda_0}}} \tag{2.2}$$

where H_0, Ω_{m_0} and Ω_{Λ_0} are the present values of Hubble constant, matter density and dark energy density respectively. Hence, what we observe is the combination of the density and the velocity field. The proper binning in redshift space to reconstruct 2D maps of the large-scale structures is constrained by both the typical size of a galaxy cluster in redshift space and redshift uncertainties. In the presence of less accurate photometric redshifts, we have two options, either using a weighted scheme to incorporate the contribution of each galaxy in all z-slices accurately or adopting wide z-slices to collect all signals from galaxies with large redshift uncertainties. Here we use the weighted approach such that the width of z-slice is constrained based on the resolution of photo-z PDFs, $\Delta z/(1+z) \sim 1\%$ (Kodra et al. in prep.). This allows us to avoid over-smoothing caused by interlopers.

The comoving size of a galaxy cluster due to the RSD effect (equation 2.2), peaks at $z = (2\Omega_{\Lambda_0}/\Omega_{m_0})^{\frac{1}{3}} - 1 \simeq 0.65$. At that redshift, a massive galaxy cluster ($\Delta v \sim 1500$ Km.s^{-1}) will be extended $\sim 18\ h^{-1}$Mpc in comoving space due to the peculiar velocity of its galaxies. In addition, the estimated redshift uncertainty ($\Delta z/(1+z) \sim 1\%$) limits the z-slice width to 35 h^{-1}Mpc ($\Delta v \sim 3000$ Km.s^{-1}). Therefore, we fix the width of redshift bins (at all redshifts) to a constant comoving size of 35 h^{-1}Mpc to satisfy both RSD effect and redshift uncertainty constraint. This results in a total of 124 z-slices spanning from $z = 0.4$ to 5. One should note that the constant comoving width does not imply a constant redshift interval. For comparison, the width of z-slice at $z = 0.4$ is 0.014, while this value is 0.096 for $z = 5$.

2.3.2 Weighted von Mises Kernel Density Estimation

The distribution of galaxies in each z-slice is analogous to a two-dimensional map where galaxies are labeled with their weights ω_j, computed from the photometric redshift PDFs. These weights are assigned by determining the fraction of redshift PDF within each z-slice. A powerful non-parametric method for density estimation is weighted Kernel Density Estimation(wKDE)(Parzen, 1962) which can be written as:

$$\sigma^j(\mathbf{X}_0) = \sum_i \widetilde{\omega}_i^j \mathrm{K}(\mathbf{X}_\mathrm{i}; \mathbf{X}_0) \tag{2.3}$$

where $\sigma^j(\mathbf{X}_0)$ is the estimated density at the position $\mathbf{X}_0$ on j^{th} z-slice and K is the kernel function. The summation is over all data points ($\mathbf{X}_i$) that exist in the desired z-slice. $\widetilde{\omega}_i^j$ is the normalized weight associated with i^{th} data point, in the j^{th} z-slice so that $\sum_i \widetilde{\omega}_i^j = 1$.

An appropriate choice of the kernel function for spherical data (RA, DEC) is the von Mises kernel (García-Portugués et al., 2013) expressed as,

$$\mathrm{K}(\mathbf{X}_\mathrm{i}; \mathbf{X}_0) = \frac{1}{4\pi \mathrm{b}^2 \sinh(1/\mathrm{b}^2)} \exp(\frac{\cos\psi}{\mathrm{b}^2}) \tag{2.4}$$

where b is the global bandwidth of the kernel function, which is the main parameter in the wKDE method and controls the smoothness of the estimate. We will explain the bandwidth selection method in Section 2.3.3. ψ is the angular distance between $\mathbf{X}_\mathrm{i} = (\mathrm{RA_i}, \mathrm{DEC_i})$ and $\mathbf{X}_0 = (\mathrm{RA}_0, \mathrm{DEC}_0)$. $\cos\psi$ can be expressed as $\sin\mathrm{DEC_i}\sin\mathrm{DEC}_0 + \cos\mathrm{DEC_i}\cos\mathrm{DEC}_0\cos(\mathrm{RA_i} - \mathrm{RA}_0)$.

It should be noted that a Gaussian kernel function cannot be used in the case of spherical data. The kernel function must integrate to unity and a Gaussian function does not satisfy this requirement on the spherical space.

2.3.3 Bandwidth selection

Bandwidth selection is a challenging problem in kernel density estimation. Choosing too narrow bandwidth leads to a high-variance estimate (under-smoothing), while too wide bandwidth leads to a high-bias estimate (over-smoothing). This bias-variance trade-off can be solved by maximizing Likelihood Cross-Validation (LCV) (Hall, 1982) which is defined by:

$$\mathrm{LCV(b)} = \frac{1}{\mathrm{N}} \sum_{\mathrm{k=1}}^{\mathrm{N}} \log \sigma_{-\mathrm{k}}(\mathbf{X}_{\mathrm{k}}) \tag{2.5}$$

where N is the total number of data points in a given z-slice and $\sigma_{-k}(\mathbf{X}_k)$ is the kernel estimator computed at position $\mathbf{X}_k$ excluding the k^{th} data point. We perform a grid search in the range of 0.0001° to 0.03° with 50 steps to find the optimized global bandwidth where LCV(b) is maximized. Figure 2.3 shows the LCV maximization results for one of the z-slices ($1.068 \leq z \leq 1.089$) in all CANDELS fields. For instance, the cross-validation method suggests $b = 0.0061°$ (a comoving distance of 0.26 h^{-1}Mpc) as the best bandwidth for the GOODS-S field at the mentioned z-slice. Figure 2.4 shows the optimized bandwidth in comoving coordinates for 124 redshift slices spanning from 0.4 to 5.

A constant bandwidth (b) over each z-slice may result in under-smoothing in regions with sparse observations and over-smoothing in crowded areas. By varying the bandwidth for each data point (i) and defining a local bandwidth (b_i), we reduce the bias in dense regions and the variance in regions with sparse data. To incorporate adaptive smoothing, we vary the local bandwidth (b_i) as (Abramson, 1982),

$$b_i = b\Big\{\frac{\sigma(\mathbf{X}_{\mathrm{i}})}{g}\Big\}^{-\alpha} \tag{2.6}$$

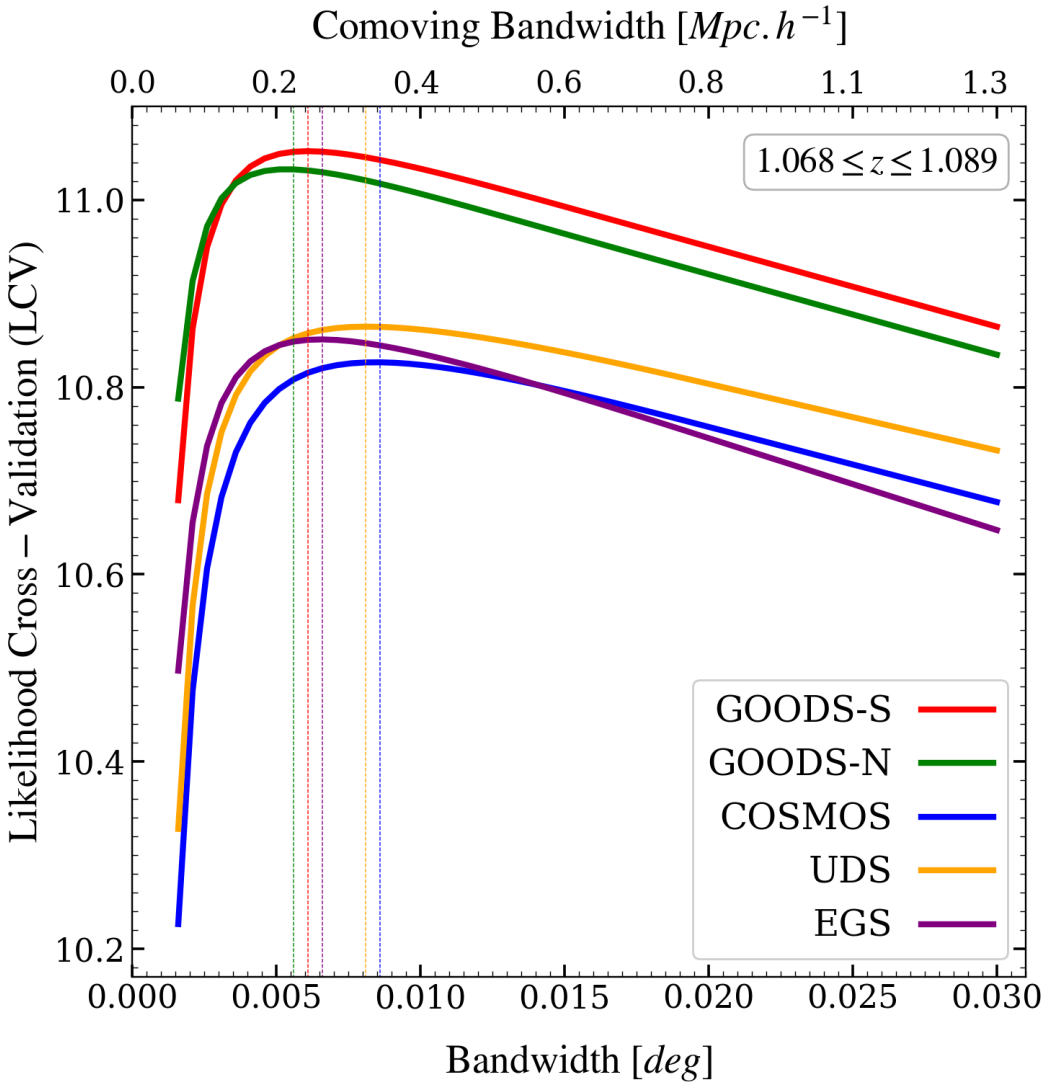


Figure 2.3: An example of Likelihood Cross-Validation optimization procedure at a given z-slice: $1.068 \leq z \leq 1.089$. We perform a grid search in the range of $0.0001^\circ \leq$ Bandwidth $\leq 0.03^\circ$ with 50 steps to maximize the LCV and find the best bandwidth (b). Dashed vertical lines show optimized bandwidths at $z \sim 1$.

where

$$\log g = \frac{1}{N}\sum_{i=1}^{N}\log \sigma(\mathbf{X}_i)$$

The sensitivity parameter, α, is a constant which satisfies $0 \leq \alpha \leq 1$ and can be fixed by simulation. In this study, we take a simple case where $\alpha = 0.5$ as the sensitivity parameter does not have a significant effect in wKDE measurement (Wang and Wang, 2007). For each redshift slice, first, we estimate the bandwidth using the cross-validation method and then we employ the adaptive bandwidth technique (equation 2.6) to reduce the variance/bias in the estimation. Finally, we need to correct the density field for the boundary effect, which is explained in the following section.

2.3.4 Boundary correction

Kernel density estimation method assumes that the density field exists in the entire space. This assumption is not valid in most cases where a survey has data only for a small area of the sky. The trade-off between the area and the depth translates into a small coverage in deep surveys (e.g., CANDELS). Missing parts of the sky not covered in the survey result in an underestimation of the density field near the edge of the survey. Different methods have been developed to remove this problem, known as the boundary effect (e.g., Reflection method (Schuster, 1985), Boundary kernel method (Müller, 1991) and Transformation method (Marron and Ruppert, 1994)). Here we use the re-normalization method to correct for this boundary effect.

The first order of the expectation value of the density fields can be written as (Jones,

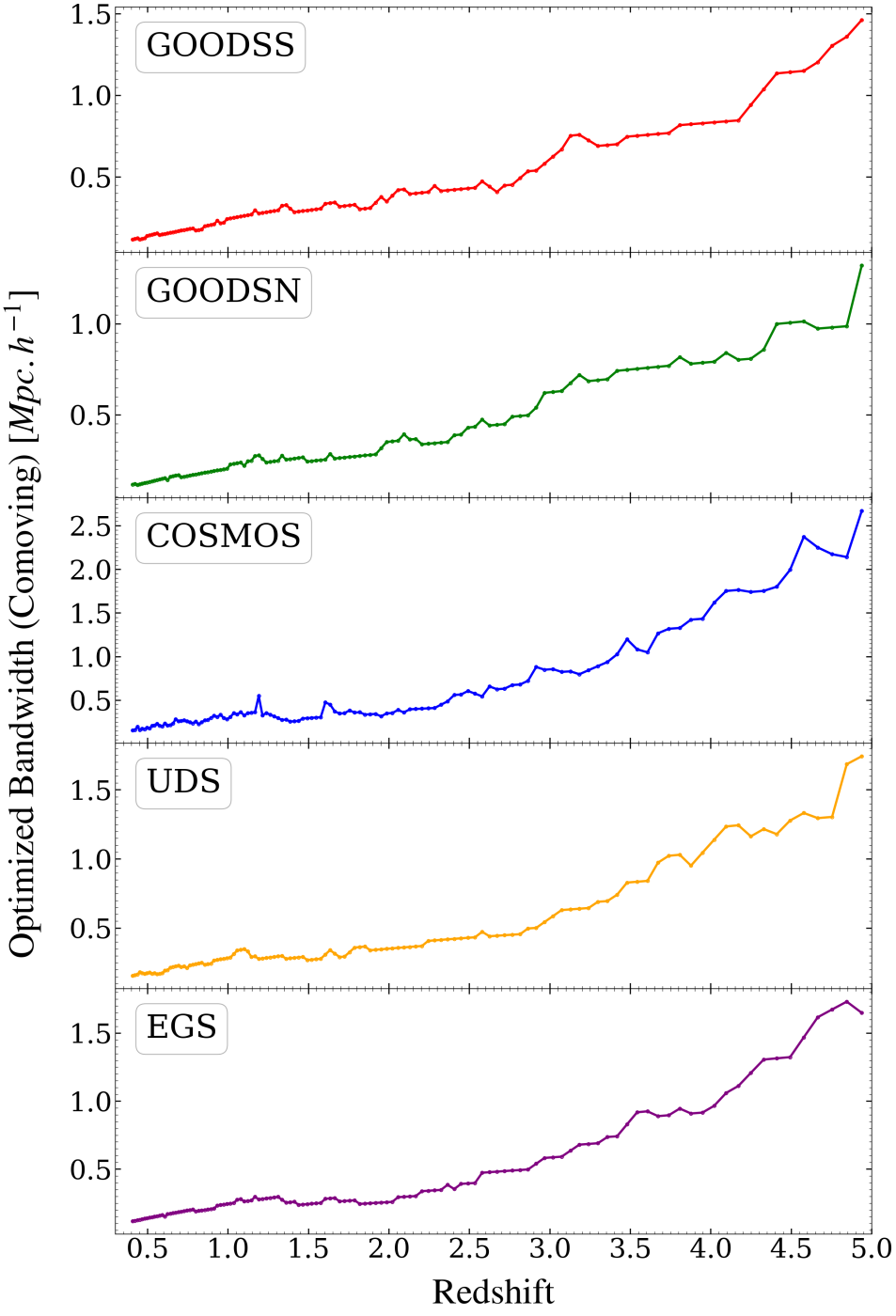


Figure 2.4: Optimized comoving global bandwidth (b) as a function of redshift for the five CANDELS fields. As we go to higher redshifts, we should increase the comoving size of the kernel function bandwidth to avoid the undesirable variance.

1993),

$$\mathbb{E}(\sigma^j(\mathbf{X}_0)) \sim \sigma^j_{\rm True}(\mathbf{X}_0) \int_S \mathrm{K}(\mathbf{X}_\mathrm{i}; \mathbf{X}_0) \tag{2.7}$$

where $\sigma^j_{\rm True}$ is the true underlying density field and the integration is performed over the survey area (S). Thus, a simple way to correct the boundary is to re-normalize the density as,

$$\sigma^j_{\rm corr}(\mathbf{X}_0) = \sigma^j(\mathbf{X}_0) n(\mathbf{X}_0) \tag{2.8}$$

where $n(\mathbf{X}_0)$ is the inverse of the kernel function integration centered at $\mathbf{X}_0$ over the survey area (S),

$$n^{-1}(\mathbf{X}_0) = \int_S \mathrm{K}(\mathbf{X}_\mathrm{i}; \mathbf{X}_0) \tag{2.9}$$

This correction results in almost unbiased estimation of the density such that $\mathbb{E}(\sigma^j_{\rm corr}(\mathbf{X}_0)) \sim \sigma^j_{\rm True}(\mathbf{X}_0)$, but it may increase the variance close to the boundary.

Figure 2.5 shows the boundary correction coefficient (n) given the bandwidth of 0.0116°. At that bandwidth, 40% of galaxies are affected by boundary problem ($n > 1$) and it is crucial to apply boundary correction to those galaxies. Not correcting for boundary effects can result in an underestimation up to a factor of ~ 3.5.

A common way to quantify the environment is defining density contrast (δ) as,

$$\delta = \frac{\sigma}{\bar{\sigma}} - 1 \tag{2.10}$$

where $\bar{\sigma}$ is the background number density, which can be evaluated using $\sum_i \omega^j_i / V$. Here, ω^j_i is the weight of i^{th} galaxy at j^{th} z-slice, and V is the volume of corresponding z-slice. The boundary problem does not affect the background number density; however, it biases σ close to the edge of the survey. Thus, boundary correction is necessary to avoid missing overdensities close to the edge.

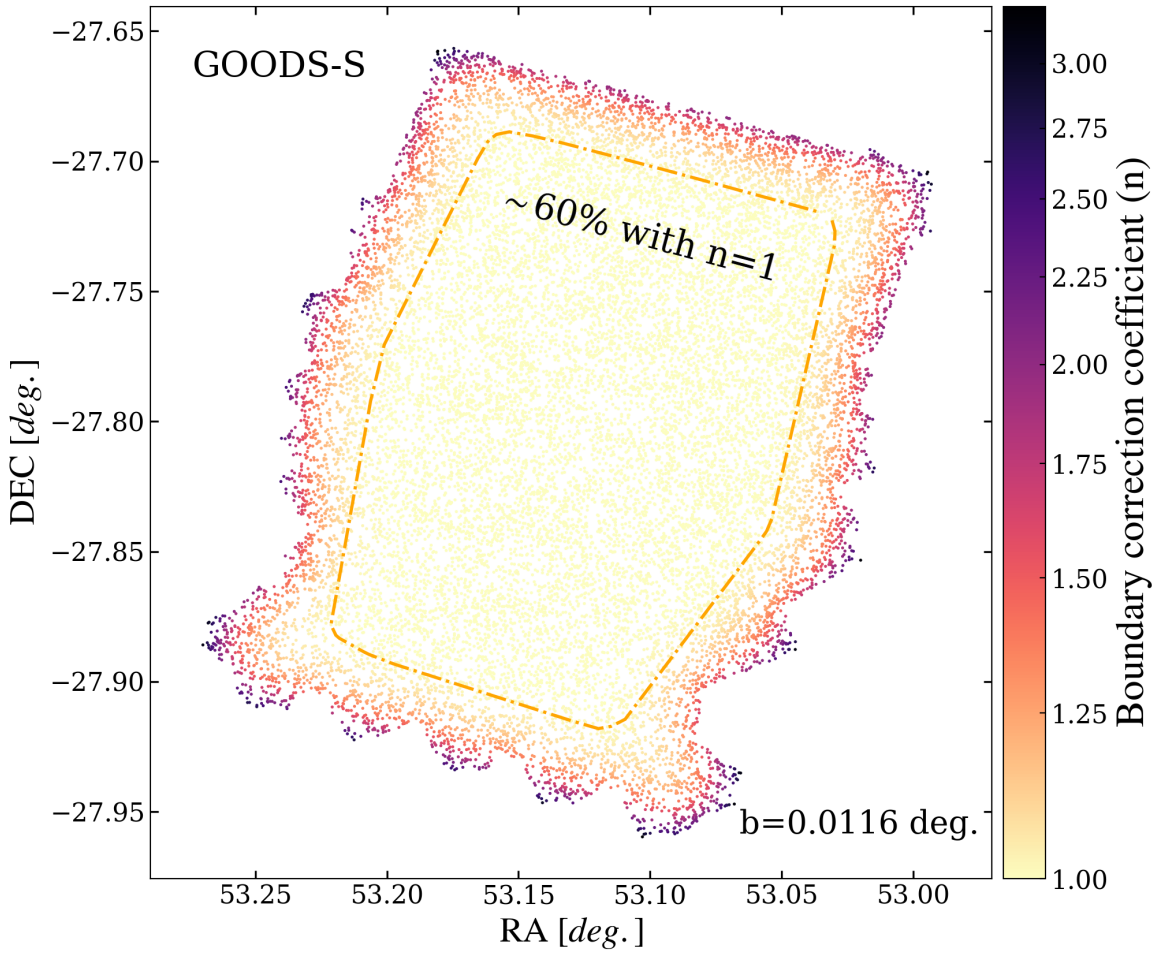


Figure 2.5: An example of the boundary correction coefficient (n) for galaxies in the GOODS-S field. It is shown that for the bandwidth of 0.0116°, the coefficient can be as large as ~ 3.5. Orange dash-dotted line encloses 60% of galaxies which are not affected by boundary problem (n=1).

2.3.5 Catalogs and density maps

We utilize the *boundary-corrected weighted von Mises kernel density estimation method* to reconstruct the density field of galaxies at $0.4 \leq z \leq 5$ in the five CANDELS fields. Details of the density measurement technique have been explained in Sections 2.3.1-2.3.4 and summarized below:

- Divide the survey into redshift slices with the comoving width of $\sim 35h^{-1}$Mpc (see Section 2.3.1).

For each z-slice:

- Weight the galaxies using their redshift PDFs to construct the two-dimensional weighted maps (see Section 2.3).
- Perform a grid-search on the bandwidth space to minimize the LCV function and find the optimum bandwidths (see Section 2.3.3).
- Compute the density field associated with each galaxy using weighted KDE with the constant bandwidth drawn from the previous step (see Section 2.3.2) and apply the boundary correction technique (see Section 2.3.4).
- Make the bandwidth to be adaptive based on the boundary-corrected densities (see Section 2.3.3) and re-run the weighted KDE with the adaptive bandwidths. Then, reapply the boundary-correction method on the adaptively derived densities.

The last step is to combine all z-slices to extract the density field associated with each galaxy.

- For each galaxy, calculate the weighted summation of its density in all z-slices (see Section 2.3) to obtain the density field of the galaxy.

The full density field catalogs are available in the electronic version. Table 2.2 shows examples of the density measurements. The first four columns show the CANDELS ID, RA, DEC and redshift (z_{best}). The last three columns give the environmental properties, including comoving/physical density and density contrast. Comoving density is the number of galaxies in a cube with a comoving volume of 1 h^{-3}Mpc^3. The physical density can be computed by scaling the comoving density by a factor of $(1+z)^3$. The density contrast indicates the number density enhancement with respect to the average density in the vicinity of the galaxy (equation 2.10).

The comoving number density and density contrast of galaxies as a function of their redshifts are shown in Figure 2.7. The limiting magnitude of the survey restricts the sources to a certain stellar mass range. Hence, the evolution of the comoving number density with redshift is an inevitable result of missing low mass galaxies at higher redshifts. In contrast, we find that the average density contrast is almost constant with redshift. This implies that the stellar mass function for a total sample of quiescent and star-forming galaxies does not change significantly with the environment. Davidzon et al. (2016) have studied the effect of the environment on the shape of the galaxy stellar mass function up to redshift $z = 0.9$, finding that the environmental dependence of the stellar mass function becomes weaker with redshift.

The histogram of the density contrast is also shown for each field in Figure 2.7. For all the fields, we find a similar distribution of density contrast, which has a dynamic range of ~ 10. For the entire sample of galaxies in all CANDELS fields, the average density contrast is 0.45 with a standard deviation of 0.75. It suggests that galaxies with a density contrast $\gtrsim$1.2 are located in an over-dense region and those with density contrast $\lesssim -0.3$ reside in a void.

Using the technique described in this paper, we estimate the density maps for all the five CANDELS fields. The evolution of the large scale structures is provided by 124

Table 2.2: Density field measurements in the GOODS-S field (Full catalogs are published in the electronic edition)

ID	RA(deg)	DEC(deg)	Redshift	Comoving Density $h^3\mathrm{Mpc}^{-3}$	Physical Density $h^3\mathrm{Mpc}^{-3}$	Density contrast
13889	53.1846685	-27.7875097	4.444	0.0112	1.7927	3.5429
13890	53.1610507	-27.7883341	1.095	0.1478	1.3476	1.7321
13893	53.1540231	-27.7876987	1.782	0.0569	1.2397	1.0454
13894	53.021758	-27.7874826	4.878	0.0099	2.0489	7.0415
13895	53.1157124	-27.7876168	1.187	0.0564	0.6183	0.269
13896	53.0802905	-27.7874276	2.407	0.0194	0.7487	0.0688
13902	53.1743111	-27.7876836	2.058	0.0223	0.6278	-0.138
13903	53.0335614	-27.7880437	1.083	0.0781	0.6991	0.4059
13911	53.1471765	-27.7884795	0.631	0.1096	0.4745	-0.2466
13912	53.1491844	-27.7878595	1.805	0.0553	1.2311	0.9543

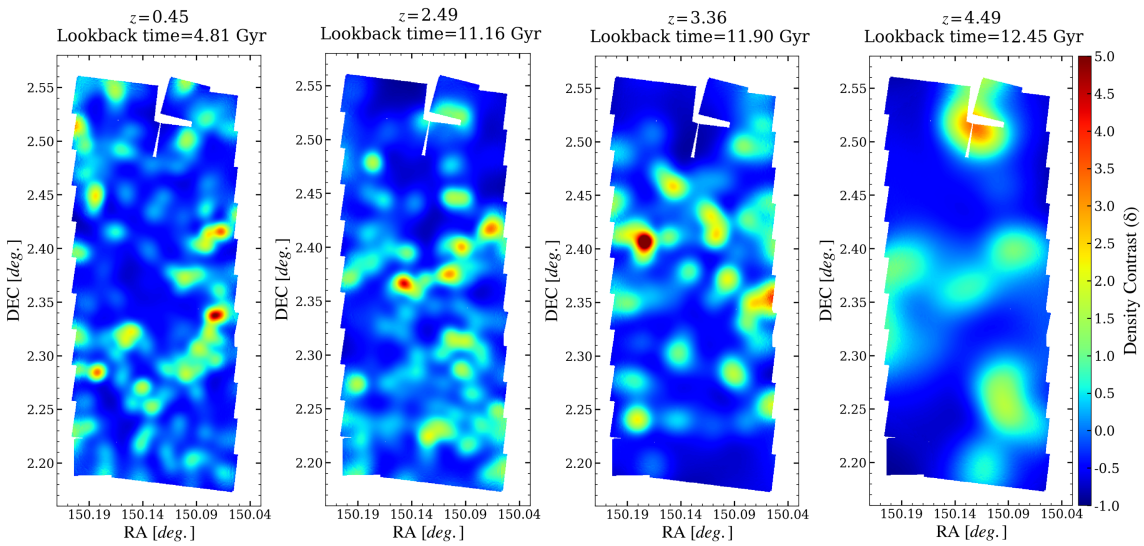
density maps covering $0.4 \leq z \leq 5$. The full density maps are available in the electronic version, with a few examples shown in Section 2.4 (Figure 2.6). In the density maps, we limited the color-bar range to 5 to get a better contrast. Therefore, any density contrast above 5 is saturated with a dark red color.

2.4 Density maps

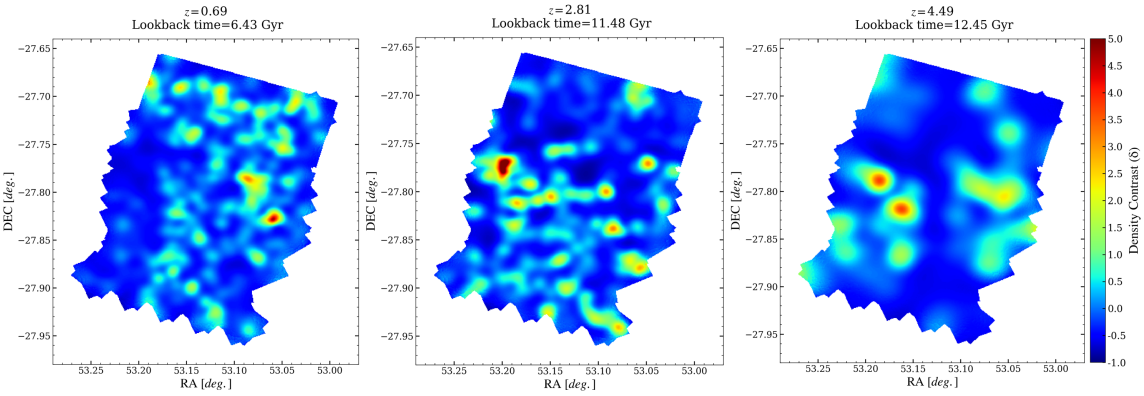
We release over-density maps of 124 z-slices ranging from $z = 0.4$ to 5 for all CANDELS fields. A few examples are provided in Figure 2.6, but the full set of plots for the 124 z-slices are available in animation in the electronic version. In density maps, the color-bar range is limited to 5 to get a better contrast. As we expect, structures on the fields with a higher declination (e.g., GOODS-N and EGS) are elongated along the right ascension axis, which is the natural effect of mapping on (RA,DEC) coordinates. It should be recalled that the density contrast of a galaxy is inferred from multiple density maps (z-slices) considering its contribution in each z-slice, which is determined by photo-z PDF.

2.5 Redshift evolution of density measurements

In this section, we investigate the correlation between density measurements and redshift. Figure 2.7 shows comoving number density and the density contrast as a function of redshift along with the distribution of density contrast separately for each field. Despite the clear evolution of comoving number density, the average density contrast, $\langle 1+\delta \rangle$, is almost constant over redshift. Although we find modest evidence of systematic trends between $\langle 1+\delta \rangle$ and redshift, especially at $z \gtrsim 3$, the variation of $\langle 1+\delta \rangle$ over redshift is limited to $\lesssim 0.3$. Thus, the study of the physical properties of galaxies (e.g.,

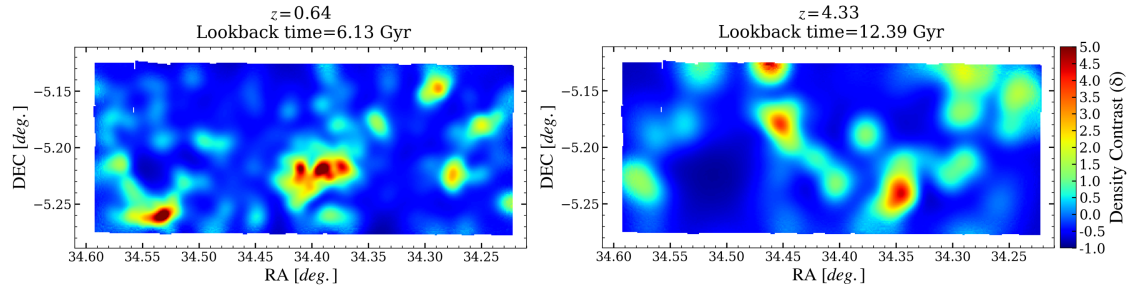


(a) COSMOS field

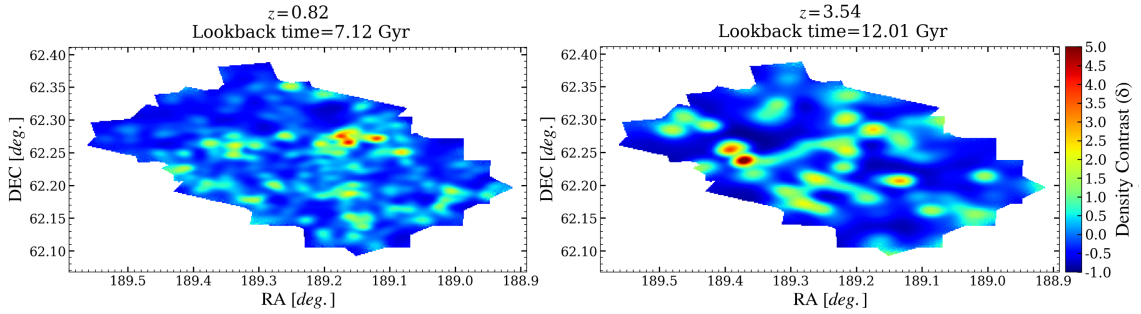


(b) GOODS-S field

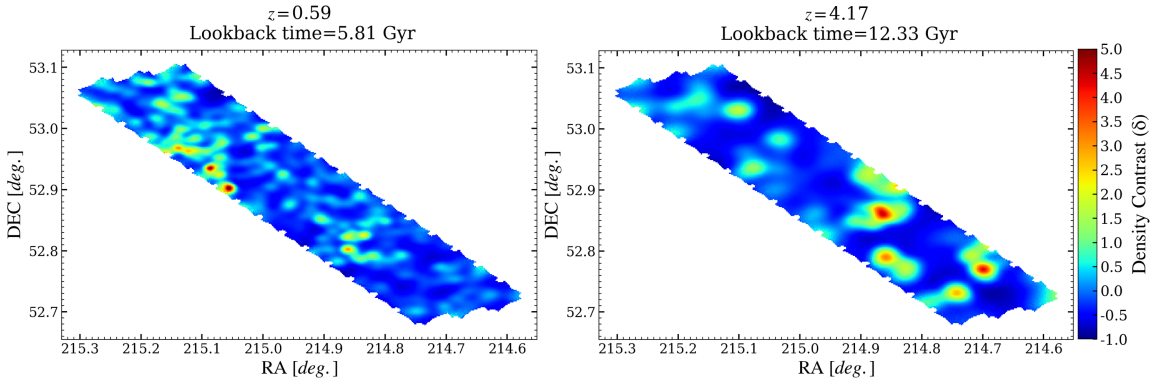
Figure 2.6: Density maps, plots for the 124 redshift slices are available in an animation format in the online version.



(c) UDS field



(d) GOODS-N field



(e) EGS field

Figure 2.6: Continued

SFR) versus density contrast (Section 2.6) is not affected by the redshift evolution of overdensity measurements. We note that the average of $1+\delta$ is slightly higher than one since we do not define the density of the background ($\bar{\sigma}$ in equation 2.10) as the average density of galaxies. We define background density as the number of galaxies (computed from their photo-z PDFs) within each z-slice divided by the volume of that z-slice.

2.6 Results

In this section, we use the estimated density fields to study the environmental effect on star formation activity of galaxies as a function of redshift. Here, we rely on the combined data from five widely separated fields to alleviate the cosmic variance effect as well as the sample size.

2.6.1 Environmental dependence of SFR and sSFR

We investigate the dependence of SFR and sSFR on the local overdensity. We build a mass-complete sample of galaxies in four redshift intervals. Each interval contains galaxies with stellar mass greater than the completeness limit at that redshift. For example, the sample at $0.4 \leq z < 0.8$ consists of 6299 galaxies with $M \geq M_{min}(0.8)$ where $M_{min}(0.8)$ is the stellar mass completeness limit at $z = 0.8$. The properties of the sample are summarized in Table 2.3. Although we have density measurements up to $z = 5$, we limit our investigation to $z \leq 3.5$. A mass-complete sample of galaxies at $3.5 < z \leq 5$, suffers from a small sample size (< 100) and may not be used to draw any statistically significant conclusions.

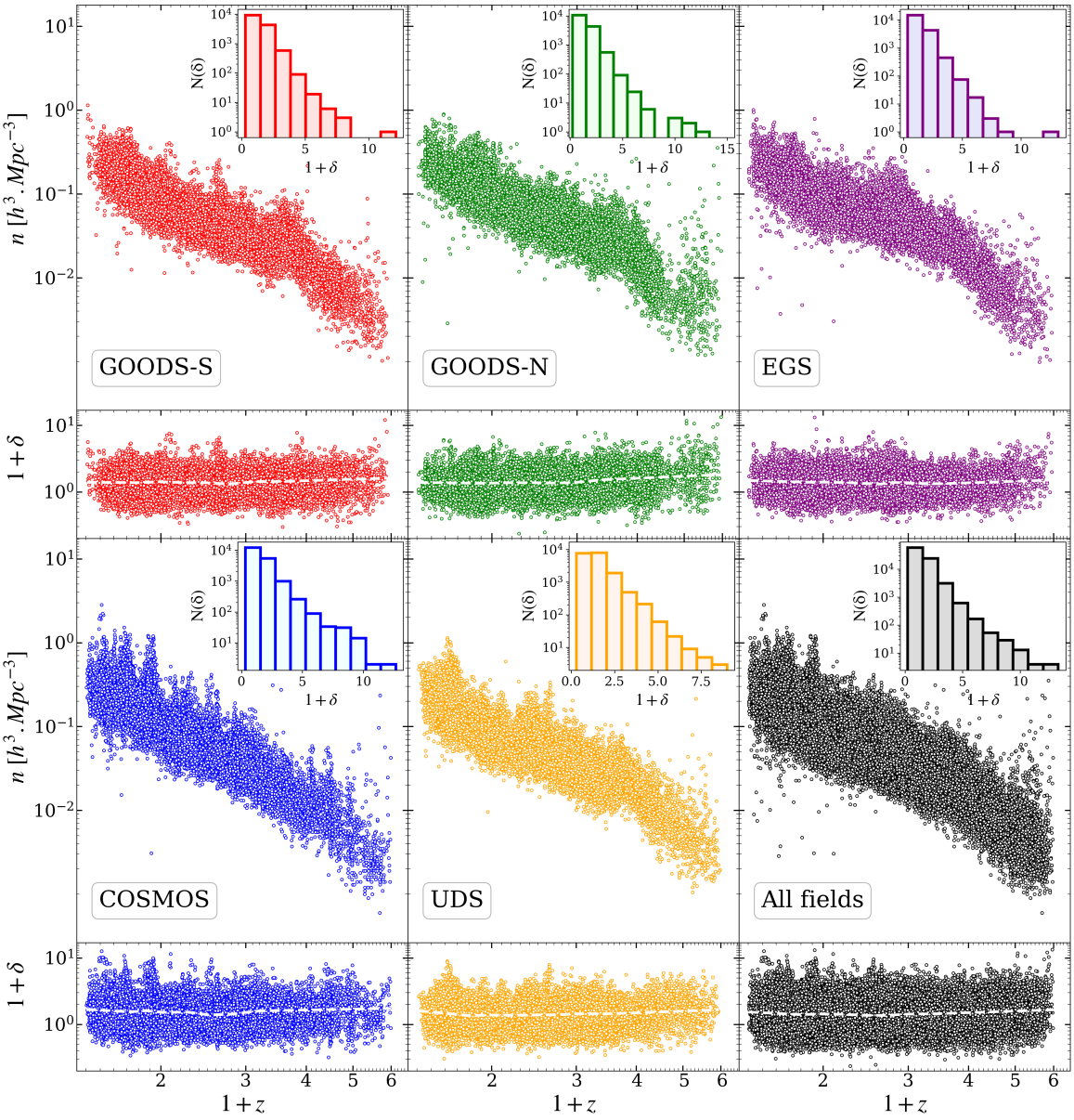


Figure 2.7: The comoving number density and the density contrast as a function of redshift as well as the density contrast histogram for each field. The comoving number density decreases with redshift due to the magnitude limit of the survey, while the average density contrast (white dashed lines) is almost constant over the cosmic time. This can be explained by the weak dependence of the stellar mass function on the environment. For all the fields, we find a similar distribution of density contrast, which has a dynamic range of ~ 10.

Figure 2.8 demonstrates the average SFR and sSFR as a function of density contrast in the four redshift intervals. It shows a clear anti-correlation between sSFR and environmental density. The same trend can be seen in SFR-density relation. At low redshift, $0.4 \leq z < 0.8$, SFR decreases by a factor of ~ 50 as the density contrast increases from $\delta \sim -0.5$ to $\delta \sim 6$. This drop is steeper (by order of magnitude) for the sSFR. At $0.8 \leq z < 1.2$, we find similar anti-correlation. These trends are in full agreement with previous studies (e.g., Patel et al., 2009; Scoville et al., 2013; Darvish et al., 2016).

At high redshift, $1.2 \leq z < 2.2$, we find that both SFR-density and sSFR-density relations follow the same trends we observe in the intermediate and low redshifts ($z < 1.2$). The average SFR and sSFR of galaxies in dense environments are significantly lower compared to those residing in under-dense regions. For example, the average sSFR decreases ~ 1.3 dex with ~ 1 dex increase in density contrast. Several studies revealed the persistence of the environmental quenching at high redshifts out to $z \sim 2$ (e.g., Grützbauch et al., 2012; Lin et al., 2012; Kawinwanichakij et al., 2017; Fossati et al., 2017; Guo et al., 2017; Ji et al., 2018). Our results confirm that anti-correlation exists in both SFR-density and sSFR-density relations at least out to $z \sim 2.2$.

Figure 2.8 also shows that the environment does play a significant role at the highest redshift bin $2.2 \leq z < 3.5$. SFR-density and sSFR-density relations behave in the same way that we observe in the local Universe. However, results need to be interpreted with caution since trends are found in different stellar mass ranges. For example, anti-correlation seen at $0.4 \leq z < 0.8$ corresponds to the galaxies with $\log(M/M_{\odot}) > 8.65$, while the relation at $2.2 \leq z < 3.5$ includes only massive galaxies with $\log(M/M_{\odot}) > 10.3$. The well-known relation between stellar mass and SFR in galaxies as a function of redshift confirms that the stellar mass plays a vital role in star formation activity of galaxies (e.g., Peng et al., 2010). It suggests that star formation activity might be affected by both the local environment and stellar mass. Therefore, we disentangle the

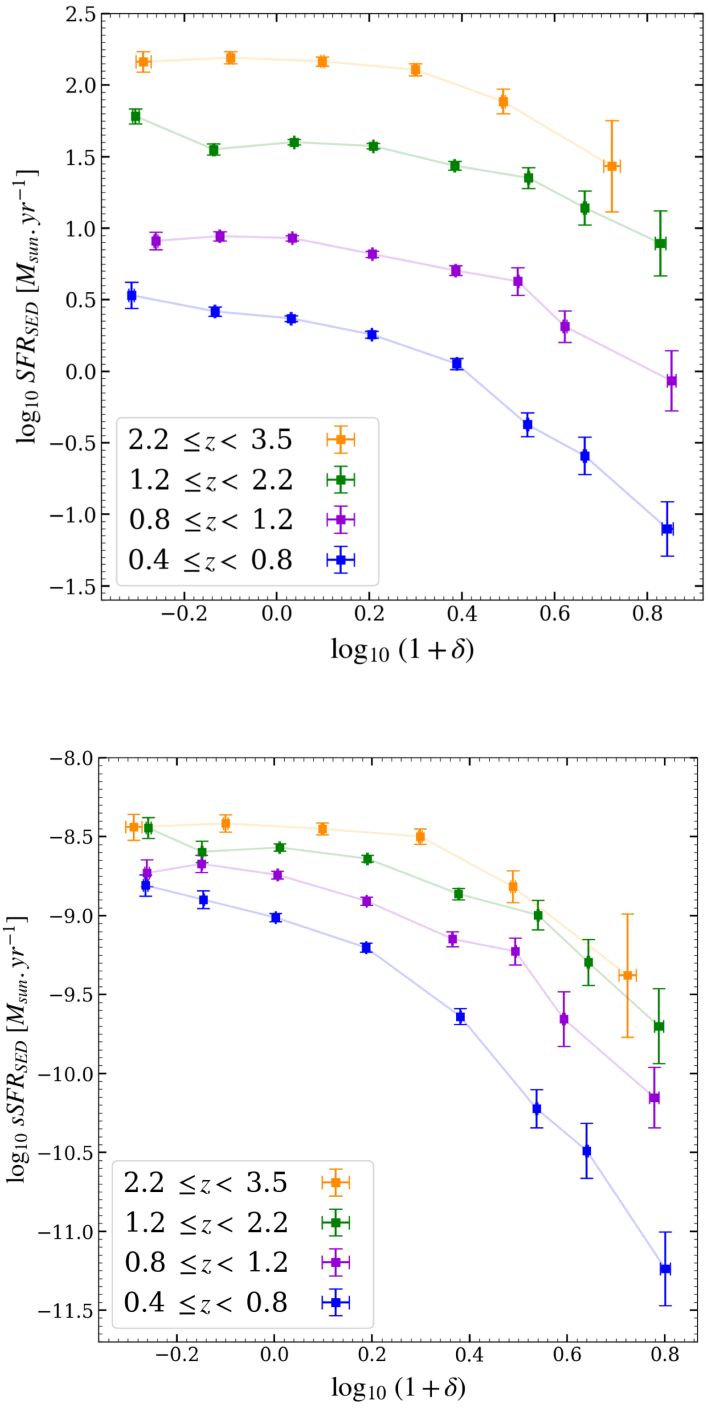


Figure 2.8: Environmental dependence of SFR and sSFR for a mass-complete sample of galaxies (Table 2.3) at four redshift bins spanning across the redshift range of $z = 0.4$ to $z = 3.5$. The average SFR and sSFR of galaxies in density contrast bins are plotted as a function of overdensity $(1 + \delta)$. Error bars show the statistical uncertainty of the average values.

influence of environment and stellar mass in the next section before proceeding to the physical interpretation of the results.

Table 2.3: Properties of mass-complete sample

Redshift Range	$\log(M_{min}/M_{\odot})$	Sample Size
$0.4 \leq z < 0.8$	8.65	6299
$0.8 \leq z < 1.2$	8.98	6279
$1.2 \leq z < 2.2$	9.67	6168
$2.2 \leq z \leq 3.5$	10.30	1047

2.6.2 SFR-Environment/Stellar Mass relation

Figure 2.9 presents the average SFR as a function of stellar mass and environment for the overall population of galaxies at the four redshift intervals. Colors indicate the average SFR in bins of environment and stellar mass. White areas show the regions with inadequate data points (< 20).

We find that SFR of massive galaxies ($M \geq 10^{11}$ $M_{\odot}$) is inversely correlated with the environment at all redshifts ($0.4 \leq z < 3.5$). For instance, at $1.2 \leq z < 2.2$, massive galaxies in dense environments on average form their stars ~ 6 times slower than galaxies with the same stellar mass located in under-dense regions. In contrast, we do not find significant environmental dependence on SFR of galaxies with lower stellar masses ($M < 10^{11}$ $M_{\odot}$) at high redshifts ($1.2 \leq z < 3.5$). It reveals that the environmental quenching for very massive galaxies persists out to $z \sim 3.5$. This concurs well with the work done by Kawinwanichakij et al. (2017), which is conducted out to $z = 2$. Moreover, Figure 2.9 demonstrates strong evidence of environmental quenching for low mass galaxies ($10^{9.5}$ $M_{\odot} < M < 10^{11}$ $M_{\odot}$) at $z < 1.2$ while it is not the case at higher redshifts ($z > 1.2$).

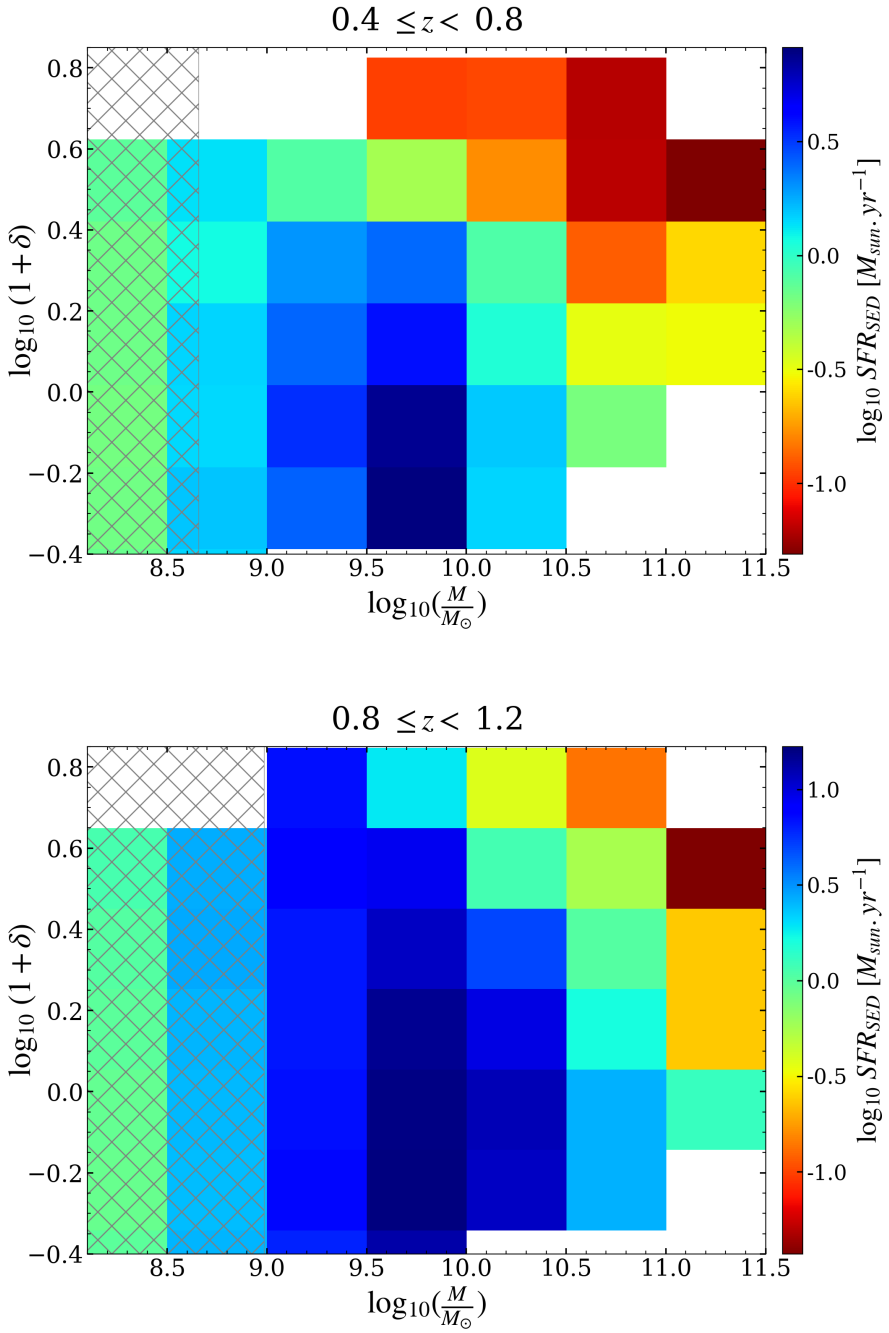


Figure 2.9: Average star formation rate in the bins of stellar mass and environment in four redshift intervals. The grey shaded regions show the incomplete stellar mass ranges. At all redshifts, we observe both environmental quenching and mass quenching for massive galaxies. We also find strong evidence of environmental quenching for low mass galaxies at low redshift.

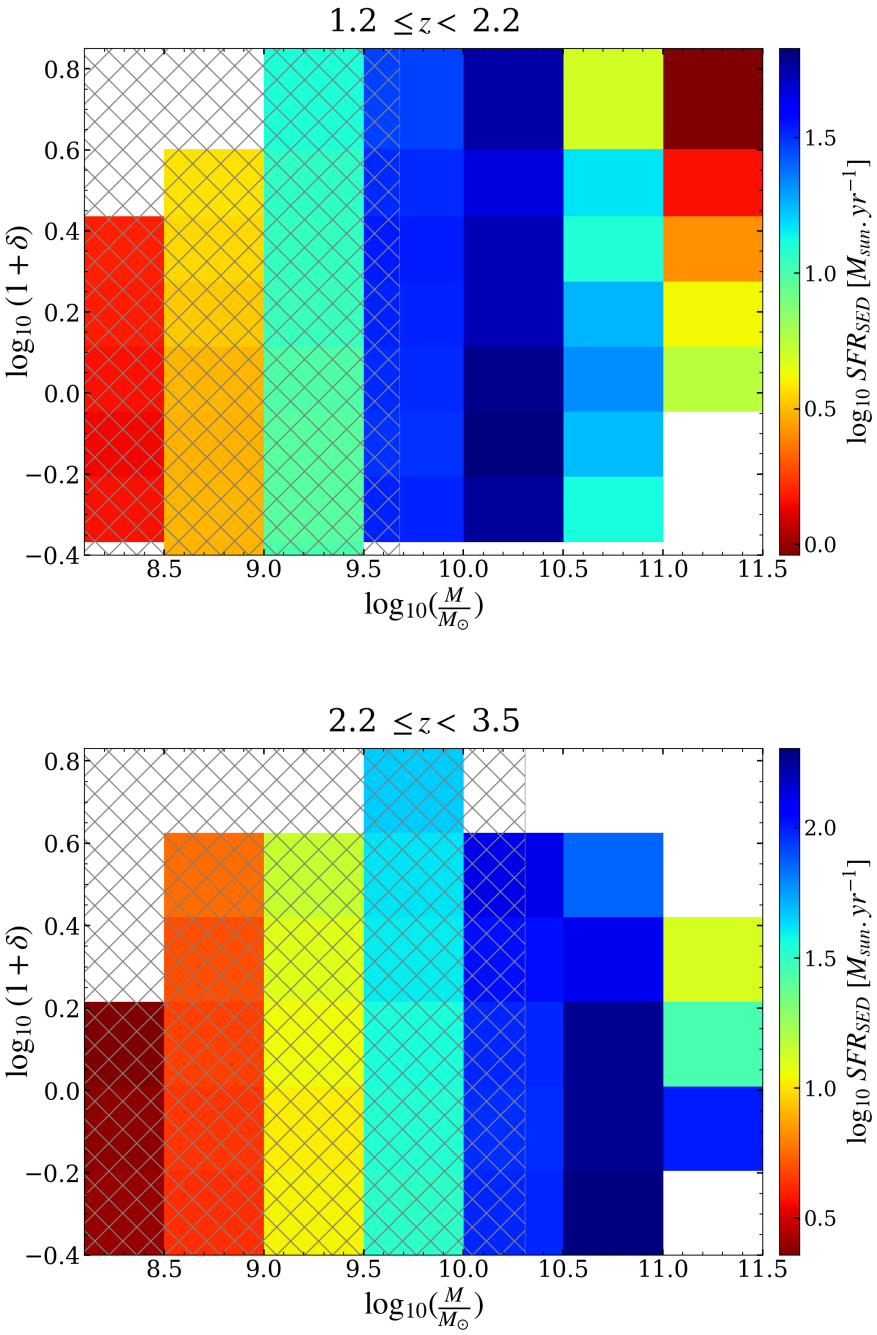


Figure 2.9: Continued

We also investigate the fraction of quiescent galaxies as a function of stellar mass and environment. Similar to Figure 2.9, we find evidence of both stellar mass and environmental quenching out to $z \sim 3.5$, such that the fraction of quiescent galaxies increases with increasing density contrast and stellar mass. In order to quantify the efficiency of environment and stellar mass in galactic quenching, we adopt the method introduced by Peng et al. (2010). We define environmental quenching efficiency, ε_{env}, as the deficiency in the fraction of star-forming galaxies in the environment with overdensity δ compared to the under-dense region,

$$\varepsilon_{env}(\delta, \delta_0, \mathrm{M}, \mathrm{z}) = 1 - \frac{\mathrm{f_s}(\delta, \mathrm{M}, \mathrm{z})}{\mathrm{f_s}(\delta_0, \mathrm{M}, \mathrm{z})} \tag{2.11}$$

where $f_s(\delta, \mathrm{M}, \mathrm{z})$ is the fraction of star-forming galaxies with stellar mass M that are located in an overdensity δ. δ_0 is the density contrast of the under-dense environment. Following Kawinwanichakij et al. (2017) we consider the lowest 25 percentile of the δ distribution (δ_{25}) as an under-dense environment (δ_0) and we compute environmental quenching efficiency for galaxies that are located in an overdensity with δ greater than the 75 percentile of the δ distribution (δ_{75}). A similar quantity can be defined for mass quenching efficiency, ε_{mass},

$$\varepsilon_{mass}(\delta, \mathrm{M}, \mathrm{M_0}, \mathrm{z}) = 1 - \frac{\mathrm{f_s}(\delta, \mathrm{M}, \mathrm{z})}{\mathrm{f_s}(\delta, \mathrm{M_0}, \mathrm{z})} \tag{2.12}$$

where $\mathrm{M_0}$ is the lowest stellar mass at any given redshift (z), which can be obtained by the stellar mass completeness limit ($\mathrm{M_{min}(z)}$). We compute mass quenching efficiency for galaxies with $\delta < \delta_{75}$.

In order to calculate the fraction of star-forming galaxies, we separate star forming and quiescent galaxies based on their rest-frame U,V and J colors along with the Muzzin

et al. (2013) criteria.

The stellar mass dependence of the mass quenching efficiency, $\varepsilon_{mass}(\delta < \delta_{75}, \mathrm{M}, \mathrm{M_{min}(z)}, \mathrm{z})$ and environmental quenching efficiency, $\varepsilon_{env}(\delta > \delta_{75}, \delta < \delta_{25}, \mathrm{M}, \mathrm{z})$ are shown in Figure 2.10. The efficiencies are calculated in stellar mass bins, $\Delta\mathrm{M} \sim 0.5$ dex, and error bars (shaded regions) are obtained considering the Poisson statistics for the number of quiescent/star-forming galaxies.

At all redshifts, mass quenching efficiency increases significantly with stellar mass, which is consistent with previous works (e.g., Peng et al., 2010). We also find that the environmental quenching efficiency is not independent of the stellar mass and it clearly increases with stellar mass, although this rise is weaker compared to the mass quenching efficiency (see also Lin et al., 2014; Papovich et al., 2018). At $z < 1.2$, the environmental quenching is dominant for low mass galaxies ($\mathrm{M} \lesssim 10^{10}\mathrm{M}_{\odot}$). For example, at $0.4 \leq z < 0.8$, the environmental quenching is $\sim$ 10 times stronger than the mass quenching. For massive galaxies ($\mathrm{M} \gtrsim 10^{10}\mathrm{M}_{\odot}$), mass quenching is the dominant quenching mechanism at all redshifts; however, environmental quenching is significant for the most massive galaxies ($\mathrm{M} \gtrsim 10^{11}\mathrm{M}_{\odot}$). For instance, at $2.2 \leq z < 3.5$, the environment and the stellar mass are almost equally responsible for the quenching of very massive galaxies ($\varepsilon_{mass} \sim \varepsilon_{env}$). This result reinforces our previous findings in Figure 2.9 that the environmental quenching of very massive galaxies exists at least out to $z \sim 3.5$. It also confirms that the environmental quenching is efficient for low mass galaxies at low and intermediate redshifts ($z < 1.2$) (see also Peng et al., 2010; Quadri et al., 2012; Scoville et al., 2013; Lin et al., 2014; Lee et al., 2015; Darvish et al., 2016; Nantais et al., 2016; Kawinwanichakij et al., 2017; Guo et al., 2017; Fossati et al., 2017; Ji et al., 2018).

2.6.3 Origin of the environmental quenching

Although most of the studies, including this work, found strong evidence of the environmental quenching out to high redshifts, the physical mechanisms that are responsible are not clearly understood. van de Voort et al. (2017) found a suppression of the cool gas accretion rate in dense environment at all redshifts, which becomes stronger at lower redshifts. This implies that a dense environment prevents the accretion of cold gas into the galaxy (cosmological starvation). As a result, the galaxy starts to consume the remaining gas reservoir in the depletion time scale, $t_{\rm depl} \propto {\rm M_{gas}/SFR}$. This scenario is known as "over-consumption" model (McGee et al., 2014; Balogh et al., 2016) and implies that the depletion time ($t_{\rm depl}$) depends on both stellar mass and redshift. The model predicts a short depletion timescale (< 100 Myrs) for massive galaxies at high redshift. Therefore, "over-consumption" scenario could explain the environmental quenching that we observe here for massive galaxies at high redshifts (Kawinwanichakij et al., 2017). Feldmann and Mayer (2015) also showed through their simulation that, at $z \sim 3.5$, sSFR of a massive galaxy ($\rm M \sim 10^{11} M_{\odot}$) drops by almost an order of magnitude within a few 100 Myrs. They found that this sudden halt at $z \sim 3.5$ is not caused by feedback processes and happens primarily due to the termination of the cool gas accretion. This provides another support that massive galaxies become quenched abruptly when their fresh gas accretion is terminated, possibly by locating in a dense environment. In addition, the lack of environmental quenching of low mass galaxies at high redshift can be explained by their low SFR, which results in longer depletion time (Balogh et al., 2016; Kawinwanichakij et al., 2017).

Furthermore, "over-consumption" model could explicate our observations at low redshift (down to $z \sim 0.4$). For galaxies with stellar mass $\rm M \sim 10^{10.5} M_{\odot}$, the depletion time ($t_{\rm depl}$) increases with decreasing redshift and reaches $\sim$ 2 Gyrs at $z \sim 0.4$

(McGee et al., 2014) which is shorter than the typical dynamical time scale at that redshift ($t_{\rm dyn} \sim 4$ Gyrs) (Balogh et al., 2016; Foltz et al., 2018). It implies that the dynamical gas stripping processes are not required to explain our observation at low redshift. Moreover, the evolution of environmental quenching efficiency with stellar mass supports "over-consumption" model where the depletion time is longer for low mass galaxies resulting in weaker quenching efficiency. Therefore, "over-consumption" picture is most likely the dominant mechanism of environmental quenching, at least in the redshift range of this study. However, it is worth highlighting that at local Universe, the depletion time grows fast, such that it reaches > 10 Gyrs at $z = 0$ for galaxies with intermediate stellar masses ($M \sim 10^{10.5} M_{\odot}$). This is longer than the dynamical time scale (Balogh et al., 2016). Consequently, "over-consumption" is likely not an effective quenching pathway at the local Universe and other dynamical processes are needed to explain the strong environmental quenching observed at $z \sim 0$ (Peng et al., 2010).

2.7 Discussion

In this work, we introduce a robust method for reconstructing the underlying number density field of galaxies. The performance of KDE has been well explored by statisticians. They found that KDE can precisely estimate underlying densities of any shape, provided that the bandwidth is selected appropriately (e.g., Silverman, 1986). We adopt a well-known Likelihood cross-validation (LCV) method to find the optimized bandwidth (e.g., Hall, 1982). Alternatively, one can use least squares cross-validation (LSCV) (Bowman, 1984), which is based on minimizing the integrated square error between the estimated and true densities. The LSCV method of bandwidth selection suffers the disadvantage of high variability (Jones et al., 1996) and a tendency to under-smooth (Chiu, 1991). We also correct densities for a systematic bias (under-estimation) near

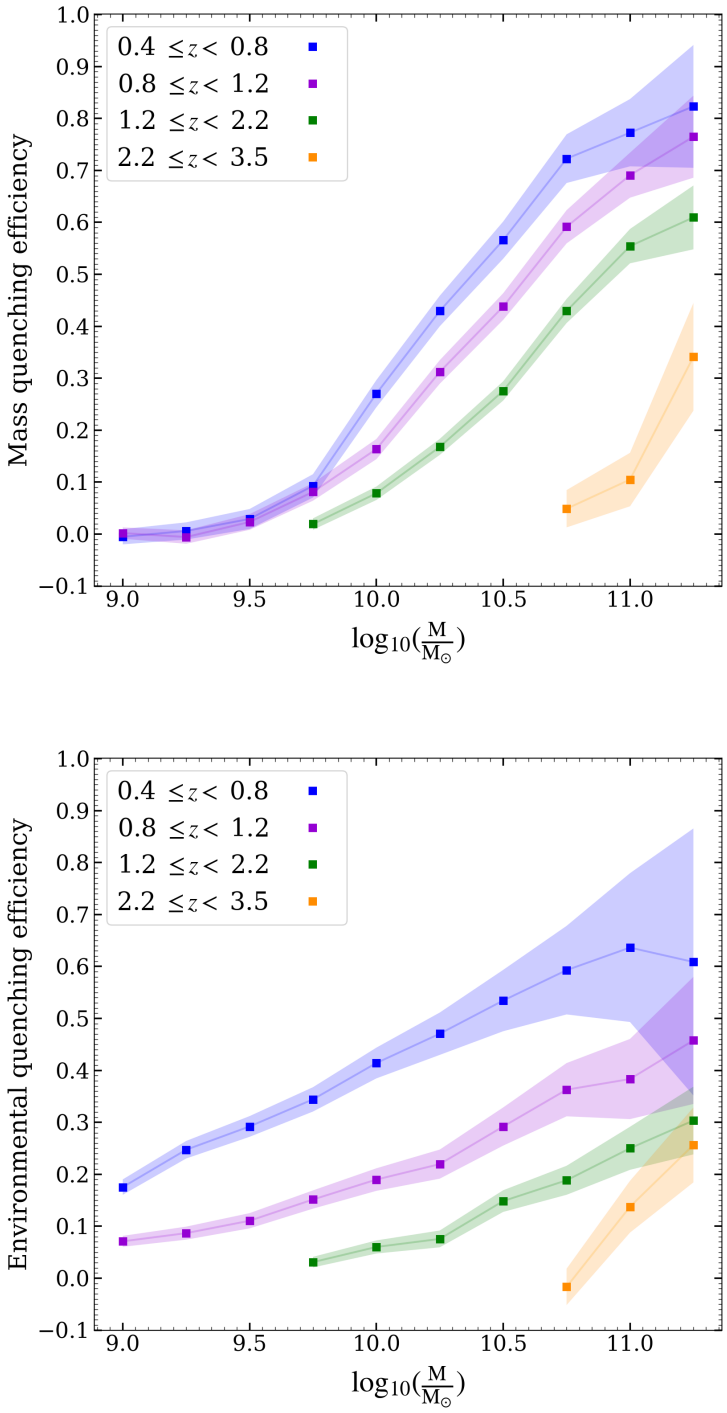


Figure 2.10: The mass quenching efficiency and environmental quenching efficiency, $\varepsilon_{env}(\delta > \delta_{75}, \delta < \delta_{25}, \mathrm{M})$ as a function of stellar mass. The efficiencies are calculated in stellar mass bins, $\Delta\mathrm{M} \sim 0.5$ dex. Shaded regions show the uncertainty of efficiencies considering the Poisson statistics for the number of quiescent/star-forming galaxies.

the edge of the survey using re-normalization. This assumes a symmetric galaxy distribution with respect to boundaries near the edges and may cause misestimated densities. This inevitable issue can be eliminated by observing as deep as CANDELS in a wider area.

Fossati et al. (2017) have measured the environmental density for a $JH_{140} \leq 24$ sample of 18,745 galaxies in the 3D-HST survey (Skelton et al., 2014) from $z = 0.5$ to 3, adopting circular aperture method (aperture radius fixed at 0.75 Mpc and width of z-slices at $\Delta v = 1500$ km.s^{-1}). We find a significant difference between their density contrasts and our measurements. Fossati et al. (2017) did not use the uniformly calculated photometric redshifts probability distributions. Instead, they assign redshifts based on the nearby galaxy with a spectroscopic redshift. They also use wider field public data for edge correction. We adopt the re-normalization method for edge-correction since we use the widest homogeneous fields with a depth of F160w = 26 AB mag.

We explore any trends that may exist between the estimated densities and redshift in Figure 2.7. This assures that the average density contrast does not evolve strongly with redshift. Otherwise, the diagram of any physical parameter (e.g., SFR, sSFR and quiescent fraction) as a function of density contrast would not be informative about the role of environment and trends could be affected by redshift evolution of physical parameters.

Furthermore, the possibility of unrealistic trends due to the different assumptions in SED fitting (e.g., star formation history) also needs to be investigated. As a test case, we repeat our analysis based on the Pacifici et al. (2012) SED fitting method, which provides a library of SEDs assuming star formation histories from a semi-analytical model. Although the trends in Figure 2.8 are more sensitive to SED fitting priors, the stellar mass dependence of environmental quenching efficiency (Figure 2.10) does not change with new measurements. This reassures that the present evidence of environ-

mental quenching at high redshift and the evolution of environmental quenching efficiency with stellar mass are not affected by SED fitting priors (Section 2.2), especially by exponentially declining star formation history.

2.8 Summary

In this work, we report measurements of the environment for a F160w $\leq$ 26 AB mag sample of 86,716 galaxies in the five CANDELS fields (GOODS-South, GOODS-North, COSMOS, EGS, UDS) at $0.4 \leq z \leq 5$. We introduce a new method, *boundary-corrected weighted von Mises kernel density estimation*, to reconstruct the underlying density field of galaxies. We find the optimal bandwidth for the von Mises kernel function in 124 z-slices spanning from $z = 0.4$ to 5 using the Likelihood Cross-Validation method. It allows us to create density field maps with the lowest bias/variance.

We then use the density measurements to investigate the role of environment in star formation activity of a mass-complete sample of galaxies at $0.4 \leq z \leq 3.5$. Our findings are summarized as follows:

1- At all redshifts, the average SFR and sSFR for a mass-complete sample of galaxies decrease with increasing density contrast. The trend is steeper at low redshift ($0.4 \leq z < 0.8$) such that the average SFR decreases by a factor of ~ 50 as the density contrast increases from $\delta \sim -0.5$ to $\delta \sim 6$.

2- We find strong evidence of environmental quenching for massive galaxies ($M \gtrsim 10^{11} M_{\odot}$) out to $z \sim 3.5$. We measured that the environmental quenching efficiency is $\gtrsim$0.2, implying that a dense environment has $\gtrsim$20% more massive quiescent galaxies compared to an under-dense region. This ratio reaches $\sim$ 60% at the lowest redshift bin of this study ($0.4 \leq z < 0.8$).

3- We find that the environmental quenching efficiency increases with stellar mass. This observation supports "over-consumption" model for environmental quenching where the gas depletion happens once the fresh gas accretion stops due to a dense environment. The gas depletion time depends on stellar mass and redshift and could explain the stellar mass dependence of the environmental quenching efficiency. The depletion time becomes longer ($>$ 10 Gyr) at lower redshifts, so it could not be a proper quenching pathway at local Universe; however, "over-consumption" is most likely the dominant environmental quenching mechanism at the redshift range of this study.

Chapter 3

The MOSDEF survey: Environmental dependence of the gas-phase metallicity of galaxies at $1.4 \leq z \leq 2.6$

Abstract

Using the near-IR spectroscopy of the MOSFIRE Deep Evolution Field (MOSDEF) survey, we investigate the role of local environment in the gas-phase metallicity of galaxies. The local environment measurements are derived from accurate and uniformly calculated photometric redshifts with well-calibrated probability distributions. Based on rest-frame optical emission lines, [NII]λ6584 and Hα, we measure gas-phase oxygen abundance of 167 galaxies at $1.37 \leq z \leq 1.7$ and 303 galaxies at $2.09 \leq z \leq 2.61$, located in diverse environments. We find that at $z \sim 1.5$, the average metallicity of galaxies in overdensities with $M_* \sim 10^{9.8} M_\odot, 10^{10.2} M_\odot$ and $10^{10.8} M_\odot$ is higher relative to their field counterparts by 0.094 ± 0.051, 0.068 ± 0.028 and 0.052 ± 0.043 dex, respectively. However, this metallicity enhancement does not exist at higher redshift, $z \sim 2.3$, where, compared to the field galaxies, we find 0.056 ± 0.043,

0.056 ± 0.028 and 0.096 ± 0.034 dex lower metallicity for galaxies in overdense environments with $M_* \sim 10^{9.8} M_\odot, 10^{10.2} M_\odot$ and $10^{10.7} M_\odot$, respectively. Our results suggest that, at $1.37 \leq z \leq 2.61$, the variation of mass-metallicity relation with local environment is small (< 0.1dex), and reverses at $z \sim 2$. Our results support the hypothesis that, at the early stages of cluster formation, owing to efficient gas cooling, galaxies residing in overdensities host a higher fraction of pristine gas with prominent primordial gas accretion, which lowers their gas-phase metallicity compared to their coeval field galaxies. However, as the Universe evolves to lower redshifts ($z \lesssim 2$), the shock-heated gas in overdensities cannot cool down efficiently, and galaxies become metal-rich rapidly due to the suppression of pristine gas inflow and re-accretion of metal-enriched outflows in overdensities.

3.1 Introduction

Over the last few years, we have made significant progress towards developing a comprehensive and self-consistent model for the formation of galaxies. At each step, however, the model is compounded by non-linear effects regarding the feedback processes involved and the parameters deriving them. Gas accretion from the intergalactic medium (IGM) supplies cold gas for a galaxy to build up its stellar population through the gravitational collapse of molecular clouds. As the stars form, heavy elements are produced in their hot cores, resulting in the chemically enriched material which will be expelled to the interstellar medium (ISM) through feedback processes such as stellar winds (Garnett, 2002; Brooks et al., 2007) and/or supernovae explosions (e.g., Steidel et al., 2010; Martin et al., 2012; Chisholm et al., 2018). Furthermore, feedback processes can remove part of the enriched material from galaxies into the IGM (e.g., Heckman et al., 1990; Tremonti et al., 2004; Chisholm et al., 2018). Thus, gas-phase

metallicity is expected to be connected with most of the galaxy evolutionary processes (e.g., gas inflow/outflow and star formation) and can be considered as one of the fundamental characteristics of galaxies that encodes information regarding galaxy evolution over cosmic time.

A tight correlation has been observed between the gas-phase metallicity and stellar mass of galaxies (Mass-Metallicity Relation, hereafter MZR) out to $z \sim 3.5$, such that galaxies with lower stellar masses have lower metallicities (e.g., Tremonti et al., 2004; Erb et al., 2006; Mannucci et al., 2009; Finkelstein et al., 2011; Steidel et al., 2014; Sanders et al., 2015, 2020). Moreover, it has been found that the MZR evolves with redshift so that galaxies at a given stellar mass have lower metallicity at high redshifts (e.g., Steidel et al., 2014; Maiolino and Mannucci, 2019; Sanders et al., 2020).

Although a tight correlation has been found between stellar mass and gas-phase metallicity spanning a wide range of stellar masses, other physical properties of galaxies could also play a role in contributing to the observed scatter in this relation. For example, the star formation rate (SFR) in galaxies controls their metallicity. At a given stellar mass, galaxies with lower SFR have higher metallicities (e.g., **?**Sanders et al., 2018). Gas accretion from the IGM adds chemically poor gas into the ISM of galaxies, which lowers the galaxy's gas-phase metallicity. On the other hand, cold gas infall that provides additional fuel for star formation increases the SFR. Therefore, cold gas accretion changes the metallicity content of the ISM in a complicated way. Lilly et al. (2013) introduced a gas regulator model that expresses the gas-phase metallicity of a galaxy in terms of the properties of the accreted gas (metallicity and rate of infalling gas) and SFR. Their model explains the observationally-confirmed dependence of the MZR on SFR.

However, the evolution of the ISM and its metal content does, to a large extent, depend on the properties of the infalling gas which, in turn, is affected by the environment

where the galaxy resides (Peng and Maiolino, 2014; Gupta et al., 2018). Chartab et al. (2020) studied the environmental dependence of star formation activity of galaxies in the five widely separated fields of the Cosmic Assembly Near-IR Deep Extragalactic Legacy Survey (CANDELS; Grogin et al., 2011; Koekemoer et al., 2011) out to $z \sim 3.5$. They found that environmental quenching efficiency evolves with stellar mass such that massive galaxies in overdense regions become quenched more efficiently than their low mass counterparts. This suggests that, besides the stellar mass quenching, which is the dominant quenching mechanism at high redshift, the environmental quenching is also effective as early as $z \sim 3$ for massive galaxies (i.e., the quiescent fraction of galaxies with $M_* \sim 10^{11} M_\odot$ are 20% higher (Chartab et al., 2020) in overdensities than that of field counterparts at $z \sim 2.8$). They also suggest that the growth of environmental quenching efficiency with stellar mass can be explained by the termination of cold gas accretion in an overdense environment. In the absence of cold gas accretion, massive galaxies exhaust their remaining gas reservoirs in shorter timescales compared to low mass galaxies (see also, Balogh et al., 2016; Kawinwanichakij et al., 2017; Fossati et al., 2017; Old et al., 2020). Using the Evolution and Assembly of GaLaxies and their Environments (EAGLE) simulation, van de Voort et al. (2017) found a suppression of the cold gas accretion rate in dense environments, mostly for satellite galaxies. Zavala et al. (2019) reported ALMA observation of 68 spectroscopically-confirmed galaxies within two proto-clusters at $z \sim 2.2$ and found that proto-clusters contain a higher fraction of massive and gas-poor galaxies compared to those residing in the field environment. They concluded that the environmental quenching exists during the early phases of cluster formation (see also, Darvish et al., 2016; Nantais et al., 2017; Ji et al., 2018; Pintos-Castro et al., 2019; Contini et al., 2020; Ando et al., 2020). However, Lemaux et al. (2020) recently observed an enhanced star formation for star-forming galaxies in overdensities at $z > 2$ (see also, Elbaz et al., 2007; Tran et al., 2010). All these obser-

vations reveal the importance of the local environment in gas accretion rate and SFR that govern the gas-phase metallicity of galaxies.

Moreover, the metallicity of the infalling gas varies in different environments. In the local Universe, it is observed that the metallicity of IGM gas in cosmic voids is $< 0.02\ Z_{\odot}$ (Stocke et al., 2007), while this number is $\sim 0.3\ Z_{\odot}$ for a cluster-like environment (e.g., Mushotzky and Loewenstein, 1997). This is consistent with results from the IllustrisTNG simulations (Gupta et al., 2018) that predict a metal-enhanced infalling gas in a dense environment out to $z \sim 1.5$. Both pieces of evidence of lower gas accretion rate and higher metallicity of IGM gas in denser environments hint that a part of the scatter on the MZR could be due to the difference in the environment of galaxies.

The metallicity of galaxies in diverse environments is extensively studied in the local Universe. Small but significant environmental dependence of the MZR is found, especially for satellite galaxies, such that at a given stellar mass, galaxies in overdensities have higher gas-phase metallicities (e.g., Cooper et al., 2008; Ellison et al., 2009; Peng and Maiolino, 2014; Wu et al., 2017; Schaefer et al., 2019). Cooper et al. (2008) found that $\gtrsim$15% of the measured scatter in the MZR is caused by environmental effects. Furthermore, Peng and Maiolino (2014) found that galaxies with high metallicities favor denser environments at $z \sim 0$. They conclude that higher metallicity of infalling gas in dense environments is responsible for the environmental dependence of the MZR.

Beyond the local Universe (at $z \gtrsim 1$), the situation is unclear. Some studies found evidence of enhanced gas-phase metallicity in low-mass cluster galaxies at $z \gtrsim 1.5$ (e.g., Kulas et al., 2013; Shimakawa et al., 2015; Maier et al., 2019). Kulas et al. (2013) used Keck/MOSFIRE observations to compare the gas-phase metallicity of 23 protocluster members ($z \sim 2.3$) with 20 field galaxies. They found that the mean metallicity of low mass galaxies ($M_{*} \sim 10^{10} M_{\odot}$) in the protocluster is ~ 0.15 dex higher than that

in the field. On the other hand, Valentino et al. (2015) reported 0.25 dex lower gas-phase metallicity for the members of a $z \sim 2$ protocluster ($M_* \sim 10^{10.5} M_\odot$) compared to field galaxies at the same redshift. Conversely, some other studies have not observed significant environmental dependence of the MZR at high redshift (e.g., Tran et al., 2015; Kacprzak et al., 2015; Namiki et al., 2019).

With the wealth of near-IR spectroscopy for the star-forming galaxies ($1.37 \leq z \leq 2.61$) in the MOSFIRE Deep Evolution Field (MOSDEF) survey (Kriek et al., 2015), combined with the local environment measurements (Chartab et al., 2020) derived from accurate and uniformly calculated photometric redshifts with well-calibrated probability distributions, in this paper we investigate the effect of the local environment on the gas-phase metallicity of galaxies at $z \sim 1.5$ and $z \sim 2.3$. In Section 3.2, we present the details of the MOSDEF sample used in this work. We then briefly describe the local number density measurements, as a proxy for the environment, based on the photometric observations of the CANDELS fields. We describe sample selection procedure in Section 3.3. In Section 3.4, we investigate the role of the environment in the MZR, followed by the physical interpretation of our observations. We discuss our results in Section 3.5 and summarize them in Section 3.6.

Throughout this work, we assume a flat ΛCDM cosmology with $H_0 = 100h$ kms^{-1}Mpc^{-1}, $\Omega_{m_0} = 0.3$ and $\Omega_{\Lambda_0} = 0.7$. All magnitudes are expressed in the AB system, and the physical parameters are measured assuming a Chabrier (2003) IMF.

3.2 Data

3.2.1 The MOSDEF survey

MOSDEF is an extensive near-IR spectroscopic survey conducted over 48.5 nights using the Keck/MOSFIRE spectrograph (McLean et al., 2012). As a part of the survey, $\sim$ 1500 galaxies in five CANDELS fields (EGS, COSMOS, GOODS-N, UDS, and GOODS-S) were observed in three redshift ranges where strong rest-optical emission lines fall within windows of atmospheric transmission (400 galaxies at $z \sim 1.5$, 750 at $z \sim 2.3$, and 400 at $z \sim 3.4$). These galaxies are selected from the 3D-HST photometric and spectroscopic catalogs (Skelton et al., 2014; Momcheva et al., 2016) to a limiting HST/F160W magnitude of 24.0, 24.5, and 25.0 at $z \sim 1.5$, 2.3, and 3.4, respectively. Based on these magnitude limits, the MOSDEF sample is roughly mass-complete down to $M_* \sim 10^{9.5} M_\odot$ (Shivaei et al., 2015). For a full description of the survey strategy, observation and data reduction, we refer readers to Kriek et al. (2015).

The spectral energy distributions (SED) of MOSDEF galaxies are fitted using the multiwavelength photometric data from the 3D-HST survey (Skelton et al., 2014; Momcheva et al., 2016). Briefly, the photometric fluxes are corrected for contamination caused by strong nebular emission lines measured from the MOSDEF rest-frame optical spectra. The flexible stellar population synthesis model of Conroy et al. (2009) is adopted to build a library of synthetic spectral energy distributions. Star formation histories are modeled with a delayed exponentially declining function (SFR $\propto t e^{-t/\tau}$), where t is the age of a galaxy and τ is the star formation timescale. Dust attenuation is applied using the Calzetti et al. (2000) law and solar stellar metallicity is assumed for all galaxies. Then, the SED fitting is performed using the FAST code (Kriek et al., 2009), which uses a χ^2 minimization method to find the best-fit stellar population model and

corresponding properties such as stellar mass and SFR. Corresponding confidence intervals are computed by perturbing the photometry using the photometric errors. Redshifts are fixed to their spectroscopic values for the SED fitting. The spectroscopic redshifts and emission line fluxes are measured from the extracted 1D spectra, presented in Kriek et al. (2015).

3.2.2 Measuring local environment of galaxies

To quantify the environment of MOSDEF galaxies, we utilize the publicly available catalog of Chartab et al. (2020) which includes measurements of local density for galaxies brighter than HST/F160W $\leq$ 26 AB mag in all the five CANDELS fields: GOODS-S (Guo et al., 2013), GOODS-N (Barro et al., 2019), COSMOS (Nayyeri et al., 2017), EGS (Stefanon et al., 2017), and UDS (Galametz et al., 2013).

As described in Chartab et al. (2020), including both the spectroscopic and photometric redshifts for density measurements can bias the estimates in favor of galaxies with spectroscopic redshifts. Thus, despite the availability of spectroscopic redshifts for $\sim$ 12% of galaxies in the CANDELS fields, the environment catalog relies on uniformly calculated photometric redshifts (normalized median absolute deviation $\sigma_{\rm NMAD} \sim 0.02$), with well-calibrated redshift probability distribution functions (PDF) (Kodra, 2019, D. Kodra et al. in prep). The environment catalog has been constructed using the full photometric redshift PDFs adopting the technique of boundary-corrected weighted von Mises kernel density estimation (Chartab et al., 2020). Here we provide a brief explanation of density measurements.

Taking advantage of the well-calibrated and tested photometric redshift PDFs of CANDELS galaxies, the position of galaxies are treated in a probabilistic way, such that all the information regarding the position of a galaxy in redshift space is embedded in

its photometric redshift PDF. The dataset within the CANDELS fields is sliced to constant comoving width, $\Delta\chi = 35h^{-1}$Mpc (e.g., $\Delta z = 0.035$ at $z \sim 2$), which is greater than both redshift space distortion (Fingers-of-God effect; Jackson, 1972) and the photometric redshift PDF resolution over the redshift range $0.4 \leq z \leq 5$. Each galaxy is distributed over all redshift slices (z-slice) based on its photometric redshift PDF, such that the galaxy has a specific weight in each z-slice. Then, to calculate the local density in each z-slice, the weighted von Mises kernel density estimation is employed. The von Mises kernel is the spherical analog of the Gaussian kernel where variables are angles (e.g., right ascension and declination) instead of linear data (García-Portugués et al., 2013). To quantify the environment, density contrast (δ) is defined as

$$\delta = \frac{\Sigma}{\bar{\Sigma}} - 1, \tag{3.1}$$

where Σ is the number density of galaxies at the desired point and $\bar{\Sigma}$ is the average density in the corresponding z-slice. Figure 3.1 shows an example of the density map for one of the z-slices at $z \sim 2.13$ in the COSMOS field. One should note that the probabilistic nature of the method allows us to take into account the contribution of all galaxies (based on their photometric redshift PDFs) when we create density maps. In Figure 3.1 we also include six spectroscopically confirmed overdensities at $z \sim 2.1$ (Yuan et al., 2014) which were initially discovered from the Magellan/FOURSTAR Galaxy Evolution Survey (ZFOURGE) (Spitler et al., 2012). We find that all six confirmed overdensities in Yuan et al. (2014) are correctly predicted in the density map.

To assign a density contrast for each galaxy, the weighted integration of the local densities over z-slices has been performed since a galaxy with a photometric redshift is not localized in redshift space but is distributed over all z-slices based on its PDF. The densities are also corrected for a systematic under-estimation near the edge of the

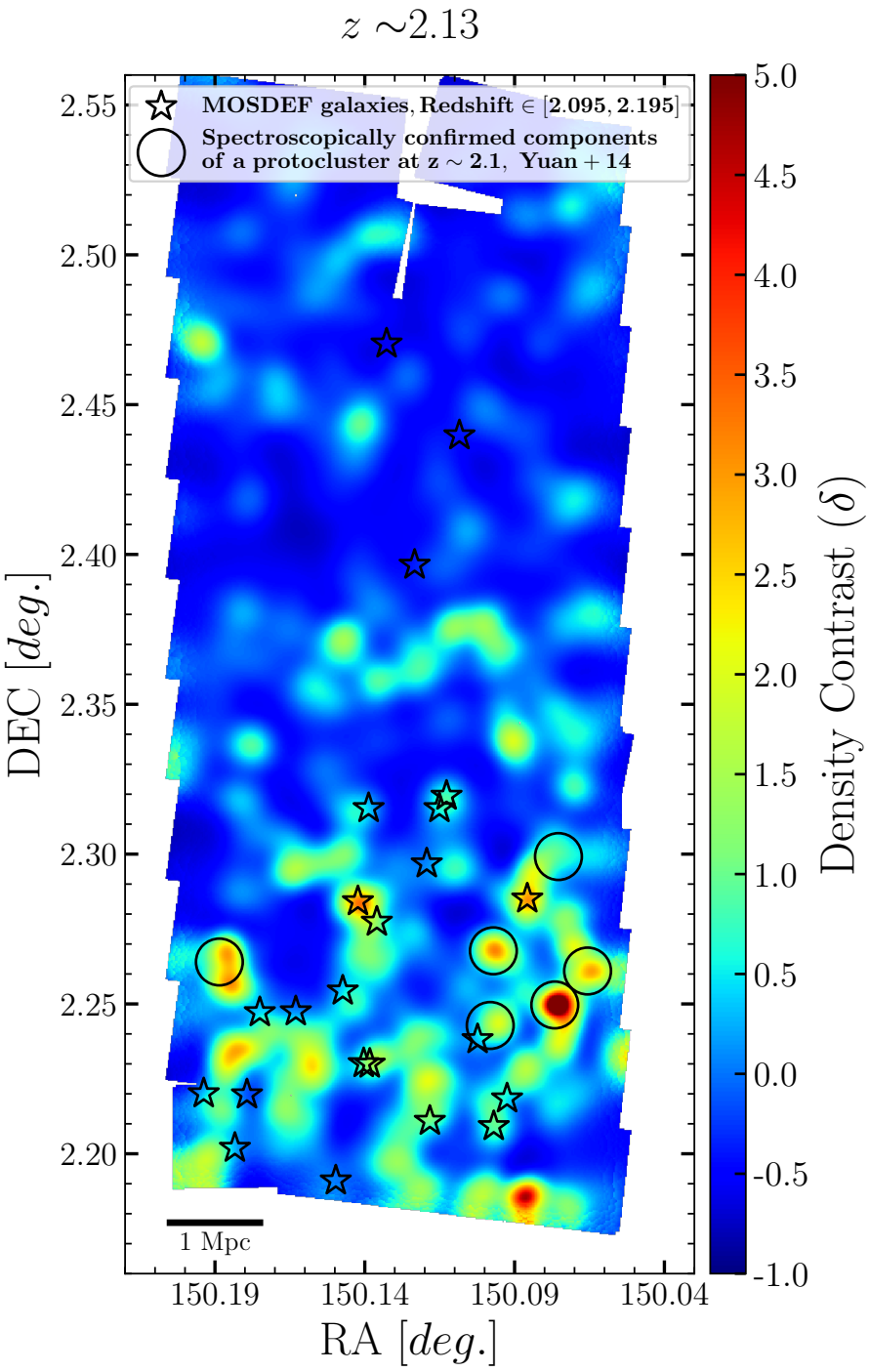


Figure 3.1: An example of the density map for the CANDELS-COSMOS field at $z \sim 2.13$ (Chartab et al., 2020). Stars show MOSDEF sources located in this z-slice. Yuan et al. (2014) spectroscopically confirmed six overdensities at this redshift which were initially identified by Spitler et al. (2012). They are denoted by open circles.

survey footprint using the re-normalization technique.

In the following section, we cross-match the MOSDEF galaxies with the CANDELS photometric catalogs to find their associated local density measurements. Although we have spectroscopic redshifts for the MOSDEF galaxies, we do not pin them in density maps based on spectroscopic redshifts to measure their environments. The density measurements trace the relative position of galaxies in the survey and are not sensitive to systematic biases that may exist in the photometric redshifts, whereas measuring densities based on the photometric redshifts but defining positions of the MOSDEF galaxies in density maps by their spectroscopic redshifts leads to inconsistencies.

3.3 Sample selection

We select MOSDEF galaxies that have Hα line luminosities with S/N $\geq$ 3. We only include the star-forming galaxies based on the UVJ rest-frame color selection (Muzzin et al., 2013) derived from SED fitting. Objects that are flagged as an active galactic nucleus (AGN) in the MOSDEF catalog are excluded. The AGNs are identified based on X-ray emission or IRAC colors. We also require that $\log([\mathrm{NII}]/\mathrm{H}\alpha) < -0.3$ to exclude optical AGNs (Coil et al., 2015; Azadi et al., 2017, 2018; Leung et al., 2017, 2019). To have a mass-complete sample, we only consider galaxies with $M_* \geq 10^{9.5} M_\odot$. These criteria result in a total of 560 galaxies at $0.78 \leq z \leq 2.64$. We then cross-match these objects with the local environment catalog of Chartab et al. (2020) within a radius equal to the FWHM size of HST/F160w band point spread function, $\sim 0.2''$. For 23 galaxies, we could not find a source within the radius of $\sim 0.2''$ in the CANDELS photometry catalogs due to the difference in the source identification and photometry extraction between the CANDELS and the 3D-HST. Also, 15 galaxies were not included in the environment catalog since the catalog is constructed based on specific selection criteria:

(1) SExtractor's stellarity parameter < 0.95, (2) requiring 95% of photometric redshift PDF of the galaxy to fall within the redshift range of $0.4 \leq z \leq 5$, and (3) brighter than 26 AB mag in HST/F160w band. 14 out of 15 missing galaxies in the environment catalog were incorrectly identified as low-z galaxies or had very broad photometric redshift PDFs which have not satisfied the second criterion, and the other missing galaxy was identified as a galaxy with HST/F160w $>$ 26 AB mag in the CANDELS photometric catalogs that has not satisfied the third criterion.

Figure 3.2 shows the comparison between the spectroscopic and the CANDELS photometric redshifts of the sample. The photometric redshift is defined as a probability-weighted expectation value of the redshift based on the photometric redshift PDF (Kodra, 2019). We find a value of 0.03 for the normalized median absolute deviation of photometric redshifts. This revalidates the accuracy of photometric redshifts, which results in reliable local density measurements. Furthermore, we removed outlier galaxies with the error > 0.5 in their photometric redshifts and galaxies out of the desired redshift range (shown with open circles in Figure 3.2). The final sample is divided into two redshift bins, 167 galaxies at $1.37 \leq z \leq 1.70$, and 303 galaxies $2.09 \leq z \leq 2.61$.

To investigate if our sample is representative of the full dynamic range of the local environment in the CANDELS data, we compare the local environmental density of CANDELS star-forming galaxies with $M_* \geq 10^{9.5} M_\odot$ at the same redshift range with our sample (the MOSDEF galaxies). We use rest-frame UVJ colors computed in Chartab et al. (2020) to select a star-forming sub-sample of CANDELS galaxies. Figure 3.3 shows the histogram of density contrast for MOSDEF galaxies and for all of the star-forming CANDELS galaxies at the same redshift ranges. We find that MOSDEF galaxies cover a wide range of environments, making them unique for studying the environmental dependence of spectroscopic properties of galaxies. There is slight evidence that MOSDEF galaxies reside in relatively denser environments than the CANDELS star-forming

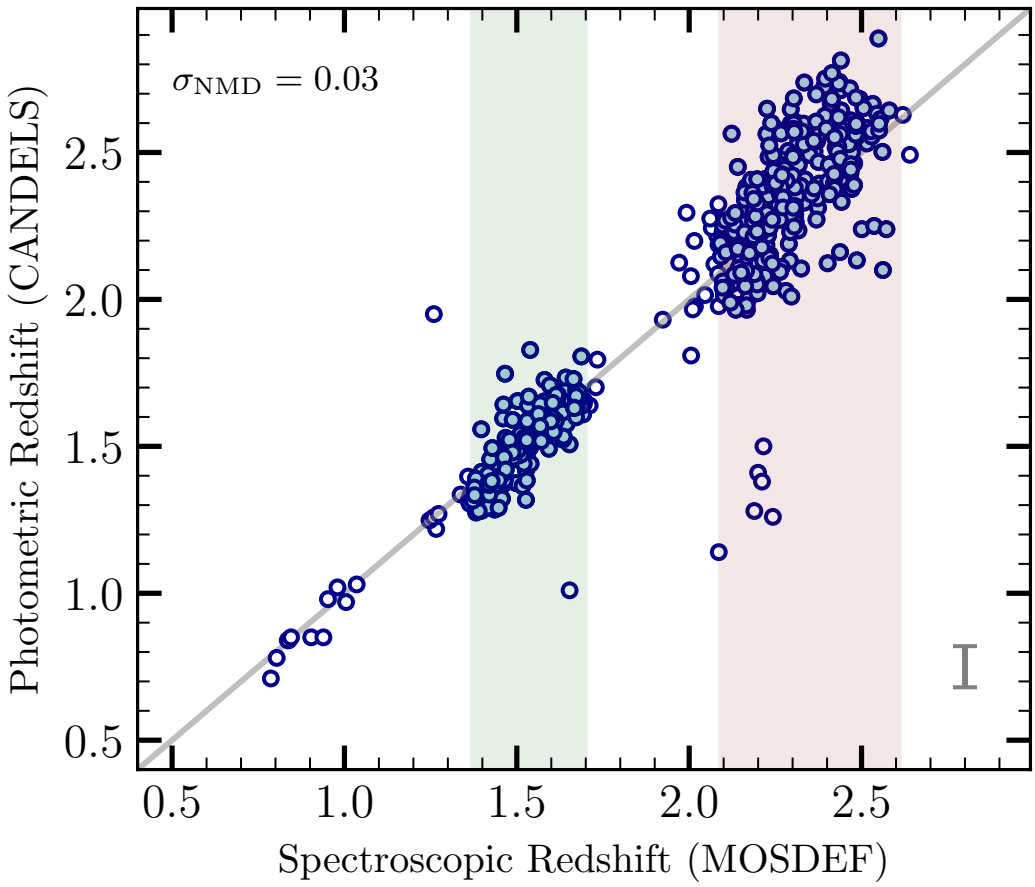


Figure 3.2: Comparison between the MOSDEF spectroscopic and CANDELS photometric redshifts of the sample. The photometric redshift is defined as a probability-weighted expectation value of the redshift based on the photometric redshift PDF. The average uncertainty on photometric redshifts is displayed in the lower right corner. The normalized median absolute deviation of the redshift for the sample is 0.03. Solid blue circles show the final sample used in this work. The green shaded region corresponds to the lower redshift bin at $z \sim 1.5$, and the pink region shows the highest redshift bin at $z \sim 2.3$. Open circles are either outliers or galaxies outside the desired redshift ranges.

sample which is expected for most spectroscopic surveys as they usually maximize the number of sources per mask. However, this effect is minimal in MOSDEF since it is an extensive program that covers ~ 600 arcmin^2 at $z \sim 2.3$ and ~ 300 arcmin^2 at $z \sim 1.5$ (Kriek et al., 2015). Given the total area of the CANDELS fields (~ 960 arcmin^2), MOSDEF covers $\sim 30\%$ and $\sim 60\%$ of the CANDELS area at $z \sim 1.5$ and $z \sim 2.3$, respectively. This wide areal coverage translates to diverse environments in our sample, as shown in Figure 3.3.

We divide the sample into three bins of environment. Although we find that the local environment distribution of our sample is well representative of the full CANDELS galaxies, we do not set the environment binning thresholds based on our sample. Instead, we calculate them from all the star-forming CANDELS galaxies at desired redshift ranges. We find the tertiles which divide the CANDELS sample into three bins of environments, each containing a third of the sample. We consider galaxies in the lowest tertile ($\delta < \delta_{\frac{1}{3}}$) as field galaxies and those within the second tertile ($\delta_{\frac{1}{3}} \leq \delta < \delta_{\frac{2}{3}}$) as intermediate-densities and the highest tertile ($\delta \geq \delta_{\frac{2}{3}}$) as overdensities. Table 3.1 shows the density contrast thresholds to define the field, intermediate and overdense samples along with the sample size and the average density contrast, $\langle 1+\delta \rangle$ at each redshift bin. Our sample includes ~ 55 and ~ 100 galaxies in each bin of the environment at $z \sim 1.5$ and $z \sim 2.3$, respectively (see Table 3.1).

Fossati et al. (2015) linked observational local number density of galaxies to their parent halo masses using a stellar mass-limited sample of galaxies ($M_* > 10^{9.5} M_\odot$) in semi-analytic models of galaxy formation. We estimate an average halo mass $M_{halo} \gtrsim 10^{13} M_\odot$, for overdensities (the highest tertile in present work) at $z \sim 2$ (see also, Fossati et al., 2017). Present-day descendants of these overdensities have $M_{halo} \gtrsim 10^{14} M_\odot$ (Behroozi et al., 2013) associated with rich clusters. Massive core halos of these present-day rich cluster progenitors at $z \sim 2$ (protoclusters) have virial radii of

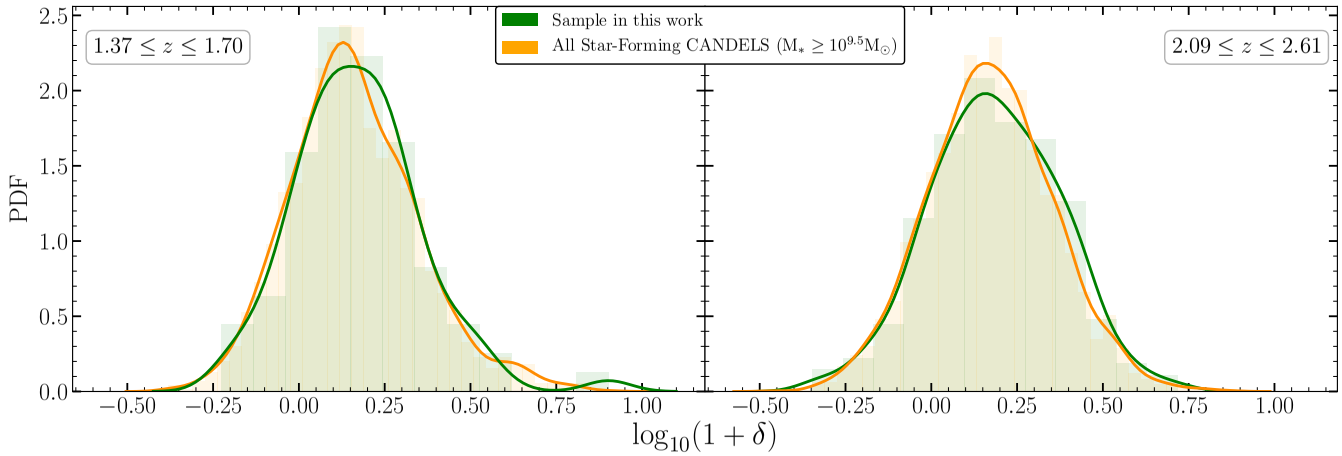


Figure 3.3: Histograms of overdensity measurements $(1+\delta)$ for all the star-forming CANDELS galaxies with $M_* \geq 10^{9.5} M_\odot$ and the sample used in this work at redshifts $z \sim 1.5$ (*left*) and $z \sim 2.3$ (*right*). Our sample covers a wide range of environments, making it unique to study the environmental dependence of spectroscopic properties of galaxies.

Table 3.1: Properties of environment-selected sample

Environment	Sample size	$1+\delta$	$\langle 1+\delta \rangle$
	$z \sim 1.5$		
Field	53	<1.19	0.97
Intermediate-density	54	1.19-1.69	1.43
Overdense	60	>1.69	2.43
	$z \sim 2.3$		
Field	97	<1.24	0.97
Intermediate-density	96	1.24-1.77	1.49
Overdense	110	>1.77	2.48

$\lesssim$1 comoving Mpc (Chiang et al., 2017), which can be observed within the CANDELS fields. Thus, our last environment bin, so-called overdensity, traces these massive cores of protoclusters at high redshift which will grow into $z = 0$ clusters. However, all the dark matter and baryons that will assemble into a $z = 0$ cluster may be very extended at $z \sim 2$, $\sim$ 50 comoving Mpc (Muldrew et al., 2015), which can not be captured in small CANDELS fields.

We also estimate an average halo mass $M_{halo} \lesssim 10^{12.5} M_{\odot}$ (Fossati et al., 2015) for the lowest tertile of the environment bins, so-called field galaxies. These halos will grow into halos with $M_{halo} \lesssim 10^{13} M_{\odot}$ at $z = 0$ (Behroozi et al., 2013). Thus, field galaxies in the present work are the progenitor of galaxies residing in very poor $z = 0$ groups/clusters (Local Group-like).

In the following section, we model the MZR for our environment-selected samples to understand how metal enrichment processes of galaxies change with their respective environments.

3.4 Results

3.4.1 The MZR in diverse environments

Here we use Hα and [NII]λ6584 lines to estimate oxygen abundances of galaxies as an indicator for their gas-phase metallicities. However, the [NII]λ6584 emission line is not detected with S/N $\geq$ 3 for 48 out of 167 and 118 out of 303 galaxies at $z \sim 1.5$ and $z \sim 2.3$, respectively. As we discuss later in Section 3.5.3, requiring [NII]λ6584-detection biases our sample toward higher gas-phase metallicities. Therefore, to include [NII]λ6584 non-detected galaxies in metallicity measurements, we create composite spectra by stacking the spectra of galaxies in bins of stellar mass and environment.

Following Shivaei et al. (2018), we shift individual spectra to the rest frame and then normalize them by the Hα luminosity. Composite spectra are computed by averaging the normalized spectra in bins of 0.5 Å considering the weights of $1/\sigma_i^2$ where σ is the standard deviation of the i^{th} spectra. The uncertainty in the weighted average is also obtained using $(\sum \frac{1}{\sigma_i^2})^{-\frac{1}{2}}$.

The resultant composite spectra are normalized to the Hα luminosity. Therefore, the average flux ratio of $\langle \frac{[\mathrm{NII}]\lambda6584}{\mathrm{H}\alpha} \rangle$ is determined by fitting a triple Gaussian function for [NII]λ6548, 6584 and Hα lines and extracting the area underneath the Gaussian function for the normalized [NII]λ6584 line. We perturb the composite spectra using their error distributions and estimate the average flux ratio for 500 trials. Also, galaxies in each bin of stellar mass and environment are bootstrap re-sampled in each trial to account for sample variance. The average and standard deviation of 500 trials are adopted as the flux ratio and its uncertainty.

We determine oxygen abundances ($12 + \log(\mathrm{O/H})$) for the composite spectra using Pettini and Pagel (2004) calibration of the $\mathrm{N2} = \frac{[\mathrm{NII}]\lambda6584}{\mathrm{H}\alpha}$ line ratio,

$$12 + \log(\mathrm{O/H}) = 8.9 + 0.57 \log(\mathrm{N2}) \tag{3.2}$$

The intrinsic uncertainty (1σ dispersion) of the oxygen abundance of an individual galaxy in the calibration is 0.18. A composite spectrum consists of N galaxies; thus, the intrinsic error in the oxygen abundance of the composite spectrum is $0.18/\sqrt{N}$ (Erb et al., 2006; Sanders et al., 2015). To include the N2 calibration error in measurements, one can calculate the total variance of metallicity by adding the intrinsic variance to the weighted average variance from the stacks. We do not include calibration error when reporting the uncertainty in metallicities. Moreover, the N2 estimator is calibrated locally and there is some debate within the literature regarding the validity of the calibration for high-redshift galaxies (Kewley et al., 2013; Steidel et al., 2014; Shapley et al., 2019; Sanders et al., 2020). However, it does not cause a problem for our work, where we study the relative metallicities to understand the effect of the local environment on the metal content of galaxies.

We further separate our sample in different environments into three bins of stellar mass such that each bin includes approximately an equal number of galaxies. We then calculate composite line luminosities for these nine bins of environments and stellar masses and calculate composite metallicities. To determine the average MZR for each environment (field, intermediate-density, and overdense), we fit a linear function to the composite metallicities and average stellar masses,

$$12 + \log(\mathrm{O/H}) = \beta + \alpha \log(\frac{\mathrm{M}_*}{\mathrm{M}_\odot}), \tag{3.3}$$

where α and β are the slope and intercept of the linear function, respectively. The best-fit parameters and errors are given in Table 3.2. To estimate errors, we perturb

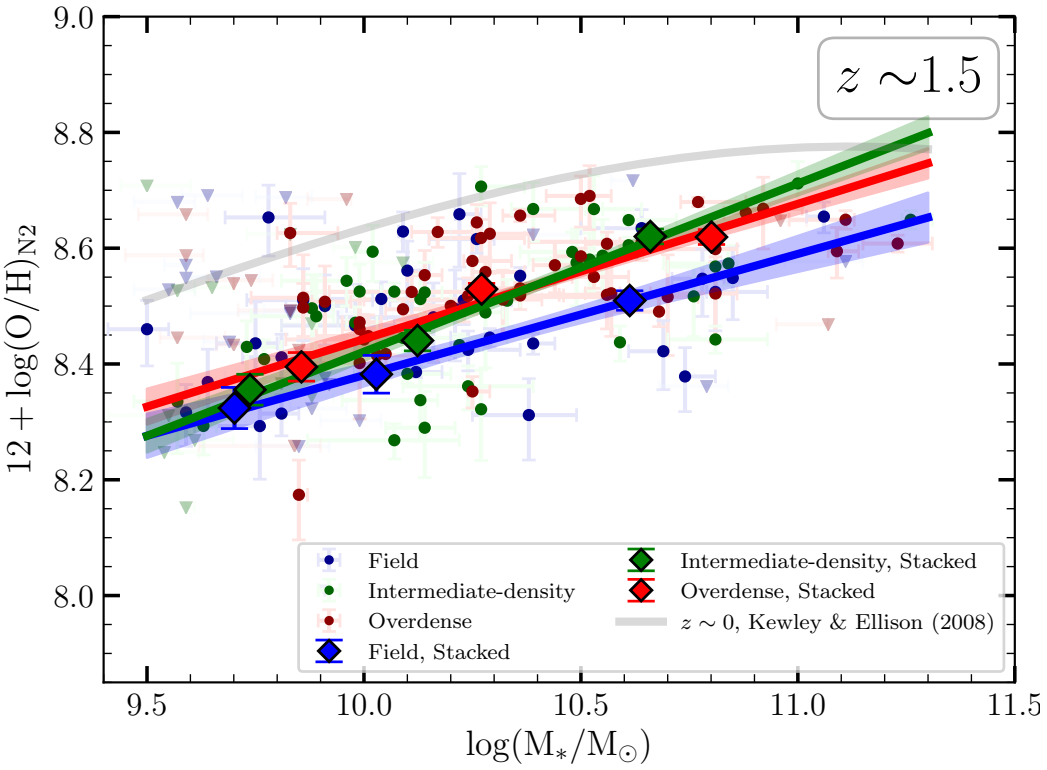


Figure 3.4: MZR at $z \sim 1.5$ for galaxies residing in 3 different environments: Overdensity (red), Intermediate-density (green), and underdensity (blue). For [NII]λ6584 non-detection, the upper limit of the metallicity is shown with inverted triangle symbols. The metallicity measurements for the composite spectra are shown with diamonds. The fitted average MZR lines (based on Table 3.2) for different environments are shown along with the 1σ error bars (the shaded regions around the best-fit models). The best-fit models suggest that the metallicity at a given stellar mass enhances in denser environments at $z \sim 1.5$ while the trend reverses at $z \sim 2.3$ such that galaxies in denser environments tend to have lower metallicities. Errors in the metallicities of composite spectra are not shown as they are smaller than the symbols (diamond). The SDSS local MZR (Kewley and Ellison, 2008) is also shown by solid grey line.

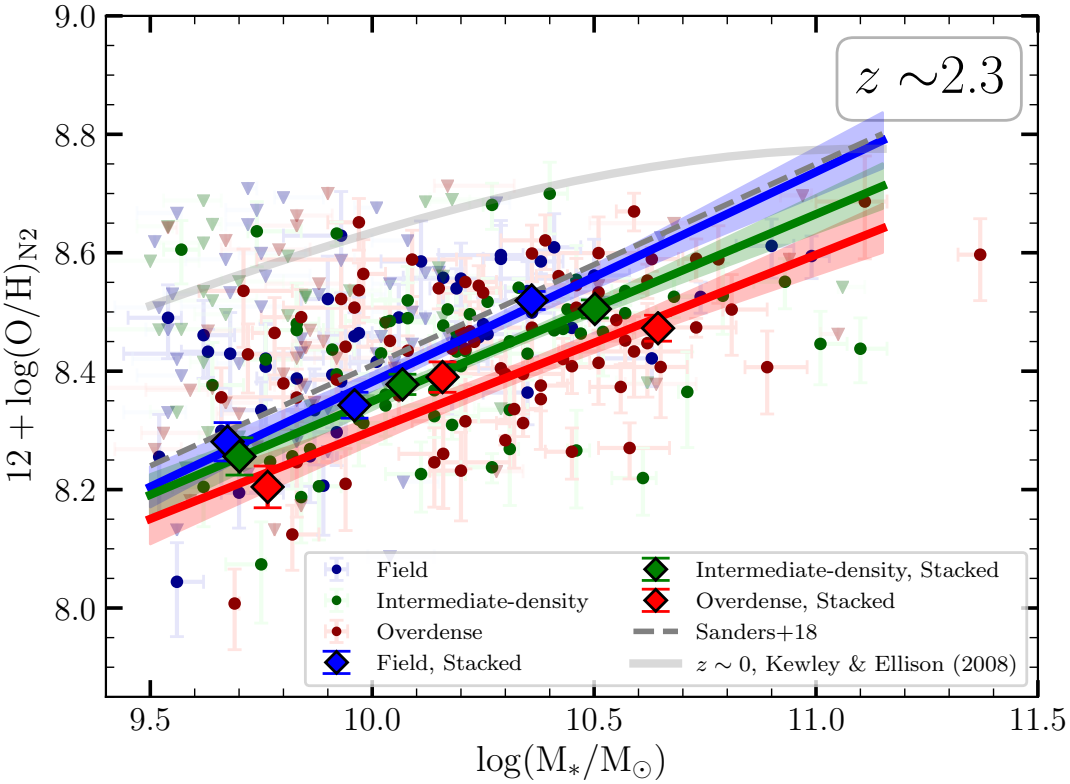


Figure 3.4: Continued: MZR at $z \sim 2.3$ for galaxies residing in 3 different environments: Overdensity (red), Intermediate-density (green), and underdensity (blue). For [NII]λ6584 non-detection, the upper limit of the metallicity is shown with inverted triangle symbols. The metallicity measurements for the composite spectra are shown with diamonds. The fitted average MZR lines (based on Table 3.2) for different environments are shown along with the 1σ error bars (the shaded regions around the best-fit models). The best-fit models suggest that the metallicity at a given stellar mass enhances in denser environments at $z \sim 1.5$ while the trend reverses at $z \sim 2.3$ such that galaxies in denser environments tend to have lower metallicities. Errors in the metallicities of composite spectra are not shown as they are smaller than the symbols (diamond). The SDSS local MZR (Kewley and Ellison, 2008) is also shown by solid grey line and the average MZR for all the MOSDEF galaxies at $z \sim 2.3$ from Sanders et al. (2018) is demonstrated by dashed grey line.

Table 3.2: Best-fit linear parameters for MZR[a]

Environment	α	β
$z \sim 1.5$		
Field	0.21 ± 0.04	6.29 ± 0.43
Intermediate-density	0.29 ± 0.03	5.51 ± 0.31
overdense	0.23 ± 0.03	6.11 ± 0.31
$z \sim 2.3$		
Field	0.35 ± 0.05	4.84 ± 0.48
Intermediate-density	0.31 ± 0.04	5.19 ± 0.44
overdense	0.30 ± 0.04	5.33 ± 0.49
Sanders + 18[b]	0.34	5.01

a Best-fit parameters for $12 + \log(\mathrm{O/H}) = \beta + \alpha \log(\mathrm{M_*/M_\odot})$

b Best-fit parameters from Sanders et al. (2018) for MOSDEF galaxies at $z \sim 2.3$ without any constraint on their environments.

metallicities according to their uncertainties and repeat the fitting process 500 times. We also include best-fit parameters from Sanders et al. (2018) for the MOSDEF galaxies at $z \sim 2.3$. Their sample is the same as ours except that they have one more constraint on the detection (S/N > 3) of the Hβ emission line.

The resulting MZRs in different environments are shown in Figure 3.4 along with the metallicity and stellar mass of individual galaxies. The metallicity measurements for the composite spectra are shown with diamonds. We find an enhancement in the gas-phase metallicity of galaxies in dense environments at $z \sim 1.5$ compared to the field galaxies, while the trend reverses at $z \sim 2.3$ such that galaxies in the dense environment tend to have lower metallicities compared to their counterparts in lower-density environments.

At $z \sim 1.5$, the average galaxies with $10^{9.5}\mathrm{M_\odot} \lesssim \mathrm{M_*} \lesssim 10^{11}\mathrm{M_\odot}$ which reside in overdensities have enhanced metallicities by ~ 0.07 dex compared to their field counterparts. At the low stellar mass end of the MZR, this enhancement is less significant, mostly due to a higher fraction of [NII]λ6584 non-detection galaxies in the low mass end. At $z \sim 2.3$, the average galaxies with $10^{9.5}\mathrm{M_\odot} \lesssim \mathrm{M_*} \lesssim 10^{11}\mathrm{M_\odot}$ in the dense environment are ~ 0.11 dex metal deficient relative to the field galaxies at the same stellar mass range.

A notable caveat for studies with a small sample size, which is the case for most of the high-redshift spectroscopic samples, is dissimilar stellar mass distributions among the galaxies in different environments. It is crucial to have the same stellar mass distributions in different environments to properly study the effect of the environment on the gas-phase metallicity of galaxies at a given stellar mass. As we bin the data to find the average MZR in diverse environments, different stellar mass distributions can result in different average metallicities. This difference can be misinterpreted as an environmental imprint on the MZR. Therefore, before proceeding to a discussion of our results, in the next section, we first perform an analysis on a mass-controlled sample of galaxies where we carefully match the stellar mass distribution of galaxies in different environments and find the composite spectra for the mass-controlled sample. This allows us to properly disentangle the effect of stellar mass from the local environment, providing an unbiased measurement of environmental trend.

3.4.2 Mass-controlled sample

The stellar mass distribution of our sample is shown in Figure 3.5. As expected, galaxies in diverse environments have different stellar mass distributions. We find that overdensities tend to have a higher fraction of massive galaxies compared to underdense regions. For instance, the median stellar mass of our field sample at $z \sim 1.5$ is $10^{9.9}M_{\odot}$, while this value for the overdense sample is $10^{10.25}M_{\odot}$. In the highest redshift bin, $z \sim 2.3$, the median stellar masses are $10^{9.93}M_{\odot}$ and $10^{10.17}M_{\odot}$ for field and overdense galaxies, respectively. It is, therefore, important to match the stellar mass distribution of galaxies in different environments to properly disentangle the effect of the local environment on the gas-phase metallicity from the stellar mass. As shown in Figure 3.5, we match the stellar mass distribution of galaxies in three environment bins by sub-

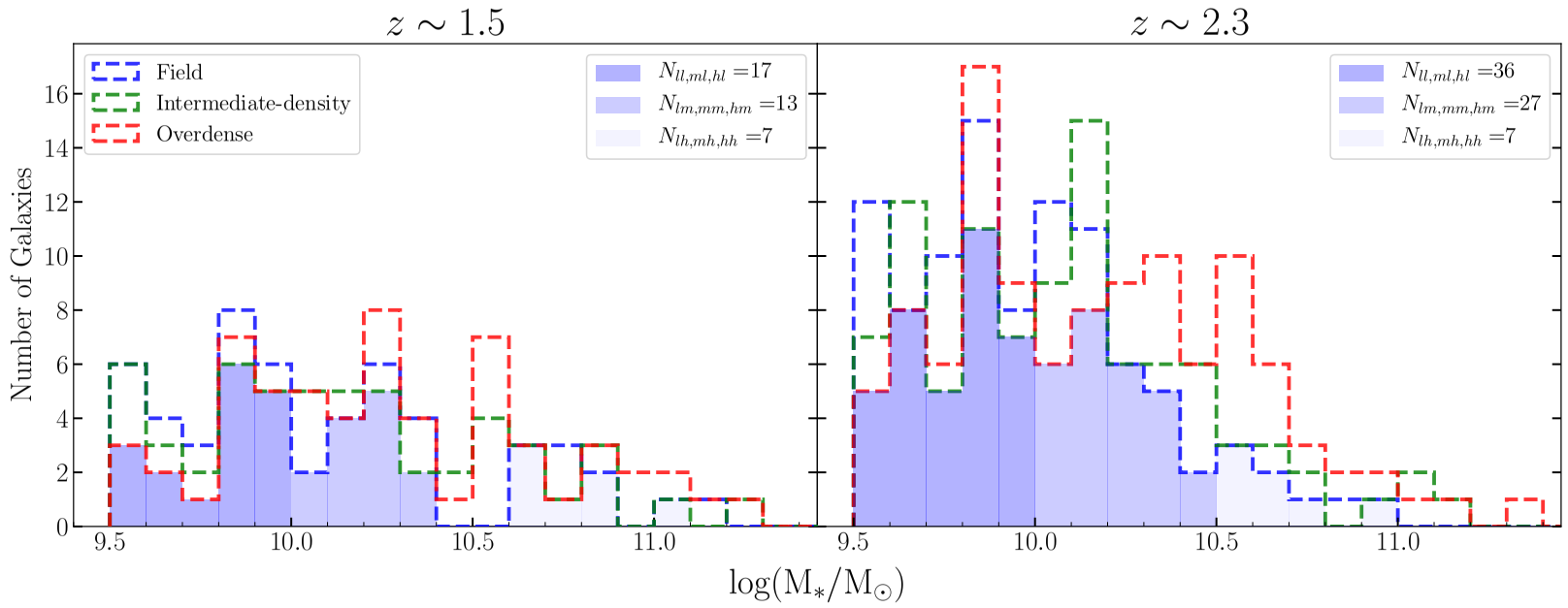


Figure 3.5: Histograms of stellar mass for the sample in different environments (dashed lines). Shaded regions show a sub-sampled data with the matched stellar mass distributions. At a given stellar mass, sub-sampled data have the same number of galaxies in every environment bin. The number of galaxies in each sub-sampled bin is shown in the format of N_{xy}, where x shows the type of environment and y shows the stellar mass range ("l": low, "m": intermediate and "h": high). For example, $N_{lh} = 7$ at $z \sim 1.5$ means that each sub-sampled data consists of 7 massive field galaxies at that redshift.

sampling the galaxies such that each environment has the same number of galaxies at a given stellar mass. We adopt a resolution of $\log(M_*/M_\odot) = 0.1$ dex when matching the stellar mass distributions. In other words, we draw a fraction of galaxies in different environments such that they have the same stellar mass distributions with 0.1 dex resolution. As there are not unique mass-matched sub-samples in different environments, we repeat sub-sampling 500 times. For each trial, we measure the composite spectra in the bins of environment and stellar mass using the same procedure described in Section 3.4.1. The stellar mass bins of $9.5 \leq \log(M_*/M\odot) < 10$, $10 \leq \log(M_*/M\odot) < 10.5$, and $\log(M_*/M\odot) \geq 10.5$ are adopted. We perturb the resultant composite spectra of each trial according to their uncertainties. Ultimately, we construct the final mass-controlled composite spectra and their errors, using the average and standard deviation of the 500 measurements, respectively.

We do not take into account stellar mass uncertainties when matching the stellar mass distribution of galaxies in different environments, since the median uncertainty of stellar masses in our sample is ~ 0.05 dex, which is smaller than the desired resolution in the mass-controlled sample, 0.1 dex.

3.4.2.1 Metallicity of mass-controlled sample

The composite spectra for the nine bins of stellar masses and environments for the mass-controlled sample are shown in Figure 3.6. The $\frac{[\mathrm{NII}]\lambda6584}{\mathrm{H}\alpha}$ estimates and gas-phase metallicity (oxygen abundance) for mass-controlled composite spectra in the nine bins of stellar masses and environments are presented in Table 3.3. We follow the same procedure as described in Section 3.4.1 to measure the metallicity of galaxies in the bins of environment and stellar mass from the composite spectra using the N2 indicator.

We also measure the average stellar masses and SFRs of the mass-matched samples in stellar mass and environment bins. We use SFRs derived from SED fitting since the

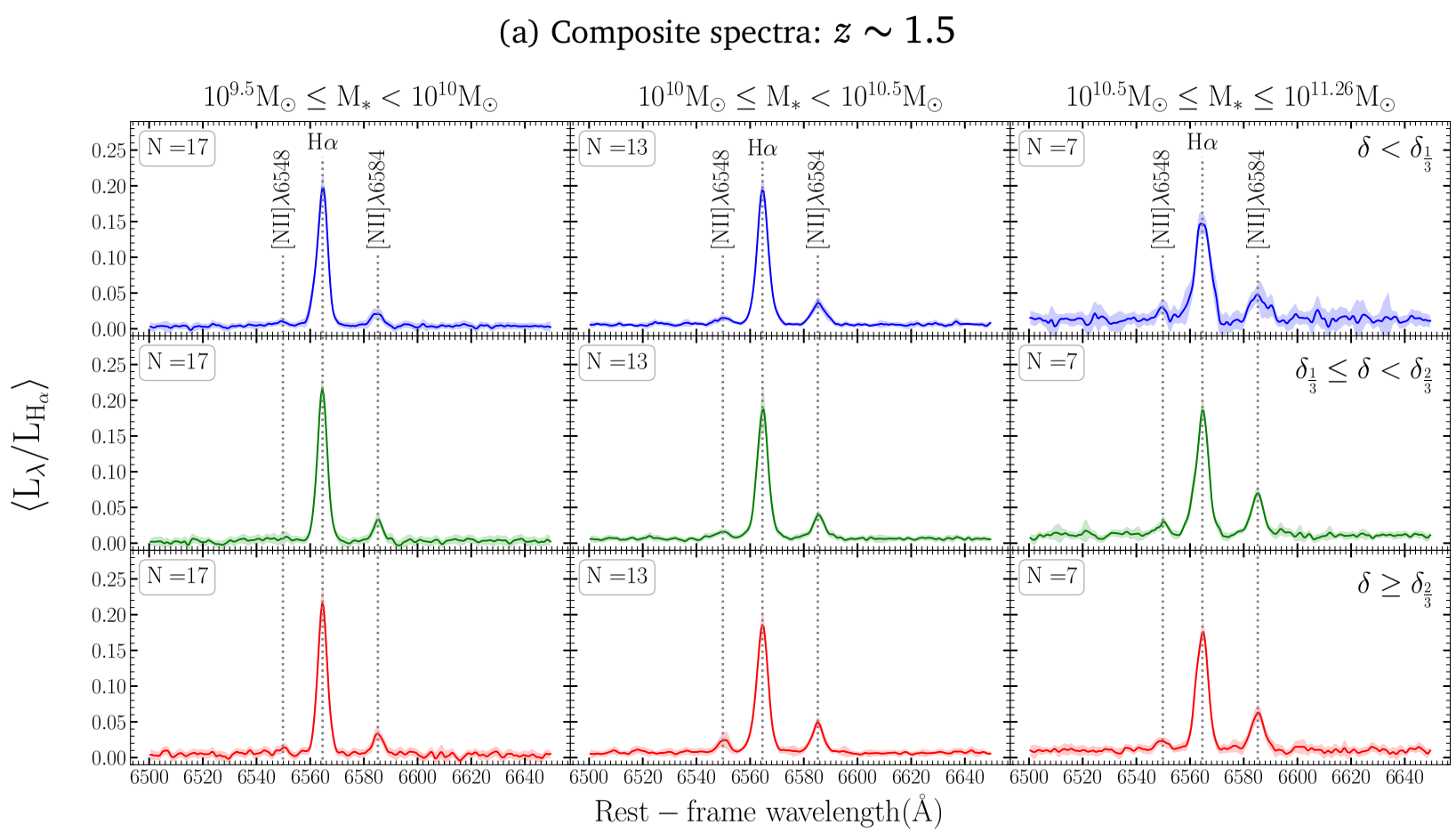


Figure 3.6: Composite spectra for the mass-controlled samples in the bins of stellar mass and environment at $z \sim 1.5$. Errors are represented with the shaded regions, which are around the weighted average spectra. To build the composite spectra for the mass-controlled samples, we use the sub-sampling technique. Every time we sub-sample the data such that they have similar stellar mass distributions in three bins of environment, we build the composite spectra and perturb them according to their uncertainties. In the end, using the average and standard deviation of 500 trials, we construct mass-controlled composite spectra and their errors, respectively.

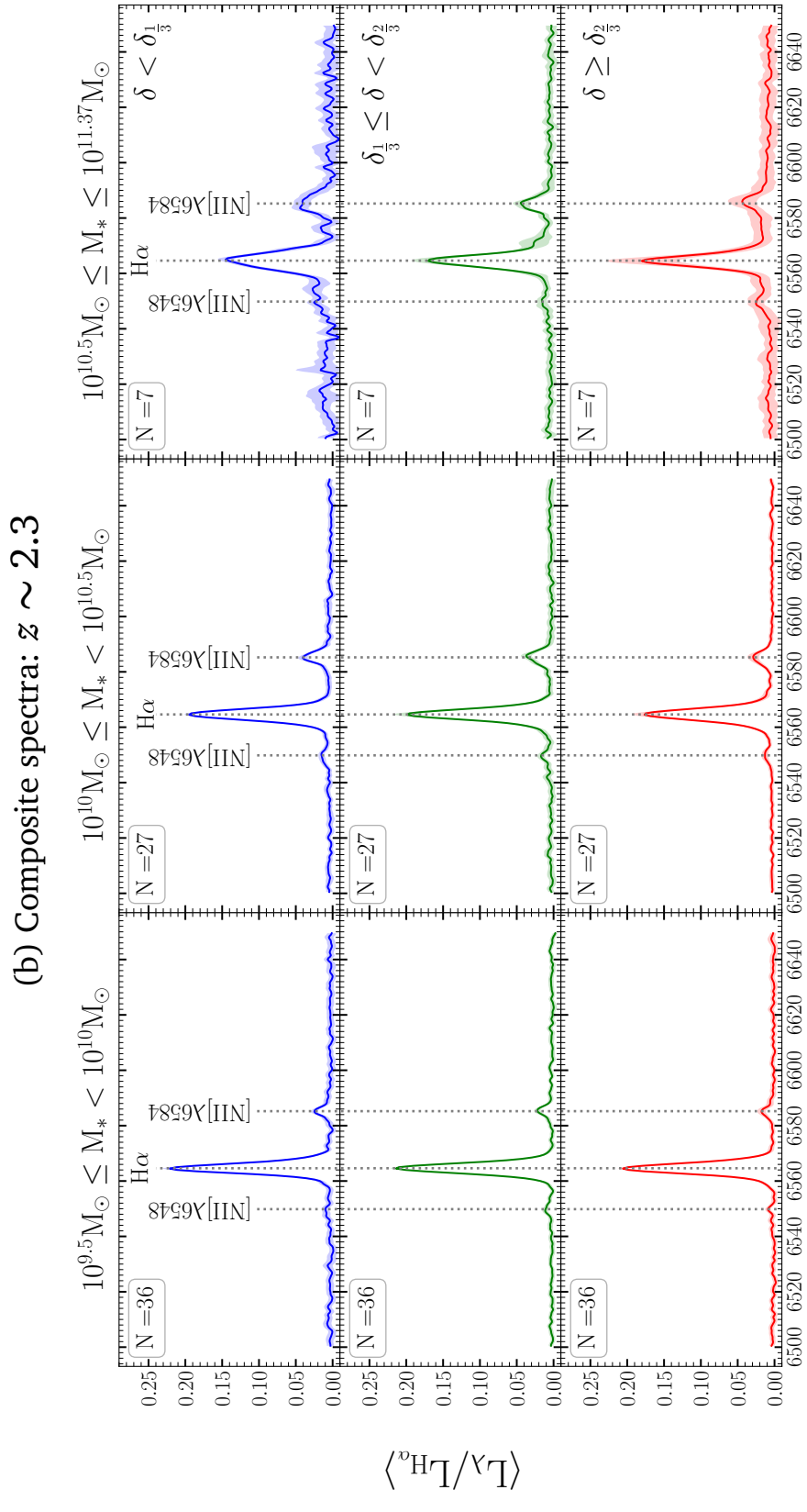


Figure 3.6: Continued

dust-corrected Hα luminosity is not available for 40% of our sample. We avoid imposing a constraint on the detection of the Hβ emission line needed for dust correction, as it decreases our sample size significantly. The variation of metallicity at fixed M_* in different environments is small (< 0.1 dex) and thus requires large sample sizes to be detected. Previous studies found that the SED-derived SFRs for MOSDEF galaxies are in general agreement with SFRs derived from dust-corrected Hα luminosities (Shivaei et al., 2016; Reddy et al., 2015). The average SED-derived SFRs listed in Table 3.3 suggest that even though there is evidence for a weak environmental dependence of the SFRs for our star-forming sample at a given stellar mass, it is not significant due to large uncertainties in the SFR measurements. Errors of SFRs listed in Table 3.3 include uncertainty in the SFR of individual galaxies as well as sample variance. A detailed study of the environmental imprints on specific SFRs of MOSDEF galaxies can properly constrain this relation. Old et al. (2020) have used [OII]-derived SFRs to investigate this relation at $1.0 < z < 1.5$ in Gemini Observations of Galaxies in Rich Early Environments Survey (GOGREEN; Balogh et al., 2017). They find no significant difference between the specific SFR of the cluster and the field sample at $z > 1.3$.

The MZR for the mass-controlled sample at $z \sim 1.5$ and $z \sim 2.3$ are presented in Figure 3.7. We emphasize that the stellar mass distributions in the environment bins are similar. This allows us to remove the effect of stellar mass on metallicity to properly investigate the environmental effects. Also, although the mass-matched sample has fewer number of galaxies than the whole sample, we sub-sample the data 500 times to incorporate the contribution of the full sample in measurements. As shown in Figure 3.7, we find that at a given stellar mass, the metallicity of galaxies changes with their respective environments at both redshift intervals considered here.

At $z \sim 1.5$, the average metallicity of galaxies in overdensities is higher than that of field galaxies, with enhancements of 0.094 ± 0.051 (1.8σ significance), 0.068 ± 0.028

(2.4σ significance) and 0.052 ± 0.043 (1.2σ significance) dex for the mass-controlled sample with $M_* \sim 10^{9.8} M_\odot, 10^{10.2} M_\odot$ and $10^{10.8} M_\odot$, respectively. For galaxies residing in the intermediate-density, the metallicity enhancements are 0.090 ± 0.052 (1.7σ significance), 0.007 ± 0.029 (insignificant) and 0.077 ± 0.042 (1.8σ significance) dex at the same stellar masses mentioned above.

In contrast, at $z \sim 2.3$, the average metallicity of galaxies in overdensities is lower than their field counterparts. Galaxies in the mass-controlled sample that reside in overdensities with $M_* \sim 10^{9.8} M_\odot, 10^{10.2} M_\odot$ and $10^{10.7} M_\odot$, are metal deficient by 0.056 ± 0.043 (1.3σ significance), 0.056 ± 0.028 (2σ significance) and 0.096 ± 0.034 (2.8σ significance) dex relative to the field sample, respectively. At this redshift, the metal deficiencies are 0.017 ± 0.032 (insignificant), 0.022 ± 0.025 (insignificant) and 0.085 ± 0.034 (2.5σ significance) dex for galaxies located in the intermediate-density with the same stellar masses as above.

Analysis of the mass-controlled samples confirms the trends already observed for the unmatched sample in Figure 3.4; however, the significance of observed trends is more reliable when we control the stellar mass distribution of the sample. Considering all stellar mass ranges ($10^{9.5} M_\odot \lesssim M_* \lesssim 10^{11} M_\odot$), at $z \sim 1.5$, on average galaxies in overdensities are ~ 0.07 dex rich in metal compared to their field counterparts. In contrast, at $z \sim 2.3$, average galaxies in overdensities have ~ 0.07 dex lower metallicities compared to the field galaxies. Before discussing the physical origin of the observed trends in detail, we compare our results with previous works in the following section.

Table 3.3: Properties of the composite spectra for the mass-controlled sample

Redshift	Environment	N_T^a	N^b	$\langle \log \frac{M_*}{M_\odot} \rangle$	$\langle \log(SFR[\frac{M_\odot}{year}]) \rangle$	$\langle N2 \rangle = \langle \frac{[NII]\lambda6584}{H\alpha} \rangle$	$\langle 12 + \log(O/H) \rangle$
1.37–1.70	Field	27	17	9.79±0.03	0.89±0.24	0.090±0.015	8.305±0.042
		16	13	10.20±0.03	1.15±0.23	0.164±0.014	8.452±0.022
		9	7	10.78±0.05	1.15±0.35	0.246±0.040	8.552±0.040
	Intermediate-density	22	17	9.81±0.04	0.75±0.18	0.130±0.016	8.395±0.030
		21	13	10.20±0.03	1.08±0.32	0.169±0.013	8.459±0.019
		7	7	10.75±0.05	1.31±0.36	0.335±0.018	8.629±0.013
	Overdense	18	17	9.81±0.03	0.78±0.20	0.132±0.016	8.400±0.030
		17	13	10.21±0.03	1.26±0.29	0.216±0.015	8.520±0.018
		9	7	10.79±0.05	1.44±0.36	0.303±0.022	8.604±0.014
2.09–2.61	Field	53	36	9.77±0.02	0.92±0.16	0.090±0.009	8.305±0.023
		36	27	10.22±0.02	1.27±0.16	0.180±0.011	8.476±0.015
		7	7	10.65±0.06	1.09±0.24	0.308±0.031	8.609±0.025
	Intermediate-density	42	36	9.76±0.02	0.91±0.14	0.084±0.008	8.288±0.022
		42	27	10.21±0.02	1.28±0.16	0.165±0.013	8.454±0.020
		9	7	10.66±0.02	1.38±0.16	0.219±0.020	8.524±0.023
	Overdense	45	36	9.77±0.02	0.99±0.13	0.072±0.011	8.249±0.036
		39	27	10.21±0.02	1.31±0.14	0.144±0.014	8.420±0.024
		21	7	10.66±0.03	1.51±0.39	0.209±0.038	8.512±0.024

a Total number of galaxies in the bins of environment and stellar mass which is sub-sampled 500 times.
b Number of galaxies in each sub-sample (shaded region in Figure 3.5).

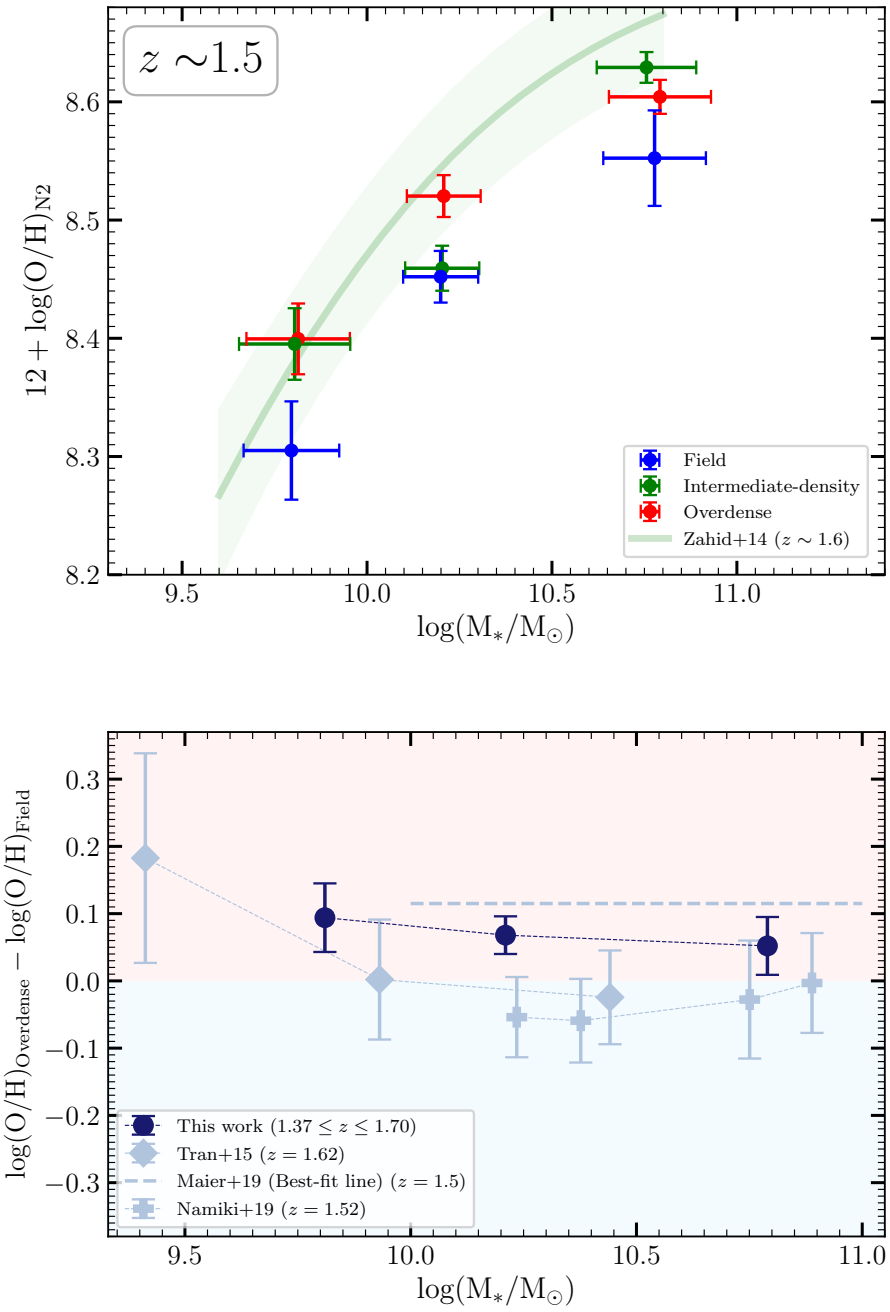


Figure 3.7: Top: MZR for mass-matched sample at $z \sim 1.5$ in three different environments: Field (blue), Intermediate-density (green) and Overdense (red). Three stellar mass bins are fixed at $9.5 \leq \log(M_*/M\odot) < 10$, $10 \leq \log(M_*/M\odot) < 10.5$ and $\log(M_*/M\odot) \geq 10.5$ and error bars in stellar masses show the 1σ scatter of the data around the average value. Bottom: The offset between the average metallicity of galaxies in overdensities (protocluster/cluster) and that of field galaxies as a function of stellar mass at $z \sim 1.5$. For comparison, previous studies in the literature at both redshifts are also included in the figure (see Section 3.5.1 for details).

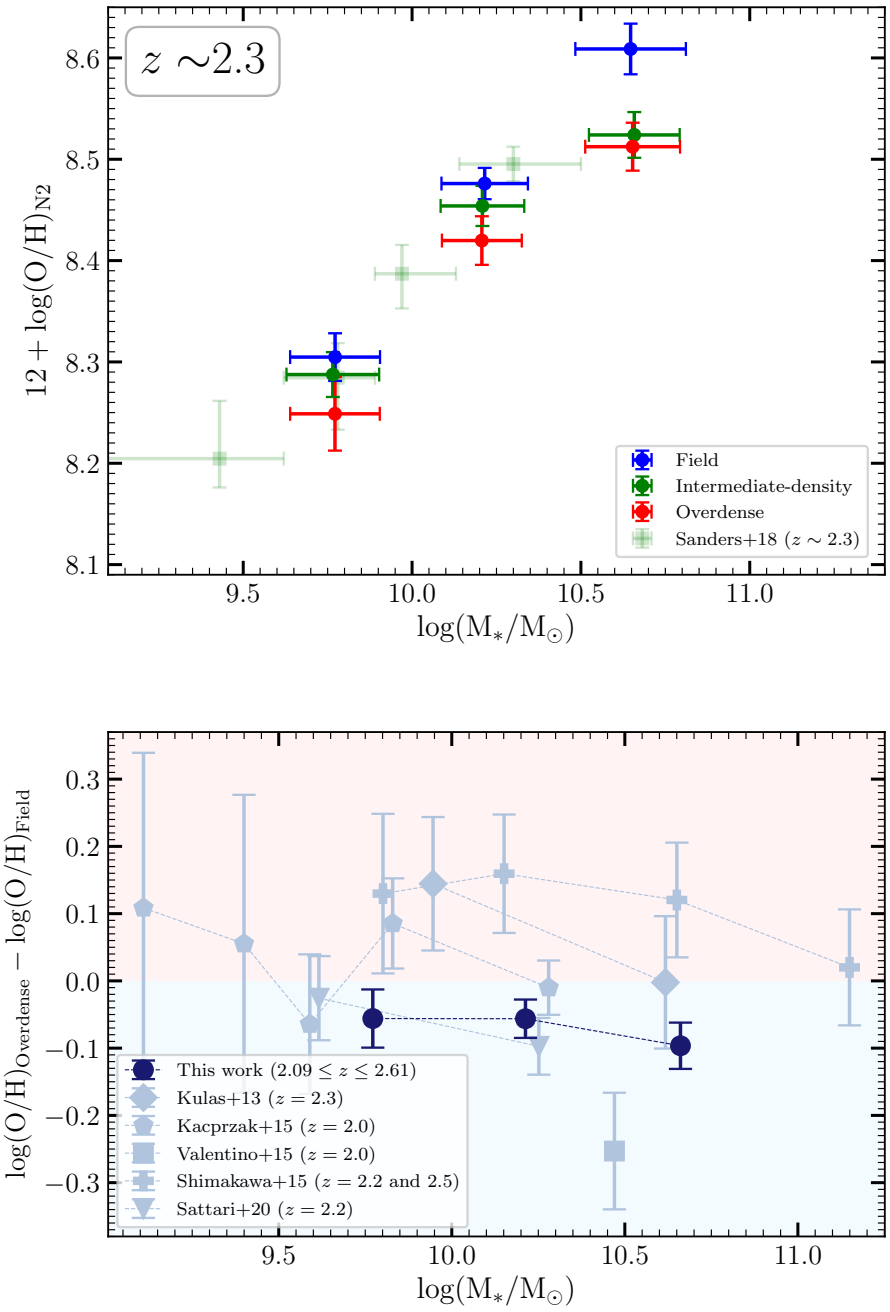


Figure 3.7: Continued: Top: MZR for mass-matched sample at $z \sim 2.3$ in three different environments: Field (blue), Intermediate-density (green) and Overdense (red). Three stellar mass bins are fixed at $9.5 \leq \log(M_*/M\odot) < 10$, $10 \leq \log(M_*/M\odot) < 10.5$ and $\log(M_*/M\odot) \geq 10.5$ and error bars in stellar masses show the 1σ scatter of the data around the average value. Bottom: The offset between the average metallicity of galaxies in overdensities (protocluster/cluster) and that of field galaxies as a function of stellar mass at $z \sim 2.3$. For comparison, previous studies in the literature at both redshifts are also included in the figure (see Section 3.5.1 for details).

3.5 Discussion

3.5.1 Comparison with previous works

Only a handful of studies have been conducted to investigate the role of the environment in the MZR at $z > 1$. The bottom panels in Figure 3.7 show the offset between the average metallicity of galaxies in overdensities (protocluster/cluster) and that of field galaxies as a function of stellar mass at both redshift bins, $z \sim 1.5$ and 2.3. In these figures, we also include the aforementioned offsets from literature. Our finding at $z \sim 1.5$ is in full agreement with the recent work of Maier et al. (2019). They studied a massive cluster at $z \sim 1.5$ and found that, at a given stellar mass, the metallicities of galaxies in the inner part of the cluster are higher by ~ 0.1 dex than those of infalling and field galaxies. They suggest that strangulation in the dense cores of clusters results in a cold gas removal that enhances the metallicity. However, Namiki et al. (2019) and Tran et al. (2015) found no significant environmental dependence of the MZR around $z \sim 1.5$. Namiki et al. (2019) compared the MZR of narrow-band Hα-selected cluster members with the field MZR of Stott et al. (2013). Tran et al. (2015) considered the field sample of Zahid et al. (2014) for the comparison. Thus, we also utilize the field samples of Stott et al. (2013) and Zahid et al. (2014) to measure the metallicity offset (shown in Figure 3.7) for the studies of Namiki et al. (2019) and Tran et al. (2015), respectively.

In the studies regarding the environmental dependence of the MZR, the sample selection bias needs to be handled properly to ensure that both field and cluster samples are selected in the same way. For example, Stott et al. (2013) found that Hα emitting galaxies selected from High-Z Emission Line Survey (HiZELS) (Sobral et al., 2013) have remarkably higher metallicity at a given stellar mass compared to a rest-frame UV-

selected sample of Erb et al. (2006). They argued that the UV-selected sample of Erb et al. (2006) tends to have a higher average SFR compared to the HiZELS narrow-band selected sample, resulting in a bias against metal-rich galaxies. Therefore, the comparison between the MZR of two different works can be biased due to the selection criteria. However, this is not an issue when the sample is selected uniformly in different environments, and measurements are performed consistently. Shimakawa et al. (2015) studied narrow-band selected galaxies in two rich overdensities at $z = 2.2$ and 2.5. They found that the metallicity of protocluster galaxies ($M_* < 10^{11}$ $M_\odot$) is ~ 0.15 dex higher than that of field galaxies, which contrasts with our result at $z \sim 2.3$. According to Stott et al. (2013) and discussions included in Shimakawa et al. (2015), selection bias is a potential concern in their study since different criteria are used for the selection of protocluster (narrow-band selected) and field (UV-selected, Erb et al., 2006) galaxies. This concern can be addressed in future studies by comparing the protocluster MZR of Shimakawa et al. (2015) with a field MZR of a narrow-band selected Hα emitters at $z \sim 2.2$.

Kacprzak et al. (2015) and Alcorn et al. (2019) studied the MZR of a protocluster (Yuan et al., 2014) at $z = 2.1$ and reported no significant environmental effect on the MZR. As shown in Figure 3.1, we also captured this protocluster in the present work. We speculate that the lack of environmental dependence of the MZR in their studies can originate from considering all the protocluster members at the same density contrast. As seen in Figure 3.1, the protocluster includes different components and considering that all the members are located in an overdensity weakens any existing environmental dependence of the MZR.

Valentino et al. (2015) found that cluster star-forming galaxies ($10^{10} M_\odot \leq M_* \leq 10^{11} M_\odot$) are ~ 0.25 dex poorer in metals than their field counterparts at $z \sim 2$. We find a similar trend at $z \sim 2.3$ with a lower average metal deficiency of ~ 0.1 dex.

Moreover, Sattari et al. *in prep.* studied a massive protocluster at $z = 2.2$ (Darvish et al., 2020) and found that protocluster galaxies with $M_* \sim 10^{9.6}$ $M_\odot$ and $M_* \sim 10^{10.2}$ $M_\odot$ are 0.03 ± 0.06 and 0.10 ± 0.04 dex metal deficient, respectively, compared to field galaxies. Our findings at $z \sim 2.3$ are in agreement with their results.

In the presence of limited sample size, unbalanced stellar mass distribution of galaxies between cluster members and the field sample can affect the environmental trends. In other words, even in each stellar mass bin (e.g., $9.5 \leq \log(M_*/M\odot) < 10$), the stacked spectra are biased toward massive galaxies, which are usually detected with higher S/N. Therefore, one needs to match the shape of stellar mass distributions in different environments as described in Section 3.4.2, and having the same average/median stellar mass in the bins of the environment does not guarantee that the effect of stellar mass on gas-phase metallicity is removed properly. Kulas et al. (2013) have studied the MZR of a protocluster at $z \sim 2.3$ and reported a 0.1 dex metallicity enhancement with respect to the field galaxies. We do not confirm such trends in the present work. Their field and protocluster samples are selected consistently (UV-selected). However, based on Figure 2 of Kulas et al. (2013), we speculate that the metallicity enhancement seen in their work can be affected by the significantly different stellar mass distributions of their protocluster and field samples.

Beyond the local Universe, all the previous works regarding environmental dependence of gas-phase metallicity are conducted by comparing MZR for a cluster/protocluster and a sample of field galaxies. However, within a given cluster/protocluster, galaxies may have different density contrasts. Thus, quantifying the environment using local density is a better approach to study environmental effect rather than considering all the members of a cluster/protocluster residing in an overdensity. For instance, Maier et al. (2019) found that the metallicity enhancement at $z \sim 1.5$ is only observed for the core (inside half of R_{200}) of a cluster, not for infalling

protocluster members. Assuming that all the cluster/protocluster members have similar local densities weakens any underlying environmental dependence of the MZR. In the present work, for the first time, we study the gas-phase metallicity of galaxies as a function of their local density and its evolution with cosmic time beyond the local Universe. Similar works have been conducted by Mouhcine et al. (2007); Cooper et al. (2008); Peng and Maiolino (2014) with SDSS sample at $z \sim 0$.

3.5.2 How does the environment affect MZR?

In this work, we find that galaxies in overdense regions have lower metallicity than their field counterparts at $z \sim 2.3$, but they become more metal-rich as they evolve to $z \sim 1.5$. In other words, the gas-phase metallicity of galaxies in a dense environment increases by ~ 0.15 dex as they evolve from $z \sim 2.3$ to $z \sim 1.5$ (~ 1.5 Gyr), but the metallicity of field galaxies are almost unchanged over this period. It implies that, at high redshift, metal enrichment processes are affected by the environment where galaxies reside.

Previous studies observed that dense environments at the early stage of galaxy cluster formation ($z \gtrsim 2$) contain a significant fraction of pristine gas (e.g., Cucciati et al., 2014). In the absence of gravitational heating processes, gas accretion in overdensities should be more prominent due to their deep potential well. However, infalling gas in overdensities with massive halos gets shock-heated and needs to radiate its kinetic energy to accrete into the halo. At high redshifts, where the average density of the Universe is higher by a factor of $(1+z)^3$, gas cooling is very efficient (van de Voort and Schaye, 2012). It is also observed that overdense regions in the early Universe are not only overdense in galaxies but also contain a large fraction of dense gas (e.g., Hennawi et al., 2015). Denser gas can cool down faster, facilitating the accretion of the metal-poor primordial gas into galaxies (Kereš et al., 2005; Dekel and Birnboim, 2006).

As a result, the prominent accretion of cold metal-poor gas in overdensities dilutes the metal content of ISM in galaxies at high redshifts. This results in the metal-poor gas in the galaxies residing in dense environments at high redshifts, as observed in the present work at $z \sim 2.3$. Valentino et al. (2015) also observed the same metal deficiency at $z \sim 2$ and concluded that the accretion of pristine gas from cluster-scale reservoirs lowers the gas-phase metallicity of galaxies in dense environments compared to their coeval field galaxies.

In contrast, at the lower redshift, $z \lesssim 2$, the gas cannot cool down efficiently in overdensities, so the galaxies in those regions start to experience cosmological starvation. The lack of pristine metal-poor gas accretion in cluster members has been observed in different simulations out to $z \sim 2$ (van de Voort et al., 2017; Gupta et al., 2018). Moreover, Gupta et al. (2018) studied chemical pre-processing of cluster galaxies in the IllustrisTNG cosmological simulation and found that at $z = 1.5$, cluster galaxies receive ~ 0.05 dex more metal-rich infalling gas than galaxies in the field. But, this metallicity enhancement disappears at higher redshifts ($z > 1.5$). At $z \sim 2$, when galaxies actively form their stars, feedback processes should be strong enough to expel part of the processed gas into the IGM through the outflows. As a result, crowded regions contain pre-processed and metal-enriched gas, which can be then re-accreted to the galaxies at lower redshifts. As pre-processed gas has higher metallicity, it can cool down faster which facilitates its accretion (Kereš et al., 2005; Dekel and Birnboim, 2006). Therefore, we speculate that both effects, suppressed primordial metal-poor gas infall and pre-processed metal-enriched gas accretion, are essential in ramping up metal production in a dense environment around $z \lesssim 2$. This can explain our result at $z \sim 1.5$, where we find metallicity enhancement for galaxies in overdensities compared to field galaxies.

In the absence of cold gas accretion, galaxies could maintain their SFR unchanged

for a period of time as they start to consume their gas reservoirs. This time ranges from a few hundred Myr for most massive galaxies up to a few Gyr for low-mass galaxies (McGee et al., 2014; Balogh et al., 2016). However, the dilution of ISM's metal content will be ceased immediately after the termination of cold gas accretion. Therefore, the absence of SFR suppression in overdense regions at $z \sim 1.5$ for our star-forming sample does not contradict the observed metal enhancement in galaxies residing in overdensities at that redshift. Given the relatively short cosmic time interval between $z \sim 2.3$ and $z \sim 1.5$ (~ 1.5 Gyr), galaxies with stellar mass range probed in the present work ($M_* < 10^{11} M_\odot$) do not have enough time to consume their remaining gas reservoirs after the halt of cold gas accretion.

Sanders et al. (2018) showed that the MZR varies with SFR at $z \sim 2.3$, such that the gas-phase metallicity of galaxies at fixed M_* is anticorrelated with their SFRs. Using SED-derived SFRs, we also find slight evidence of enhanced SFR for star-forming galaxies located in overdensities. Therefore, the prominent gas accretion in overdense regions at high redshifts can explain both lower gas-phase metallicity and higher SFR of galaxies residing in overdensities at $z \sim 2.3$. It is worth noting that gas outflows can also play a significant role in lowering the gas-phase metallicity of galaxies located in overdensities and actively forming stars. These outflows can re-accrete into galaxies at lower redshifts and increase the gas-phase metallicity of galaxies as found in the present work at $z \sim 1.5$.

3.5.3 The MZR with [NII]λ6584-detection requirement

We perform linear regression for individual galaxies (non-stacked) to calculate the best-fit MZR for galaxies with significant detection (S/N > 3) in [NII]λ6584. We note that requiring detection in [NII]λ6584 introduces a bias to our sample toward higher gas-

phase metallicities, especially in the low-mass end of the MZR where non-detections are prevalent; however, it is worth investigating the offset between the best-fit MZRs in different environments without considering the contribution of [NII]λ6584 non-detections. We fit a linear model, $12 + \log(\mathrm{O/H}) = \mathrm{Z}_0 + \alpha[\log(\mathrm{M}_*/\mathrm{M}_\odot) - 10]$, to the MZR of the samples in two extreme environment bins (underdensity and overdensity) considering both measurement errors in stellar mass and gas-phase metallicity. Figure 3.8 shows the best-fit lines along with 2D-posterior distributions of the slope and the intercept for the bins of environment. A significant distinction ($> 2\sigma$) between the posterior of fit parameters for the field galaxies and those located in overdensities suggests that the metallicity is enhanced at $z \sim 1.5$ and suppressed at $z \sim 2.3$ for galaxies in overdensities compared to field counterparts. Based on the best-fit models and their corresponding uncertainties, on average, galaxies in overdensities at $z \sim 1.5$ have 0.050 ± 0.024 dex higher gas-phase metallicity compared to coeval field galaxies. The trend reverses at higher redshift, $z \sim 2.3$, such that galaxies residing in overdensities are metal deficient by 0.055 ± 0.025 dex than the field counterparts. These results are in general agreement with our findings using mass-matched stacked spectra (Figure 3.7). Moreover, we estimate the intrinsic scatter of the MZR (σ_{int}) in both extreme environments at $z \sim 1.5$ and $z \sim 2.3$. We assume that the observed scatter (σ_{obs}) around the best-fit MZR is $\sigma^2_{\mathrm{obs}} = \sigma^2_{\mathrm{int}} + \sigma^2_{\mathrm{meas}}$, where σ_{meas} is the average measurement uncertainty. We estimate that the intrinsic scatter of the MZR at $z \sim 1.5$ ($z \sim 2.3$) is 0.07 (0.07) and 0.07 (0.10) dex for the field galaxies and those residing in overdensities, respectively, which are consistent with the intrinsic scatter of $z \sim 0$ MZR in different environments (Cooper et al., 2008). We note that, although ignoring [NII]λ6584 non-detections does not change our conclusions, the bias is evident in the low-mass end of the MZR by comparing best-fit lines in Figure 3.8 with gas-phase metallicities derived from mass-matched stacked spectra (Figure 3.7). Therefore, the stacking technique

employed in Section 3.4 is the preferred method as it provides an unbiased MZR where the contribution of [NII]λ6584 non-detections are taken into account properly.

3.6 Summary

Using a large near-IR spectroscopic sample drawn from the MOSDEF survey, combined with the local density measurements from the CANDELS photometric survey, we study the environmental dependence of the MZR at $z \sim 1.5$ and $z \sim 2.3$. We cross-match MOSDEF galaxies with the publicly available catalog of local density measurements in five CANDELS fields (Chartab et al., 2020), and use the N2 $= \frac{[\mathrm{NII}]\lambda6584}{\mathrm{H}\alpha}$ indicator to measure the gas-phase oxygen abundances of 167 galaxies at $1.37 \leq z \leq 1.7$ and 303 galaxies at $2.09 \leq z \leq 2.61$.

The samples are labeled as overdense, intermediate-density, and field based on their local density measurements. We match the stellar mass distribution of our sample in three different environments to properly disentangle the effects of stellar mass from those related to the environment. Massive galaxies are mostly found in overdensities and unmatched underlying stellar mass distributions between different environments can affect the strength of the trends or even change the observed trends in the presence of a limited sample size. For the mass-matched sample, our findings can be summarized as follows:

- At $z \sim 1.5$, the average metallicity of galaxies in overdensities with $\mathrm{M}_* \sim 10^{9.8}\mathrm{M}_\odot$, $10^{10.2}M_\odot$ and $10^{10.8}\mathrm{M}_\odot$ is higher relative to their field counterparts by 0.094 ± 0.051 (1.8σ significance), 0.068 ± 0.028 (2.4σ significance) and 0.052 ± 0.043 (1.2σ significance) dex, respectively. Also, the metallicity enhancements for $\mathrm{M}_* \sim 10^{9.8}\mathrm{M}_\odot$ and $10^{10.8}\mathrm{M}_\odot$ galaxies in intermediate-densities are 0.090 ± 0.052 (1.7σ

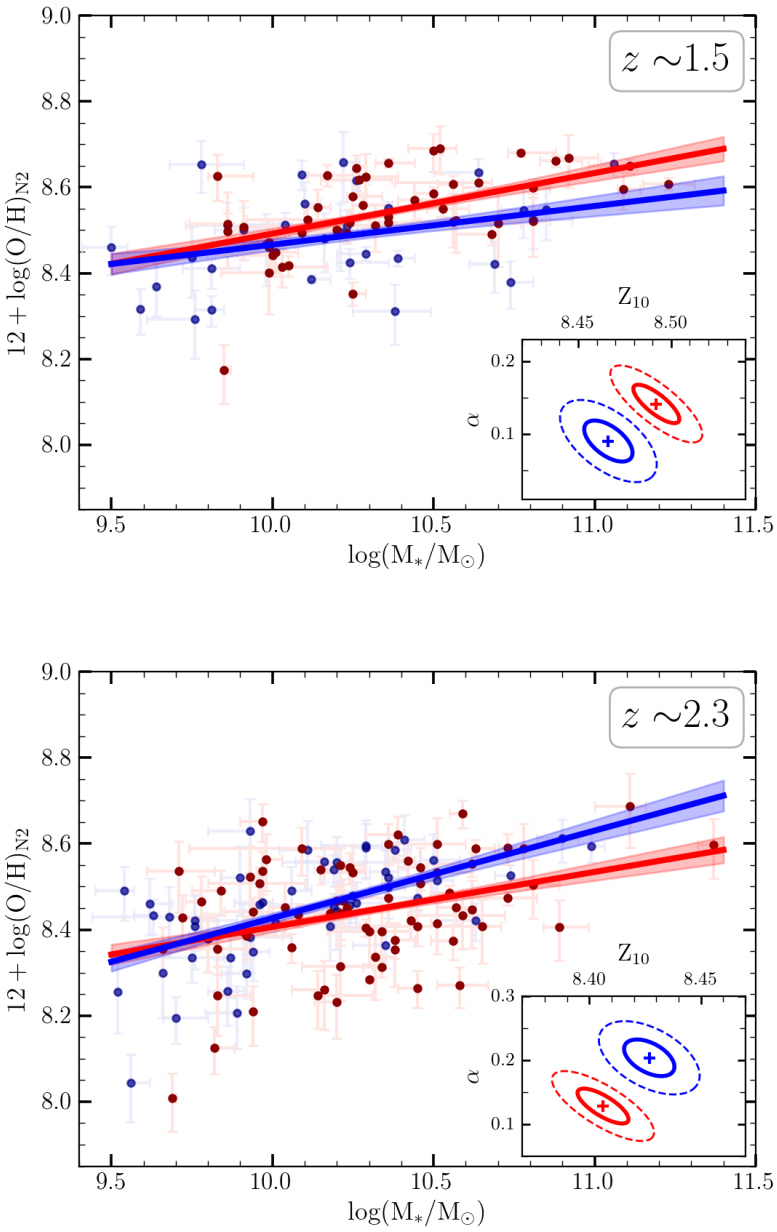


Figure 3.8: Similar to Figure 3.4, but here we ignore [NII]λ6584 non-detection galaxies. The best-fit lines to MZRs are shown for galaxies with S/N $>$ 3 detection in [NII]λ6584, residing in two extreme environments, overdensity (red) and underdensity (blue). The shaded regions around the best-fit models show 1σ error in the best-fit lines. Both gas-phase metallicity and stellar-mass errors are taken into account in regression analysis. Sub-panels show the best value ("+"), 1σ (solid) and 2σ (dashed) confidence intervals for the two-dimensional posterior distribution of the slope and the intercept, which are considered to be free parameters of the linear model, $12 + \log(O/H) = Z_0 + \alpha[\log(M_*/M_\odot) - 10]$.

significance) and 0.077 ± 0.042 (1.8σ significance) dex, respectively, with being insignificant for $M_* \sim 10^{10.2} M_\odot$ galaxies.

- At $z \sim 2.3$, galaxies that reside in overdensities with $M_* \sim 10^{9.8} M_\odot, 10^{10.2} M_\odot$ and $10^{10.7} M_\odot$, have lower gas-phase metallicity by 0.056 ± 0.043 (1.3σ significance), 0.056 ± 0.028 (2σ significance) and 0.096 ± 0.034 (2.8σ significance) dex compared to their coeval field sample, respectively. This metal deficiency is insignificant for galaxies residing in intermediate-densities except for the massive galaxies ($M_* \sim 10^{10.7} M_\odot$), where we found 0.085 ± 0.034 (2.5σ significance) dex metal deficiency compared to field counterparts.

- Our results suggest that the efficient gas cooling mechanisms at high redshifts result in the prominent accretion of primordial metal-poor gas into the galaxies in overdensities. This cold metal-poor gas can dilute the metal content of ISM gas and lowers the gas-phase metallicity of galaxies, as seen in the present work ($z \sim 2.3$). However, as galaxies evolve to the lower redshifts ($z \lesssim 2$), the shock-heated gas in overdensities with massive halos cannot cool down efficiently, which prevents it from accreting into the galaxy. The termination of pristine gas accretion in overdensities along with the accretion of pre-processed gas due to the strong outflows increase the metallicity of galaxies at lower redshifts, $z < 2$. This scenario can explain our result at $z \sim 1.5$, where we find metallicity enhancement for galaxies in overdensities compared to coeval field galaxies.

Chapter 4

A machine learning approach to predict missing wave-band photometry

Abstract

Using machine learning we can transfer the knowledge gained in fields with wealth of observations to those which lack such extensive observations. As a proof of concept, using the recently updated COSMOS catalog (COSMOS2020), we train a model to predict Euclid-like observations in near-IR Y-, J-, H-bands from the Hawaii Two-0 (H20) data. H20 is an ongoing survey covering ugriz and Spitzer/IRAC observations over Euclid deep fields. We find that magnitudes in Y-, J-, H-bands can be predicted with remarkable accuracy (1σ mag scatter $\lesssim$0.1 for galaxies brighter than 24 AB mag) from H20-like observations. Our results suggest that in the presence of a limited number of bands, a machine learning model trained over the population of galaxies with extensive spectral coverage outperforms template-fitting. Such machine learning model maximally comprises the information acquired over previous large surveys and breaks degeneracies in the parameter space of template-fitting inevitable in the presence of a few bands. We show that including near-IR observations in the training of a random

forest model to predict either the photometric redshifts or stellar masses based on H20 data does not improve model performance significantly, meaning that near-IR bands are predictable and H20-like bands capture most of the information needed to measure photometric redshifts and stellar masses. One should note that our results and models are limited to $z < 5.5$ due to the redshift range of the training sample.

4.1 Introduction

Future space telescopes such as Euclid and The Roman Space Telescope (formerly called WFIRST) will provide broad-band imaging for millions of galaxies, pivotal to better understanding of dark sectors of the Universe (dark matter and dark energy) as well as co-evolution of galaxies and large-scale structures over cosmic time. The challenge, however, is to obtain wide waveband coverage to constrain spectral energy distribution (SED) of millions of galaxies and measure their photometric redshifts and physical parameters such as stellar mass and star formation rate. For instance, Euclid will rely on ground-based optical photometry to measure photometric redshifts (Euclid Collaboration et al., 2020). Recently, a new ground-based survey, Hawaii Two-0 (H20; McPartland et al. in prep), has been designed to provide complementary photometric data for future space missions. The H20 will provide u-band observations from MegaCam instrument in Canada-France-Hawaii telescope (CFHT) and g-,r-,i-, z-band imaging from Hyper Suprime-Cam instrument on the Subaru telescope over 20 square degrees of the Euclid deep fields. Spitzer/IRAC observations from the Spitzer Legacy Survey (SLS; Capak et al., 2016) are also available in the same fields. The question then is how to extract maximum information from the acquired data? Can we use the existing data to simulate future surveys to optimize their design?

Template fitting is widely used to infer photometric redshifts of galaxies and their physical properties (e.g., Arnouts et al., 1999; Bolzonella et al., 2000; Ilbert et al., 2006). However, theoretical synthetic templates may not be representative of real parameter space for galaxies. For example, templates can include spectral energy distributions which do not have an observational analog. This will cause degeneracy in parameter measurement, especially when we reconstruct SEDs with few bands. Many of these degeneracies break by obtaining data with wide spectral coverage (i.e. a larger number of wavebands). An example of such data set is the COSMOS Evolution Survey (COSMOS; Scoville et al., 2007) that has been observed in more than 40 bands from X-ray to far-infrared wavelengths. The wealth of information in this field provides very well constrained SEDs for galaxies. However, not all surveys have as many photometric bands as the COSMOS field. Using machine learning techniques, we can transfer this information gained in the COSMOS field to the fields such as Euclid deep fields, where such numerous wave-bands photometry do not exist. Using the existing information, if we could develop techniques to predict fluxes in the wavebands that are not available in galaxy surveys, we could 1) further constrain their SEDs and hence, improve the accuracy of physical parameter measurements; (2) carefully design future surveys; (3) significantly save in the observing time; (4) identify and only observe in selected wavebands that are essential.

Machine learning has become popular in recent years to build models based on spectroscopic redshifts (Carrasco Kind and Brunner, 2014; Masters et al., 2017) and train models based on synthetic templates (Hemmati et al., 2019) or mock catalogs generated from galaxy simulations (Davidzon et al., 2019; Simet et al., 2021). These methods are particularly useful as machine learning algorithms can learn more complicated relations given that a sample of accurate and large training data is available (Mucesh et al., 2021). Moreover, these models speed up parameter measurement, which is crucial with

the advent of a flood of data from upcoming surveys (Hemmati et al., 2019).

In this paper, we develop a technique to predict fluxes in missing bands and hence, improve the wavelength resolution of existing photometric data. We evaluate the technique using the latest version of the COSMOS data. We predict fluxes in the Euclid wavebands from the ground-based observations of the H20 fields. We then test how the Euclid data will help in improving measurements of the photometric redshifts and stellar masses of galaxies, using COSMOS as a test sample.

In Section 4.2, we briefly introduce the updated COSMOS photometry catalog (COSMOS2020; J. Weaver et al. in prep.), and use that to build a sample of H20-like and H20+Euclid-like galaxies. In Section 4.3, we use dimensionality reduction techniques to visualize photometry of data in 2-dimensional space. This is followed by Section 4.4 where we train a machine learning algorithm, Random Forest model, to predict fluxes in Euclid-like wavebands using data in wavebands similar to the existing H20. In Section 4.5, we investigate the accuracy of the photometric redshifts and stellar masses given the limited number of bands available in H20-like and H20+Euclid data. We discuss and summarize our results in Section 4.6.

Throughout this work, we assume flat ΛCDM cosmology with $H_0 = 70\,\mathrm{kms^{-1}Mpc^{-1}}$, $\Omega_{m_0} = 0.3$ and $\Omega_{\Lambda_0} = 0.7$. All magnitudes are expressed in the AB system, and the physical parameters are measured assuming a Chabrier (2003) IMF.

4.2 Data

We use the updated version of the COSMOS catalog, COSMOS2020 (J. Weaver et al. in prep.). This provides much deeper near-infrared and mid-IR (Spitzer) photometric data as well as two independent methods for photometric extraction - the conventional and a profile-fitting (The Farmer; J. Weaver et al., in prep.) methods. The catalog

contains consistent photometric data in 39 bands from FUV to mid-infrared consisting of broad-, medium and narrow filters. All the data are reduced to the same scale with appropriate PSFs. The photometric redshifts are calculated using the LePhare (Arnouts et al., 1999; Ilbert et al., 2006) code with similar configuration described in Ilbert et al. (2013). Given the large number of bands with deep observations, photometric redshift solutions are accurate, reaching a normalized median absolute deviation ($\sigma_{\rm NMAD}$) of 0.03 for galaxies as faint as $i = 27$ AB mag (J. Weaver et al. in prep.). The redshifts of galaxies were then fixed on their estimated photometric redshifts and the stellar masses were estimated. In this paper, we consider COSMOS2020 photometric redshifts and stellar masses as a "ground truth" since spectroscopic redshifts are available for a limited number of galaxies and are biased towards a specific population of galaxies (mostly star-forming with strong emission lines).

We use two sets of wave-bands: 1) H20-like bands: $\mathbf{A} := \{\mathrm{u, g, r, i, z, ch1, ch2}\}$, 2) H20+Euclid-like bands: $\mathbf{B} := \{\mathrm{u, g, r, i, z, Y, J, H, ch1, ch2}\}$. u-band observations are conducted by MegaCam instrument at CFHT, and other optical bands (g,r,i and z) are available from Subaru's Hyper Suprime-Cam (HSC) imaging. Spitzer/IRAC channel 1,2 (ch1,ch2) data are compiled from all the IRAC observations of the COSMOS field (Moneti et al., in prep.). Near-IR photometry in Y, J and H bands are obtained from the UltraVista survey (McCracken et al., 2012). We select a subset of the COSMOS2020 galaxies that are observed in all the aforementioned bands and have i-band magnitude ≤ 25. We further exclude galaxies with the best-fit $\chi^2 > 2$ to have a clean sample of galaxies with well-constrained photometric redshifts and physical parameters. These selection criteria result in 121531 galaxies out to $z \sim 5.5$. Some sources have negative fluxes in the desired bands, which is likely due to the variation of background across the image. We set these fluxes to zero.

4.3 Data Visualization

Photometric redshifts and physical parameter measurements for galaxies in a photometric survey are solely based on their fluxes in N wave-bands. For instance, the H20-like data with N=7 bands occupy a 7-dimensional space, where the position of each galaxy is determined by its fluxes in 7 bands. Therefore, galaxies with similar positions in N-dimensional space are expected to have similar redshifts and physical parameters if we believe that N is large enough to fully sample the observed SED of galaxies. They even will have similar fluxes in other wave-bands where observations are not available. However, showing galaxy fluxes in a high-dimensional space (e.g., 7-dimensional space) is impossible and thus, we use dimensionality reduction techniques to present them in 2D space such that the information of higher dimension is maximally preserved. In this work, we use Uniform Manifold Approximation and Projection (UMAP; McInnes et al., 2018) technique to visualize our sample in a 2-dimensional space. UMAP is a non-linear dimensionality reduction technique that estimates the topology of the high-dimensional data and uses this information to construct a low-dimensional representation of data that preserves structure information on local scales. It also outperforms other dimensional reduction algorithms such as t-SNE (t-Distributed Stochastic Neighbor Embedding; van der Maaten and Hinton, 2008) since it preserves structures on global scales as well. In a simple sense, UMAP constructs a high-dimensional weighted graph by extending a radius around each data point and connecting points when their radii overlap. This radius varies locally based on the distance to the n^{th} nearest neighbor of each point. The number of the nearest neighbor (n) is the hyper-parameter in UMAP that should be fixed to construct high-dimensional graph. Small (large) value for n will preserve more local (global) structures. Once the high-dimensional weighted graph is constructed, UMAP optimizes the layout of a low-dimensional map to be as

similar as possible to the high-dimensional graph.

We use the UMAP Python library [1] to map 7-dimensional flux space of H20-like data to 2 dimensions considering nearest neighbor value of 50 to provide a balance between preserving local and global structures. We do not map magnitudes or colors since non-detected values cannot be handled properly when using them. Multi-waveband fluxes contain all the information regarding colors, but using colors misses information regarding fluxes or magnitudes. Therefore, mapping fluxes of galaxies from that space to 2-dimension is a better way than using colors. Since fluxes in different bands have fairly similar distributions, no normalization is needed before applying UMAP. In the case of significantly distinct distributions, normalization is needed to avoid the dominance of a wave-band with larger dynamic range. Figure 4.1 shows a 2-D visualization of the sample with H20-like bands using the UMAP algorithm. As an example, the mapped data are color-coded by the H-band fluxes (not present in H20 photometry) in micro Jansky. The smooth transition of the H-band fluxes in the 2D representation in Figure 4.1 reassures us that galaxies with similar fluxes in H20-like bands also have similar H-band fluxes. We note that the H20-like data set does not include H-band data (that is going to be provided by Euclid).

Visualized data in Figure 4.1 show qualitatively that the H-band fluxes are predictable using H20-like data. To perform a quantitative assessment on how accurately one can predict fluxes in Euclid-like JHK bands given the H20-like observations, we train a Random Forest (RF; Breiman, 2001) model with half of our sample and evaluate the model's performance with the other half. A random forest consists of an ensemble of regression trees. The algorithm picks a subsample of the dataset, builds a regression tree based on the subsample and repeats this procedure numerous times. The final value is the average of all the values predicted by all the trees in the forest. Having numerous

[1] https://github.com/lmcinnes/umap

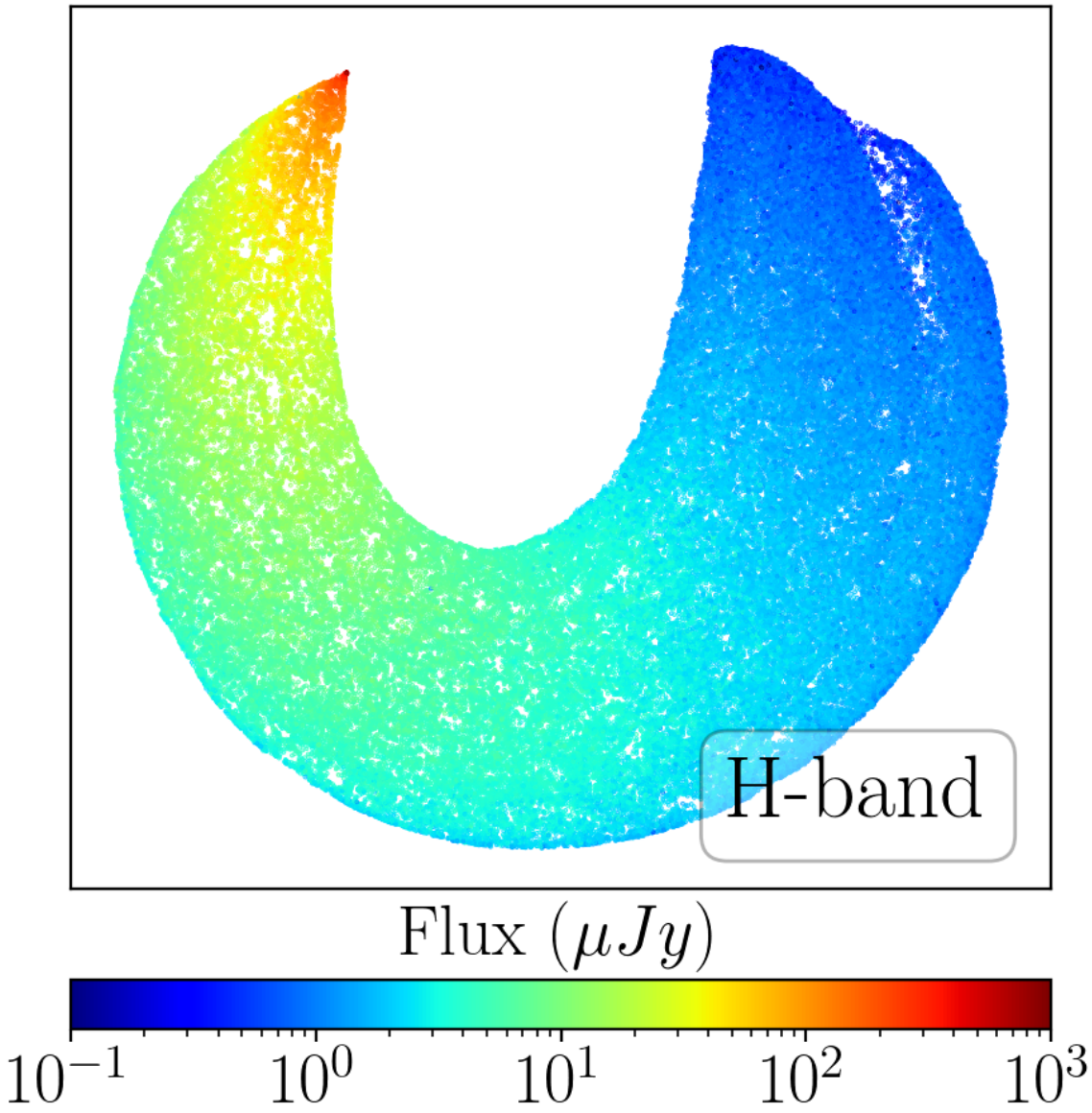


Figure 4.1: 2-D visualization of the sample with H20-like bands using the UMAP technique. The mapped data are color-coded by the H-band fluxes. Smooth gradient of H-band fluxes in the 2-D representation reassures us that galaxies with similar fluxes in H20-like bands have similar H-band fluxes as well.

decision trees based on subsampled data makes this algorithm unbiased and unaffected by overfitting. Another advantage of this method is that the inputs do not need to be scaled before feeding into the model. In the following section, we train a random forest model and evaluate its accuracy.

4.4 Flux predictions

We split the sample into a training and a test sample. With 60,766 galaxies as a training sample, we build a Random Forest model with 100 regression trees to predict Euclid-like YJH bands from the H20-like band fluxes. We use Python implementation of the algorithm (Scikit-learn; Pedregosa et al., 2011) [2] with its default parameters to build the model. The true (observed) fluxes in the YJH bands are available in the COSMOS2020 catalog. Using the trained Random Forest model, we then predict the expected fluxes for galaxies not included in the training set, with the results compared in Figure 4.2. For each band, we compare the predicted magnitudes ($\mathrm{Mag_{Predicted}}$) with the true observed magnitudes ($\mathrm{Mag_{True}}$). We find that the Random Forest model can predict unbiased YJH fluxes with high accuracy. The bottom panel in each figure shows the scatter of the $\mathrm{Mag_{Predicted}} - \mathrm{Mag_{True}}$ as a function of true magnitudes. With median magnitude discrepancy (Δ) of ~ 0.004, we find that the offset is comparable with discrepancies that arise from different methods of photometric data reduction. Weaver et al. (in preparation) found that the median tension between the magnitudes derived from aperture photometry and profile-fitting extraction is $\Delta \sim 0.03$ in YJ bands and $\Delta \sim 0.02$ in H-band. Thus, such small offsets in the random forest regressor are within the intrinsic uncertainties of the data reduction techniques. Green solid and dashed lines in the sub-panels of Figure 4.2 show the median of Δ and its 68% confidence intervals, re-

[2] https://scikit-learn.org/stable

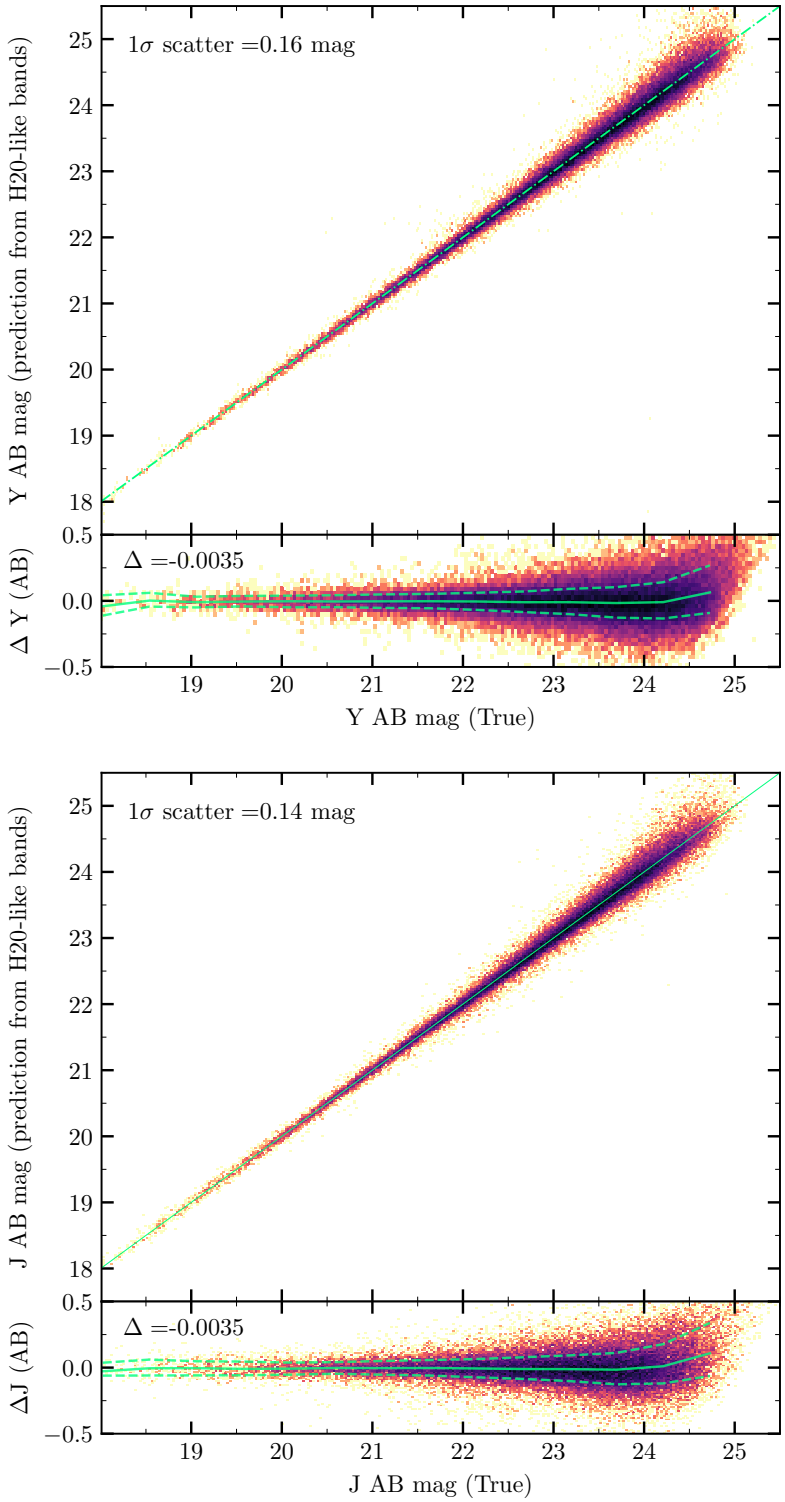


Figure 4.2: The performance of the Random Forest model on the 60,766 test galaxies not used for the training of the model. The model is trained based on H20-like bands (u,g,r,i,z,ch1,ch2) and predicts Euclid-like YJH bands. Bottom panels show the scatter of $\mathrm{Mag_{Predicted}} - \mathrm{Mag_{True}}$ as a function of true magnitudes and Δ is the median offset in these scatter plots.

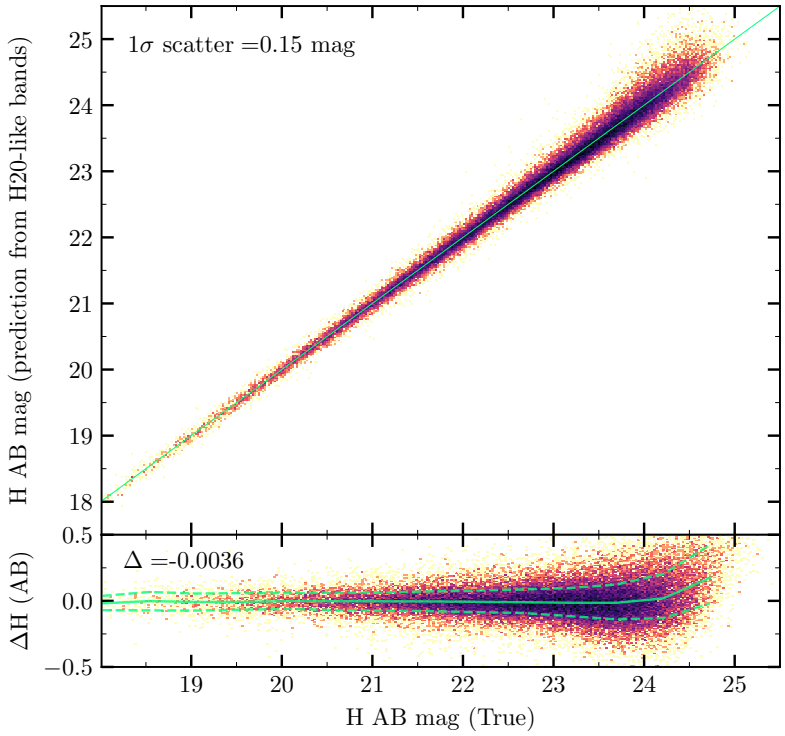


Figure 4.2: Continued

spectively. The scatter in the prediction is < 0.15 mag for galaxies brighter than 24 AB mag. This shows that near-IR observations of Euclid can be accurately predicted from the available observations of H20. While our focus in this paper is on the Euclid and H20 bands, the method we present is general and directly applicable to other surveys.

Furthermore, we repeat our analysis by excluding Spitzer/IRAC bands as the Euclid wide survey covers areas that are not fully covered by the Spitzer/IRAC. We find that, solely based on optical bands (ugriz), near-IR fluxes in Y-, J- and H-band can be predicted with the median offset of $\Delta = 0.002, 0.009$ and 0.021, respectively. The scatter around the median trend is less than $0.15, 0.16$ and 0.2 in YJH bands respectively for galaxies as faint as 24 AB (the 5σ magnitude limit of the Euclid wide survey in YJH bands (Laureijs et al., 2011)). This implies that even without Spitzer/IRAC coverage, near-IR fluxes can be constrained using optical observations.

4.5 Photometric redshift and stellar mass

In the previous section we showed that, given the observations of the H20 survey, near-IR observations of Euclid are well constrained. In other words, observations of the COSMOS field provide valuable information regarding the distribution of galaxies in the flux space even if we do not observe galaxies as extensively as it is done in the COSMOS field in terms of the spectral coverage. When we use template fitting code with synthetic templates, we usually do not take into account this valuable information. There are two approaches to incorporate this information in the photometric redshifts or physical parameters measurements. First, add a prior to fluxes in the bands that are not observed in the survey. For instance, when we perform SED fitting using H20-like bands, we can add priors to the YJH bands based on a Random Forest model, which is trained over the population of galaxies from the COSMOS observations. Second, train a model based on SED-fitting results calculated with a large number of bands. In this case, when we feed our model with H20-like data, it will decide about the best value of a parameter based on both the existence of similar observations in the COSMOS field (information from galaxy populations) and the SED-fitting solution for that galaxy.

In this section, we employ the latter approach to train a model to predict the photometric redshifts and the stellar masses of galaxies based on H20-like and H20+Euclid-like bands. We train a random forest model based on a training sample of observed galaxies. The inputs of the model are H20-like fluxes and the output is either photometric redshift or stellar mass computed from SED fitting over 39 bands available in the COSMOS2020 catalog. We also train another similar model where the inputs are H20+Euclid-like bands. Figure 4.3 shows the performance of trained models on the test sample with 60,766 galaxies. We find that both models recover photometric redshifts and stellar masses with similar accuracy with being slightly accurate using H20+Euclid-

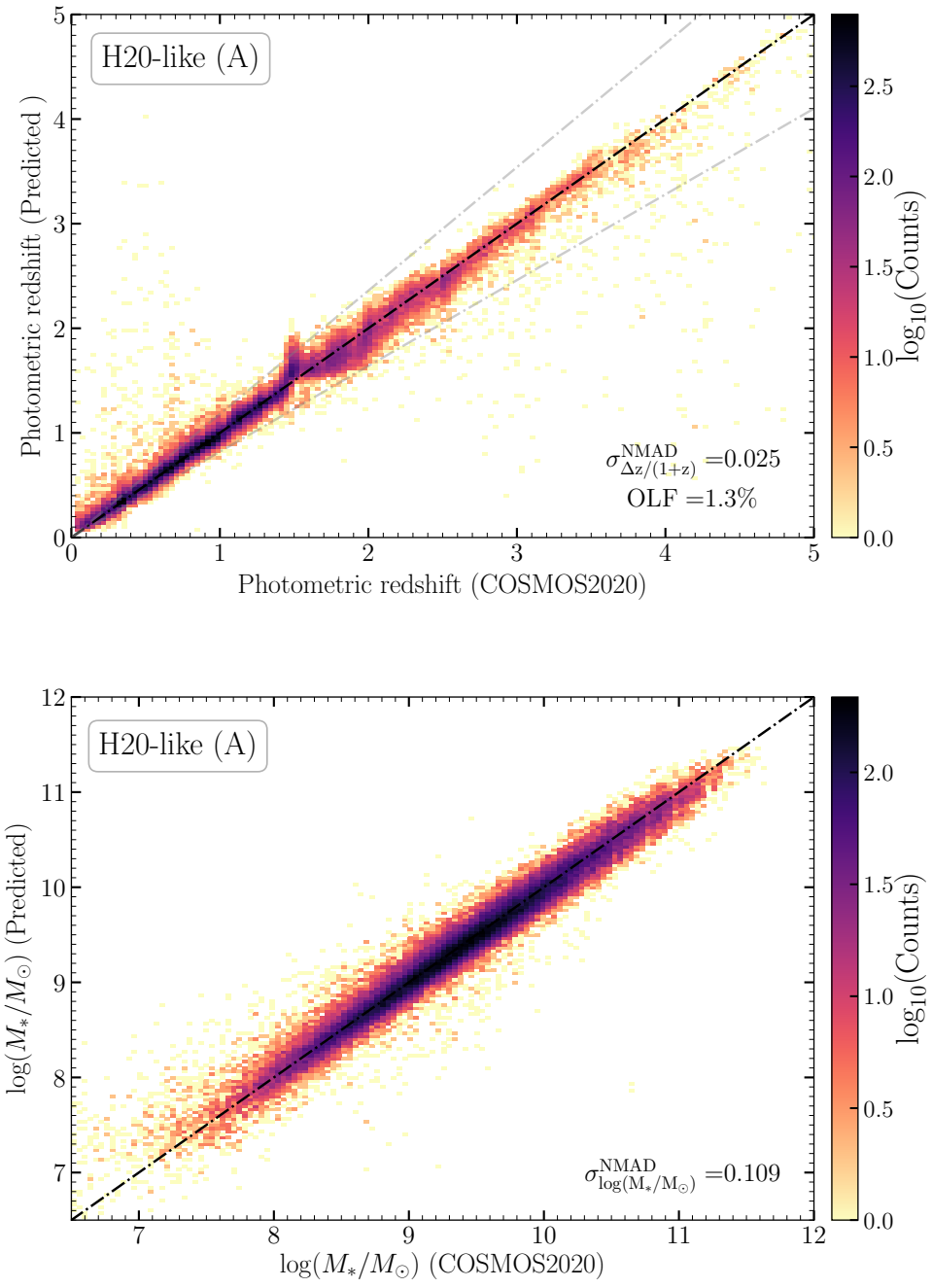


Figure 4.3: Performance of the random forest model to predict photometric redshifts and stellar masses when the model is trained by H20-like bands. The trained model recovers photometric redshifts and stellar masses with high accuracy. The similar performance of the model with and without YJH bands originates from the fact that the H20-like bands capture most of the information available in YJH bands as shown in Figure 4.2. The black dashed lines show one-to-one relation, and the gray dashed lines correspond to the predicted redshifts at $\pm 0.15(1+z)$ (outlier definition boundaries).

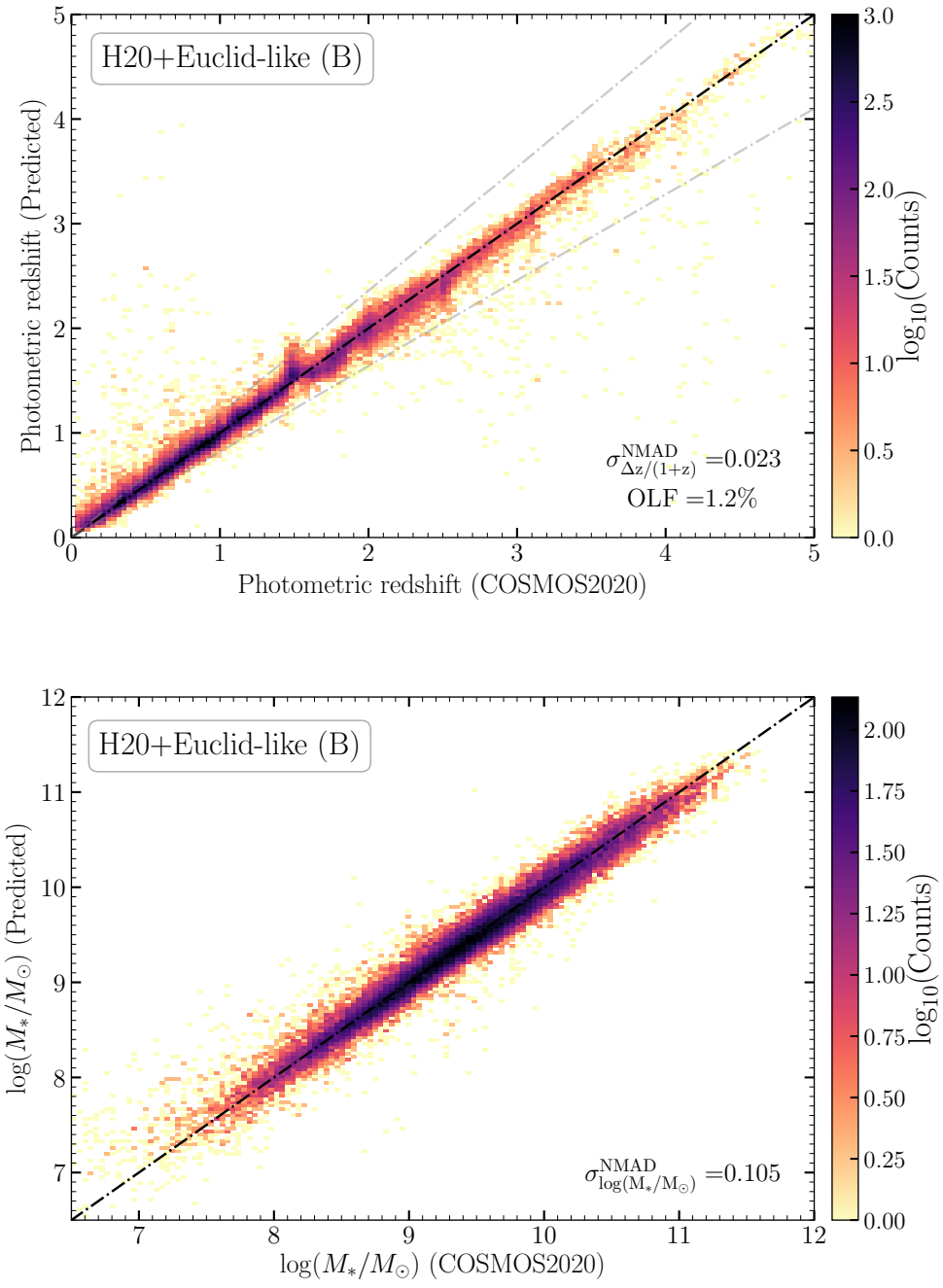


Figure 4.3: Performance of the random forest model to predict photometric redshifts and stellar masses when the model is trained by H20+Euclid-like bands (right panels). The trained model recovers photometric redshifts and stellar masses with high accuracy. The similar performance of the model with and without YJH bands originates from the fact that the H20-like bands capture most of the information available in YJH bands as shown in Figure 4.2. The black dashed lines show one-to-one relation, and the gray dashed lines correspond to the predicted redshifts at $\pm 0.15(1+z)$ (outlier definition boundaries).

like inputs. Normalized median absolute deviation ($\sigma_{\rm NMAD}$) of $\Delta z/(1+z)$ is ~ 0.02 for both models with $\sim 1\%$ outlier fraction. Outlier galaxies are defined as galaxies with $\Delta z/(1+z) > 0.15$. The median absolute deviation of $\log(M_*/M_\odot)$ is ~ 0.1 dex for both models. We explain this similar performance using the results of Section 4.4. The random forest model with H20-like bands comprises all the information regarding Euclid bands as we trained the model with the population of observed COSMOS galaxies. Therefore, it should recover photometric redshifts and stellar masses as accurately as the model which includes near-IR JHK observations. Slight under-performance of the model with H20-like bands can be explained by the existence of a scatter in the predictions of JHK bands fluxes (Figure 4.2).

4.5.1 Synthetic templates

In the following, we use UMAP to visualize photometry of synthetic SED models commonly used in template-fitting procedure. We build a set of theoretical templates using 2016 version of a library of Bruzual and Charlot (2003), considering Chabrier (2003) initial mass function. Star formation histories are modeled with an exponentially declining function (SFR $\propto e^{-t/\tau}$), where τ is the star formation timescale. Dust attenuation is applied using the Calzetti et al. (2000) law and solar stellar metallicity is assumed for all templates. We build $\sim 750,000$ theoretical templates assuming $\tau \in (0.1, 10)$ Gyr, $t \in (0.1, 13.7)$ Gyr, $A_v \in (0, 2)$ mag and $z \in (0, 5.5)$. t and A_v are the stellar age and the extinction in the visual band, respectively. We then calculate the syntetic photometry in both H20-like and H20+Euclid-like bands by applying the corresponding filter response.

As we learned the topology of fluxes in the H20-like bands for real observed galaxies in COSMOS2020 catalog (Figure 4.1), we can transform H20-like band fluxes of syn-

tetic photometry into the learned space. Figure 4.4 shows the 2-D visualization of the theoretical templates with H20-like bands in that learned space. As an example, data points in the reduced dimension are color-coded by their syntetic H-band fluxes in micro Jansky. Comparing theoretical templates with the observed data shown in Figure 4.1 reveals that model galaxies encounter degeneracies. In this specific example, we show that templates with similar H20-like fluxes have more diverse H-band fluxes than real observations, which can produce degenerate results when template fitting is performed based on H20-like bands. Adding information of the COSMOS2020 observations as a prior imposes a strong correlation between the observed and missing bands and makes the theoretical templates less degenerate as shown in Figure 4.1. For example, the dark blue arc in the left side of Figure 4.4 mismatches with the observational counterpart. In other words, synthetic templates predict H-band flux of $\sim 0.1\ \mu Jy$ for galaxies in that vicinity (i.e., the dark blue arc), but real observations show that they have in fact H-band flux of $\sim 10\ \mu Jy$. This shows that extra information that exists in the previous observations can add valuable information to template fitting analysis.

In the present work, we focus on predicting the Euclid bands from optical and/or IR observations and investigate how accurate other missing bands can be recovered; however, the method could be applied to many other bands. If one adds a predicted band in the template-fitting procedure, the errors should be assigned based on the 1σ scatter of the predicted flux (dashed green lines in Figure 4.2). It is particularly important to properly take into account the systematic scatter of the predicted bands in template-fitting and ensure that the predicted bands are not over-weighted in best-template selection. However, it is worth highlighting that the better approach would be using a machine learning model which is trained based on template-fitting results of a galaxy population with well-constrained SEDs such as COSMOS2020 (Figure 4.3).

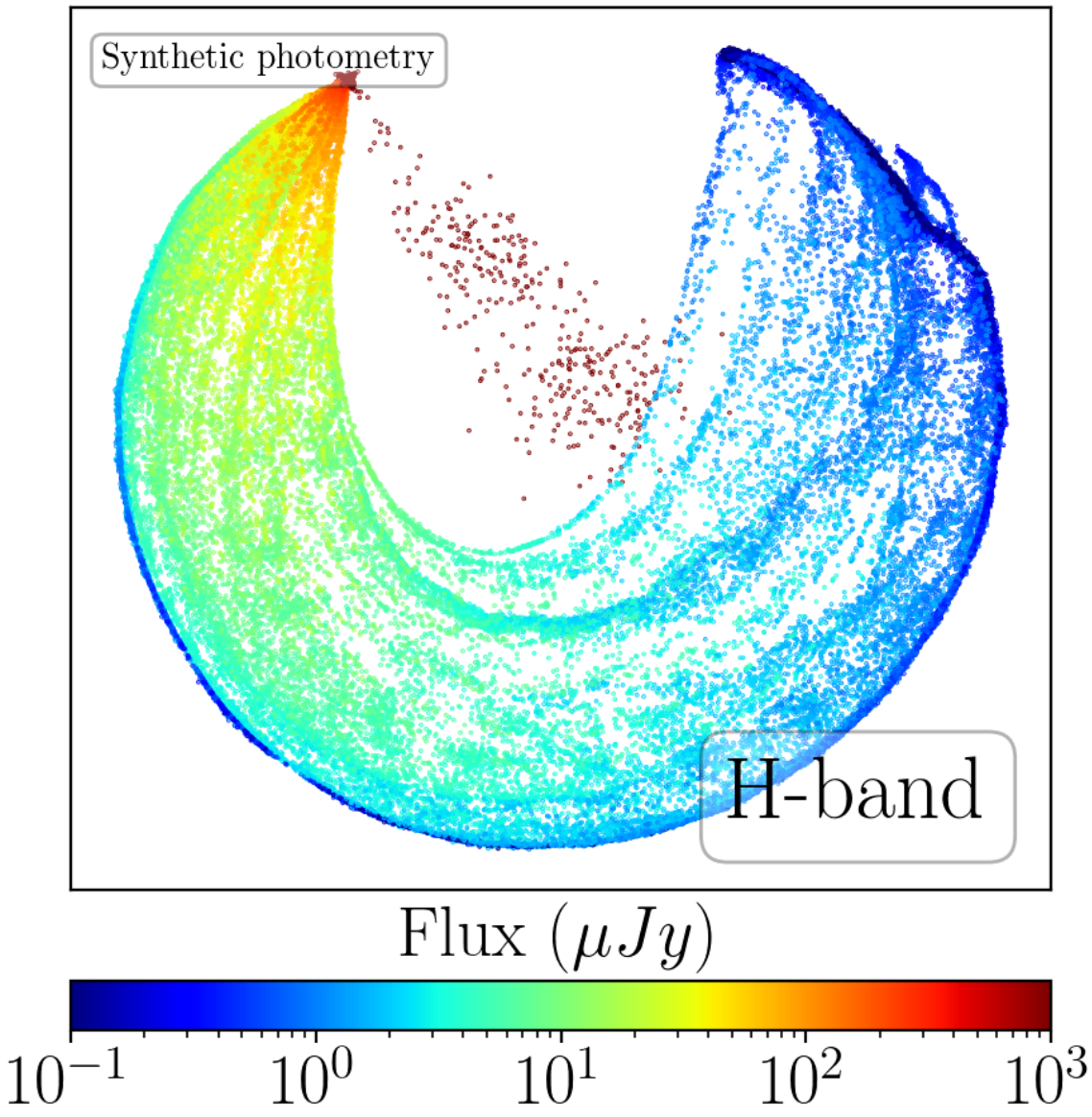


Figure 4.4: Similar to Figure 4.1, but for synthetic photometric data. The high-dimensional synthetic H20-like data are transformed to the space learned in Figure 4.1. The map is color-coded by the synthetic H-band fluxes. Existing dissimilarities between this figure and Figure 4.1 show that synthetic models lack the observed information.

4.6 Discussion and Summary

In this paper, we present a machine learning technique that can be used to predict missing fluxes from a survey. To prove the concept, we applied the method trained on the COSMOS data to predict Euclid-like near-IR observations based on H20-like survey data, which include ugriz and Spitzer/IRAC observations. We demonstrate that theoretical templates lack such valuable information already observed through numerous bands in the COSMOS field. We conclude that degeneracies in template-fitting can be alleviated if one trains a model based on template-fitting solutions for observed galaxies with extensive observations instead of using some arbitrary large parameter space. We show that a model trained on H20-like bands is as accurate as a model which is trained over H20+Euclid-like bands, given that the model is trained over the observed galaxy population with a vast number of wave-bands.

Masters et al. (2015) mapped high-dimensional color space of COSMOS galaxies in Euclid-like bands using the self-organizing map (SOM) technique (Kohonen, 1982) and proposed a spectroscopy survey to fully cover regions in reduced color space with no spectroscopic redshifts. This survey, C3R2, is awarded 44.5 nights on Keck telescope to map the color-redshift relation necessary for week lensing cosmology (Masters et al., 2017, 2019). Later on, Hemmati et al. (2019) used SOM to map the color space of theoretical models and used the reduced map as a fast template-fitting technique. In the present work, we use a new technique, UMAP, to find a 2-dimensional representation of a high-dimensional flux distribution. This technique is superior to other techniques as it preserves both local and global structures. Points far away from each other in high dimension remain far away in the low dimensional representation of data construed by UMAP. It is also significantly faster than other dimensionality-reduction techniques such as t-SNE. Therefore, UMAP is a powerful method that can be utilized instead of

the SOM technique, which is commonly used in the literature.

Acquiring data for galaxy surveys over wide areas and a range of wavelengths with a large number of wave-bands is costly. In this paper, we develop a new method based on machine learning algorithms to transfer information from previous extensive surveys (e.g., COSMOS) to complement the present and future surveys in their missing bands with no additional telescope time. The method generates comprehensive data similar to those expected from future surveys (e.g., Euclid) which can be used to simulate and optimize their observations. The technique can also enhance photometric catalogs that are already available in different fields (e.g., GOODS-South) where there are only a limited number of wavebands observed.

Chapter 5

Summary and Conclusions

This thesis is mainly focused on studying observed properties of galaxies as a function of their environment out to high redshifts. The number density maps of galaxies provide valuable information regarding the environment in which a galaxy resides. However, density map measurements can be challenging due to the lack of accurate and unbiased redshifts, especially at high redshift universe. This problem motivated us to develop a fully probabilistic method that uses the probability distribution of photometric redshifts to produce density maps.

5.1 Summary of Chapter 2

Using the photometric redshift distribution function, we estimated each galaxy's contribution to the density map at a given redshift slice. Therefore, each galaxy contributes to different redshift slices depending on the width of its respective redshift probability distribution. We employed weighted von-Mises kernel density estimation where the bandwidth of the kernel function is optimized by likelihood cross-validation. The von-Mises kernel is the spherical analog of the Gaussian kernel where variables are angles (e.g.,

right ascension and declination) instead of linear data. In Chapter 2, we introduced details of the method and we applied them to the CANDELS fields. Using photometric redshift probability distributions, we measured accurate density maps at different redshifts for all the CANDELS fields out to $z = 5$. We corrected density maps for systematic underestimation caused by the edge of the survey fields. We made a publicly available catalog of environment measurements for all the CANDELS fields (Chartab et al., 2020). We measured mass quenching efficiency (i.e., the fraction of massive quenched galaxies that would be star-forming if they were low mass in the same environment) and environmental quenching efficiency (i.e., the fraction of quenched galaxies in overdensities that would be star-forming if they were in the underdense region) out to $z \sim 3.5$ and found that both efficiencies depend on redshift and stellar mass. The environmental quenching efficiency decreases with increasing redshift and decreasing stellar mass. Besides the stellar mass quenching, which is the dominant quenching mechanism at high redshift ($z > 1$), the environmental quenching is efficient for massive galaxies at high redshifts (i.e., the quiescent fraction of galaxies with $M_* \sim 10^{11} M_\odot$ are 20% higher in overdensities than that of field counterparts at $z \sim 2.5$). We also found that the environmental quenching is more efficient than the stellar mass quenching for low mass galaxies ($M \lesssim 10^{10} M_\odot$) at low and intermediate redshifts, $z \lesssim 1.2$ (Chartab et al., 2020).

5.2 Summary of Chapter 3

One of the highly debated questions in extragalactic astronomy is how the environment of high-redshift galaxies affects their gas-phase metallicity. In the third chapter, we used the near-IR spectra of galaxies in the MOSFIRE Deep Evolution Field (MOSDEF) survey, combined with our density maps, to investigate this question. We used the $N2 = \frac{[NII]\lambda 6584}{H\alpha}$ indicator to measure the gas-phase oxygen abundances of 167 galaxies

at $1.37 \leq z \leq 1.7$ and 303 galaxies at $2.09 \leq z \leq 2.61$. Using our density estimates, we divided the sample into overdense, intermediate-density, and field based on their local density measurements. Furthermore, we matched the stellar mass distribution of the sample in three different environments to properly disentangle the effects of stellar mass from those related to the environment. We found that, at $1.37 \leq z \leq 2.61$, the variation of mass-metallicity relation with the local environment is small (< 0.1 dex), and its effect reverses at $z \sim 2$ (Chartab et al., 2021).

Our results in this chapter supported the hypothesis that, at the early stages of cluster formation, owing to efficient gas cooling, galaxies residing in overdensities host a higher fraction of pristine gas with prominent primordial gas accretion, which lowers their gas-phase metallicity compared to their coeval field galaxies. However, at lower redshifts ($z < 2$), the shock-heated gas in overdensities cannot cool down efficiently, and galaxies become metal-rich rapidly due to the suppression of pristine gas inflow and re-accretion of metal-enriched outflows in overdensities.

5.3 Summary of Chapter 4

Future observations will broaden our understanding by providing cosmologically significant volume needed for large-scale structure studies. Hawaii Two-0 (H20) survey is a wide (~ 20 deg^2) and deep survey with spectral coverage of 2–5 μm essential to estimate photometric redshifts and physical parameters of the high-redshift galaxies. The survey covers the North Ecliptic Pole (NEP) and Chandra Deep Field South (CDFS) Fields, which are selected as calibration fields for the Euclid mission. In chapter 4, we have developed a technique to predict fluxes in photometric bands that are absent from galaxy surveys, using machine learning algorithms trained on real data. For the training set, we used COSMOS2020, the latest COSMOS photometric catalog. For this purpose,

we trained a random forest regression model, which is powerful in learning non-linear relations as we have in the color space of galaxies. Also, we used Uniform Manifold Approximation and Projection (UMAP) to project high dimensional flux space of both real COSMOS2020 and synthetic template photometries commonly used in SED-fitting. We found that these predictions improve SED-derived photometric redshifts and physical parameters of H20 galaxies by imposing data-driven prior from COSMOS galaxies. Furthermore, this method can help optimize future missions (e.g., Euclid, Roman Space Telescope) as it provides simulated broadband photometry in missing bands based on the available observations of galaxies in NEP and CDFS fields.

References

Abramson, I. S.: 1982, *Ann. Statist.* **10(4)**, 1217

Alcorn, L. Y., Gupta, A., Tran, K.-V., Kacprzak, G. G., Yuan, T., Cohn, J., Forrest, B., Glazebrook, K., Harshan, A., Kewley, L. J., Labbé, I., Nanayakkara, T., Papovich, C., Spitler, L. R., and Straatman, C. M. S.: 2019, *ApJ* **883(2)**, 153

Alt, H. and Godau, M.: 1995, *Int. J. Comput. Geometry Appl.* **5**, 75

Ando, M., Shimasaku, K., and Momose, R.: 2020, *MNRAS* **496(3)**, 3169

Arnouts, S., Cristiani, S., Moscardini, L., Matarrese, S., Lucchin, F., Fontana, A., and Giallongo, E.: 1999, *MNRAS* **310**, 540

Azadi, M., Coil, A., Aird, J., Shivaei, I., Reddy, N., Shapley, A., Kriek, M., Freeman, W. R., Leung, G. C. K., Mobasher, B., Price, S. H., Sanders, R. L., Siana, B., and Zick, T.: 2018, *ApJ* **866(1)**, 63

Azadi, M., Coil, A. L., Aird, J., Reddy, N., Shapley, A., Freeman, W. R., Kriek, M., Leung, G. C. K., Mobasher, B., Price, S. H., Sanders, R. L., Shivaei, I., and Siana, B.: 2017, *ApJ* **835(1)**, 27

Balogh, M., Eke, V., Miller, C., Lewis, I., Bower, R., Couch, W., Nichol, R., Bland-Hawthorn, J., Baldry, I. K., Baugh, C., Bridges, T., Cannon, R., Cole, S., Colless, M., Collins, C., Cross, N., Dalton, G., de Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Gomez, P., Gray, A., Hawkins, E., Jackson, C., Lahav, O., Lumsden, S., Maddox, S., Madgwick, D., Norberg, P., Peacock, J. A., Percival, W., Peterson, B. A., Sutherland, W., and Taylor, K.: 2004, *MNRAS* **348**, 1355

Balogh, M. L., Gilbank, D. G., Muzzin, A., Rudnick, G., Cooper, M. C., Lidman, C., Biviano, A., Demarco, R., McGee, S. L., Nantais, J. B., Noble, A., Old, L., Wilson, G.,

Yee, H. K. C., Bellhouse, C., Cerulo, P., Chan, J., Pintos-Castro, I., Simpson, R., van der Burg, R. F. J., Zaritsky, D., Ziparo, F., Alonso, M. V., Bower, R. G., De Lucia, G., Finoguenov, A., Lambas, D. G., Muriel, H., Parker, L. C., Rettura, A., Valotto, C., and Wetzel, A.: 2017, *MNRAS* **470(4)**, 4168

Balogh, M. L., McGee, S. L., Mok, A., Muzzin, A., van der Burg, R. F. J., Bower, R. G., Finoguenov, A., Hoekstra, H., Lidman, C., Mulchaey, J. S., Noble, A., Parker, L. C., Tanaka, M., Wilman, D. J., Webb, T., Wilson, G., and Yee, H. K. C.: 2016, *MNRAS* **456(4)**, 4364

Barro, G., Pérez-González, P. G., Cava, A., Brammer, G., Pandya, V., Eliche Moral, C., Esquej, P., Domínguez-Sánchez, H., Alcalde Pampliega, B., Guo, Y., Koekemoer, A. M., Trump, J. R., Ashby, M. L. N., Cardiel, N., Castellano, M., Conselice, C. J., Dickinson, M. E., Dolch, T., Donley, J. L., Espino Briones, N., Faber, S. M., Fazio, G. G., Ferguson, H., Finkelstein, S., Fontana, A., Galametz, A., Gardner, J. P., Gawiser, E., Giavalisco, M., Grazian, A., Grogin, N. A., Hathi, N. P., Hemmati, S., Hernán-Caballero, A., Kocevski, D., Koo, D. C., Kodra, D., Lee, K.-S., Lin, L., Lucas, R. A., Mobasher, B., McGrath, E. J., Nandra, K., Nayyeri, H., Newman, J. A., Pforr, J., Peth, M., Rafelski, M., Rodríguez-Munoz, L., Salvato, M., Stefanon, M., van der Wel, A., Willner, S. P., Wiklind, T., and Wuyts, S.: 2019, *ApJl Supp.* **243(2)**, 22

Behroozi, P. S., Wechsler, R. H., and Conroy, C.: 2013, *ApJ* **770(1)**, 57

Birnboim, Y. and Dekel, A.: 2003, *MNRAS* **345(1)**, 349

Bolzonella, M., Miralles, J. M., and Pelló, R.: 2000, *Astronomy and Astrophysics* **363**, 476

Boselli, A. and Gavazzi, G.: 2014, *Astronomy and Astrophysics Reviews* **22**, 74

Bowman, A. W.: 1984, *Biometrika* **71(2)**, 353

Breiman, L.: 2001, in *Machine Learning*, pp 5–32

Brooks, A. M., Governato, F., Booth, C. M., Willman, B., Gardner, J. P., Wadsley, J., Stinson, G., and Quinn, T.: 2007, *ApJl* **655(1)**, L17

Bruzual, G. and Charlot, S.: 2003, *MNRAS* **344**, 1000

Calzetti, D., Armus, L., Bohlin, R. C., Kinney, A. L., Koornneef, J., and Storchi-Bergmann, T.: 2000, *ApJ* **533**, 682

Capak, P., Abraham, R. G., Ellis, R. S., Mobasher, B., Scoville, N., Sheth, K., and Koekemoer, A.: 2007, *ApJl Supp.* **172(1)**, 284

Capak, P., Arendt, R., Arnouts, S., Bartlett, J., Bouwens, R., Brinchman, J., Brodwin, M., Carollo, M., Castander, F., Charlot, S., Chary, R. R., Cohen, J., Cooray, A., Conselice, C., Coupon, J., Cuby, J. G., Culliandre, J., Davidzon, I., Dole, H., Dunlop, J., Eisenhardt, P., Ferrara, A., Gardner, J., Hasinger, G., Hildebrandt, H., Ho, S., Ilbert, O., Jouvel, S., Kashlinsky, A., LeFevre, O., LeFloc'h, E., Maraston, C., Masters, D., McCracken, H. J., Mei, S., Mellier, Y., Mitchell-Wynn, K., Moustakas, L., Nayyeri, H., Paltani, S., Rhodes, J., Salvato, M., Sanders, D., Scaramella, R., Scarlata, C., Scoville, N., Silverman, J., Speagle, J., Stanford, S., Stern, D., Teplitz, H., and Toft, S.: 2016, *The Euclid/WFIRST Spitzer Legacy Survey*, Spitzer Proposal

Carrasco Kind, M. and Brunner, R. J.: 2014, *MNRAS* **438(4)**, 3409

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

Chartab, N., Mobasher, B., Darvish, B., Finkelstein, S., Guo, Y., Kodra, D., Lee, K.-S., Newman, J. A., Pacifici, C., Papovich, C., Sattari, Z., Shahidi, A., Dickinson, M. E., Faber, S. M., Ferguson, H. C., Giavalisco, M., and Jafariyazani, M.: 2020, *ApJ* **890(1)**, 7

Chartab, N., Mobasher, B., Shapley, A. E., Shivaei, I., Sanders, R. L., Coil, A. L., Kriek, M., Reddy, N. A., Siana, B., Freeman, W. R., Azadi, M., Barro, G., Fetherolf, T., Leung, G., Price, S. H., and Zick, T.: 2021, *ApJ* **908(2)**, 120

Chiang, Y.-K., Overzier, R. A., Gebhardt, K., and Henriques, B.: 2017, *ApJl* **844(2)**, L23

Chisholm, J., Tremonti, C., and Leitherer, C.: 2018, *MNRAS* **481(2)**, 1690

Chiu, S.-T.: 1991, *Ann. Statist.* **19(4)**, 1883

Coil, A. L., Aird, J., Reddy, N., Shapley, A. E., Kriek, M., Siana, B., Mobasher, B., Freeman, W. R., Price, S. H., and Shivaei, I.: 2015, *ApJ* **801**, 35

Conroy, C., Gunn, J. E., and White, M.: 2009, *ApJ* **699**, 486

Contini, E., Gu, Q., Ge, X., Rhee, J., Yi, S. K., and Kang, X.: 2020, *ApJ* **889(2)**, 156

Cooper, M. C., Newman, J. A., Weiner, B. J., Yan, R., Willmer, C. N. A., Bundy, K., Coil, A. L., Conselice, C. J., Davis, M., Faber, S. M., Gerke, B. F., Guhathakurta, P., Koo, D. C., and Noeske, K. G.: 2008, *MNRAS* **383**, 1058

Crowl, H. H., Kenney, J. D. P., van Gorkom, J. H., and Vollmer, B.: 2005, *The Astronomical Journal* **130(1)**, 65

Cucciati, O., Zamorani, G., Lemaux, B. C., Bardelli, S., Cimatti, A., Le Fèvre, O., Cassata, P., Garilli, B., Le Brun, V., Maccagni, D., Pentericci, L., Tasca, L. A. M., Thomas, R., Vanzella, E., Zucca, E., Amorin, R., Capak, P., Cassarà, L. P., Castellano, M., Cuby, J. G., de la Torre, S., Durkalec, A., Fontana, A., Giavalisco, M., Grazian, A., Hathi, N. P., Ilbert, O., Moreau, C., Paltani, S., Ribeiro, B., Salvato, M., Schaerer, D., Scodeggio, M., Sommariva, V., Talia, M., Taniguchi, Y., Tresse, L., Vergani, D., Wang, P. W., Charlot, S., Contini, T., Fotopoulou, S., López-Sanjuan, C., Mellier, Y., and Scoville, N.: 2014, *Astronomy and Astrophysics* **570**, A16

Darvish, B., Mobasher, B., Sobral, D., Rettura, A., Scoville, N., Faisst, A., and Capak, P.: 2016, *ApJ* **825**, 113

Darvish, B., Mobasher, B., Sobral, D., Scoville, N., and Aragon-Calvo, M.: 2015, *ApJ* **805**, 121

Darvish, B., Scoville, N. Z., Martin, C., Sobral, D., Mobasher, B., Rettura, A., Matthee, J., Capak, P., Chartab, N., Hemmati, S., Masters, D., Nayyeri, H., O'Sullivan, D., Paulino-Afonso, A., Sattari, Z., Shahidi, A., Salvato, M., Lemaux, B. C., Fèvre, O. L., and Cucciati, O.: 2020, *ApJ* **892(1)**, 8

Davidzon, I., Cucciati, O., Bolzonella, M., De Lucia, G., Zamorani, G., Arnouts, S., Moutard, T., Ilbert, O., Garilli, B., and Scodeggio, M.: 2016, *Astronomy and Astrophysics* **586**, A23

Davidzon, I., Laigle, C., Capak, P. L., Ilbert, O., Masters, D. C., Hemmati, S., Apostolakos, N., Coupon, J., de la Torre, S., Devriendt, J., Dubois, Y., Kashino, D., Paltani, S., and Pichon, C.: 2019, *MNRAS* **489(4)**, 4817

Dekel, A. and Birnboim, Y.: 2006, *MNRAS* **368(1)**, 2

Donzelli, C. J. and Pastoriza, M. G.: 1997, *ApJl Supp.* **111(1)**, 181

Dressler, A.: 1980, *ApJ* **236**, 351

Elbaz, D., Daddi, E., Le Borgne, D., Dickinson, M., Alexander, D. M., Chary, R.-R., Starck, J.-L., Brandt, W. N., Kitzbichler, M., MacDonald, E., Nonino, M., Popesso, P., Stern, D., and Vanzella, E.: 2007, *Astronomy and Astrophysics* **468**, 33

Ellison, S. L., Simard, L., Cowan, N. B., Baldry, I. K., Patton, D. R., and McConnachie, A. W.: 2009, *MNRAS* **396(3)**, 1257

Erb, D. K., Shapley, A. E., Pettini, M., Steidel, C. C., Reddy, N. A., and Adelberger, K. L.: 2006, *ApJ* **644(2)**, 813

Euclid Collaboration, Desprez, G., Paltani, S., Coupon, J., Almosallam, I., Alvarez-Ayllon, A., Amaro, V., Brescia, M., Brodwin, M., Cavuoti, S., De Vicente-Albendea, J., Fotopoulou, S., Hatfield, P. W., Hartley, W. G., Ilbert, O., Jarvis, M. J., Longo, G., Rau, M. M., Saha, R., Speagle, J. S., Tramacere, A., Castellano, M., Dubath, F., Galametz, A., Kuemmel, M., Laigle, C., Merlin, E., Mohr, J. J., Pilo, S., Salvato, M., Andreon, S., Auricchio, N., Baccigalupi, C., Balaguera-Antolínez, A., Baldi, M., Bardelli, S., Bender, R., Biviano, A., Bodendorf, C., Bonino, D., Bozzo, E., Branchini, E., Brinchmann, J., Burigana, C., Cabanac, R., Camera, S., Capobianco, V., Cappi, A., Carbone, C., Carretero, J., Carvalho, C. S., Casas, R., Casas, S., Castander, F. J., Castignani, G., Cimatti, A., Cledassou, R., Colodro-Conde, C., Congedo, G., Conselice, C. J., Conversi, L., Copin, Y., Corcione, L., Courtois, H. M., Cuby, J. G., Da Silva, A., de la Torre, S., Degaudenzi, H., Di Ferdinando, D., Douspis, M., Duncan, C. A. J., Dupac, X., Ealet, A., Fabbian, G., Fabricius, M., Farrens, S., Ferreira, P. G., Finelli, F., Fosalba, P., Fourmanoit, N., Frailis, M., Franceschi, E., Fumana, M., Galeotta, S., Garilli, B., Gillard, W., Gillis, B., Giocoli, C., Gozaliasl, G., Graciá-Carpio, J., Grupp, F., Guzzo, L., Hailey, M., Haugan, S. V. H., Holmes, W., Hormuth, F., Humphrey, A., Jahnke, K., Keihanen, E., Kermiche, S., Kilbinger, M., Kirkpatrick, C. C., Kitching, T. D., Kohley, R., Kubik, B., Kunz, M., Kurki-Suonio, H., Ligori, S., Lilje, P. B., Lloro, I., Maino, D., Maiorano, E., Marggraf, O., Markovic, K., Martinet, N., Marulli, F., Massey, R., Maturi, M., Mauri, N., Maurogordato, S., Medinaceli, E., Mei, S., Meneghetti, M., Metcalf, R. B., Meylan, G., Moresco, M., Moscardini, L., Munari, E., Niemi, S., Padilla, C., Pasian, F., Patrizii, L., Pettorino, V., Pires, S., Polenta, G., Poncet, M., Popa, L., Potter, D., Pozzetti, L., Raison, F., Renzi, A., Rhodes, J., Riccio, G., Rossetti, E., Saglia, R., Sapone, D., Schneider, P., Scottez, V., Secroun, A., Serrano, S., Sirignano, C., Sirri,

G., Stanco, L., Stern, D., Sureau, F., Tallada Crespí, P., Tavagnacco, D., Taylor, A. N., Tenti, M., Tereno, I., Toledo-Moreo, R., Torradeflot, F., Valenziano, L., Valiviita, J., Vassallo, T., Viel, M., Wang, Y., Welikala, N., Whittaker, L., Zacchei, A., Zamorani, G., Zoubian, J., and Zucca, E.: 2020, *Astronomy and Astrophysics* **644**, A31

Feldmann, R. and Mayer, L.: 2015, *MNRAS* **446(2)**, 1939

Finkelstein, S. L., Hill, G. J., Gebhardt, K., Adams, J., Blanc, G. A., Papovich, C., Ciardullo, R., Drory, N., Gawiser, E., Gronwall, C., Schneider, D. P., and Tran, K.-V.: 2011, *ApJ* **729(2)**, 140

Foltz, R., Wilson, G., Muzzin, A., Cooper, M. C., Nantais, J., van der Burg, R. F. J., Cerulo, P., Chan, J., Fillingham, S. P., and Surace, J.: 2018, *ApJ* **866(2)**, 136

Förster Schreiber, N. M., Übler, H., Davies, R. L., Genzel, R., Wisnioski, E., Belli, S., Shimizu, T., Lutz, D., Fossati, M., Herrera-Camus, R., Mendel, J. T., Tacconi, L. J., Wilman, D., Beifiori, A., Brammer, G. B., Burkert, A., Carollo, C. M., Davies, R. I., Eisenhauer, F., Fabricius, M., Lilly, S. J., Momcheva, I., Naab, T., Nelson, E. J., Price, S. H., Renzini, A., Saglia, R., Sternberg, A., van Dokkum, P., and Wuyts, S.: 2019, *ApJ* **875(1)**, 21

Fossati, M., Wilman, D. J., Fontanot, F., De Lucia, G., Monaco, P., Hirschmann, M., Mendel, J. T., Beifiori, A., and Contini, E.: 2015, *MNRAS* **446(3)**, 2582

Fossati, M., Wilman, D. J., Mendel, J. T., Saglia, R. P., Galametz, A., Beifiori, A., Bender, R., Chan, J. C. C., Fabricius, M., Bandara, K., Brammer, G. B., Davies, R., Förster Schreiber, N. M., Genzel, R., Hartley, W., Kulkarni, S. K., Lang, P., Momcheva, I. G., Nelson, E. J., Skelton, R., Tacconi, L. J., Tadaki, K., Übler, H., van Dokkum, P. G., Wisnioski, E., Whitaker, K. E., Wuyts, E., and Wuyts, S.: 2017, *ApJ* **835**, 153

Galametz, A., Grazian, A., Fontana, A., Ferguson, H. C., Ashby, M. L. N., Barro, G., Castellano, M., Dahlen, T., Donley, J. L., Faber, S. M., Grogin, N., Guo, Y., Huang, K.-H., Kocevski, D. D., Koekemoer, A. M., Lee, K.-S., McGrath, E. J., Peth, M., Willner, S. P., Almaini, O., Cooper, M., Cooray, A., Conselice, C. J., Dickinson, M., Dunlop, J. S., Fazio, G. G., Foucaud, S., Gardner, J. P., Giavalisco, M., Hathi, N. P., Hartley, W. G., Koo, D. C., Lai, K., de Mello, D. F., McLure, R. J., Lucas, R. A., Paris, D., Pentericci, L., Santini, P., Simpson, C., Sommariva, V., Targett, T., Weiner, B. J., Wuyts, S., and the CANDELS Team: 2013, *ApJl Supp.* **206**, 10

García-Portugués, E., Crujeiras, R. M., and González-Manteiga, W.: 2013, *J. Multivar. Anal.* **121**, 152

Garnett, D. R.: 2002, *ApJ* **581(2)**, 1019

Grogin, N. A., Kocevski, D. D., Faber, S. M., Ferguson, H. C., Koekemoer, A. M., Riess, A. G., Acquaviva, V., Alexander, D. M., Almaini, O., Ashby, M. L. N., Barden, M., Bell, E. F., Bournaud, F., Brown, T. M., Caputi, K. I., Casertano, S., Cassata, P., Castellano, M., Challis, P., Chary, R.-R., Cheung, E., Cirasuolo, M., Conselice, C. J., Roshan Cooray, A., Croton, D. J., Daddi, E., Dahlen, T., Davé, R., de Mello, D. F., Dekel, A., Dickinson, M., Dolch, T., Donley, J. L., Dunlop, J. S., Dutton, A. A., Elbaz, D., Fazio, G. G., Filippenko, A. V., Finkelstein, S. L., Fontana, A., Gardner, J. P., Garnavich, P. M., Gawiser, E., Giavalisco, M., Grazian, A., Guo, Y., Hathi, N. P., Häussler, B., Hopkins, P. F., Huang, J.-S., Huang, K.-H., Jha, S. W., Kartaltepe, J. S., Kirshner, R. P., Koo, D. C., Lai, K., Lee, K.-S., Li, W., Lotz, J. M., Lucas, R. A., Madau, P., McCarthy, P. J., McGrath, E. J., McIntosh, D. H., McLure, R. J., Mobasher, B., Moustakas, L. A., Mozena, M., Nandra, K., Newman, J. A., Niemi, S.-M., Noeske, K. G., Papovich, C. J., Pentericci, L., Pope, A., Primack, J. R., Rajan, A., Ravindranath, S., Reddy, N. A., Renzini, A., Rix, H.-W., Robaina, A. R., Rodney, S. A., Rosario, D. J., Rosati, P., Salimbeni, S., Scarlata, C., Siana, B., Simard, L., Smidt, J., Somerville, R. S., Spinrad, H., Straughn, A. N., Strolger, L.-G., Telford, O., Teplitz, H. I., Trump, J. R., van der Wel, A., Villforth, C., Wechsler, R. H., Weiner, B. J., Wiklind, T., Wild, V., Wilson, G., Wuyts, S., Yan, H.-J., and Yun, M. S.: 2011, *ApJl Supp.* **197**, 35

Grützbauch, R., Bauer, A. E., Jørgensen, I., and Varela, J.: 2012, *MNRAS* **423(4)**, 3652

Grützbauch, R., Conselice, C. J., Bauer, A. E., Bluck, A. F. L., Chuter, R. W., Buitrago, F., Mortlock, A., Weinzirl, T., and Jogee, S.: 2011, *MNRAS* **418**, 938

Gunn, J. E. and Gott, J. Richard, I.: 1972, *ApJ* **176**, 1

Guo, Y., Bell, E. F., Lu, Y., Koo, D. C., Faber, S. M., Koekemoer, A. M., Kurczynski, P., Lee, S.-K., Papovich, C., Chen, Z., Dekel, A., Ferguson, H. C., Fontana, A., Giavalisco, M., Kocevski, D. D., Nayyeri, H., Pérez-González, P. G., Pforr, J., Rodríguez-Puebla, A., and Santini, P.: 2017, *ApJl* **841(2)**, L22

Guo, Y., Ferguson, H. C., Giavalisco, M., Barro, G., Willner, S. P., Ashby, M. L. N., Dahlen, T., Donley, J. L., Faber, S. M., Fontana, A., Galametz, A., Grazian, A., Huang, K.-H., Kocevski, D. D., Koekemoer, A. M., Koo, D. C., McGrath, E. J., Peth, M., Salvato, M.,

Wuyts, S., Castellano, M., Cooray, A. R., Dickinson, M. E., Dunlop, J. S., Fazio, G. G., Gardner, J. P., Gawiser, E., Grogin, N. A., Hathi, N. P., Hsu, L.-T., Lee, K.-S., Lucas, R. A., Mobasher, B., Nandra, K., Newman, J. A., and van der Wel, A.: 2013, *ApJl Supp.* **207**, 24

Gupta, A., Yuan, T., Torrey, P., Vogelsberger, M., Martizzi, D., Tran, K.-V. H., Kewley, L. J., Marinacci, F., Nelson, D., Pillepich, A., Hernquist, L., Genel, S., and Springel, V.: 2018, *MNRAS* **477(1)**, L35

Hall, P.: 1982, *Biometrika* **69(2)**, 383

Heckman, T. M., Armus, L., and Miley, G. K.: 1990, *ApJl Supp.* **74**, 833

Hemmati, S., Capak, P., Pourrahmani, M., Nayyeri, H., Stern, D., Mobasher, B., Darvish, B., Davidzon, I., Ilbert, O., Masters, D., and Shahidi, A.: 2019, *ApJl* **881(1)**, L14

Hennawi, J. F., Prochaska, J. X., Cantalupo, S., and Arrigoni-Battaia, F.: 2015, *Science* **348(6236)**, 779

Ilbert, O., Arnouts, S., McCracken, H. J., Bolzonella, M., Bertin, E., Le Fèvre, O., Mellier, Y., Zamorani, G., Pellò, R., Iovino, A., Tresse, L., Le Brun, V., Bottini, D., Garilli, B., Maccagni, D., Picat, J. P., Scaramella, R., Scodeggio, M., Vettolani, G., Zanichelli, A., Adami, C., Bardelli, S., Cappi, A., Charlot, S., Ciliegi, P., Contini, T., Cucciati, O., Foucaud, S., Franzetti, P., Gavignaud, I., Guzzo, L., Marano, B., Marinoni, C., Mazure, A., Meneux, B., Merighi, R., Paltani, S., Pollo, A., Pozzetti, L., Radovich, M., Zucca, E., Bondi, M., Bongiorno, A., Busarello, G., de La Torre, S., Gregorini, L., Lamareille, F., Mathez, G., Merluzzi, P., Ripepi, V., Rizzo, D., and Vergani, D.: 2006, *Astronomy and Astrophysics* **457**, 841

Ilbert, O., Capak, P., Salvato, M., Aussel, H., McCracken, H. J., Sanders, D. B., Scoville, N., Kartaltepe, J., Arnouts, S., Le Floc'h, E., Mobasher, B., Taniguchi, Y., Lamareille, F., Leauthaud, A., Sasaki, S., Thompson, D., Zamojski, M., Zamorani, G., Bardelli, S., Bolzonella, M., Bongiorno, A., Brusa, M., Caputi, K. I., Carollo, C. M., Contini, T., Cook, R., Coppa, G., Cucciati, O., de la Torre, S., de Ravel, L., Franzetti, P., Garilli, B., Hasinger, G., Iovino, A., Kampczyk, P., Kneib, J.-P., Knobel, C., Kovac, K., Le Borgne, J. F., Le Brun, V., Le Fèvre, O., Lilly, S., Looper, D., Maier, C., Mainieri, V., Mellier, Y., Mignoli, M., Murayama, T., Pellò, R., Peng, Y., Pérez-Montero, E., Renzini, A., Ricciardelli, E., Schiminovich, D., Scodeggio, M., Shioya, Y., Silverman, J., Surace, J., Tanaka, M., Tasca, L., Tresse, L., Vergani, D., and Zucca, E.: 2009, *ApJ* **690**, 1236

Ilbert, O., McCracken, H. J., Le Fèvre, O., Capak, P., Dunlop, J., Karim, A., Renzini, M. A., Caputi, K., Boissier, S., Arnouts, S., Aussel, H., Comparat, J., Guo, Q., Hudelot, P., Kartaltepe, J., Kneib, J. P., Krogager, J. K., Le Floc'h, E., Lilly, S., Mellier, Y., Milvang-Jensen, B., Moutard, T., Onodera, M., Richard, J., Salvato, M., Sanders, D. B., Scoville, N., Silverman, J. D., Taniguchi, Y., Tasca, L., Thomas, R., Toft, S., Tresse, L., Vergani, D., Wolk, M., and Zirm, A.: 2013, *Astronomy and Astrophysics* **556**, A55

Jackson, J. C.: 1972, *MNRAS* **156**, 1P

Ji, Z., Giavalisco, M., Williams, C. C., Faber, S. M., Ferguson, H. C., Guo, Y., Liu, T., and Lee, B.: 2018, *ApJ* **862(2)**, 135

Jones, M. C.: 1993, *Statistics and Computing* **3(3)**, 135

Jones, M. C., Marron, J. S., and Sheather, S. J.: 1996, *Journal of the American Statistical Association* **91(433)**, 401

Kacprzak, G. G., Yuan, T., Nanayakkara, T., Kobayashi, C., Tran, K.-V. H., Kewley, L. J., Glazebrook, K., Spitler, L., Taylor, P., Cowley, M., Labbe, I., Straatman, C., and Tomczak, A.: 2015, *ApJl* **802(2)**, L26

Kauffmann, G., White, S. D. M., Heckman, T. M., Ménard, B., Brinchmann, J., Charlot, S., Tremonti, C., and Brinkmann, J.: 2004, *MNRAS* **353**, 713

Kawinwanichakij, L., Papovich, C., Quadri, R. F., Glazebrook, K., Kacprzak, G. G., Allen, R. J., Bell, E. F., Croton, D. J., Dekel, A., Ferguson, H. C., Forrest, B., Grogin, N. A., Guo, Y., Kocevski, D. D., Koekemoer, A. M., Labbé, I., Lucas, R. A., Nanayakkara, T., Spitler, L. R., Straatman, C. M. S., Tran, K.-V. H., Tomczak, A., and van Dokkum, P.: 2017, *ApJ* **847(2)**, 134

Kennicutt, Jr., R. C.: 1998, *Annual Reviews of Astronomy and Astrophysics* **36**, 189

Kereš, D., Katz, N., Weinberg, D. H., and Davé, R.: 2005, *MNRAS* **363(1)**, 2

Kewley, L. J., Dopita, M. A., Leitherer, C., Davé, R., Yuan, T., Allen, M., Groves, B., and Sutherland, R.: 2013, *ApJ* **774**, 100

Kewley, L. J. and Ellison, S. L.: 2008, *ApJ* **681(2)**, 1183

Kodra, D.: 2019, *Ph.D. thesis*, University of Pittsburgh

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., Grogin, N. A., Kocevski, D. D., Koo, D. C., Lai, K., Lotz, J. M., Lucas, R. A., McGrath, E. J., Ogaz, S., Rajan, A., Riess, A. G., Rodney, S. A., Strolger, L., Casertano, S., Castellano, M., Dahlen, T., Dickinson, M., Dolch, T., Fontana, A., Giavalisco, M., Grazian, A., Guo, Y., Hathi, N. P., Huang, K.-H., van der Wel, A., Yan, H.-J., Acquaviva, V., Alexander, D. M., Almaini, O., Ashby, M. L. N., Barden, M., Bell, E. F., Bournaud, F., Brown, T. M., Caputi, K. I., Cassata, P., Challis, P. J., Chary, R.-R., Cheung, E., Cirasuolo, M., Conselice, C. J., Roshan Cooray, A., Croton, D. J., Daddi, E., Davé, R., de Mello, D. F., de Ravel, L., Dekel, A., Donley, J. L., Dunlop, J. S., Dutton, A. A., Elbaz, D., Fazio, G. G., Filippenko, A. V., Finkelstein, S. L., Frazer, C., Gardner, J. P., Garnavich, P. M., Gawiser, E., Gruetzbauch, R., Hartley, W. G., Häussler, B., Herrington, J., Hopkins, P. F., Huang, J.-S., Jha, S. W., Johnson, A., Kartaltepe, J. S., Khostovan, A. A., Kirshner, R. P., Lani, C., Lee, K.-S., Li, W., Madau, P., McCarthy, P. J., McIntosh, D. H., McLure, R. J., McPartland, C., Mobasher, B., Moreira, H., Mortlock, A., Moustakas, L. A., Mozena, M., Nandra, K., Newman, J. A., Nielsen, J. L., Niemi, S., Noeske, K. G., Papovich, C. J., Pentericci, L., Pope, A., Primack, J. R., Ravindranath, S., Reddy, N. A., Renzini, A., Rix, H.-W., Robaina, A. R., Rosario, D. J., Rosati, P., Salimbeni, S., Scarlata, C., Siana, B., Simard, L., Smidt, J., Snyder, D., Somerville, R. S., Spinrad, H., Straughn, A. N., Telford, O., Teplitz, H. I., Trump, J. R., Vargas, C., Villforth, C., Wagner, C. R., Wandro, P., Wechsler, R. H., Weiner, B. J., Wiklind, T., Wild, V., Wilson, G., Wuyts, S., and Yun, M. S.: 2011, *ApJl Supp.* **197**, 36

Kohonen, T.: 1982, *Biological Cybernetics* **43(1)**, 59

Kriek, M., Shapley, A. E., Reddy, N. A., Siana, B., Coil, A. L., Mobasher, B., Freeman, W. R., de Groot, L., Price, S. H., Sanders, R., Shivaei, I., Brammer, G. B., Momcheva, I. G., Skelton, R. E., van Dokkum, P. G., Whitaker, K. E., Aird, J., Azadi, M., Kassis, M., Bullock, J. S., Conroy, C., Davé, R., Kereš, D., and Krumholz, M.: 2015, *ApJl Supp.* **218(2)**, 15

Kriek, M., van Dokkum, P. G., Labbé, I., Franx, M., Illingworth, G. D., Marchesini, D., and Quadri, R. F.: 2009, *ApJ* **700**, 221

Kulas, K. R., McLean, I. S., Shapley, A. E., Steidel, C. C., Konidaris, N. P., Matthews, K., Mace, G. N., Rudie, G. C., Trainor, R. F., and Reddy, N. A.: 2013, *ApJ* **774(2)**, 130

Larson, R. B.: 1974, *MNRAS* **169**, 229

Larson, R. B. and Tinsley, B. M.: 1978, *ApJ* **219**, 46

Laureijs, R., Amiaux, J., Arduini, S., Augueres, J. L., Brinchmann, J., Cole, R., Cropper, M., Dabin, C., Duvet, L., Ealet, A., Garilli, B., Gondoin, P., Guzzo, L., Hoar, J., Hoekstra, H., Holmes, R., Kitching, T., Maciaszek, T., Mellier, Y., Pasian, F., Percival, W., Rhodes, J., Saavedra Criado, G., Sauvage, M., Scaramella, R., Valenziano, L., Warren, S., Bender, R., Castander, F., Cimatti, A., Le Fèvre, O., Kurki-Suonio, H., Levi, M., Lilje, P., Meylan, G., Nichol, R., Pedersen, K., Popa, V., Rebolo Lopez, R., Rix, H. W., Rottgering, H., Zeilinger, W., Grupp, F., Hudelot, P., Massey, R., Meneghetti, M., Miller, L., Paltani, S., Paulin-Henriksson, S., Pires, S., Saxton, C., Schrabback, T., Seidel, G., Walsh, J., Aghanim, N., Amendola, L., Bartlett, J., Baccigalupi, C., Beaulieu, J. P., Benabed, K., Cuby, J. G., Elbaz, D., Fosalba, P., Gavazzi, G., Helmi, A., Hook, I., Irwin, M., Kneib, J. P., Kunz, M., Mannucci, F., Moscardini, L., Tao, C., Teyssier, R., Weller, J., Zamorani, G., Zapatero Osorio, M. R., Boulade, O., Foumond, J. J., Di Giorgio, A., Guttridge, P., James, A., Kemp, M., Martignac, J., Spencer, A., Walton, D., Blümchen, T., Bonoli, C., Bortoletto, F., Cerna, C., Corcione, L., Fabron, C., Jahnke, K., Ligori, S., Madrid, F., Martin, L., Morgante, G., Pamplona, T., Prieto, E., Riva, M., Toledo, R., Trifoglio, M., Zerbi, F., Abdalla, F., Douspis, M., Grenet, C., Borgani, S., Bouwens, R., Courbin, F., Delouis, J. M., Dubath, P., Fontana, A., Frailis, M., Grazian, A., Koppenhöfer, J., Mansutti, O., Melchior, M., Mignoli, M., Mohr, J., Neissner, C., Noddle, K., Poncet, M., Scodeggio, M., Serrano, S., Shane, N., Starck, J. L., Surace, C., Taylor, A., Verdoes-Kleijn, G., Vuerli, C., Williams, O. R., Zacchei, A., Altieri, B., Escudero Sanz, I., Kohley, R., Oosterbroek, T., Astier, P., Bacon, D., Bardelli, S., Baugh, C., Bellagamba, F., Benoist, C., Bianchi, D., Biviano, A., Branchini, E., Carbone, C., Cardone, V., Clements, D., Colombi, S., Conselice, C., Cresci, G., Deacon, N., Dunlop, J., Fedeli, C., Fontanot, F., Franzetti, P., Giocoli, C., Garcia-Bellido, J., Gow, J., Heavens, A., Hewett, P., Heymans, C., Holland, A., Huang, Z., Ilbert, O., Joachimi, B., Jennins, E., Kerins, E., Kiessling, A., Kirk, D., Kotak, R., Krause, O., Lahav, O., van Leeuwen, F., Lesgourgues, J., Lombardi, M., Magliocchetti, M., Maguire, K., Majerotto, E., Maoli, R., Marulli, F., Maurogordato, S., McCracken, H., McLure, R., Melchiorri, A., Merson, A., Moresco, M., Nonino, M., Norberg, P., Peacock, J., Pello, R., Penny, M., Pettorino, V., Di Porto, C., Pozzetti, L., Quercellini, C., Radovich, M., Rassat, A., Roche, N., Ronayette, S., Rossetti, E., Sartoris, B., Schneider, P., Semboloni, E., Serjeant, S., Simpson, F., Skordis, C., Smadja, G., Smartt, S., Spano, P., Spiro, S., Sullivan, M., Tilquin, A., Trotta, R., Verde, L., Wang, Y., Williger, G., Zhao, G., Zoubian, J., and Zucca, E.: 2011, *arXiv e-prints* p. arXiv:1110.3193

Lee, S.-K., Im, M., Kim, J.-W., Lotz, J., McPartland, C., Peth, M., and Koekemoer, A.: 2015, *ApJ* **810(2)**, 90

Lemaux, B. C., Cucciati, O., Le Fèvre, O., Zamorani, G., Lubin, L. M., Hathi, N., Ilbert, O., Pelliccia, D., Amorín, R., Bardelli, S., Cassata, P., Gal, R. R., Garilli, B., Guaita, L., Giavalisco, M., Hung, D., Koekemoer, A., Maccagni, D., Pentericci, L., Ribeiro, B., Schaerer, D., Shen, L., Talia, M., Tomczak, A. R., Vanzella, E., Vergani, D., and Zucca, E.: 2020, *arXiv e-prints* p. arXiv:2009.03324

Leung, G. C. K., Coil, A. L., Aird, J., Azadi, M., Kriek, M., Mobasher, B., Reddy, N., Shapley, A., Siana, B., Fetherolf, T., Fornasini, F. M., Freeman, W. R., Price, S. H., Sanders, R. L., Shivaei, I., and Zick, T.: 2019, *ApJ* **886(1)**, 11

Leung, G. C. K., Coil, A. L., Azadi, M., Aird, J., Shapley, A., Kriek, M., Mobasher, B., Reddy, N., Siana, B., Freeman, W. R., Price, S. H., Sanders, R. L., and Shivaei, I.: 2017, *ApJ* **849(1)**, 48

Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., and Peng, Y.: 2013, *ApJ* **772**, 119

Lin, L., Dickinson, M., Jian, H.-Y., Merson, A. I., Baugh, C. M., Scott, D., Foucaud, S., Wang, W.-H., Yan, C.-H., Yan, H.-J., Cheng, Y.-W., Guo, Y., Helly, J., Kirsten, F., Koo, D. C., Lagos, C. d. P., Meger, N., Messias, H., Pope, A., Simard, L., Grogin, N. A., and Wang, S.-Y.: 2012, *ApJ* **756(1)**, 71

Lin, L., Jian, H.-Y., Foucaud, S., Norberg, P., Bower, R. G., Cole, S., Arnalte-Mur, P., Chen, C.-W., Coupon, J., Hsieh, B.-C., Heinis, S., Phleps, S., Chen, W.-P., Lee, C.-H., Burgett, W., Chambers, K. C., Denneau, L., Draper, P., Flewelling, H., Hodapp, K. W., Huber, M. E., Kaiser, N., Kudritzki, R. P., Magnier, E. A., Metcalfe, N., Price, P. A., Tonry, J. L., Wainscoat, R. J., and Waters, C.: 2014, *ApJ* **782(1)**, 33

Machacek, M., Nulsen, P. E. J., Jones, C., and Forman, W. R.: 2006, *ApJ* **648(2)**, 947

Maier, C., Hayashi, M., Ziegler, B. L., and Kodama, T.: 2019, *Astronomy and Astrophysics* **626**, A14

Maiolino, R. and Mannucci, F.: 2019, *Astronomy and Astrophysics Reviews* **27(1)**, 3

Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., and Gnerucci, A.: 2010, *MNRAS* **408**, 2115

Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., Pastorini, G., Pozzetti, L., Gnerucci,

A., Risaliti, G., Schneider, R., Lehnert, M., and Salvati, M.: 2009, *MNRAS* **398(4)**, 1915

Marron, J. S. and Ruppert, D.: 1994, *Journal of the Royal Statistical Society. Series B (Methodological)* **56(4)**, 653

Martin, C. L., Shapley, A. E., Coil, A. L., Kornei, K. A., Bundy, K., Weiner, B. J., Noeske, K. G., and Schiminovich, D.: 2012, *ApJ* **760(2)**, 127

Masters, D., Capak, P., Stern, D., Ilbert, O., Salvato, M., Schmidt, S., Longo, G., Rhodes, J., Paltani, S., Mobasher, B., Hoekstra, H., Hildebrandt, H., Coupon, J., Steinhardt, C., Speagle, J., Faisst, A., Kalinich, A., Brodwin, M., Brescia, M., and Cavuoti, S.: 2015, *ApJ* **813(1)**, 53

Masters, D. C., Stern, D. K., Cohen, J. G., Capak, P. L., Rhodes, J. D., Castander, F. J., and Paltani, S.: 2017, *ApJ* **841(2)**, 111

Masters, D. C., Stern, D. K., Cohen, J. G., Capak, P. L., Stanford, S. A., Hernitschek, N., Galametz, A., Davidzon, I., Rhodes, J. D., Sanders, D., Mobasher, B., Castander, F., Pruett, K., and Fotopoulou, S.: 2019, *ApJ* **877(2)**, 81

McCracken, H. J., Milvang-Jensen, B., Dunlop, J., Franx, M., Fynbo, J. P. U., Le Fèvre, O., Holt, J., Caputi, K. I., Goranova, Y., Buitrago, F., Emerson, J. P., Freudling, W., Hudelot, P., López-Sanjuan, C., Magnard, F., Mellier, Y., Møller, P., Nilsson, K. K., Sutherland, W., Tasca, L., and Zabl, J.: 2012, *Astronomy and Astrophysics* **544**, A156

McGee, S. L., Bower, R. G., and Balogh, M. L.: 2014, *MNRAS* **442**, L105

McInnes, L., Healy, J., and Melville, J.: 2018, *arXiv e-prints* p. arXiv:1802.03426

McLean, I. S., Steidel, C. C., Epps, H. W., Konidaris, N., Matthews, K. Y., Adkins, S., Aliado, T., Brims, G., Canfield, J. M., Cromer, J. L., Fucik, J., Kulas, K., Mace, G., Magnone, K., Rodriguez, H., Rudie, G., Trainor, R., Wang, E., Weber, B., and Weiss, J.: 2012, *MOSFIRE, the multi-object spectrometer for infra-red exploration at the Keck Observatory*, Vol. 8446 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 84460J

Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., Skelton, R. E., Whitaker, K. E.,

Nelson, E. J., Fumagalli, M., Maseda, M. V., Leja, J., Franx, M., Rix, H.-W., Bezanson, R., Da Cunha, E., Dickey, C., Förster Schreiber, N. M., Illingworth, G., Kriek, M., Labbé, I., Ulf Lange, J., Lundgren, B. F., Magee, D., Marchesini, D., Oesch, P., Pacifici, C., Patel, S. G., Price, S., Tal, T., Wake, D. A., van der Wel, A., and Wuyts, S.: 2016, *ApJl Supp.* **225(2)**, 27

Mouhcine, M., Baldry, I. K., and Bamford, S. P.: 2007, *MNRAS* **382(2)**, 801

Mucesh, S., Hartley, W. G., Palmese, A., Lahav, O., Whiteway, L., Bluck, A. F. L., Alarcon, A., Amon, A., Bechtol, K., Bernstein, G. M., Carnero Rosell, A., Carrasco Kind, M., Choi, A., Eckert, K., Everett, S., Gruen, D., Gruendl, R. A., Harrison, I., Huff, E. M., Kuropatkin, N., Sevilla-Noarbe, I., Sheldon, E., Yanny, B., Aguena, M., Allam, S., Bacon, D., Bertin, E., Bhargava, S., Brooks, D., Carretero, J., Castander, F. J., Conselice, C., Costanzi, M., Crocce, M., da Costa, L. N., Pereira, M. E. S., De Vicente, J., Desai, S., Diehl, H. T., Drlica-Wagner, A., Evrard, A. E., Ferrero, I., Flaugher, B., Fosalba, P., Frieman, J., García-Bellido, J., Gaztanaga, E., Gerdes, D. W., Gschwend, J., Gutierrez, G., Hinton, S. R., Hollowood, D. L., Honscheid, K., James, D. J., Kuehn, K., Lima, M., Lin, H., Maia, M. A. G., Melchior, P., Menanteau, F., Miquel, R., Morgan, R., Paz-Chinchón, F., Plazas, A. A., Sanchez, E., Scarpine, V., Schubnell, M., Serrano, S., Smith, M., Suchyta, E., Tarle, G., Thomas, D., To, C., Varga, T. N., Wilkinson, R. D., and DES Collaboration: 2021, *MNRAS* **502(2)**, 2770

Muldrew, S. I., Hatch, N. A., and Cooke, E. A.: 2015, *MNRAS* **452(3)**, 2528

Mushotzky, R. F. and Loewenstein, M.: 1997, *ApJl* **481(2)**, L63

Muzzin, A., Marchesini, D., Stefanon, M., Franx, M., McCracken, H. J., Milvang-Jensen, B., Dunlop, J. S., Fynbo, J. P. U., Brammer, G., Labbé, I., and van Dokkum, P. G.: 2013, *ApJ* **777**, 18

Muzzin, A., Wilson, G., Yee, H. K. C., Gilbank, D., Hoekstra, H., Demarco, R., Balogh, M., van Dokkum, P., Franx, M., Ellingson, E., Hicks, A., Nantais, J., Noble, A., Lacy, M., Lidman, C., Rettura, A., Surace, J., and Webb, T.: 2012, *ApJ* **746(2)**, 188

Müller, H.-G.: 1991, *Biometrika* **78(3)**, 521

Namiki, S. V., Koyama, Y., Hayashi, M., Tadaki, K.-i., Kashikawa, N., Onodera, M., Shimakawa, R., Kodama, T., Tanaka, I., Förster Schreiber, N. M., Kurk, J., and Genzel, R.: 2019, *ApJ* **877(2)**, 118

Nantais, J. B., Muzzin, A., van der Burg, R. F. J., Wilson, G., Lidman, C., Foltz, R., DeGroot, A., Noble, A., Cooper, M. C., and Demarco, R.: 2017, *MNRAS* **465(1)**, L104

Nantais, J. B., van der Burg, R. F. J., Lidman, C., Demarco, R., Noble, A., Wilson, G., Muzzin, A., Foltz, R., DeGroot, A., and Cooper, M. C.: 2016, *Astronomy and Astrophysics* **592**, A161

Nayyeri, H., Hemmati, S., Mobasher, B., Ferguson, H. C., Cooray, A., Barro, G., Faber, S. M., Dickinson, M., Koekemoer, A. M., Peth, M., Salvato, M., Ashby, M. L. N., Darvish, B., Donley, J., Durbin, M., Finkelstein, S., Fontana, A., Grogin, N. A., Gruetzbauch, R., Huang, K., Khostovan, A. A., Kocevski, D., Kodra, D., Lee, B., Newman, J., Pacifici, C., Pforr, J., Stefanon, M., Wiklind, T., Willner, S. P., Wuyts, S., Castellano, M., Conselice, C., Dolch, T., Dunlop, J. S., Galametz, A., Hathi, N. P., Lucas, R. A., and Yan, H.: 2017, *ApJl Supp.* **228**, 7

Old, L. J., Balogh, M. L., van der Burg, R. F. J., Biviano, A., Yee, H. K. C., Pintos-Castro, I., Webb, K., Muzzin, A., Rudnick, G., Vulcani, B., Poggianti, B., Cooper, M., Zaritsky, D., Cerulo, P., Wilson, G., Chan, J. C. C., Lidman, C., McGee, S., Demarco, R., Forrest, B., De Lucia, G., Gilbank, D., Kukstas, E., McCarthy, I. G., Jablonka, P., Nantais, J., Noble, A., Reeves, A. M. M., and Shipley, H.: 2020, *MNRAS* **493(4)**, 5987

Pacifici, C., Charlot, S., Blaizot, J., and Brinchmann, J.: 2012, *MNRAS* **421(3)**, 2002

Papovich, C., Kawinwanichakij, L., Quadri, R. F., Glazebrook, K., Labbé, I., Tran, K.-V. H., Forrest, B., Kacprzak, G. G., Spitler, L. R., Straatman, C. M. S., and Tomczak, A. R.: 2018, *ApJ* **854(1)**, 30

Parzen, E.: 1962, *Ann. Math. Statist.* **33(3)**, 1065

Patel, S. G., Holden, B. P., Kelson, D. D., Illingworth, G. D., and Franx, M.: 2009, *ApJl* **705(1)**, L67

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: 2011, *Journal of Machine Learning Research* **12**, 2825

Peng, Y.-j., Lilly, S. J., Kovač, K., Bolzonella, M., Pozzetti, L., Renzini, A., Zamorani, G.,

Ilbert, O., Knobel, C., Iovino, A., Maier, C., Cucciati, O., Tasca, L., Carollo, C. M., Silverman, J., Kampczyk, P., de Ravel, L., Sanders, D., Scoville, N., Contini, T., Mainieri, V., Scodeggio, M., Kneib, J.-P., Le Fèvre, O., Bardelli, S., Bongiorno, A., Caputi, K., Coppa, G., de la Torre, S., Franzetti, P., Garilli, B., Lamareille, F., Le Borgne, J.-F., Le Brun, V., Mignoli, M., Perez Montero, E., Pello, R., Ricciardelli, E., Tanaka, M., Tresse, L., Vergani, D., Welikala, N., Zucca, E., Oesch, P., Abbas, U., Barnes, L., Bordoloi, R., Bottini, D., Cappi, A., Cassata, P., Cimatti, A., Fumana, M., Hasinger, G., Koekemoer, A., Leauthaud, A., Maccagni, D., Marinoni, C., McCracken, H., Memeo, P., Meneux, B., Nair, P., Porciani, C., Presotto, V., and Scaramella, R.: 2010, *ApJ* **721**, 193

Peng, Y.-j. and Maiolino, R.: 2014, *MNRAS* **438(1)**, 262

Pettini, M. and Pagel, B. E. J.: 2004, *MNRAS* **348(3)**, L59

Pintos-Castro, I., Yee, H. K. C., Muzzin, A., Old, L., and Wilson, G.: 2019, *ApJ* **876(1)**, 40

Pozzetti, L., Bolzonella, M., Zucca, E., Zamorani, G., Lilly, S., Renzini, A., Moresco, M., Mignoli, M., Cassata, P., Tasca, L., Lamareille, F., Maier, C., Meneux, B., Halliday, C., Oesch, P., Vergani, D., Caputi, K., Kovač, K., Cimatti, A., Cucciati, O., Iovino, A., Peng, Y., Carollo, M., Contini, T., Kneib, J.-P., Le Févre, O., Mainieri, V., Scodeggio, M., Bardelli, S., Bongiorno, A., Coppa, G., de la Torre, S., de Ravel, L., Franzetti, P., Garilli, B., Kampczyk, P., Knobel, C., Le Borgne, J.-F., Le Brun, V., Pellò, R., Perez Montero, E., Ricciardelli, E., Silverman, J. D., Tanaka, M., Tresse, L., Abbas, U., Bottini, D., Cappi, A., Guzzo, L., Koekemoer, A. M., Leauthaud, A., Maccagni, D., Marinoni, C., McCracken, H. J., Memeo, P., Porciani, C., Scaramella, R., Scarlata, C., and Scoville, N.: 2010, *Astronomy and Astrophysics* **523**, A13

Quadri, R. F., Williams, R. J., Franx, M., and Hildebrandt, H.: 2012, *ApJ* **744(2)**, 88

Reddy, N. A., Kriek, M., Shapley, A. E., Freeman, W. R., Siana, B., Coil, A. L., Mobasher, B., Price, S. H., Sanders, R. L., and Shivaei, I.: 2015, *ApJ* **806(2)**, 259

Sanders, R. L., Shapley, A. E., Kriek, M., Freeman, W. R., Reddy, N. A., Siana, B., Coil, A. L., Mobasher, B., Davé, R., Shivaei, I., Azadi, M., Price, S. H., Leung, G., Fetherolf, T., de Groot, L., Zick, T., Fornasini, F. M., and Barro, G.: 2018, *ApJ* **858(2)**, 99

Sanders, R. L., Shapley, A. E., Kriek, M., Reddy, N. A., Freeman, W. R., Coil, A. L., Siana, B., Mobasher, B., Shivaei, I., Price, S. H., and de Groot, L.: 2015, *ApJ* **799(2)**, 138

Sanders, R. L., Shapley, A. E., Reddy, N. A., Kriek, M., Siana, B., Coil, A. L., Mobasher, B., Shivaei, I., Freeman, W. R., Azadi, M., Price, S. H., Leung, G., Fetherolf, T., de Groot, L., Zick, T., Fornasini, F. M., and Barro, G.: 2020, *MNRAS* **491(1)**, 1427

Sattari, Z., Mobasher, B., Chartab, N., Darvish, B., Shivaei, I., Scoville, N., and Sobral, D.: 2021, *arXiv e-prints* p. arXiv:2102.05637

Schaefer, A. L., Tremonti, C., Pace, Z., Belfiore, F., Argudo-Fernandez, M., Bershady, M. A., Drory, N., Jones, A., Maiolino, R., Stark, D., Wake, D., and Yan, R.: 2019, *ApJ* **884(2)**, 156

Schuster, E. F.: 1985, *Communications in Statistics - Theory and Methods* **14(5)**, 1123

Scoville, N., Arnouts, S., Aussel, H., Benson, A., Bongiorno, A., Bundy, K., Calvo, M. A. A., Capak, P., Carollo, M., Civano, F., Dunlop, J., Elvis, M., Faisst, A., Finoguenov, A., Fu, H., Giavalisco, M., Guo, Q., Ilbert, O., Iovino, A., Kajisawa, M., Kartaltepe, J., Leauthaud, A., Le Fèvre, O., LeFloch, E., Lilly, S. J., Liu, C. T.-C., Manohar, S., Massey, R., Masters, D., McCracken, H. J., Mobasher, B., Peng, Y.-J., Renzini, A., Rhodes, J., Salvato, M., Sanders, D. B., Sarvestani, B. D., Scarlata, C., Schinnerer, E., Sheth, K., Shopbell, P. L., Smolčić, V., Taniguchi, Y., Taylor, J. E., White, S. D. M., and Yan, L.: 2013, *ApJl Supp.* **206**, 3

Scoville, N., Aussel, H., Brusa, M., Capak, P., Carollo, C. M., Elvis, M., Giavalisco, M., Guzzo, L., Hasinger, G., and Impey, C.: 2007, *ApJl Supp.* **172(1)**, 1

Shapley, A. E., Sanders, R. L., Shao, P., Reddy, N. A., Kriek, M., Coil, A. L., Mobasher, B., Siana, B., Shivaei, I., Freeman, W. R., Azadi, M., Price, S. H., Leung, G. C. K., Fetherolf, T., de Groot, L., Zick, T., Fornasini, F. M., and Barro, G.: 2019, *ApJl* **881(2)**, L35

Shimakawa, R., Kodama, T., Tadaki, K.-i., Hayashi, M., Koyama, Y., and Tanaka, I.: 2015, *MNRAS* **448(1)**, 666

Shivaei, I., Kriek, M., Reddy, N. A., Shapley, A. E., Barro, G., Conroy, C., Coil, A. L., Freeman, W. R., Mobasher, B., Siana, B., Sanders, R., Price, S. H., Azadi, M., Pasha, I., and Inami, H.: 2016, *ApJl* **820(2)**, L23

Shivaei, I., Reddy, N. A., Siana, B., Shapley, A. E., Kriek, M., Mobasher, B., Freeman, W. R., Sanders, R. L., Coil, A. L., Price, S. H., Fetherolf, T., Azadi, M., Leung, G., and Zick, T.: 2018, *ApJ* **855(1)**, 42

Shivaei, I., Reddy, N. A., Steidel, C. C., and Shapley, A. E.: 2015, *ApJ* **804(2)**, 149

Silk, J.: 1977, *ApJ* **211**, 638

Silverman, B. W.: 1986, *Density Estimation for Statistics and Data Analysis*, Springer

Simet, M., Chartab, N., Lu, Y., and Mobasher, B.: 2021, *ApJ* **908(1)**, 47

Skelton, R. E., Whitaker, K. E., Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., Labbé, I., Franx, M., van der Wel, A., Bezanson, R., Da Cunha, E., Fumagalli, M., Förster Schreiber, N., Kriek, M., Leja, J., Lundgren, B. F., Magee, D., Marchesini, D., Maseda, M. V., Nelson, E. J., Oesch, P., Pacifici, C., Patel, S. G., Price, S., Rix, H.-W., Tal, T., Wake, D. A., and Wuyts, S.: 2014, *ApJl Supp.* **214(2)**, 24

Sobral, D., Smail, I., Best, P. N., Geach, J. E., Matsuda, Y., Stott, J. P., Cirasuolo, M., and Kurk, J.: 2013, *MNRAS* **428(2)**, 1128

Spitler, L. R., Labbé, I., Glazebrook, K., Persson, S. E., Monson, A., Papovich, C., Tran, K.-V. H., Poole, G. B., Quadri, R., van Dokkum, P., Kelson, D. D., Kacprzak, G. G., McCarthy, P. J., Murphy, D., Straatman, C. M. S., and Tilvi, V.: 2012, *ApJl* **748(2)**, L21

Stefanon, M., Yan, H., Mobasher, B., Barro, G., Donley, J. L., Fontana, A., Hemmati, S., Koekemoer, A. M., Lee, B., Lee, S.-K., Nayyeri, H., Peth, M., Pforr, J., Salvato, M., Wiklind, T., Wuyts, S., Ashby, M. L. N., Castellano, M., Conselice, C. J., Cooper, M. C., Cooray, A. R., Dolch, T., Ferguson, H., Galametz, A., Giavalisco, M., Guo, Y., Willner, S. P., Dickinson, M. E., Faber, S. M., Fazio, G. G., Gardner, J. P., Gawiser, E., Grazian, A., Grogin, N. A., Kocevski, D., Koo, D. C., Lee, K.-S., Lucas, R. A., McGrath, E. J., Nandra, K., Newman, J. A., and van der Wel, A.: 2017, *ApJl Supp.* **229**, 32

Steidel, C. C., Erb, D. K., Shapley, A. E., Pettini, M., Reddy, N., Bogosavljević, M., Rudie, G. C., and Rakic, O.: 2010, *ApJ* **717(1)**, 289

Steidel, C. C., Rudie, G. C., Strom, A. L., Pettini, M., Reddy, N. A., Shapley, A. E., Trainor, R. F., Erb, D. K., Turner, M. L., Konidaris, N. P., Kulas, K. R., Mace, G., Matthews, K., and McLean, I. S.: 2014, *ApJ* **795(2)**, 165

Stocke, J. T., Danforth, C. W., Shull, J. M., Penton, S. V., and Giroux, M. L.: 2007, *ApJ*

671(1), 146

Stott, J. P., Sobral, D., Bower, R., Smail, I., Best, P. N., Matsuda, Y., Hayashi, M., Geach, J. E., and Kodama, T.: 2013, *MNRAS* **436(2)**, 1130

Tran, K.-V. H., Nanayakkara, T., Yuan, T., Kacprzak, G. G., Glazebrook, K., Kewley, L. J., Momcheva, I., Papovich, C. J., Quadri, R., Rudnick, G., Saintonge, A., Spitler, L. R., Straatman, C., and Tomczak, A.: 2015, *ApJ* **811(1)**, 28

Tran, K.-V. H., Papovich, C., Saintonge, A., Brodwin, M., Dunlop, J. S., Farrah, D., Finkelstein, K. D., Finkelstein, S. L., Lotz, J., McLure, R. J., Momcheva, I., and Willmer, C. N. A.: 2010, *ApJl* **719(2)**, L126

Tremonti, C. A., Heckman, T. M., Kauffmann, G., Brinchmann, J., Charlot, S., White, S. D. M., Seibert, M., Peng, E. W., Schlegel, D. J., Uomoto, A., Fukugita, M., and Brinkmann, J.: 2004, *ApJ* **613(2)**, 898

Valentino, F., Daddi, E., Strazzullo, V., Gobat, R., Onodera, M., Bournaud, F., Juneau, S., Renzini, A., Arimoto, N., Carollo, M., and Zanella, A.: 2015, *ApJ* **801(2)**, 132

van de Voort, F., Bahé, Y. M., Bower, R. G., Correa, C. A., Crain, R. A., Schaye, J., and Theuns, T.: 2017, *MNRAS* **466(3)**, 3460

van de Voort, F. and Schaye, J.: 2012, *MNRAS* **423(4)**, 2991

van der Maaten, L. and Hinton, G.: 2008, *Journal of Machine Learning Research* **9**, 2579

Wang, B. and Wang, X.: 2007, *arXiv e-prints* p. arXiv:0709.1616

White, S. D. M. and Rees, M. J.: 1978, *MNRAS* **183**, 341

Woo, J., Dekel, A., Faber, S. M., Noeske, K., Koo, D. C., Gerke, B. F., Cooper, M. C., Salim, S., Dutton, A. A., Newman, J., Weiner, B. J., Bundy, K., Willmer, C. N. A., Davis, M., and Yan, R.: 2013, *MNRAS* **428(4)**, 3306

Wu, P.-F., Zahid, H. J., Hwang, H. S., and Geller, M. J.: 2017, *MNRAS* **468(2)**, 1881

Yuan, T., Nanayakkara, T., Kacprzak, G. G., Tran, K.-V. H., Glazebrook, K., Kewley, L. J., Spitler, L. R., Poole, G. B., Labbé, I., Straatman, C. M. S., and Tomczak, A. R.: 2014, *ApJl* **795(1)**, L20

Zahid, H. J., Dima, G. I., Kudritzki, R.-P., Kewley, L. J., Geller, M. J., Hwang, H. S., Silverman, J. D., and Kashino, D.: 2014, *ApJ* **791(2)**, 130

Zavala, J. A., Casey, C. M., Scoville, N., Champagne, J. B., Chiang, Y., Dannerbauer, H., Drew, P., Fu, H., Spilker, J., Spitler, L., Tran, K. V., Treister, E., and Toft, S.: 2019, *ApJ* **887(2)**, 183