Title
Understanding Large-Scale Structure, Galaxies, and Ionized Gas at z~2−3

Author
Topping, Michael Weber

Publication Date
2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Understanding Large-Scale Structure, Galaxies, and Ionized Gas at $z \sim 2 - 3$

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Astronomy

by

Michael Weber Topping

2020

ABSTRACT OF THE DISSERTATION

Understanding Large-Scale Structure, Galaxies, and Ionized Gas at $z \sim 2-3$

by

Michael Weber Topping
Doctor of Philosophy in Astronomy
University of California, Los Angeles, 2020
Professor Alice Eve Shapley, Chair

Understanding the nature and evolution of the high-redshift universe is crucial in forming complete models of galaxy evolution and large-scale structure formation. In this dissertation I investigate several aspects of the high-redshift universe, including the structure of a massive galaxy protocluster, understanding the nature and interplay between massive stars and ionized gas in high-redshift galaxies. First, I present an analysis of densely sampled spectroscopic observations of galaxies within the SSA22 protocluster at $z \sim 3.09$ which reveal two distinct structures separated in redshift space and segregated on the sky. An analysis of similar structures within cosmological N-body simulations reveals that such a distribution of galaxies can only be explained as two nearby overdensities which will remain distinct as they evolve to $z = 0$. Based on the N-body simulations, I find that the opportunity to observe such a phenomenon is incredibly rare, with an occurrence rate of $7.4h^3\mathrm{Gpc}^3$. In this dissertation I also investigate the differences between local and high-redshift galaxies suggested by the offset towards higher [OIII]λ5007/Hβ and [NII]λ6584/Hα on the 'BPT' diagram. I analyze combined rest-UV and rest-optical spectra of high-redshift galaxies. Crucially, rest-UV spectra provide a powerful constraint on the population of massive stars within high-redshift galaxies, which is an important driver powering the observed rest-optical emission lines.

To investigate this origin of the offset on the BPT diagram, I construct two composite spectra composed of high-redshift galaxies at different locations on the BPT diagram. This analysis demonstrates that high-redshift galaxies that are more offset typically have younger stellar ages, lower stellar metallicities, higher ionization parameters, and are more α-enhanced compared to high-redshift galaxies that lie along the local sequence. In addition, this analysis reveals that even galaxies that are entirely consistent with the local nebular excitation sequence appear to be α-enhanced. This suggests that a similarity in the location of high-redshift and local galaxies in the BPT diagram may not be indicative of a similarity in their physical properties. I further expand upon this analysis by fitting the joint rest-UV and rest-optical properties of individual galaxies in the sample. An important aspect of analyzing individual galaxy spectra is a quantitative limit on how well we can fit the spectra. By introducing noise to model galaxy spectra which has known properties, and binding its best-fit properties, I determine that galaxy properties can be accurately reproduced if the spectrum has a SNR > 4. The best-fit properties of individual galaxies in our sample reveals that they have comparable ionization parameters to those of local HII regions the share the same nebular metallicity. In addition, I find that all galaxies in the sample show evidence for being α-enhanced resulting in harder ionizing spectra compared to local galaxies. These results point toward the observed offset on the BPT diagram being primarily caused by a harder ionizing spectra at fixed nebular metallicity.

The dissertation of Michael Weber Topping is approved.

Matthew A. Malkan

Smadar Naoz

Tommaso L. Treu

Alice Eve Shapley, Committee Chair

University of California, Los Angeles

2020

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I wish to acknowledge, in appreciation, my research advisor Professor Alice Shapley for the educational environment she has fostered. I wish to extend special thanks to those of Hawaiian ancestry on whose sacred mountain we are privileged to be guests. Without their generous hospitality, most of the observations presented herein would not have been possible.

Chapter 2 of this dissertation is a version of Topping et al. (2016) and is reproduced by permission of the AAS. Support for this work was provided by NSF grants AST-0908805 and AST-1313472 and the David & Lucile Packard Foundation.

Chapter 3 of this dissertation is a version of Topping et al. (2018) and is reproduced by permission of the AAS. Support for this work was provided by the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) and the Partnership for Advanced Supercomputing in Europe (PRACE, www.prace-ri.eu) for funding the MultiDark simulation project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre (LRZ, www.lrz.de). The Bolshoi simulations have been performed within the Bolshoi project of the University of California High-Performance AstroComputing Center (UC-HiPACC) and were run at the NASA Ames Research Center.

Chapter 4 of this dissertation is a version of Topping et al. (2019) and is reproduced by permission of MNRAS.

Chapter 5 of this dissertation is a version of a soon to be published manuscript reproduced by permission of MNRAS.

VITA

2014	B.A., Astrophysics, University of Colorado, Boulder.
2016	M.S., Astronomy and Astrophysics, University of California, Los Angeles.
2014-2019	Teaching Associate/Assistant/Fellow, Department of Physics and Astronomy, University of California, Los Angeles.
2014-Present	Graduate Student Researcher, Department of Physics and Astronomy, University of California, Los Angeles.

PUBLICATIONS

"The Efficiency of Stellar Reionization: Effects of Rotation, Metallicity, and Initial Mass Function." **Topping, Michael W.**, Shull, J. Michael The Astrophysical Journal, Volume 800, Issue 2, article id. 97, 15 pp. (2015)

"Substructure within the SSA22 Protocluster at z$\sim$ 3.09." **Topping, Michael W.**, Shapley, Alice E.; Steidel, Charles C. The Astrophysical Journal Letters, Volume 824, Issue 1, article id. L11, 6 pp. (2016)

"Understanding Large-scale Structure in the SSA22 Protocluster Region Using Cosmological Simulations." **Topping, Michael W.**, Shapley, Alice E.; Steidel, Charles C.; Naoz, Smadar, Primack, Joel R. The Astrophysical Journal, Volume 852, Issue 2, article id. 134, 12 pp. (2018)

"The MOSDEF-LRIS Survey: The Interplay Between Massive Stars and Ionized Gas in High-Redshift Star-Forming Galaxies." **Topping, Michael W.**, Alice

E. Shapley, Naveen A. Reddy, Ryan L. Sanders, Alison L. Coil, Mariska Kriek, Bahram Mobasher, Brian Siana; *Submitted to MNRAS* (2019)

CHAPTER 1

Introduction

The high-redshift universe hosts much more activity compared to the local universe. At the peak epoch of star formation $z \sim 2-3$, typical star-forming galaxies are forming stars at a rate $1-2$ orders-of-magnitude greater than in local galaxies (Madau & Dickinson 2014). It is at this time of rapid evolution when many galaxy property relations are set. Furthermore, galaxy clusters, the largest bound objects in the local universe, are beginning to collapse, providing a view into their formation. In this dissertation I present an investigation into apparent substructure seen within a progenitor of a massive galaxy cluster. In addition, I present an analysis of the internal properties of high-redshift star-forming galaxies including the interplay between their massive star populations and ionized interstellar medium (ISM). I then examine how these properties differ between high-redshift and local galaxies, and across the high-redshift sample.

1.1 Large-scale structures at high redshift

The most massive galaxy clusters represent the largest gravitationally bound structures in the universe. At higher redshift, before they have fully virialized, these clusters exist as protoclusters, consisting of a large number, but diffuse collection of galaxies. Despite their loose configuration, and due to the high overdensities in which protoclusters exists, protoclusters have been identified out to $z \sim 6$ (Toshikawa et al. 2014). There are currently many techniques for finding high-redshift protoclusters, including the serendipitous identification of redshift

overdensities within spectroscopic surveys of Lyman Break Galaxies (LBGs), Lyα emitters (LAEs) or other magnitude-limited galaxy samples (Steidel et al. 1998, 2003, 2005; Harikane et al. 2017; Chiang et al. 2015; Lemaux et al. 2014), targeted searches for LAEs around radio galaxies (e.g. Venemans et al. 2007), and Lyα forest tomography (Lee et al. 2016). Based on these several methods, the number of known $z > 2$ protoclusters has grown dramatically over the past decade. Studying the key high-redshift epoch of structure formation, when the clusters are still collapsing, helps to give us a more complete picture of massive galaxy clusters and their environments at $z = 0$.

Protoclusters are ideal laboratories for studying not only the formation end evolution of large-scale structure, but also a vast array of extragalactic phenomena across key cosmic epochs. Among these are LAEs and LBGs, Quasars, submillimeter galaxies (Umehata et al. 2015; Geach et al. 2016; Hayatsu et al. 2017), Lyα Blobs (LABs; Matsuda et al. 2011; Geach et al. 2005, 2016), radio galaxies, as well as some of the earliest detections of diffuse intercluster light (Wang et al. 2016). It is important to study the environments in which these objects live in order to understand the effects that impact their formation and evolution. The study of environmental effects on galaxies in protoclusters is aided by the large number of coeval galaxies they contain.

In addition to allowing the study of protocluster members, the protoclusters themselves inform our understanding of the nature and evolution of large-scale structure. Because protoclusters are the largest gravitationally bound structures, and observable to high redshift, they provide a powerful constraint on structure formation models. These models make predictions for observed properties such as mass and size, as well as the spatial distribution of progenitor halos within the protocluster. An interesting implication of significant protocluster substructure is how its presence affects the detection of massive overdensities. The presence of nearby structure may boost the significance of observed overdensities, and thus bias the samples of discovered protoclusters.

1.2 Massive stars and ISM conditions at high redshift

In order to have a complete picture of galaxy evolution, we must have an understanding of the physical conditions within high-redshift galaxies. Large 8-10m class telescopes with sensitive multiplexed instruments have allowed for the detailed study of the galaxy population at high redshift. Observations of local galaxies revealed a tight sequence of increasing [NII]λ6584/Hα and decreasing [OIII]λ5007/Hβ. Similar results of high-redshift galaxies suggested that there were differences within the physical conditions of high-redshift galaxies compared to their local counterparts. In particular, observations using Keck/NIRSPEC found that high-redshift galaxies trace a sequence toward higher [NII]λ6584/Hα and decreasing [OIII]λ5007/Hβ on the BPT diagram. Several potential caueses, such as increased densities, changes in abundance patterns, contribution of AGNs and shocks, and a higher ionizing spectrum at fixed oxygen abundance.

While statistical samples ($\sim$ 1000) of high-redshift galaxies observed in the rest-optical have revealed a wealth of information about the physical conditions within galaxies at high redshift, the exact differences between the internal properties of high-redshift and local star-forming galaxies are still unknown. As the exact shape of the ionizing spectrum is intimately linked with the observables surrounding HII regions, it is imperative to obtain a solid understanding of the sources of ionizing radiation, massive stars.

Directly observing the ionizing spectrum within high-redshift galaxies is challenging, however much can be learned about it with knowledge of the massive star population. In star-forming galaxies, the UV spectrum ($\sim 1000\mathring{A} - 3000\mathring{A}$ is dominated by massive stars. In the local universe atmospheric absorption removes the ability to observe this regime from the ground, making observations more restricted. However, at high redshift the situation is improved because this region of the spectrum is redshifted into optical wavelengths, making it readily available to multiplexed instruments on large ground-based telescopes.

In this dissertation I analyze combined deep rest-UV and rest-optical spectra for 260 galaxies at $2.0 \leq z \leq 2.7$. With this data I am able to directly constrain the ionizing radiation field within these galaxies, allowing us to accurately model the observed rest-optical emission lines using photoionization models. In addition, I investigate the underlying cause for the offset of high-redshift galaxies on the BPT diagram by testing how high-redshift galaxies are different from their local counterparts, as well as finding differences within the high-redshift galaxy population.

This dissertation is organized as follows. In Chapter 2, I present the detection of apparent substructure within a massive protocluster at $z = 3.09$ in the SSA22 field. Chapter 3 further investigates the substructure in the context of the ΛCDM model, and attempts to characterize the substructure and predict its properties as it evolves to $z = 0$. Chapter 4 presents a combined rest-UV and rest-optical analysis of composite spectra at $2.0 \leq z \leq 2.7$ in order to understand the connection between massive stars and ionized gas, and examines the differences between high-redshift and local galaxies. Chapter 5 expands the combined rest-UV and rest-optical analysis aimed at understanding the differences between internal properties of local and high-redshift galaxies by studying these properties within individual high-redshift galaxies. Finally, Chapter 6 summarizes the results of these studies and presents some future work.

CHAPTER 2

Substructure Within the Concentration of Redshift $Z \sim 3.09$ Galaxies in the SSA22A Field

2.1 Introduction

Determining the evolution of the largest clusters of galaxies is crucial to understanding the formation of large scale structure of the universe. Steidel et al. (1998) originally found a concentration of 16 galaxies at $z \sim 3.09$ within observations of 67 Lyman break galaxies (LBGs) at $z \geq 2$ in the SSA22a field. This concentration corresponds to an overdensity of $\delta_{gal} = 5 \pm 1.2$ (Steidel et al. 2000). From this overdensity, calculated the mass of the protocluster to be $M \sim 10^{15} M_{\odot}$, consistent with being a progenitor of a Coma-like cluster at $z = 0$. Cosmological structure formation theory predicts how massive galaxy clusters form, whether it is either slowly accreting halos or the results of major mergers (White & Rees 1978; Lacey & Cole 1993; White & Frenk 1991). Additionally, the existence and rarity of massive galaxy clusters is an important metric in order to determine parameters of large scale structure evolution (Mandelbaum et al. 2013; Kauffmann et al. 1993). Simulations of structure evolution provide a useful comparison to observations of the most massive $M > 10^{15} M_{\odot}$ galaxy cluster progenitor candidates.

LBGs provide a useful tracer of structure at higher redshifts, as they have a well defined photometric selection function and thus a measured overdensity is quantifiable (Kauffmann et al. 1999). Furthermore, LBGs trace out structure within which one can look for Lyα emitters (LAEs) using finely tuned narrowband

filters. Previous studies of this region have shown a similar overdensity for both LAEs and LBGs, despite the objects having been selected in different manners (Steidel et al. 2000). In addition, the region hosts an overdensity of active galactic nuclei (AGNs), as well as a large number of Lyα blobs (Matsuda et al. 2004; Lehmer et al. 2009a,c).

We investigate the structure of this overdensity with high spatial density spectroscopic observations, in order to better characterize the environment of an expected $M \gtrapprox 10^{15} M_{\odot}$ cluster progenitor. Our sample includes spectroscopic observations of more objects than previous studies, which allows us to better determine the significance of substructures present in the field. In order to obtain a detailed view of the dynamical state of the protocluster, it is important to trace the redshift distribution in detail. We find that the galaxies in this overdensity are arranged in a bimodal distribution in redshift space, suggesting that two separate structures are present. Additionally, observing the distribution on the sky shows segregation between objects in each of the two peaks of the distribution.

In § 2.2, we first describe our observations and data reduction, as well as the measurement of redshifts for our objects. In § 2.3 we describe the structures seen in redshift and physical space of the region. Finally, in § 2.4 we present our conclusions and a discuss implications of our work for galaxy protocluster evolution. We use cosmological parameters of $H_0 = 67.8$ km s^{-1} Mpc^{-1}, $\Omega_M = 0.3089$, and $\Omega_\Lambda = 0.6911$. (Planck Collaboration et al. 2015).

2.2 Observations/Methods

2.2.1 Data

The objects in our sample are found using methods to select both LAEs and LBGs. The LAEs were selected using deep broadband BV imaging from Subaru/Suprime-cam and narrowband 4980 Å imaging from Keck/LRIS (Oke et al. 1995) and

Subaru/Suprime-cam. A full description of these observations, including the methods of reduction can be found in Nestor et al. (2011, 2013). We obtain a list of LBGs from catalogs presented in Steidel et al. (2003), who created the catalogs using UGR imaging and photometric selection criteria to obtain LBGs at $z \sim 3$.

Keck/LRIS spectroscopy for the LBGs in our sample was obtained from several observing campaigns utilizing different instrumental setups. These include the original LBG survey of Steidel et al. (2003), for which galaxies were observed using the 300 line mm^{-1} grating prior to the installation of the dichroic capability in the LRIS spectrograph. Other spectroscopy was obtained using the 400 line mm^{-1} grism while using the dichroic capabilities of LRIS (Shapley et al. 2006). Further spectroscopy on LBG and LAE photometric candidates was obtained utilizing the setups described in Nestor et al. (2013). In brief, these setups include seven slit masks using the 300 line mm^{-1} grism, two slit masks using the 600 line mm^{-1} grism, and one mask of deep observations using the 400 line mm^{-1} grism. All of these setups also include the use of a dichroic beam splitter.

2.2.2 Redshift Measurements

We measured the redshifts of the objects in our sample using both Lyα emission, and if present, metal absorption lines. For measuring the redshifts, we developed code to systematically fit a line profile to either emission or absorption lines. If present, we first made an estimate for the position of the Lyα emission. This estimate is used by the code to cut out part of the spectrum near the line, and becomes an initial guess for the fit of the line. We first fit a single Gaussian function to the Lyα emission. In some cases, this did not provide an adequate fit, and we were required to add in a second Gaussian function to produce an acceptable fit. This is the case for a Lyα emission that exhibits a double-peaked morphology. In total, we measured redshifts for 202 galaxies, including 116 LAEs and 86 LBGs. Due to our interest in the $z \sim 3.09$ structure, we only select galaxies that are

within $3.05 \leq z \leq 3.12$. All other galaxies are considered to be not associated with the structure. This cut reduces our sample to 146 galaxies, comprising 106 LAEs and 40 LBGs.

Objects showing double-peaked Lyα emission typically shows one of two different morphologies. The emission either showed two emission lines with comparable amplitude, or presented a secondary peak at bluer wavelengths that was much smaller. We used the same method to measure the redshift for both types of double-peaked Lyα emission, which was to fit the trough between the peaks, and the redshift was calculated using the wavelength of the center of the trough. Some objects also exhibited absorption lines. To measure the redshift of the absorption lines, we measure the redshift of each line individually, then averaged them to obtain a more accurate redshift. Some objects had repeat observations, and in this case we use the redshift measured from the observation with the highest signal to noise ratio (SNR). Additionally, combining repeated observations of the same object did not alter the redshift measurements by greater than our measured error.

In order to investigate the substructures that these galaxies populate, we want to find their systemic redshifts relative to each other. The Lyα emission typically traces gas that is outflowing from the galaxies, which will shift the redshift of these objects away from their actual positions. We calculate the systemic redshifts of objects in our sample by shifting the redshift of measured spectral features to the rest frame of the galaxies. The shift required to calculate the systemic redshift has been measured to be different for LAEs and LBGs, and depends on which spectral features are observed. The adjustments we used for each of these cases are described below. We shift the redshift of objects that are classified as LAEs with only Lyα emission by $\delta v = -200$ km s^{-1} (Trainor et al. 2015). We shift the redshift of LAEs showing both Lyα emission and interstellar absorption lines by $\delta v = 0.114\Delta v_{abs,em} - 230$ km s^{-1}, where $\Delta v_{abs,em}$ is the velocity difference between the Lyα and interstellar absorption redshifts. We shift the redshift of

objects classified as LBGs showing Lyα emission by $\delta v = -310$ km s^{-1}, and finally, we shift the redshift of LBGs showing only interstellar absorption lines by $\delta v = 150$ km s^{-1} (Adelberger et al. 2003). To apply the systemic correction, we convert the shifts from velocity space to a redshift space using: $\delta v/c = \delta z/(1+z)$. Once these systemic corrections are made, we have the final list of objects with systemic redshifts, along with the positions on the sky.

2.3 Results

2.3.1 Redshift Histogram

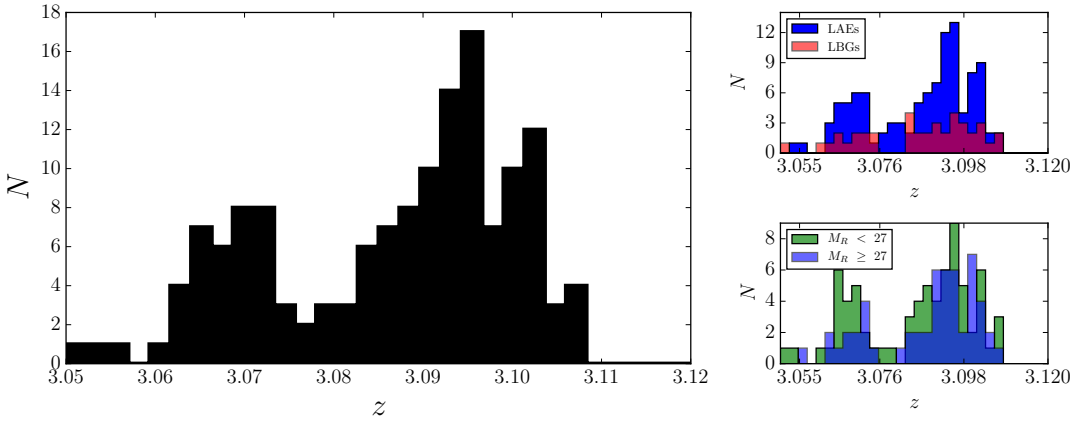

Figure 2.1: (left) Redshift histogram for objects at $z \sim 3.09$ in the SSA22a field. In this histogram we can see the double peaked structure, with one peak (red) centered at $z = 3.095$, and the other peak (blue) centered at $z = 3.069$. The histogram contains a total of 146 galaxies comprising 106 LAEs and 40 LBGs. (top right) Distribution of LAEs and LBGs in the substructure of the SSA22a field. (bottom right) Distribution of brighter ($M < 27$) and fainter ($M \geq 27$) galaxies in the substructure of the SSA22a field. We find no significant difference in the distributions of the brighter and fainter galaxies within the structures.

With the finalized list of redshifts in the region, we construct a histogram from the finalized list of redshifts in the SSA22a field (Figure 2.1). The histogram consists of 146 galaxies with measured redshifts. We see that the spike discovered in Steidel et al. (1998) at $z \approx 3.09$ contains two distinct peaks, one with a central redshift of $z \approx 3.07$ (blue peak), and another with central redshift of $z \approx 3.095$ (red peak). We fit a Gaussian to each of the two peaks. We define the boundary

between the two peaks to be the bottom of the trough of the distribution. With this boundary, we find that the blue peak contains of 42 galaxies, and the red peak contains 104 galaxies. In velocity space, the width of the blue peak is: $\sigma_b = 350$ km s^{-1}, and the width of the red peak is: $\sigma_r = 540$ km s^{-1}. The extent of the whole structure spans a velocity range of $\Delta_{z,tot} = 0.05$, which corresponds to: $\Delta_{tot} = 3670$ km s^{-1}.

We also investigate the distributions of LAEs and LBGs in the two peaks. Figure 2.1 displays the separate histograms of both LAEs and LBGs, and both distributions show evidence of double peaked structure. LBGs [LAEs] comprise 27% [73%] of the red peak, and 28% [72%] of the blue peak. Since LBGs and LAEs typically exist in different luminosity space, this suggests that the luminosity function in the two peaks are consistent. We have V-band data of $\sim$ 85% of galaxies from Subaru.

We classify each galaxy as either brighter or fainter, depending on whether it has $M_V < 27$, or $M_V \geq 27$ respectively. We perform a similar analysis, by plotting a histogram of brighter and fainter galaxies separately (Figure 2.1). We find that there is no evidence that brighter or fainter galaxies preferentially appear in either of the two redshift peaks. Subsequently, because more luminous galaxies are more strongly clustered (Ouchi et al. 2004; Lee et al. 2006), we will consider the absence of segregation between brighter and fainter galaxies in the context of galaxy modeling. As a final comparison, we find that galaxy exhibiting absorption lines, Lyα, or both are equally proportioned across the peaks, as well as the distribution of Lyα showing a double peaked morphology.

We calculate the virial mass of each peak in the redshift distribution.

$$M^{vir} \approx \frac{5 < R > \sigma^2}{G} \tag{2.1}$$

We find a virial mass for the red peak of $M_r^{vir} \approx 10^{15.39} M_\odot$, and a virial mass of the blue peak of $M_b^{vir} \approx 10^{14.92} M_\odot$. For this calculation we adopt the field

Table 2.1. Redshift Histogram Fit Parameters

Peak	N	Center $[z]$	σ $[\Delta z]$	σ $[\Delta v$ km s$^{-1}]$
Red	104	3.095 ± 0.001	$(7.37 \pm 0.54) \times 10^{-3}$	540
Blue	42	3.069 ± 0.001	$(4.73 \pm 0.72) \times 10^{-3}$	350

of view of our observations for the radius, and values for the velocity dispersion are measured from the redshift histogram and are given in Table 2.1. This is not necessarily an accurate calculation for the mass of each distribution for several reasons. Most importantly, it is likely that this structure is not virialized. Due to the early evolutionary state of the cluster, the velocity dispersion σ will provide an incorrect value for the mass. We expect that the structure is still collapsing, and as a result, the velocity dispersion is lower than what we expect for this protocluster when it is virialized. This leads to an underestimation of the mass. Another error with this calculation is the determination of the size of the protocluster. Yamada et al. (2012) show that the overdensity of LAEs extends beyond the area of our observed field.

2.3.2 Scatter Plots

In addition to the redshift information, we obtain spatial information of the objects using Subaru/Suprime-Cam observations (Nestor et al. 2011; Miyazaki et al. 2002). Using `iraf` tasks, we centroid the positions of all objects to get accurate positions of our objects using the same world coordinate system (WCS) solution.

The color of each point represents the redshift, where red points are members of the red peak, and blue points are members of the blue peak. The redshift of the trough between the two peaks is represented by the color white. Galaxies in the blue [red] peaks are represented by triangles [squares].

Figure 2.2 (bottom left) displays the positions of objects on the sky. We

determine the physical separation of objects on the sky by calculating the angular scale at the redshift of each object. We calculate the angular scale using the angular distance (Hogg 1999). We calculated the comoving Mpc axes in Figure 2.2 using the angular scale at $z = 3.09$, and results in a scale of: 7.79 kpc/″.In comoving space, our observations span $\approx$ 10 comoving Mpc on a side in R.A. and decl., however in the line-of-sight direction the objects appear to be spread out over $\sim$ 40 Mpc. This difference could be due to one of a few factors. First, the field-of-view of our observations does not span the entire space of the cluster (Yamada et al. 2012). Additionally, these observations could suffer from redshift space distortions, which would cause the appearance of elongation in space along the line-of-sight direction (Kaiser 1987).

In order to search for a spatial separation of galaxies in different redshift peaks, we split the positions of objects in our sample into quadrants, represented by grey dashed lines in Figure 2.2 (bottom left). We count the number of galaxies in each quadrant that are within either the red or blue redshift peak. The number of galaxies in the red or blue peak is listed next to each quadrant in Figure 2.2. We find that the upper left quadrant is strongly deficient in lower redshift galaxies, and the bottom left quadrant has a slight deficiency in lower redshift galaxies. This result suggests that galaxies in either the red or blue peak do not cover the same area. This could suggest that we are seeing the edge of the structure containing the galaxies in the blue peak. The top and right panel of Figure 2.2 show the R.A. vs. z and z vs. decl. respectively. On both panels the absence of galaxies in redshift space between the two peaks is apparent. In the top panel, we observe a slight increase of objects towards lower R.A.

2.4 Discussion

We have observed 146 LAEs and LBGs at $z \approx 3.09$ present in the high-density field SSA22. We recognize that these galaxies exhibit a double peaked structure

in redshift space, suggesting that this could be the merger of two massive galaxy clusters. The presence of peculiar velocities makes the translation between redshift and physical space non-trivial. One question is whether the substructure in the SSA22a redshift histogram actually suggests physical substructure. However, the segregation on the sky of galaxies in each redshift peak suggests that the two peaks in the SSA22a redshift distribution are physically distinct structures.

As stated previously, our observations do not cover the full extent of the SSA22a protocluster. Yamada et al. (2012) survey 4980 arcmin^2 in the SSA22a field, and find 91 LAEs at a redshift consistent with the $z \sim 3.09$ structure. There are 19 LAEs that overlap with our sample, resulting in 72 new objects for us to add to our sample. Figure 2.3 displays objects from both our observations, and those published in Yamada et al. (2012). With the inclusion of additional galaxies, the trends in redshift across the field are still present, namely, galaxies toward higher R.A. typically have larger redshifts compared to objects toward lower R.A. We also add those galaxies in Yamada et al. (2012), that are not already in our sample to our redshift histogram. Figure 2.3 shows the resulting distribution. We find that the double peaked structure is still present with the inclusion of additional galaxies. Due to the large spatial extent that these additional galaxies span, the presence of double peaked structure suggests that this structure extends throughout the protocluster, and is not localized to the field-of-view of our observations.

In addition to a separation of high and low redshift objects in the field, Matsuda et al. (2005) finds large filamentary structures in SSA22a. The filaments appear to connect at the region of highest galaxy density. We expect galaxies to follow the filamentary structures, and the orientation of the filaments would create a redshift gradient along the length of a filament. Figure 2.2 shows the presence of a gradient, resulting in changing number of low or high redshift galaxies in different regions of the field. The redshift gradient is also apparent with the inclusion of Yamada et al. (2012) objects. This provides further evidence of the existence of a filamentary structure in the field. We find a lack of galaxies at $z \approx 3.08$, resulting

in the trough in the histogram in Figure 2.1.

The filamentary structure in Matsuda et al. (2005) exhibits a gradient in redshift space, spanning $3.088 \leq z \leq 3.108$. The galaxies in this filament are contained within the red peak in Figure 2.1. Matsuda et al. (2005) notes a smaller structure at lower redshift ($z \sim 3.075$), which could correspond to galaxies in our blue peak. The offset in this smaller structure from our blue peak is likely due to their lack of systemic redshift correction, which would lower the redshifts of the galaxies by ~ 0.003 and put this structure into closer agreement with our blue peak. Figure 2.4 shows the positions of galaxies in our red peak on the sky. We find that galaxies at higher R.A. and decl. are at lower redshift. This is the opposite trend that is seen in Matsuda et al. (2005). This is likely due to our sample having more galaxies, therefore we have higher spatial resolution of any present structures.

We focused on the densest region of the protocluster, which is likely sufficient to determine the existence of structures because nearly all observed galaxies are members of the structures (Muldrew et al. 2015). We obtained results by observing the central $\sim 5'.5 \times 7'.6$ of the protocluster, which results in a high purity, $> 90\%$, of our sample; that is, most of our galaxies are members of the protocluster, and not coincident field galaxies at this redshift (Muldrew et al. 2015). A further study of this region would investigate protocluster members further away from the core of this cluster, and would allow a better characterization of the dynamics of structures in the protocluster (Hayashino et al. 2004). Additionally, observing galaxies across the full extent of the cluster would help characterize filamentary structures in the field.

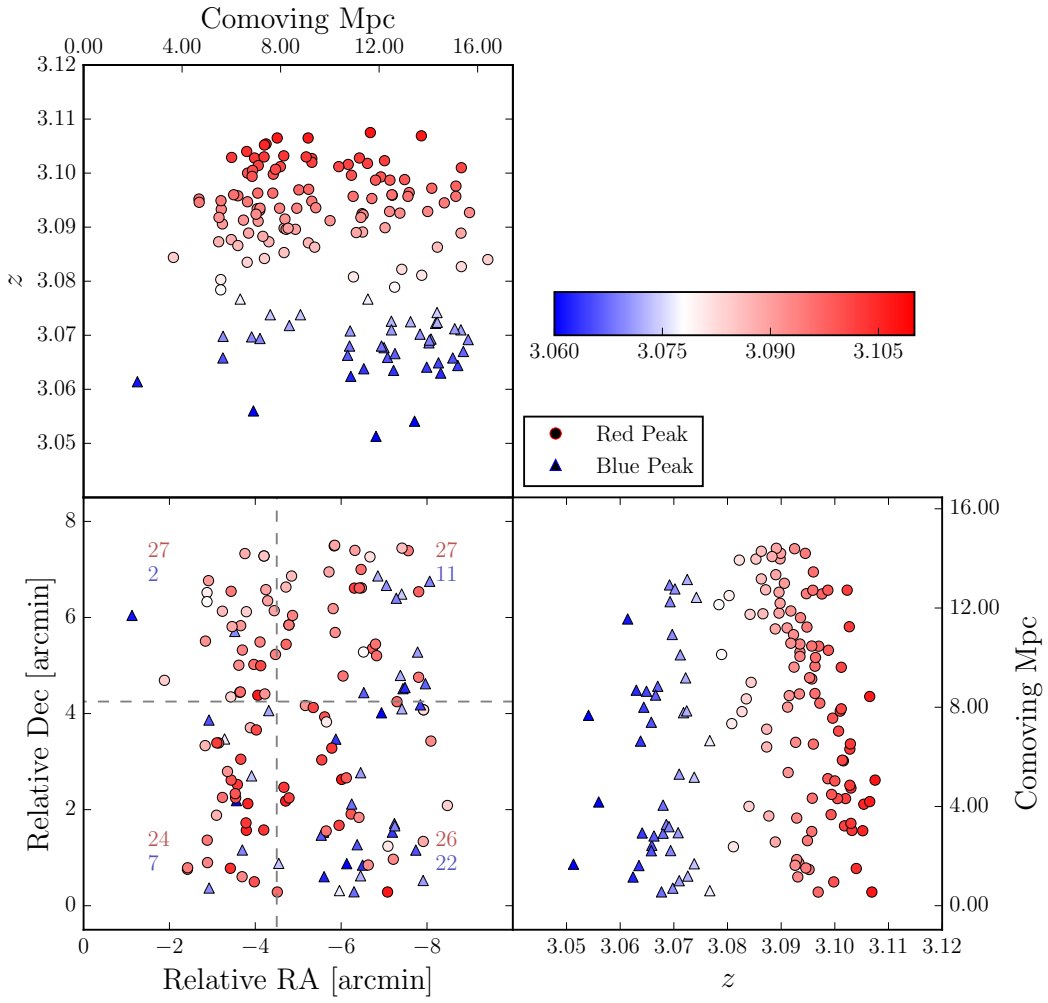

Figure 2.2: Bottom left panel: Locations on the sky of all of our observed galaxies. The blue triangles represent galaxies contained in the peak on the histogram centered at $z \approx 3.07$, and the red circles are galaxies contained within the peak centered at $z \approx 3.09$. The field of view is approximately $5'.5 \times 7'.6$ which corresponds to $10.5 \times 14.5\ h^{-1}$ Mpc. The color of each point corresponds to the redshift of each object, with red points being at a higher redshift, and blue points at a lower redshift. White on the colorbar corresponds to the bottom of the trough between the peaks in the redshift histogram, and is meant to be a rough demarcation between the objects in each peak. The galaxies are separated into four quadrants, displayed as the grey dashed lines. The number of galaxies in either the red or blue peak, in each quadrant is displayed by a red or blue number respectively. Top panel: Galaxies plotted as R.A. vs. z. Here the absence of galaxies between the red and blue peak is clearly visible. There is a slight preference for blue galaxies to be present at lower R.A.. Right panel: Galaxies plotted as z vs decl.. There is no significant difference in red or blue galaxies to lower or higher decl.

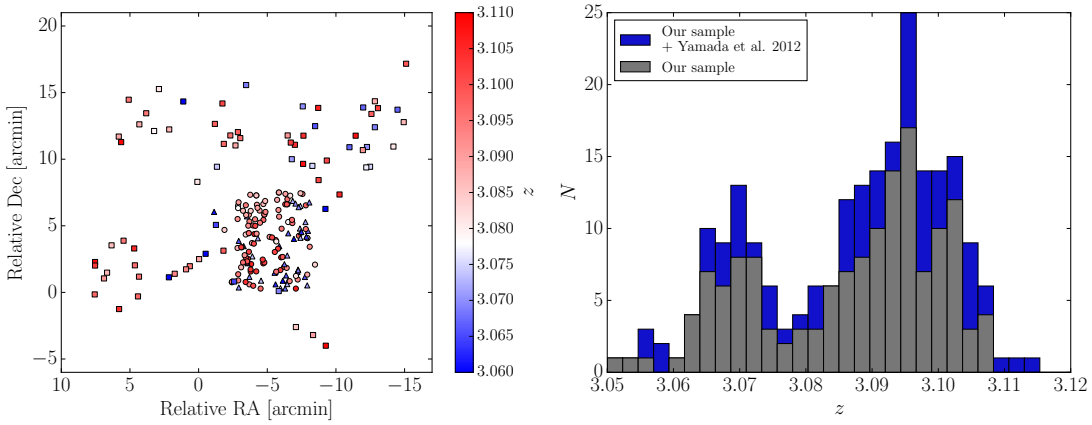

Figure 2.3: (left) Scatter plot of objects in our sample, combined with those presented in Yamada et al. (2012). The absence of lower redshift (blue points) galaxies toward the left side of the figure is apparent with both objects in our sample, as well as that of Yamada et al. (2012). (right) Redshift histogram including objects reported in Yamada et al. (2012). The double-peaked structure is present in the distribution of both samples.

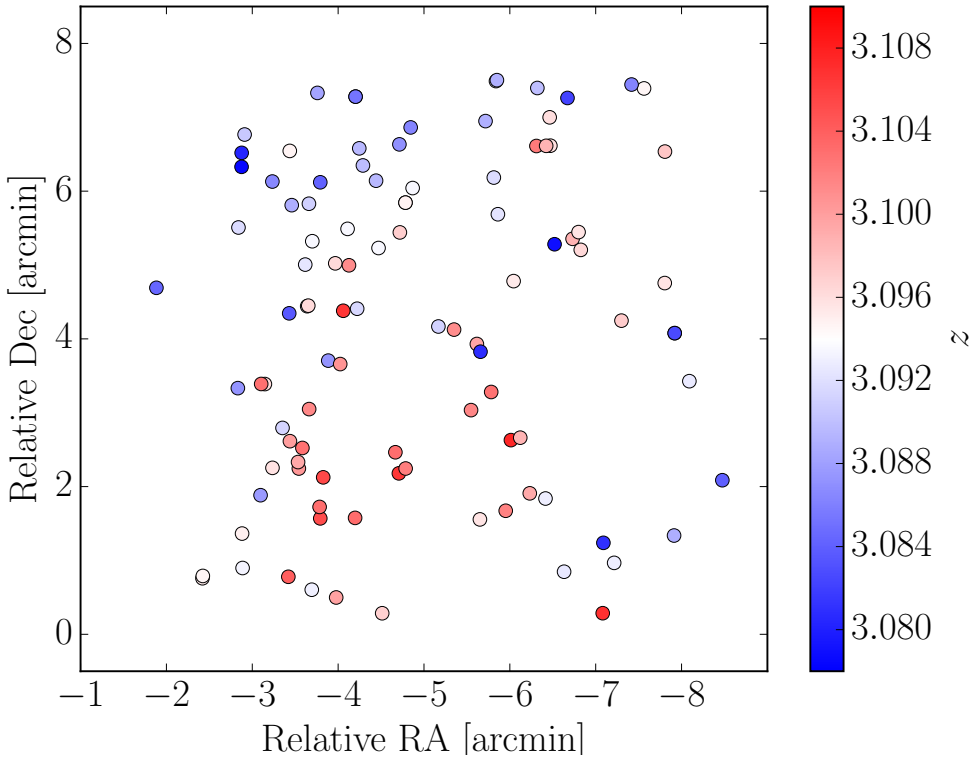

Figure 2.4: Galaxies contained in the red peak of the redshift distribution. These galaxies exhibit a gradient in redshift space from the upper left to the bottom right.

CHAPTER 3

Understanding Large-scale Structure in the SSA22 Protocluster Region Using Cosmological Simulations

3.1 Introduction

As the largest gravitationally bound structures, galaxy clusters are ideal objects for probing the formation of large scale structure in the universe. Due to their extreme nature, galaxy clusters and protoclusters are an optimal setting to study the effects of environment on galaxy formation and evolution. The progenitors of todays galaxy clusters, i.e. "protoclusters" have been identified all the way out to $z \sim 6$, using a variety of techniques (Toshikawa et al. 2014). The study of galaxy clusters and protoclusters is further aided by the multiple techniques that have been developed in order to find them.

There are currently many techniques for finding high-redshift protoclusters, including the serendipitous identification of redshift overdensities within spectroscopic surveys of Lyman Break Galaxies (LBGs), Lyα emitters (LAEs) or other magnitude-limited galaxy samples (Steidel et al. 1998, 2003, 2005; Harikane et al. 2017; Chiang et al. 2015; Lemaux et al. 2014), targeted searches for LAEs around radio galaxies (e.g. Venemans et al. 2007), and Lyα forest tomography (Lee et al. 2016). Based on these several methods, the number of known $z > 2$ protoclusters has grown dramatically over the past decade. Studying the key high-redshift epoch of structure formation, when the clusters are still collapsing, helps to give

us a more complete picture of massive galaxy clusters and their environments at $z = 0$.

Steidel et al. (1998) reported the discovery of the SSA22 galaxy protocluster at $z = 3.09$ within a large survey of $z \sim 3$ LBGs, and measured an overdensity of $\delta_{gal} = 3$, with the expectation of the overdensity evolving into a massive Coma-like cluster with a mass of $M \sim 10^{15} M_{\odot}$ by $z = 0$. Based on an expanded dataset, Steidel et al. (2000) obtained a revised estimate for the overdensity of $\delta_{gal} = 6.0 \pm 1.2$. Since then, the area surrounding the $z \sim 3.09$ overdensity has been observed through multiple observing campaigns spanning from radio to X-ray wavelengths. These studies have revealed tens of Lyman alpha blobs (Matsuda et al. 2011; Geach et al. 2005, 2016), and multiple X-ray sources (Lehmer et al. 2009b; Geach et al. 2009). Additional studies include deep ALMA observations in the central region of the protocluster (Umehata et al. 2015; Geach et al. 2016; Hayatsu et al. 2017), near-infrared spectroscopic observations of massive red K-band-selected galaxies (Kubo et al. 2015), and high-resolution *Hubble Space Telescope* imaging (Chapman et al. 2004).

In addition to the extensive multi-wavelength studies of SSA22, followup spectroscopic observations have revealed details about structure within the overdensity. Matsuda et al. (2005) mapped the three-dimensional structure of LAEs in and around the protocluster, and reported evidence for large-scale filamentary structure. Topping et al. (2016) showed the existence of two distinct groups of galaxies, both LAEs and LBGs, separated both on the sky and in redshift space, and observed as a double-peaked redshift histogram. This structure was discovered by focusing on the highest density region of the protocluster, but remains persistent when the observed region is expanded (Topping et al. 2016; Yamada et al. 2012). From these studies it is unclear what the evolution and fate of the $z \sim 3.09$ protocluster and its surrounding structure will be down to $z = 0$. In particular, we would like to understand if these structures will coalesce, or remain distinct throughout their evolution.

Cosmological N-body simulations provide a useful tool for studying the evolution of large scale structure. Recently, the increase in computational power leads to cosmological simulations with both higher resolution extending down to lower-mass halos, and larger volumes including the largest, rarest structures in the universe. These advances, combined with the availability of easily searchable halo catalogs and merger trees, enable us to use simulations to understand the underlying physical structures observed in SSA22, and how they evolve to the present day.

In this chapter, we further investigate the nature of the large-scale structure presented by Topping et al. (2016). We utilize the halo catalogs and merger tree information from the Small MultiDark Planck (SMDPL) dark matter simulation (Rodríguez-Puebla et al. 2016; Klypin et al. 2016; Behroozi et al. 2013b), which has sufficient resolution and simulation volume to compare multiple simulated protoclusters with our observations. We examine massive overdensities at the redshift of the SSA22 protocluster in order to understand the intrinsic physical structure giving rise to the observed structure at $z \sim 3$, and what such structure evolves into by $z = 0$. Section 3.2 describes our observations and the calculation of an updated overdensity and mass estimate for the SSA22 protocluster based on current spectroscopic data. Section 3.3 describes the cosmological simulation used to interpret the SSA22 observations, and the methods used to compare it to the observations. Section 3.4 shows the results of a comparison between the observations and simulations. Finally, Section 3.5 discusses an analytic approach to understanding the results from the simulation, and a calculation of the cosmic abundance of large-scale structure similar to the observed structure in SSA22. This paper adopts a cosmology of $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $n_s = 0.96$, $\sigma_8 = 0.8228$, and $h = 1.0$, unless otherwise stated. We also use the abbreviation cMpc for "comoving Mpc."

3.2 SSA22 Field

3.2.1 Data

Our sample consists of LBGs and LAEs with spectroscopic measurements within a $9' \times 9'$ region of the SSA22 field, centered on R.A.=22:17:34, decl.=00:15:04 (J2000), as described by Steidel et al. (1998). The LBGs in our sample were selected as part of the survey of $z \sim 3$ star-forming galaxies presented in Steidel et al. (2003). The LAEs were first identified using broadband BV imaging from Subaru/Suprime-cam in addition to narrowband imaging from Keck/LRIS and Subaru/Suprime-cam using a filter tuned to the wavelength of Lyα at $z = 3.09$ (centered on 4985Å with a bandwidth of 80Å). The LAEs were selected based on BV$-$NB4985 colors indicating a narrowband excess, which ensures a sample of galaxies with large (> 20Å) Lyα EWs at redshifts coincident with the central density peak of SSA22 ($3.05 \lesssim z \lesssim 3.12$). The spectroscopic measurements for galaxies in the SSA22 field were obtained using the LRIS spectrograph at the Keck observatory across multiple observing campaigns and instrumental configurations (Steidel et al. 2003; Nestor et al. 2011, 2013). A more detailed description of the redshift determinations can be found in Topping et al. (2016), and further details about the observations and data reduction can be found in Steidel et al. (2003) and Nestor et al. (2011, 2013).

We determined the systemic redshift of galaxies in the SSA22 field by measuring the redshift of Lyα emission, interstellar metal absorption lines, or both, and removing the effects of large-scale gas outflows. We applied the formulas presented in Trainor et al. (2015) for LAEs and Adelberger et al. (2003) for LBGs, to translate from the observed rest-UV emission and absorption redshifts to the true, systemic redshifts. We compiled the resulting systemic redshifts of galaxies within SSA22 into a redshift histogram (Figure 3.1). Galaxies in the SSA22 redshift histogram are clearly separated into peaks centered at $z = 3.069$ (blue) and $z = 3.095$ (red) with widths $\sigma_{z,b} = 0.0047$ and $\sigma_{z,r} = 0.0074$ respectively.

Hereafter, we describe the total, blue, and red regions using the subscripts t, b, and r respectively.

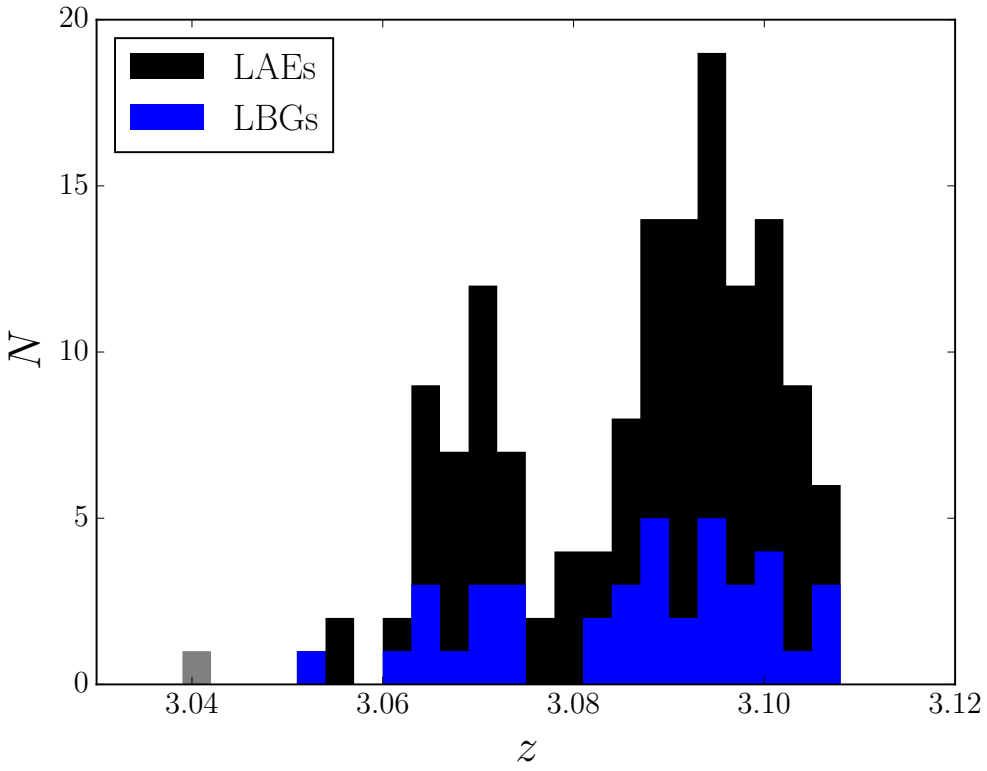

Figure 3.1: Redshift histogram of LAEs and LBGs in the SSA22 field. The double peaked morphology is clearly present with peaks at $z = 3.069$ (blue peak) and $z = 3.095$ (red peak). The blue histogram shows the contribution from the LBGs, and the remaining black histogram is the contribution from LAEs.

3.2.2 Galaxy Overdensity Calculation

The significance of the SSA22 overdensity has been calculated in past work (Steidel et al. 1998, 2000). However, given our significantly larger sample of spectroscopic redshifts in SSA22 (Topping et al. 2016), and the updated LBG redshift selection function (Steidel et al. 2003), it is worth revisiting this calculation. To estimate the galaxy overdensity qualitatively, we compared the number of galaxies contained in the SSA22 redshift spike ($N_{\rm obs}$) with the number of galaxies expected in the same redshift interval from the LBG average selection function ($N_{\rm expect}$). For this

calculation, we restricted $N_{\rm obs}$ to the LBGs in our observed sample and did not consider LAEs, since the LBGs have a well-defined redshift selection function. We define the galaxy overdensity, δ_{gal}, as:

$$\delta_{gal} = \frac{N_{\rm obs}}{N_{\rm expect}} - 1. \tag{3.1}$$

The observed sample used for this calculation includes 82 LBGs in the redshift interval $2.6 \leq z \leq 3.4$. The redshift histogram of these galaxies is shown in Figure 3.2, where the well-known overdensity at $z \sim 3.09$ is clearly visible.

To construct the LBG selection function, we used the sample of LBGs from Steidel et al. (2003), with one key difference. The inclusion of SSA22 galaxies in the sample would increase the value of the selection function within the $z = 3.09$ spike interval, thus biasing the inferred overdensity towards lower values. Therefore, we excluded these galaxies, with 883 redshifts remaining. We fit a spline to the histogram of the remaining galaxies, which resulted in a smooth selection function. Finally, we normalized the selection to the SSA22 redshift histogram, which allowed us to directly compare the number of LBGs in a given redshift interval. Determining the correct normalization is a key step in calculating the galaxy overdensity. Specifically, we normalized the LBG selection function such that its integral over the redshift ranges $2.6 \leq z \leq 3.03$ and $3.12 \leq z \leq 3.4$, was equal to the number of observed galaxies in the SSA22 field in the same redshift intervals. These ranges were chosen to match the number of "field" galaxies in SSA22 and the overall LBG selection function. The resulting selection function is displayed in Figure 3.2 overlaid on the SSA22 LBG redshift histogram.

Using the redshift histogram and the LBG selection function, we computed the galaxy overdensity of SSA22. In detail, we calculated the galaxy overdensity for three components of the SSA22 protocluster: the blue peak, the red peak, and the total volume. We carefully determined the boundaries of the redshift intervals in order to accurately calculate the overdensity. In contrast to previous work, here we

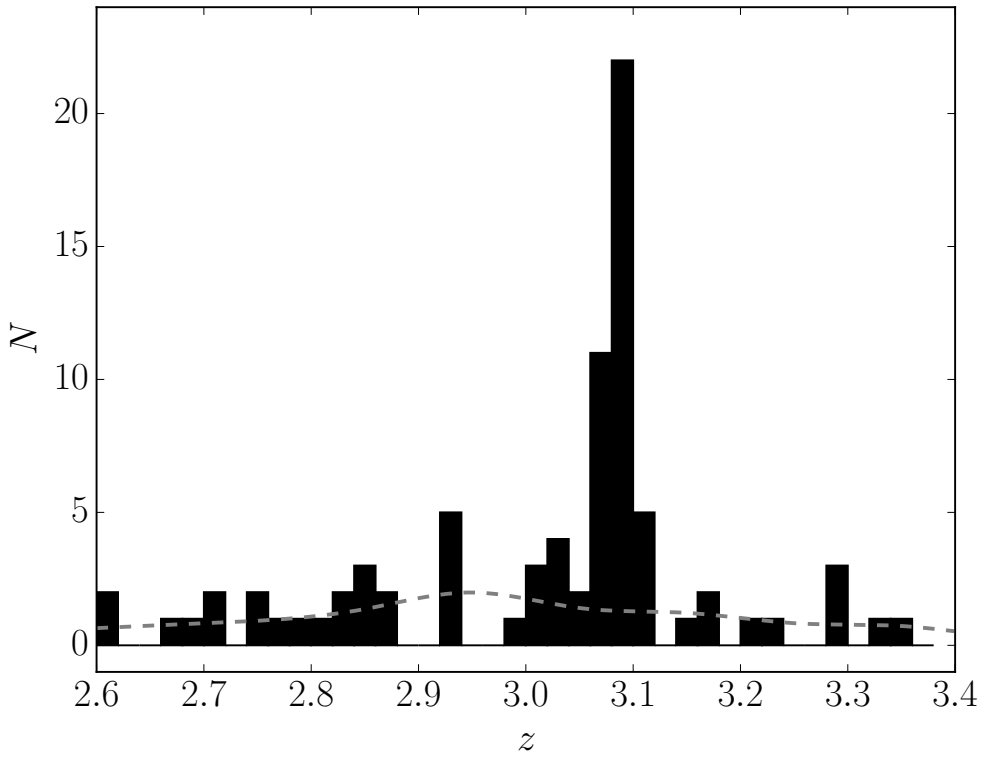

Figure 3.2: Redshift histogram of LBGs observed in the SSA22 field. The grey dashed line shows the LBG selection function determined using 883 LBGs, and normalized using the method described in the text. The bin size used in this redshift histogram is too coarse to observe the double-peaked structure near $z = 3.09$.

found that the low and high redshift boundaries of the total SSA22 interval were self-evident, as defined by a large gap on either side of the redshift distribution, with the boundaries occurring at the redshift of the last galaxy on each side of the overdensity. Therefore, we set the low and high redshift boundaries to $z = 3.0598$, and $z = 3.1048$ respectively, and removed the galaxies that define these boundaries from our future calculations. To find the boundary that separates the red and blue peaks, we fit the sum of two Gaussians to the redshift histogram, and determined the redshift at the minimum of the trough between the two peaks. We measured this boundary to be at $z = 3.0788$.

Due to the effects of redshift-space distortions (Kaiser 1987), and the fact that the SSA22 protocluster is collapsing, the redshift intervals we defined are con-

tracted compared to the ranges defined by the physical size of the protocluster in the Hubble flow. We used a correction factor C (Padmanabhan 1993) to quantify this effect, as defined by:

$$C = 1 + f - f(1 - \delta_m)^{\frac{1}{3}}, \tag{3.2}$$

where

$$f = \frac{d \ln D}{d \ln a}, \tag{3.3}$$

D is the linear growth factor, a is the cosmological scale factor, and δ_m is the matter overdensity, related to δ_{gal} through:

$$1 + b\delta_m = C(1 + \delta_{gal}). \tag{3.4}$$

We defined the LBG bias factor, b (Equation 3.5), by comparing $\sigma_{8,gal}$, the LBG number fluctuations, and $\sigma_{8,CDM} = 0.8228$, which corresponds to $\sigma_{8,CDM}|_{z=3.09} = 0.254$ at $z = 3.09$ (Planck Collaboration et al. 2014).

$$b^2 = \left. \frac{\sigma^2_{8,gal}}{\sigma^2_{8,CDM}} \right|_{z=3.09} \tag{3.5}$$

We calculated the value of $\sigma_{8,gal}$ using the correlation length, r_0, and the slope, γ, from the LBG autocorrelation function:

$$\sigma_{8,gal} = \frac{72(\frac{r_0}{8 \text{ cMpc}})^\gamma}{2^\gamma(3-\gamma)(4-\gamma)(6-\gamma)} \tag{3.6}$$

(Peebles 1980). For these calculations, we found a value of $f = 0.986$, and adopted values of $r_0 = 6.0 \pm 0.5 h^{-1}$ Mpc and $\gamma = 1.5$ from Trainor & Steidel (2012), which result in a bias of $b = 3.84 \pm 0.25$. We estimated the errors of the bias from the uncertainties of the autocorrelation function parameters, r_0 and γ, and $\sigma_{8,CDM}$. Table 3.1 shows the values of these parameters resulting from our calculation.

Neglecting the effects of large-scale redshift-space distortions (i.e., infall) in estimating the number of LBGs expected from the LBG selection function causes us to underestimate the relevant redshift interval, and therefore the expected number of galaxies, N_{expect}, as well. We corrected for this effect by increasing the redshift interval by a factor of $1/C$ (see Table 3.1) when integrating the LBG selection function, and recalculating the number of galaxies expected within the interval, as well as the associated galaxy overdensity. One subtlety lies in the fact that our correction factor, C, was initially calculated based on an overdensity that was overestimated due to the underestimate of the selection function, resulting in a correction that is too large. We therefore recomputed the correction factor using the updated overdensity, and repeated the procedure of correcting the redshift interval of the selection function, and recalculating the overdensity. We iterated this process until the galaxy overdensity converged to its true value, which we adopted as our final value for the overdensity. We obtained overdensities of $\delta_{t,gal} = 7.6 \pm 1.4$, $\delta_{r,gal} = 9.5 \pm 2.0$, and $\delta_{b,gal} = 4.8 \pm 1.8$, for the total, red, and blue regions respectively. Our updated total overdensity is larger than the value previously reported in Steidel et al. (1998) ($\delta_{t,gal} = 3.6^{+1.4}_{-1.2}$) but consistent with the value reported in Steidel et al. (2000) ($\delta_{t,gal} = 6.0 \pm 1.2$).

3.2.3 Mass Calculation

Using the updated estimates of the galaxy overdensity and appropriate volume for each section of the protocluster, corrected for the effects of redshift distortion, we computed the total, blue-peak and red-peak protocluster masses using:

$$M = \bar{\rho} V_{\rm true} (1 + \delta_m), \tag{3.7}$$

where $\bar{\rho}$ is the mean density of the universe, and $V_{\rm true} = V_{\rm apparent}/C$.

We calculated the mass overdensity, δ_m, of each region using Equation 3.4, utilizing the values for the correction factors, C (see Table 3.1), that we obtained

at the end of the iterative process described above. Using these correction factors, we calculated mass overdensities of $\delta_{t,m} = 1.3 \pm 0.4$, $\delta_{r,m} = 1.5 \pm 0.4$, and $\delta_{b,m} = 0.9 \pm 0.3$, for the total cluster, red peak, and blue peak respectively.

In order to estimate $V_{\rm apparent}$ (and the corresponding $V_{\rm true}$) for each region, we multiplied its line-of-sight extent and on-sky area. In the line-of-sight dimension, the spatial extent is represented by the difference in the radial comoving distance between the two redshift boundaries. We used the on-sky coverage of our observations, as described in Topping et al. (2016), as the area in the transverse dimensions, corresponding to a value of 12×14 h^{-2} cMpc^2 for the area on the sky. For the blue peak, we reduced the area on the sky because the galaxies contained within this peak cover only $\sim 75\%$ of the observing area (Topping et al. 2016). Our observations, and therefore the area used in our calculations, did not cover the full extent of the protocluster, as probed by e.g., Matsuda et al. (2005) and Yamada et al. (2012). Therefore, increasing the volume to enclose the entire protocluster may result in an increased mass estimate. On the other hand, our observations were centered on the highest density region of the protocluster, so expanding the protocluster volume may dilute the overdensity, therefore negating the expected mass increase caused by using a larger volume. For example, using the positions presented in Hayashino et al. (2004) we determined that the average surface density of LAEs decreases by $\sim 20\%$ if our observing window size is doubled. Analysis of protocluster membership in the Millennium Simulation shows that only $\sim 50\%$ of the galaxies within this area will be gravitationally bound to the main cluster by $z = 0$ (Muldrew et al. 2015). The net result of these two effects is a predicted $z = 0$ mass higher than our estimate, but much more uncertain.

Based on the δ_m and $V_{\rm true}$ values described above, we calculated the mass of the total cluster to be $(3.19 \pm 0.40) \times 10^{15}$ h^{-1} $M_\odot$, and calculated the mass of the red (blue) peak to be $(2.15 \pm 0.32) \times 10^{15}$ h^{-1} $M_\odot$ ($(0.76 \pm 0.17) \times 10^{15}$ h^{-1} $M_\odot$). We determined the errors on our mass calculation based on our uncertainties of the mass overdensity. The volumes encompassing the red and blue peaks do not

fill the entire space of the total overdensity, so the sum of the red and blue peak masses is less than the mass of the entire structure.

Table 3.1. Protocluster peak mass calculation

	Blue	Red	Total
z_{min}	3.0598	3.0788	3.0598
z_{max}	3.0788	3.1048	3.1048
z_{peak}	3.069 ± 0.001	3.095 ± 0.001	–
N_{expect}	1.71	2.66	4.40
N_{obs}	10	28	38
δ_{gal}	4.83 ± 1.84	9.51 ± 1.99	7.64 ± 1.40
δ_m	0.9	1.509	1.285
C	0.765	0.647	0.688
Dimensions $[h^{-3}\ cMpc \times cMpc \times cMpc]$	$\frac{3}{4} \times 12 \times 14 \times 18.38$	$12 \times 14 \times 24.95$	$12 \times 14 \times 43.338$
$V\ [h^{-3}\ cMpc^3]$	2315.9	4191.6	7280.8
M	$(0.757 \pm 0.171) \times 10^{15} h^{-1} M_\odot$	$(2.146 \pm 0.324) \times 10^{15} h^{-1} M_\odot$	$(3.194 \pm 0.401) \times 10^{15} h^{-1} M_\odot$

3.3 Methods and Simulations

We use cosmological N-body simulations in order to better understand the underlying physical structures giving rise to the observed properties of the SSA22 protocluster, as well as its evolution in the context of structure formation. In this section we present a description of the simulations we used, our technique for identifying protoclusters, and finally the methods that we used to search for analogs of the observed SSA22 structures in the simulation.

3.3.1 SMDPL Description

We use halo catalog and merger tree information drawn from the Small MultiDark Planck (SMDPL) simulation data set[1] (Klypin et al. 2016) in order to compare the observed structure in SSA22 to what is found in cosmological N-body simulations (Klypin et al. 2016; Behroozi et al. 2013b,a; Rodríguez-Puebla et al. 2016). We chose this simulation because its box size (400 h^{-1} Mpc) allows for a large enough sample ($N = 19$) of clusters that are within the estimated 3σ uncertainty of the mass of the red peak in SSA22 (i.e., $10^{15}\ h^{-1}\ M_{\odot} \leq M \leq 1.7 \times 10^{15}\ h^{-1}\ M_{\odot}$). Hereafter, we describe masses of halos using their virial mass, $M_{\rm vir}$, defined by Rodríguez-Puebla et al. (2016). The SMDPL simulation is also characterized by the following cosmological parameters: $\Omega_m = 0.307$, $\Omega_{\Lambda} = 0.693$, $h = 0.678$, $n_s = 0.96$, and $\sigma_8 = 0.829$. These parameters are consistent with current Planck results (Planck Collaboration et al. 2014), as opposed to those adopted for the Millennium simulation (Springel et al. 2005, $\Omega_m = 0.25$, $\Omega_{\Lambda} = 0.75$, $h = 0.73$, $\sigma_8 = 0.9$). In addition, with a particle mass of $M_{part} = 9.63 \times 10^7\ h^{-1}{\rm M}_{\odot}$, the mass resolution of the SMDPL simulation allows us to identify robust halos down to the mass that may host galaxies similar to ones in our observations ($M \sim 10^{10.6}\ h^{-1}{\rm M}_{\odot}$). The halo catalogs are saved in a series of 117 snapshots,

[1] http://hipacc.ucsc.edu/Bolshoi/MergerTrees.html

starting at Snapshot Number 0 (called *snapnum* in the catalogs) at $z = 18.56$, and ending with $snapnum = 116$ at $z = 0$. The snapshots are saved with a time resolution of $\Delta z \approx 0.16$ at $z \sim 3$. This time resolution allows us to perform our analysis on halos at the epoch of the SSA22 protocluster observations. The difference between the cosmological parameters used in the SMDPL simulation and our analysis in Section 3.2 is not significant, and therefore our inferences based on the results are valid.

3.3.2 Protocluster Identification

Based on the mass calculations presented in Section 3.2.3, we expect the SSA22 protocluster to evolve into a massive ($M \sim 10^{15}\ h^{-1}M_{\odot}$) cluster at $z = 0$, so we start by selecting all $z = 0$ halos, determined using the ROCKSTAR spherical overdensity method (Behroozi et al. 2013a), in the simulation with masses $M > 10^{15}\ h^{-1}\mathrm{M}_{\odot}$ from the SMDPL halo catalog. We identify 19 systems that meet this criterion. After identifying these halos, we follow their histories through the merger trees constructed from the simulation (Behroozi et al. 2013b), in order to select the progenitor halos at a given epoch ($z = 3.03$, $snapnum = 31$). We chose this snapshot as it has the closest redshift to that of the SSA22 protocluster.

3.3.3 Methods for Comparison

We present two methods to search for SSA22 analogs in the SMDPL simulation. First, we start by assuming that the observed structure in SSA22 will collapse into a massive cluster at $z = 0$. To mimic this regime in our analysis of the SMDPL simulation, we limit our sample to halos that collapse to a single massive structure at $z = 0$. We also employ an alternate, complementary approach in which we construct a sample of halos within a volume surrounding each of the $z \sim 3$ protoclusters with no requirement on their status as a member of the descendant cluster at $z = 0$. We then identify what kind of structures form from

these halos by $z = 0$, and compare them to the current predictions for the fate of the SSA22 protocluster.

3.3.3.1 Progenitors Only

We begin by describing the method that selects our parent sample of halos based on their membership in a single massive structure at $z = 0$. In order to compare any structure present in the simulated protoclusters to the structure observed in SSA22, we constructed redshift histograms from the sample of cluster progenitor halos. We created redshift histograms by viewing each protocluster from multiple sight lines. By observing through many sight lines we obtained a comprehensive view of each protocluster, and a better chance of detecting any structure that may be present. We expect adjacent sight lines to show similar evidence of structure, and since each sight line is a different random realization of the protocluster, sampling many sight lines allows us to differentiate between real structure and statistical flukes. For each protocluster we observed 3600 sight lines, each of which is separated by 6° in the azimuthal $\theta \in [0, 2\pi)$ direction, and 3° in the polar $\phi \in [0, \pi)$ direction.

For a given sight line, the simulated redshift histogram consists of calculated redshifts for 146 halos that are progenitors of a particular protocluster. We chose this number of halos to be the same as the number of galaxies (both LBGs and LAEs) that have spectroscopic redshifts in SSA22. To select these halos, we first narrowed down the sample based on their projected positions in the protocluster. We required that selected halos be within the observed area of SSA22, $\sim 12 \times 14\ h^{-2}$ cMpc^2, centered on the highest density peak. To choose the 146 halos whose redshifts make up the redshift histogram for a given sight line, we first randomly selected 40 halos out of all cluster progenitor halos with masses above $M > 10^{11.55}\ h^{-1}\mathrm{M}_{\odot}$(Trainor & Steidel 2012), corresponding to LBGs in our observed SSA22 sample. We then randomly selected 106 halos from among the

remaining cluster progenitor halos with masses $M > 10^{10.6}\ h^{-1}\mathrm{M}_{\odot}$(Gawiser et al. 2007), which represent the LAEs in our simulated redshift histogram. This selection process typically results in a sample that contains $\sim 10\%$ of the total cluster progenitors. This analysis assumes that the LBGs and LAEs in our sample are the central galaxies of their host dark matter halos, as opposed to satellites. The similar number densities and clustering strengths of LBGs and their host halos (Conroy et al. 2008; Trainor & Steidel 2012), in addition to the low halo occupation fraction of LAEs (1-10%; Gawiser et al. 2007), suggest that this assumption is valid.

To calculate the observed redshift of a halo, we first required its 3D position and velocity, given in the SMDPL halo catalog. We defined the center of the protocluster as the center-of-mass of all cluster progenitor halos, and set the center of each protocluster to be at $z = 3.09$. We calculated the redshift of each halo by determining its line-of-sight distance away from the protocluster center, and the corresponding velocity using the Hubble flow. In addition, we adjusted the estimated redshift to take into account the line-of-sight peculiar velocity, Δv, of each halo using $\Delta z = \Delta v/c \times (1 + z_H)$, where z_H is the redshift of the halo after taking into account the Hubble flow. We then collected the redshifts of all 146 halos into a redshift histogram.

We began by using the two-sample Kolmogorov-Smirnov (KS) test as a metric for comparison between each simulated protocluster redshift distribution and the observed SSA22 distribution. For each KS test, we determined the probability that the two distributions were drawn from the same parent distribution, a p-value. We introduced a p-value cutoff of $p \geq 0.4$, which distinguished redshift histograms that were well represented by two peaks, and those that presented only a single peak. We determined the value for this cutoff by trial-end-error. We adjusted the cutoff and visually inspected each qualifying histogram and its best fit models to determine at what p-value the histograms are typically double peaked. This cutoff allowed us to exclude those redshift histograms from further

analysis that did not show similar structure to that in the SSA22 protocluster.

After we determined the existence of structure in a given sight line, we compared the simulated redshift histogram to the one observed in SSA22. We first fit the sum of two Gaussians to the simulated redshift histogram. We then required the associated best fit parameters to be comparable to parameters found for the SSA22 redshift histogram. The requirements for the parameters of the larger (l) and smaller (s) peaks are as follows:

$$\begin{aligned} 0.341 \leq \frac{N_s}{N_l} \leq 0.493 \\ 0.004 \leq \sigma_l \leq 0.01 \\ 0.004 \leq \sigma_s \leq 0.01 \\ 0.02 \leq \Delta z \leq 0.032. \end{aligned} \tag{3.8}$$

In these expressions, N_s and N_l are the number of galaxies in the smaller and larger peaks, respectively. We determined a boundary at the trough between the two peaks, and counted the number of galaxies on either side. We define σ_l and σ_s as the best-fit standard deviations, in redshift units, of the large and small peaks, respectively, and Δz as the redshift difference between the centers of the two peaks. The existence of redshift histograms that fit these criteria would suggest that the observed structure in SSA22 may collapse into a single massive cluster at $z = 0$.

3.3.3.2 Halos in Surrounding Volume

In addition to searching for structure within the distribution of the $z \sim 3$ progenitors of a single massive $z = 0$ cluster, we also investigated halos in a volume surrounding each protocluster, regardless of their membership in a particular $z = 0$ structure. The full width covered by the SSA22 redshift histogram corresponds to a distance of $\sim 42\ h^{-1}$ cMpc along the line of sight. Accordingly, to isolate

a comparable volume in the simulation, we began by selecting all halos within a 42 h^{-1} cMpc radius from the center for each of the 19 identified protoclusters. We then followed the procedure described in Section 3.3.3.1 of selecting 146 halos, calculating redshifts, creating redshift histograms, and determining the similarity of the simulated and observed SSA22 redshift histograms, for 3600 sight lines of each protocluster.

We used the halo merger trees to determine the $z = 0$ structures formed from galaxies present in the $z \sim 3$ redshift distribution selected in this volume-limited manner. Accordingly, the underlying nature and evolution of double-peaked structure in a protocluster at $z \sim 3$, identified with this method as being analogous to the SSA22 protocluster, will then shed light on the potential fate of the observed structures in SSA22.

3.4 Results

3.4.1 Protocluster Members

We first tested the assumption that the double-peaked redshift histogram is representative of the progenitors of a single massive ($M \geq 10^{15}$ $h^{-1}\mathrm{M}_{\odot}$) protocluster at $z = 0$. Under this assumption, we expect that the majority of the galaxies in SSA22 will collapse into a massive cluster at $z = 0$. By investigating the $z = 3$ cluster progenitors of massive clusters at $z = 0$, we are able to identify which, if any, parts of the structure will be a component of the cluster once it has collapsed.

Using the methods described in Section 3.3.3.1 we determined whether there is any structure comparable to that of the SSA22 protocluster, in any of the 19 massive protoclusters in the SMDPL simulation. We found that none of the protoclusters had any sight lines that show evidence for a double-peaked morphology with best-fit parameters similar to those in SSA22, as defined in Equation 3.8. Figure 3.3 displays the redshift histogram that, out of all sight lines of the 19

protoclusters, shows the greatest similarity with SSA22 as defined by the KS-test p-value ($p = 1.8 \times 10^{-4}$). Even this distribution does not show a double-peaked morphology. By observing the spatial distribution of progenitor halos we can understand why there is very little extended structure present. The range of redshifts present in the SSA22 protocluster corresponds to a spatial separation of $\Delta z = 0.045$ ($\sim 40\ h^{-1}$ cMpc), while the $z \sim 3$ halo progenitors of a single massive $z = 0$ cluster in the SMDPL simulation typically extend over $\Delta z = 0.015$ ($\sim 13\ h^{-1}$ cMpc). Sufficiently high peculiar velocities could perturb the redshifts outside the primary structure, however the collapsing nature of these protoclusters tends to compress the redshift distribution on such scales, not expand it. In summary, comparison with the SMDPL simulation demonstrates that the double-peaked morphology observed in the SSA22 redshift histogram does not comprise the coalescing progenitors of a single $z = 0$ structure.

3.4.2 Surrounding Volume Halos

The approach described in the previous section was based on a starting assumption that the entire double-peaked structure in SSA22 corresponds to the progenitor of a single $M \geq 10^{15}\ h^{-1}\mathrm{M}_{\odot}$ cluster at $z = 0$. Therefore we restricted our analysis to include only the $z \sim 3$ progenitor halos of such $z = 0$ clusters. Using the alternative approach described in Section 3.3.3.2, we attempt to find structures within the volumes surrounding protoclusters in the simulation at $z \sim 3$ that, when "observed" (as we observe the SSA22 field), produce redshift histograms at $z \sim 3$ that are similar to that in SSA22. We then used the SMDPL simulation to characterize the evolution of such structures to $z = 0$.

When examining the distributions of halos in the more extended volumes surrounding massive cluster progenitors, we do find sight lines yielding redshift histograms similar to that of the SSA22 protocluster based on the SMDPL merger trees (Klypin et al. 2016; Behroozi et al. 2013b). Figure 3.4 shows an example of

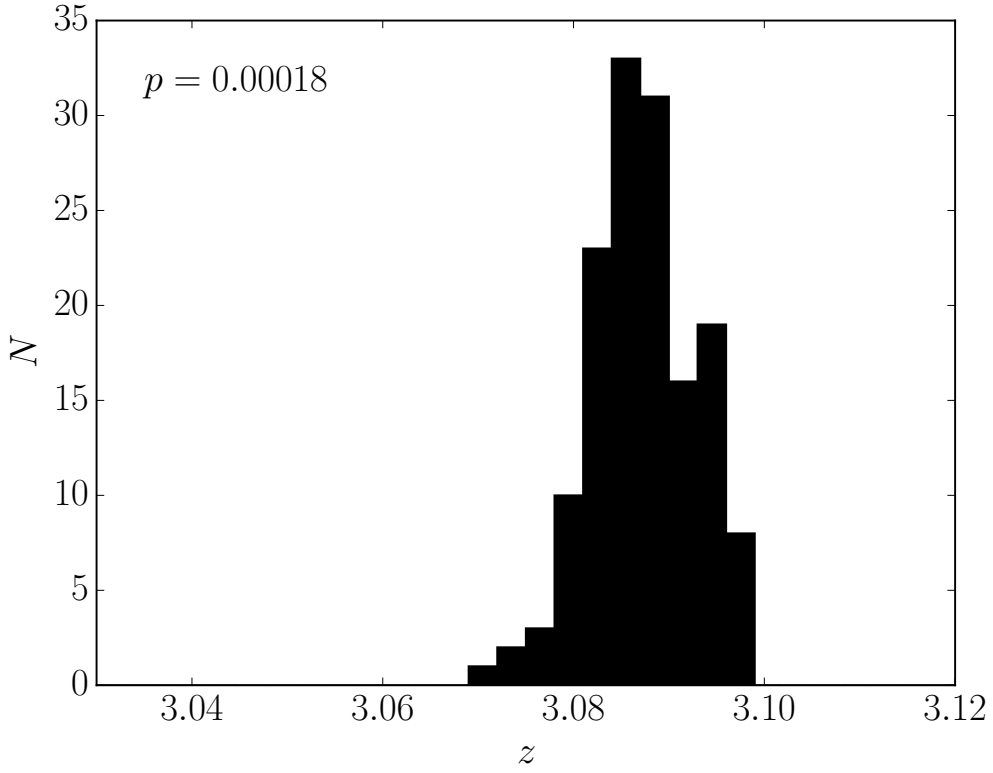

Figure 3.3: Redshift histogram calculating using only halos that are cluster progenitors. This is the redshift histogram that is the most similar to that of SSA22 across all sight lines in each of our 19 protoclusters. Even in this case the p-value suggests that the observed and simulated redshift distributions are significantly different.

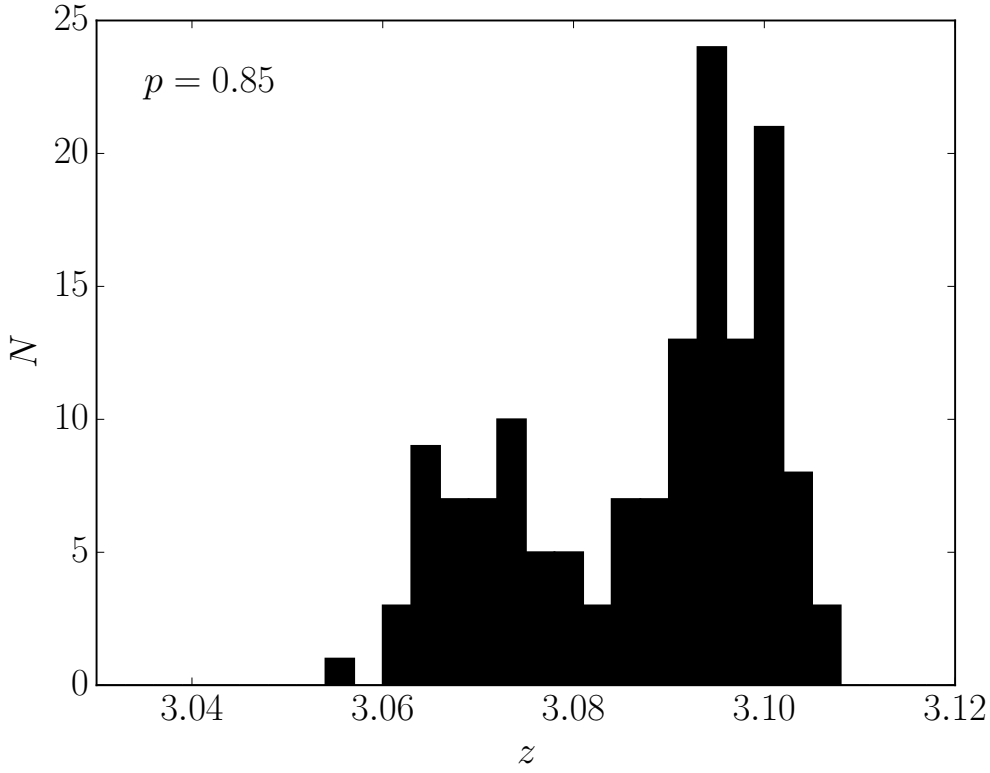

Figure 3.4: Example of a double-peaked redshift histogram computed by selecting cluster progenitor halos, as well as halos in the volume surrounding the protocluster. We determined this redshift histogram to fit the SSA22 similarity criteria given in Equation 3.8.

a redshift histogram ($p = 0.85$) computed for a single sight line of one simulated protocluster that fits our criteria for similarity to the SSA22 redshift histogram. In addition to finding sight lines that satisfy our similarity criteria stated in Equation 3.8, we find that many of these 'good' sight lines occur from similar viewing angles, suggesting that they are due to real structure, and not statistical flukes.

We separate the 19 protoclusters into three categories based on the number of distinct sight line groups present in each protocluster. The three categories are: "no sight lines", "single sight line group", and "multiple sight line groups". To assign a protocluster to one of the categories we first looked at the p-values distributed throughout the sight-lines. Figures 3.5(a)-3.7(a) show a projection of the p-value distribution as a function of sight-line. We then looked in detail at the redshift histogram produced when observing along a sight-line with a high p-value that passes through a possible cluster progenitor to confirm that it satisfied the similarity criteria of Equation 3.8. Figure 3.7(c) shows an example where the double-peaked structure of the redshift distribution can clearly be seen. If several of these sight-lines are clustered around a specific viewing angle, we consider the viewing angles to be a sight-line group. Finally, we categorize each protocluster volume based on the number of sight-line groups. Below we describe the three categories to which we assign each protocluster, with an example from each category detailing the important features in each case.

3.4.2.1 No sight lines

One subset of protocluster volumes in the SMDPL simulation that we investigated did not give rise to a double-peaked redshift histogram from any of the sight line viewing angles. An example of a protocluster in this category is shown in Figure 3.5. Figure 3.5(a) shows the KS $p-$value calculated from the redshift histogram of each sight line. While there are some sight lines with elevated p-values, there are not multiple adjacent sight lines with elevated $p-$values at any

particular viewing angle. The absence of double peaked histograms in this cluster is expected given the lack of any nearby massive cluster in the SMDPL simulation volume. This lack of adjacent structure is shown Figure 3.5(b), where the descendant mass of each halo is displayed. We find that $2/19 = 11\%$ protoclusters in our sample fall into this category.

3.4.2.2 Single sight line group

Another subset of protocluster volumes each yield a single group of closely packed sight lines that produce double-peaked histograms. Figure 3.6 shows an example of a protocluster in this category. In this example, many sight lines near the southern pole have high p-values suggesting that there is some structure arising in the redshift histograms. In addition, many of these sight lines also fit our criteria for similarity to the SSA22 redshift histogram, given in Equation 3.8. The viewing angle of these sight lines is coincident with the progenitor of a second, massive ($M = 10^{14.4}\ h^{-1}\mathrm{M}_{\odot}$) protocluster. Figure 3.6(b) shows this protocluster toward the bottom of the panel. We display the positions of halos from one sight line that shows a similar redshift histogram to that of SSA22 in Figure 3.6(c). At $z \approx 3$, the main and adjacent structure appear as two separate groups of halos, separated by a lower-density gap. In many cases, the smaller group of halos is the progenitor of a cluster with mass comparable to the expected mass of the blue (smaller) redshift peak in SSA22 at $z = 0$ ($M \sim 0.7 \times 10^{15}\ h^{-1}\mathrm{M}_{\odot}$). Most of the halos that make up the larger and smaller redshift peaks are progenitors of either the main or neighboring cluster. At $z = 0$, the two structures have collapsed into two distinct clusters. We find that $9/19 = 47\%$ protoclusters in our sample fall into this category.

3.4.2.3 Multiple sight line groups

The last category consists of protocluster volumes that each contain more than one distinct group of adjacent sight lines. Each of these groups is composed of many closely packed sight lines that produce a double peaked histogram. Figure 3.7 shows an example of a protocluster in this category. The KS p-value distribution (Figure 3.7, a) shows similar properties to the distribution presented in the 'single sight line' case. Protoclusters in this category however, show multiple separate viewing angles comprised of many sight lines with elevated p-values, as seen by the different groups of green points. Each one of these separate viewing angles corresponds to the presence of another nearby massive protocluster. Similar to the adjacent structures in the 'single sight line' group, many of the neighboring structures in the 'multiple sight line' category have masses comparable to the predicted $z = 0$ mass of the blue redshift peak in SSA22. The centers of these neighboring protoclusters typically lie 10-20 h^{-1} cMpc away from the main protocluster. All neighboring protoclusters are separate from each other at $z \sim 3$, and the majority of halos that make up double peaked histograms are members of the main protocluster, and a single neighboring protocluster, as no sight lines intersect multiple neighboring protoclusters. The neighbors, in addition to the main protocluster, all remain distinct as they collapse to separate structures at $z = 0$. We find that $8/19 = 42\%$ of the protoclusters in our sample fall into this category.

3.5 Discussion

3.5.1 Comparison with Analytic Predictions

We found that the double-peaked redshift histogram of the SSA22 protocluster is the reflection of the presence of less-massive ($> 10^{14}$ $h^{-1}\mathrm{M}_{\odot}$) protoclusters in its vicinity. We can use a simple analytic approach to explain quantitatively the prevalence of neighboring, less-massive clusters around $> 10^{15}$ $h^{-1}\mathrm{M}_{\odot}$. Due

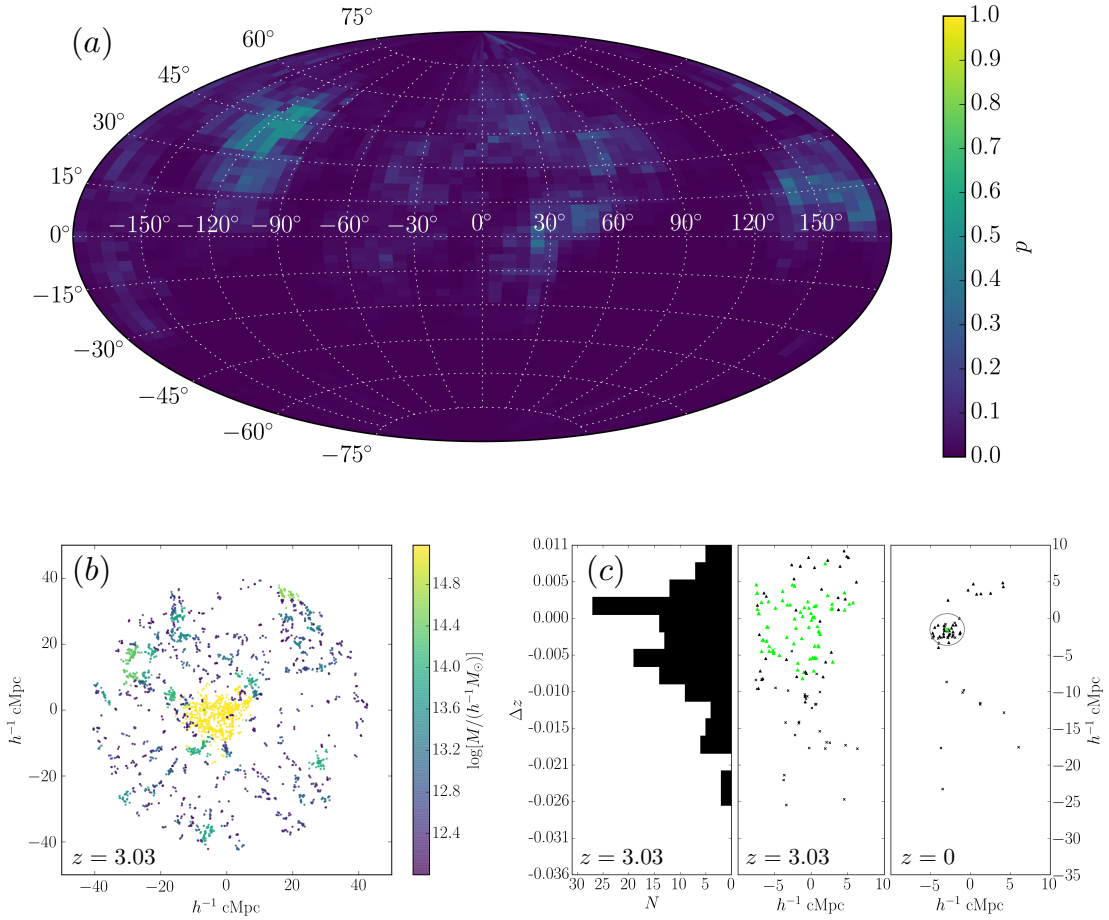

Figure 3.5: Example simulation results for a protocluster volume in the "no sight lines" category. (a): Mollweide projection of KS p-values calculated for redshift histograms created from each sight line toward the protocluster. Each 'pixel' in the projection represents a single sight line. 'Pixels' with higher p-values are sight lines that have redshift histograms comparable to that observed in SSA22, however none of the sight lines in category meet our criteria for similarity (Equation 3.8). (b): Scatter plot of halos in the volume surrounding a protocluster, colored by their $z = 0$ descendant mass. The yellow points in the center are the $M \sim 10^{15}$ $h^{-1}\mathrm{M}_{\odot}$ cluster progenitors. (c): Spatial positions of halos selected from a typical sight line at $z = 3.03$ (middle), their corresponding redshift histogram at $z = 3.03$ without including sightline-dependent peculiar velocity corrections (left), and their descendant positions at $z = 0$ (right). In the middle panel, $z = 3.03$ halos contained in $z = 0$ halos with $M > 10^{14}$ $h^{-1}\mathrm{M}_{\odot}$ are colored based on their cluster membership. The green points at $z = 3.03$, which make up the main protocluster, have merged into a single halo at $z = 0$. Halos in the right panel with masses $M > 10^{14}$ $h^{-1}\mathrm{M}_{\odot}$ are drawn with their corresponding R_{200} radii.

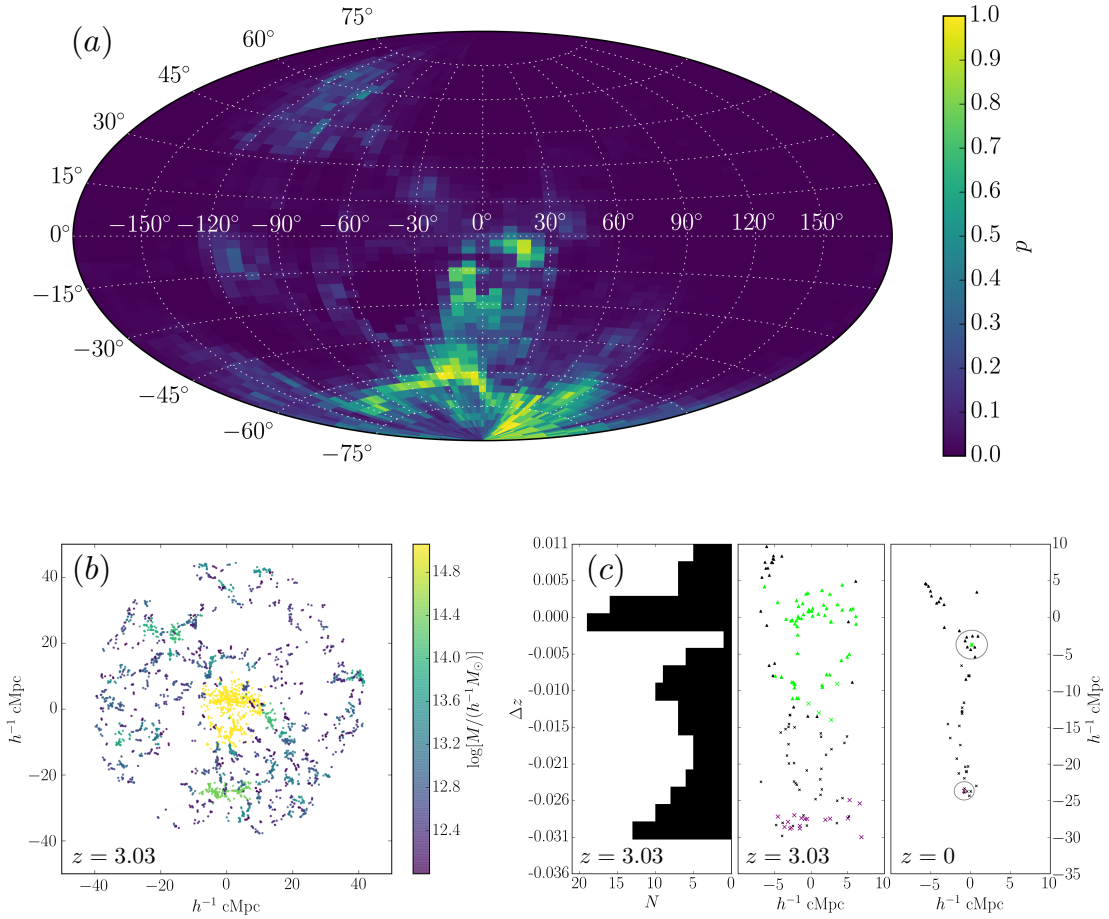

Figure 3.6: Example simulation results for a protocluster volume in the "single sight line" category. (a): Same as Figure 3.5(a). The group of sight lines near $\phi = -90$ (south pole) all have similar redshift histograms, suggesting that their double peaks are not due to random variance. The sight lines whose redshift histograms satisfy Equation 3.8 are a small subset of the bright 'pixels' in this panel. (b): Same as Figure 3.5(b). A $M > 10^{14}\ h^{-1}\mathrm{M}_{\odot}$ protocluster can be seen as a collection of green points at $(0, -25)$. (c): Same as Figure 3.5(c). In the middle panel, halos with descendant masses $> 10^{14}\ h^{-1}\mathrm{M}_{\odot}$ are colored based on their cluster membership. Points are also displayed as a triangle or a '$\times$' for their membership in the larger or smaller peak respectively determined after adjusting their redshifts due to their peculiar velocities. Results from one sight line that produced a double-peaked redshift histogram (left) based on the velocities and positions of halos at $z = 3.03$ (middle). The two protoclusters that give rise to the double-peaked redshift histogram remain distinct to $z = 0$ (right). At $z = 0$ (right) these two groups of halos have each collapsed to a single point. At the $z = 3.03$ epoch, an absence of halos is present at $\Delta z \sim -0.015$ (left, center). At $z = 0$ (right) the two groups of halos have collapsed to form distinct clusters. Halos with masses $M > 10^{14}\ h^{-1}\mathrm{M}_{\odot}$ are drawn with their corresponding R_{200} radii. The two groups are also easily seen in the redshift histogram (left), which does not include corrections based on halo peculiar velocities.

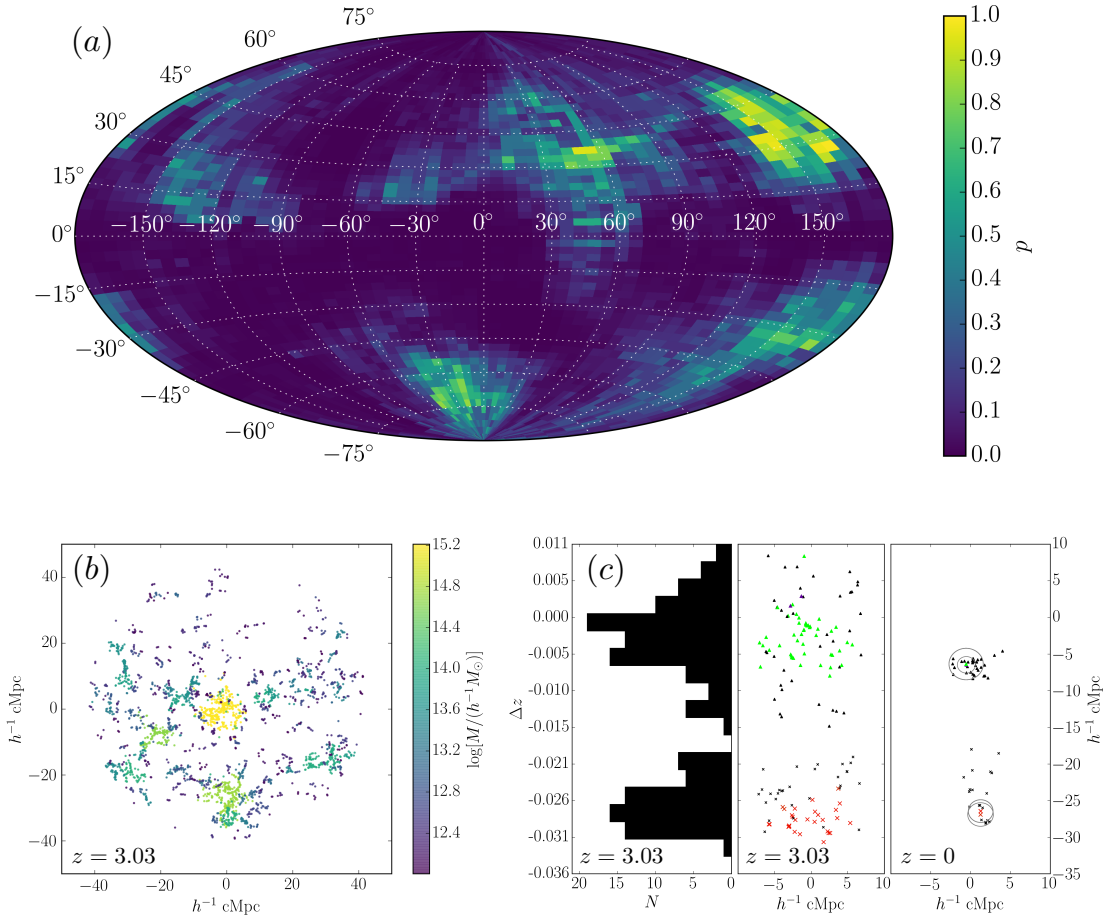

Figure 3.7: Example simulation results for a protocluster volume in the "multiple sight lines" category. (a): Same as Figure 3.6(a). Several distinct groups of sight lines are visible at different viewing angles. (b): Same as Figure 3.6(b). In the volume surrounding this protocluster, multiple other protoclusters can be seen as groups of green points at $(0, -30)$ and $(-20, 10)$. (c): Same as Figure 3.6(c). At the $z = 3.03$ epoch (left, middle), the two distinct groups of halos can be clearly seen at $\Delta z = -0.026$ and $\Delta z = 0.0$. The halos in each of these two groups have different $z = 0$ descendants. Massive halos at $z = 0$ that appear in the same position are separated in the into-the-page direction.

to halo biasing, we expect the most massive clusters at $z = 0$, which lie on an enhanced density peak, to be surrounded by smaller, but still massive, nearby clusters (Kaiser 1984; Barkana & Loeb 2004). Using the halo-halo correlation function and the halo mass function, we calculated the number of clusters at a given distance away from some of the most massive clusters. In this section, in order to more accurately compare to simulations, we adopt a cosmology consistent with the SMDPL simulation: $\Omega_m = 0.308$, $\Omega_\Lambda = 0.692$, $h = 0.677$, $\sigma_8 = 0.8228$, (Planck Collaboration et al. 2014).

We define the halo-halo correlation as the excess probability of finding a neighbor at a distance r and in the volume δV as

$$\delta P = n\delta V(1 + \xi(r)), \tag{3.9}$$

where n is the average number density of halos (Peebles 1980). We use linear bias to relate the linear matter correlation function, $\xi_{lin}(r)$, to the two-point correlation function of halos with masses M_1 and M_2, $\xi_{hh}(M_1, M_2, r)$, by

$$\xi_{hh}(M_1, M_2, r) = b(M_1)b(M_2)\xi_{lin}(r). \tag{3.10}$$

To calculate the linear bias factor, $b(M)$, we adopt the definition given by Quadri et al. (2007):

$$b_h(M) = 1 + \frac{1}{\delta_c}\left[\nu'^2 + b\nu'^{2(1-c)} - \frac{\nu'^{2c}/\sqrt{a}}{\nu'^{2c} + b(1-c)(1-c/2)}\right], \tag{3.11}$$

where $\nu' = \sqrt{a}\delta_c/\sigma(M, z)$, $\sigma(M, z)$ is the mass variance on scales of $R = \left(\frac{3M}{4\pi\bar{\rho}}\right)^{1/3} h^{-1}$ Mpc, and $\bar{\rho}$ is the mean matter density of the universe. As in Quadri et al. (2007), we use values of $\delta_c = 1.686$, $a = 0.707$, $b = 0.5$, and $c = 0.6$.

We calculate the linear mass correlation function from the power spectrum of

fluctuations, $P(k)$, using:

$$\xi_{lin}(R) = \frac{1}{2\pi^2} \int_0^\infty P(k) \frac{\sin kR}{kR} k^2 dk \tag{3.12}$$

and derive a power spectrum based on the methods described in Naoz & Barkana (2005).

Using the halo-halo correlation function, we predicted the mean number of halos, with masses $M \geq M_2$, within a surrounding volume centered on a halo with mass M_1 using:

$$\langle N(R) \rangle = n \int_0^R 4\pi r^2 \left[1 + \xi_{hh}(r)\right] dr. \tag{3.13}$$

Where n is the average number density of halos, calculated using the halo mass function of Sheth & Tormen (1999).[2]

Using the method described here, we obtain an analytic prediction for the prevalence of $> 10^{14}\ h^{-1}\mathrm{M}_\odot$ clusters as a function of distance from a $> 10^{15}\ h^{-1}\mathrm{M}_\odot$ cluster at $z = 0$. We then compare our analytic prediction with the results from the SMDPL simulations, and finally with our observations of the SSA22 protocluster. We calculated the number of halos of a given mass within a given distance, R from the center of a cluster with a mass corresponding to the the mass of one of the 19 clusters present in the SMDPL simulation. This process was repeated for each of the 19 $M > 10^{15}\ h^{-1}\mathrm{M}_\odot$ clusters in the simulation, and then we averaged the resulting total number. Figure 3.8 shows this analytic result, calculated using the method described above (dashed lines). For comparison, Figure 3.8 displays the number of halos with a given mass and within a given radius, R computed directly from the SMDPL simulation by counting the average number of halos in a sphere with radius R centered on each of the 19 $M > 10^{15}\ h^{-1}M_\odot$ clusters at $z = 0$ (solid lines).

[2]We obtained the same results when repeating this analysis adopting the halo mass function described in Tinker et al. (2008), with parameters: $A = 0.144$, $a = 1.351$, $b = 3.113$, and $c = 1.187$ provided by Rodríguez-Puebla et al. (2016).

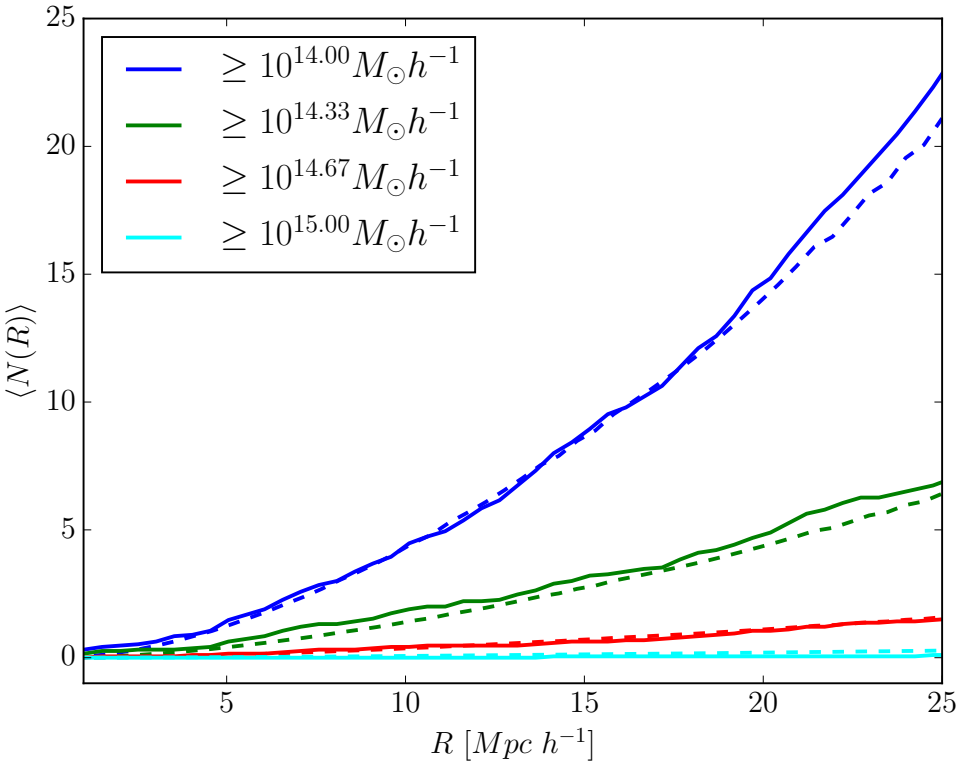

Figure 3.8: Average number of halos of a given mass within a sphere of radius R centered on a $10^{15} h^{-1} \mathrm{M}_\odot$cluster at $z = 0$. Shown here are the analytic predictions (dashed lines) calculated using the method described in Section 3.5.1, compared to the number $N(R)$ measured directly from the SMDPL simulation using $\geq 10^{15} h^{-1} \mathrm{M}_\odot$ halos as the central halo (solid lines).

In the case of the SSA22 protocluster, the adjacent structure lies at a distance of $D \sim 20\ h^{-1}$ cMpc calculated from the difference in peak redshifts neglecting the effects of infall. Within this distance, our analysis predicts $\sim 1-2$ clusters with a mass comparable to the mass of the blue peak of SSA22. This number increases by $\sim 20\%$ when the mass of the central cluster is doubled. This analytic prediction is consistent with our results which place more protoclusters in the "single sight-line" category, compared to the other categories. We also predict ~ 10 clusters with masses $\sim 10^{14}\ h^{-1}\mathrm{M}_{\odot}$ within this distance, which is again consistent with the simulations. However, the neighboring clusters that give rise to double-peaked redshift histograms typically have masses of $\geq 3 \times 10^{14}\ h^{-1}\mathrm{M}_{\odot}$.

3.5.2 Observing Frequency

We have determined which, if any, sight lines in a given simulated protocluster produce redshift histograms that present a double-peaked morphology, and whether they are similar to the observed redshift histogram in SSA22. In this section, we discuss the probability of observing a double-peaked redshift histogram, based on our analysis of protoclusters in the SMDPL simulation. For this analysis, we calculated the density of protoclusters that, when observed, would result in a redshift histogram that contains two peaks similar to that of SSA22, or any structure beyond a single redshift peak. By searching through sight lines across all protoclusters in the simulation, we determined the frequency at which observations of massive protoclusters would yield double-peaked redshift histograms.

We started by calculating the covering fraction of sight lines that produced double-peaked redshift histograms. For an individual protocluster, we calculated the total covering fraction by summing up the contribution from each sight line that we have determined to be double peaked. The area on the sky covered by a

single sight line is given by:

$$\Delta\Omega = \int_{\theta}^{\theta+\Delta\theta} \int_{\phi}^{\phi+\Delta\phi} d\theta d\phi, \tag{3.14}$$

where $\Delta\theta = 6\,\mathrm{deg}$, $\Delta\phi = 3\,\mathrm{deg}$, and (θ, ϕ) is the angle of the sight line. For a given protocluster, the covering fraction of sight lines with redshift histograms similar to that of SSA22 is:

$$F = \frac{\Omega}{4\pi} \tag{3.15}$$

where $\Omega = \sum \Delta\Omega$ is the total solid angle covered by the relevant sight lines.

On average, the covering fraction of sight lines for a given protocluster is $F = 0.025 \pm 0.017$, with values for individual protoclusters ranging from $F = 0$ for protoclusters in the "no sight lines" category, to $F = 0.065$ for a protocluster in the "multiple sight lines" category. We also consider counting sight lines that are better fit by two peaks, but whose fitting parameters may not fit the criteria presented in Equation 3.8. Such sight lines contain evidence of structure beyond the main protocluster, but, when observed, do not produce redshift histograms similar to that of SSA22. The average covering fraction of such additional sight lines is $F = 0.13$. To determine the occurrence rate of structures similar to those observed in SSA22, we multiply the covering fraction of sight lines producing double peaked histograms by the number density of massive protoclusters in the SMDPL simulation, $19/400^3\ h^3\ \mathrm{cMpc}^{-3} = 296\ h^3\ \mathrm{Gpc}^{-3}$. We therefore calculated the cosmic abundance of observing structure similar to that of the SSA22 protocluster to be $n = 7.4\ h^3\ \mathrm{Gpc}^{-3}$. This density suggests that the observed structure in the SSA22 protocluster is rare, and its discovery unexpected within the $1.07 \times 10^{-3}\ h^3\ \mathrm{Gpc}^{-3}$ volume of the survey that discovered it (Steidel et al. 2003). Even placing a less stringent similarity requirement for the simulated redshift histograms (i.e., some evidence for structure ($p > 0.4$), as defined in Section 3.3.3.1 without strictly satisfying Equation 3.8), we find a cosmic abundance of only $n = 38\ h^3\ \mathrm{Gpc}^{-3}$, which still makes the discovery of SSA22 extremely fortuitous within the LBG survey

volume. Hints of bimodality have been seen in other protoclusters (e.g.; Kuiper et al. 2011; Venemans et al. 2007). However, better spectroscopic sampling as well as evidence of a spatial offset between redshift peaks are required to determine the similarity of these structures to the observed large-scale structure in SSA22.

3.6 Summary and Conclusion

We have used an updated spectroscopic sample to measure the overdensity and mass of the SSA22 protocluster, and its associated structure. We then attempted to understand these results using the SMDPL cosmological simulation, and a simple analytic approach. In detail:

1. We used an updated sample of spectroscopic redshifts of LBGs in the SSA22 field to measure the overdensities of the total SSA22 region ($\delta_{t,gal} = 7.6 \pm 1.4$), and the blue and red peaks present in its redshift histogram ($\delta_{b,gal} = 4.8 \pm 1.8$, $\delta_{r,gal} = 9.5 \pm 2.0$). We utilized updated overdensity measurements to calculate the masses of the total region ($M_t = (3.19 \pm 0.40) \times 10^{15} h^{-1} M_{\odot}$), the blue redshift peak ($M_b = (0.76 \pm 0.17) \times 10^{15} h^{-1} M_{\odot}$), and red redshift peak ($M_r = (2.15 \pm 0.32) \times 10^{15} h^{-1} M_{\odot}$).

2. Using our updated predictions for the masses of these two peaks, we made use of the Small MultiDark Planck simulation to determine the nature of the double-peaked redshift distribution. First, we tested the scenario that the structure in SSA22 is all contained in the progenitor of a single massive cluster. For this analysis, we looked in the simulation only at halos that would eventually collapse into a single massive ($M > 10^{15}\ h^{-1} M_{\odot}$) structure at $z = 0$. From these we created simulated redshift histograms and compared their morphology to the observed redshift distribution in SSA22. In the 19 $M > 10^{15}\ h^{-1} M_{\odot}$ protoclusters in the simulation that we observed, none had progenitor halo distributions that alone produced a redshift histogram

consistent with the double-peaked shape observed in SSA22.

3. We performed a complementary approach that considered all halos within a certain distance of each individual $M > 10^{15}\ h^{-1}M_{\odot}$ protocluster in the simulation, regardless of membership in the associated descendant cluster at $z = 0$. Using this method, we found that 17/19 of the simulated protoclusters had configurations that, when observed from at least some lines of sight, produced redshift histograms similar to that of the observed distribution in SSA22. For each of these 17 protoclusters, the viewing angles that produced the matching redshift histograms contain the main overdensity along with a neighboring aligned, but less massive, overdensity. Following these adjacent protoclusters through time in the simulation, we saw that the two structures in the volume remained distinct to $z = 0$, demonstrating that the second peak in the redshift histogram can be caused by a separate virialized structure from the main protocluster.

4. We further investigated the results from the simulation using a simple analytic approach. Using the halo-halo correlation function derived from the dark matter power spectrum, we predicted the number of halos of a given mass within a distance R from a massive cluster. The results from this analysis are consistent with what we have seen in the simulation, predicting $\sim 1 - 2$ massive halos surrounding each main cluster capable of producing a second peak in the redshift distribution.

5. Finally, using the covering fraction of sight lines of simulated protoclusters that produced double-peaked redshift histograms, and the number density of massive protoclusters, we predicted the occurrence of a structure similar to that observed in SSA22 to be 7.4 h^3 Gpc^{-3}.

Previous estimates of the mass of the SSA22 overdensity have been produced by considering the volume containing the red and blue peak as a single massive

protocluster. By treating the entire region as a single overdensity, previous studies have overestimated the mass of the main, $M \sim 10^{15}\ h^{-1} M_{\odot}$ protocluster. The existence of the second (blue) peak must be considered in order to obtain an accurate measurement of the mass.

Due to the limited area that our observations cover, we are restricted to observing structure coincident with the line-of-sight to the main protocluster. In order to fully understand the connection between the structure and the main protocluster, deep and densely sampled spectroscopic observations must be performed in an area extending at least $\sim 20\ h^{-1}$ Mpc ($\sim 11'$) away from the center of the protocluster. This approach would allow us to not only fully map the structure already observed, but also find other massive nearby structures, if present. In addition to wider-field observations of the SSA22 protocluster, an in-depth analysis of the structure present in additional known protoclusters (e.g., HS1700+643 at $z = 2.299$, and HS1549+195 at $z = 2.842$; Steidel et al. 2005, 2011) found in larger cosmic volumes will demonstrate the variety of environments of the most massive structures in the universe as they formed.

CHAPTER 4

The MOSDEF-LRIS Survey: The Interplay Between Massive Stars and Ionized Gas in High-Redshift Star-Forming Galaxies

4.1 Introduction

Rest-optical spectroscopy is a powerful tool that can be used to determine a wealth of information on the physical conditions within the interstellar medium (ISM) of galaxies. Measurements of optical nebular emission lines from local star-forming galaxies demonstrate that they trace a tight sequence of increasing [NII]λ6584/Hα and decreasing [OIII]λ5007/Hβ emission-line ratios (e.g., Veilleux & Osterbrock 1987; Kauffmann et al. 2003). The observed variation in emission-line ratios along the star-forming sequence reflects the increasing oxygen abundance and stellar mass and decreasing HII-region excitation of its constituent galaxies (Masters et al. 2016). Early observations with Keck/NIRSPEC suggested possible differences in the emission-line properties of high-redshift galaxies in the [OIII]λ5007/Hβ vs. [NII]λ6584/Hα "BPT" diagram (Baldwin et al. 1981; Shapley et al. 2005; Erb et al. 2006; Liu et al. 2008). New, statistical samples from the MOSFIRE Deep Evolution Field (MOSDEF; Kriek et al. 2015) survey and the Keck Baryonic Structure Survey (KBSS; Steidel et al. 2014) show that typical high-redshift galaxies are offset towards higher [OIII]λ5007/Hβ and/or [NII]λ6584/Hα on average relative to local galaxies.

There are many possible causes for this observed difference between local and

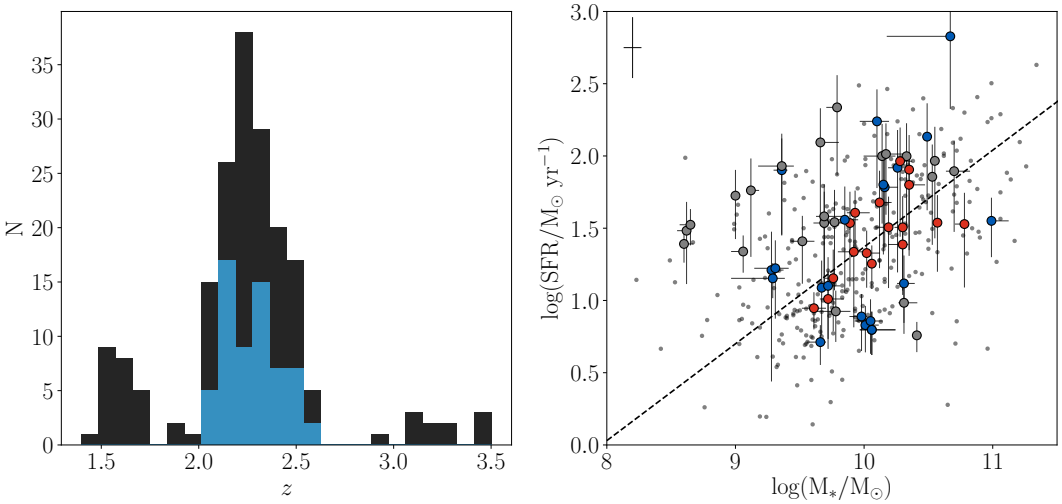

Figure 4.1: Left: Redshift histogram for all objects with redshifts measured from LRIS spectra (black), totalling 188 galaxies. The blue redshift histogram comprises all objects within $2.09 \leq z \leq 2.61$ ($z_{\rm med} = 2.28$) with all four BPT lines (Hβ, [OIII]λ5007, Hα, [NII]λ6584) detected at $\geq 3\sigma$ in MOSFIRE spectra from the MOSDEF survey, totalling 62 galaxies. Right: SFR calculated from the dust-corrected Balmer lines vs. M$_*$ for all objects with LRIS redshifts at $2.0 \leq z \leq 2.7$ (large circles). Blue and red points indicate galaxies included, respectively, in the *high* and *low* composite spectra described in Section 4.3. Galaxies from the MOSDEF survey within $2.0 \leq z \leq 2.7$ that have both Hα and Hβ detected with $\geq 3\sigma$ are depicted by small grey points. The median errorbar is shown in the top left corner. The dashed line shows the SFR- M$_*$ relation of $z \sim 2.3$ galaxies from the MOSDEF survey calculated by Sanders et al. (2018).

$z \sim 2$ galaxies, including higher ionization parameters, harder ionizing spectra at fixed nebular metallicity, higher densities, variations in gas-phase abundance patterns, and enhanced contributions from AGNs and shocks at high redshift (see e.g., Kewley et al. 2013, for a review). Early results from the MOSDEF survey suggested that the offset of high-redshift galaxies on the BPT diagram is caused in part by the order-of-magnitude higher physical density in $z \sim 2$ star-forming regions, but is primarily a result of an enhanced N/O ratio abundance at fixed oxygen abundance in offset $z \sim 2$ star-forming galaxies relative to local systems (Masters et al. 2014; Shapley et al. 2015; Sanders et al. 2016b). Furthermore, there is evidence that the BPT offset is strongest among low-mass, young galaxies (Shapley et al. 2015; Strom et al. 2017). Results from KBSS were used to argue instead that the observed offset is more likely driven by a harder stellar ionizing spectrum at fixed nebular metallicity, which can also explain at least some of the observed emission-line patterns (Steidel et al. 2016; Strom et al. 2017).Recently, updated results from the MOSDEF survey corroborate these results favoring a harder stellar ionizing spectrum at fixed nebular metallicity (Sanders et al. 2019; Shapley et al. 2019), which arises naturally due to the super-solar O/Fe values of the massive ionizing stars that excite the HII regions in these $z \sim 2$ star-forming galaxies. Such α-enhancement would naturally exist in high-redshift galaxies due to rapid formation timescales, resulting in enrichment by a larger fraction of Type II relative to Type Ia supernova explosions.

In star-forming galaxies, massive stars are the predominant sources of ionizing radiation driving the nebular emission lines included in the BPT diagram. As such, studying the properties of massive stars enables us to address the origin of the observed rest-optical spectroscopic differences between local and high-redshift galaxies. The formation and evolution of massive stars is intimately linked with the evolving properties of the ionized ISM. Specifically, the formation of massive stars is driven by the accretion of gas onto galaxies, and, in turn, massive stars regulate the chemical enrichment of the ISM by driving galaxy-scale outflows, and

polluting the ISM when they explode as core-collapse supernovae. Additionally, due to the short-lived nature of these stars, they provide a probe of star-forming galaxies on timescales shorter to or equal to the typical dynamical timescale. One avenue for studying the properties of the massive star populations in high-redshift galaxies is directly observing their light using rest-UV spectroscopy.

Rest-UV spectra of star-forming galaxies contain many features tracing the massive, young stars that supply the ionizing luminosity exciting the gas in star-forming regions. These features, such as the CIV$\lambda\lambda$1548, 1550 (Crowther et al. 2006; Leitherer et al. 2001) and HeIIλ1640 (Brinchmann et al. 2008) stellar wind lines, and a host of stellar photospheric features (Rix et al. 2004), provide information on the population of massive stars. In particular, using a given set of model assumptions, these features can be used to establish the form of the initial mass function (IMF), the abundance of Wolf-Rayet (WR) stars, and the nature of the ionizing spectrum in star-forming regions. The features of rest-UV spectra have also been used to estimate stellar abundances (i.e., Fe/H) in high-redshift galaxies. Halliday et al. (2008) used the FeIII-sensitive 1978 Å index defined by Rix et al. (2004) to measure a stellar metallicity of $Z_*/Z_\odot = 0.267$ in a composite spectrum of 75 $z \sim 2$ star-forming galaxies. Sommariva et al. (2012) employed a similar approach, and investigated new photospheric absorption line indicators suitable as calibrations of the stellar metallicity in high-redshift galaxies. They applied these calibrations to the rest-UV spectra of nine $z \sim 3.3$ individual galaxies, and one composite spectrum to construct the $z \geq 2.5$ M_*-Z_* relation. Compared to the previously mentioned works, Cullen et al. (2019) instead fit models to the full rest-UV spectrum, an approach that uses all of the stellar-metallicity sensitive spectral features simultaneously. They applied this method to composite spectra to constrain the stellar metallicity of star-forming galaxies spanning a redshift range of $2.5 < z < 5.0$ and a stellar mass range of $8.5 < \log(M_*/M_\odot) < 10.2$.

Expanding on previous work, recent studies have made use of rest-UV spectra in combination with rest-optical spectra of high-redshift galaxies (Steidel et al.

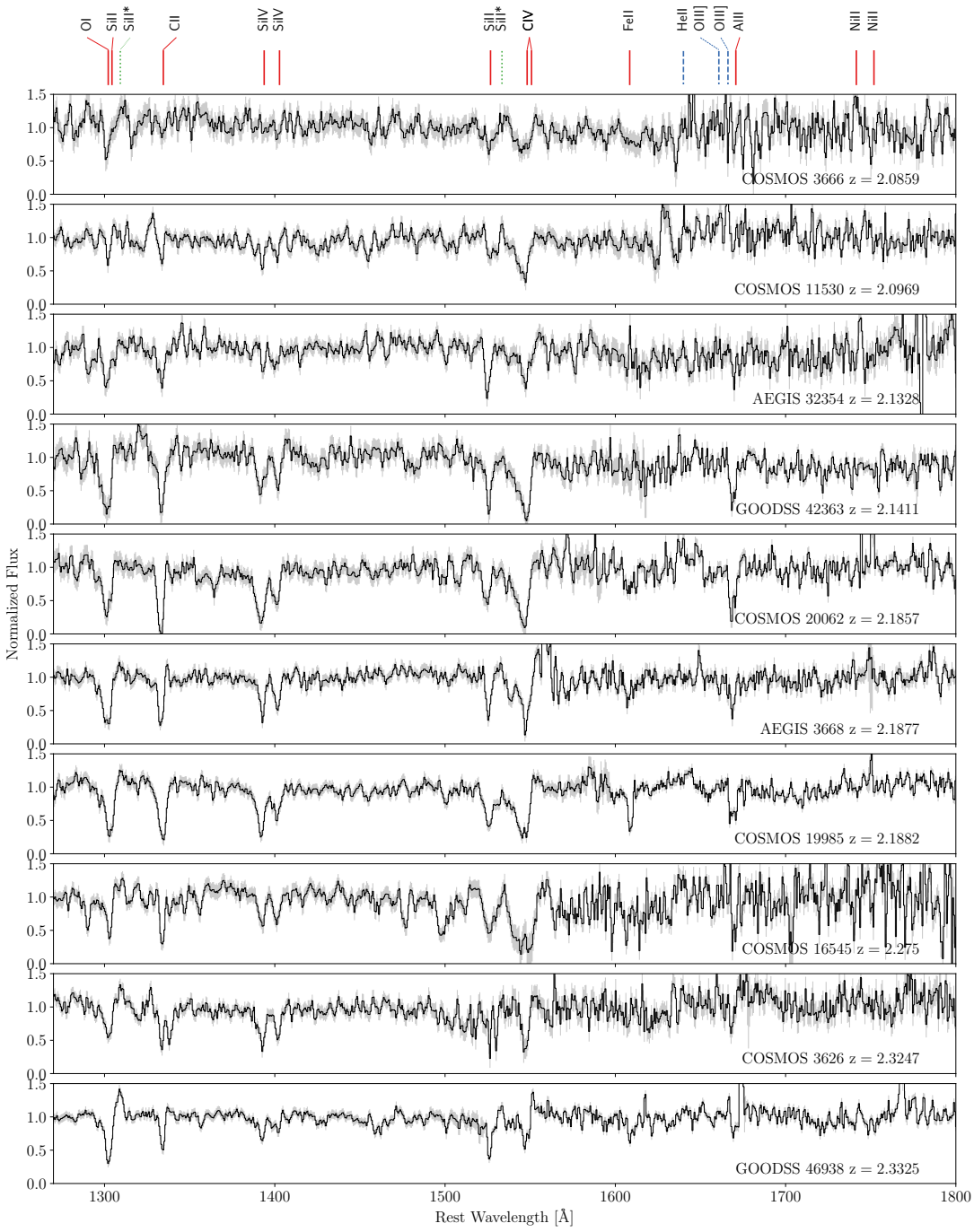

Figure 4.2: Ten continuum-normalized individual rest-UV spectra from our sample ordered by redshift. These spectra have the highest continuum signal-to-noise ratio (SNR) from our sample with median SNR/pixel = 7 ($4.5 \leq$ SNR/pixel ≤ 12), measured over the wavelength range $1425\text{Å} \leq \lambda \leq 1500$ Å. Labels on the top of the figure indicate several spectral features including stellar absorption lines (solid red lines), nebular emission lines (dashed dark blue lines), and fine structure lines (dotted green lines). At these redshifts, the dichroic cutoff between the red- and blue-side spectra occurs at a typical rest-frame wavelength of ~ 1500 Å. The 1σ error spectrum is depicted by the shaded region surrounding each spectrum.

2016; Chisholm et al. 2019). Using composite rest-UV and rest-optical spectra of 30 star-forming galaxies at $z \sim 2.4$, Steidel et al. (2016) found that the observed properties constrained by their composite spectra can be reproduced only by models that include binary stars, have low stellar metallicities ($Z_*/Z_\odot \sim 0.1$) and moderate nebular metallicities ($Z_{\rm neb}/Z_\odot \sim 0.5$). These results indicate α-enhancement for the $z \sim 2$ galaxies in Steidel et al. (2016) relative to the solar abundance pattern, given that Z_* is primarily tracing Fe/H and $Z_{\rm neb}$ is tracing O/H. By analyzing a single composite rest-UV spectrum, Steidel et al. (2016) only probed average properties of their high-redshift galaxy sample. With single rest-UV and rest-optical composite spectra it is not possible to probe the average rest-UV spectral properties as a function of the location in the BPT diagram. In this paper we expand upon the important initial work of Steidel et al. (2016) by utilizing combined rest-UV and rest-optical spectra of 62 $z \sim 2.3$ galaxies spanning a broad range of physical properties. With this large sample, we investigate how the rest-UV spectral properties of the massive star population, including the inferred ionizing radiation field, vary for galaxies with different rest-optical emission-line properties in order to uncover the origin of differences between high-redshift and local galaxies in the BPT diagram.

The organization of this paper is as follows: Section 4.2 describes our observations, data reduction, and methods. Section 4.3 presents the results of our analysis, Section 4.4 provides a discussion of our results, and Section 4.5 gives a summary of our key results. Throughout this paper we assume a cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $H_0 = 70$ km s^{-1} Mpc^{-1}, and adopt solar abundances from Asplund et al. (2009) (i.e., $Z_\odot = 0.014$).

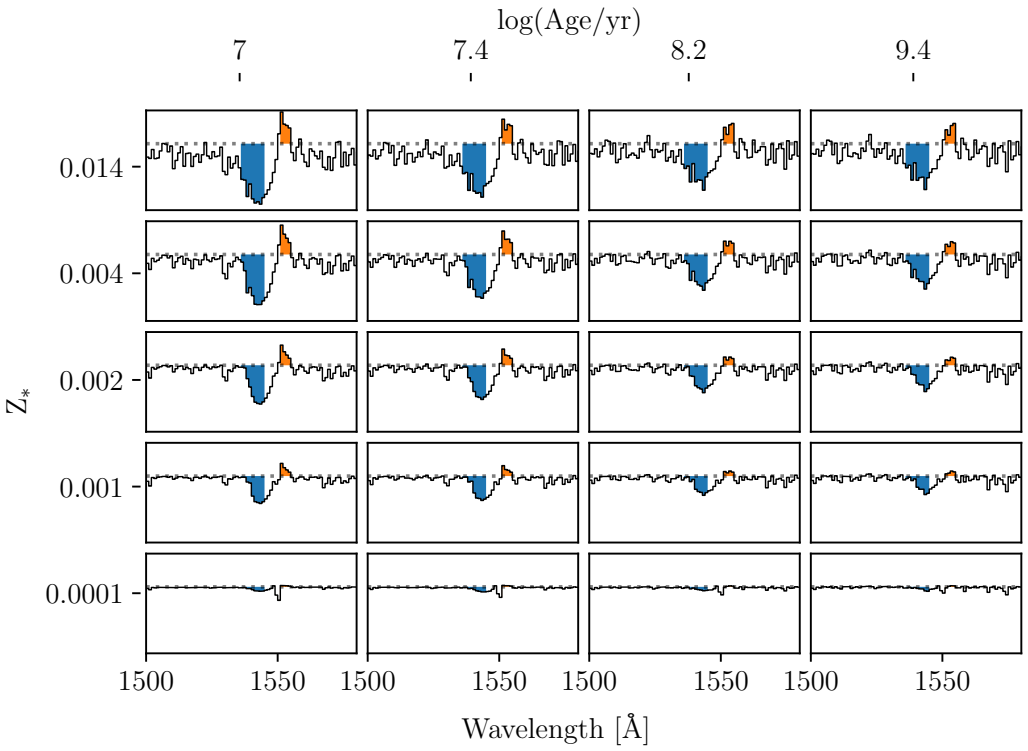

Figure 4.3: Zoomed-in regions around the CIV$\lambda\lambda$1548, 1550 profile for several of the BPASS+Cloudy models of different stellar metallicities and ages used in our analysis. The blueshifted wing of the broad wind absorption, and the redshifted wind emission feature of the CIV$\lambda\lambda$1548, 1550 profile are highlighted in blue and orange respectively, showing how the overall CIV$\lambda\lambda$1548, 1550 profile varies with age and metallicity. Both the absorption (blue) and emission (orange) increase with strength toward higher stellar metallicity and younger age. Both of these shaded regions are outside of the range where interstellar absorption is important and which we mask for our stellar population fitting analysis.

4.2 Methods and Observations

4.2.1 Rest-Optical Spectra and the MOSDEF survey

Our analysis utilizes rest-optical spectroscopy of $z \sim 2.3$ galaxies from the MOSDEF survey (Kriek et al. 2015). The MOSDEF survey consists of moderate resolution (R $\sim$ 3500) near-infrared spectra of $\sim$ 1500 H-band selected galaxies observed over 48.5 nights during 2012–2016 and targeted to lie within three distinct redshift intervals ($1.37 \leq z \leq 1.70$, $2.09 \leq z \leq 2.61$, and $2.95 \leq z \leq 3.80$) near the epoch of peak star formation ($1.4 \leq z \leq 3.8$) using the Multi-Object Spectrometer for Infra-Red Exploration (MOSFIRE; McLean et al. 2012). The actual redshift intervals are slightly different from our initial target ranges, based on the scatter between photometric and spectoscopic redshifts, and we redefine them as $1.40 \leq z \leq 1.90$, $1.90 \leq z \leq 2.65$, and $2.95 \leq z \leq 3.80$. Additionally, the MOSDEF survey targeted galaxies in the *Hubble Space Telescope* extragalactic legacy fields in regions covered by the CANDELS (Grogin et al. 2011) and 3D-HST (Momcheva et al. 2016) surveys, which have assembled extensive ancillary multi-wavelength datasets. MOSDEF spectra were used to measure fluxes and redshifts of all rest-optical emission lines detected within the Y, J, H, and K bands, the strongest of which are: [OII]λ3727, Hβ, [OIII]$\lambda\lambda$4959, 5007, Hα, [NII]λ6584, and [SII]$\lambda\lambda$6717, 6731.

4.2.2 LRIS Observations and Data

In order to characterize how galaxy properties vary across the BPT diagram, we selected a subset of MOSDEF galaxies for rest-UV spectroscopic followup based on the following criterion. We prioritized selecting galaxies drawn from the MOSDEF survey for which all four BPT emission lines (Hβ, [OIII], Hα, [NII]) were detected with $\geq 3\sigma$. Next highest priority was given to objects where Hα, Hβ, and [OIII] were detected, and an upper limit on [NII] was available. Finally, in order

of decreasing priority, the remaining targets were selected based on: availability of spectroscopic redshift measurement from MOSDEF (with higher priorities given to those objects at $1.90 \leq z \leq 2.65$ than those at $1.40 \leq z \leq 1.90$ or $2.95 \leq z \leq 3.80$), objects observed as part of the MOSDEF survey without successful redshift measurements, and objects not observed on MOSDEF masks but contained within the 3D-HST survey catalog (Momcheva et al. 2016) and lying within the MOSDEF target photometric redshift and apparent magnitude range. These targets comprise $\sim$ 260 observed galaxies with redshifts $1.4 \leq z \leq 3.8$[1], which is large and diverse enough to create bins across multiple galaxy properties (e.g., location in the BPT diagram, stellar mass, SFR). For this analysis, we do not include the small fraction of objects identified as AGN based on their X-ray and rest-IR properties. Figure 4.1 displays the redshift histogram and distributions of Hα-based SFR and M_* derived from SED fitting (Kriek et al. 2015) of the objects in our sample. A more detailed description of our method for SED fitting is described in Section 4.3.

A detailed description of the LRIS data acquisition and data reduction procedures will be presented elsewhere, however a brief summary is provided here. The data were obtained using the Low-Resolution Imaging Spectrograph (LRIS; Oke et al. 1995) during five observing runs totalling ten nights between January 2017 and June 2018. We observed 9 multi-object slit masks with $1\farcs2$ slits in the COSMOS, AEGIS, GOODS-S, and GOODS-N fields targeting 259 distinct galaxies. We used the d500 dichroic, the 400 lines mm^{-1} grism blazed at 3400Å on the blue side, and the 600 lines mm^{-1} grating blazed at 5000Å on the red side. This setup provided continuous wavelength coverage from the atmospheric cut-off at 3100 Å up to a typical red wavelength limit of 7650 Å. The blue side yielded a

[1]Of the 260 observed galaxies, 214 galaxies had a redshift from the MOSDEF survey, with 32, 162, and 20 in the redshift intervals $1.40 \leq z \leq 1.90$, $1.90 \leq z \leq 2.65$, and $2.95 \leq z \leq 3.80$ respectively. The remaining 46 galaxies had either a spectroscopic redshift prior to the MOSDEF survey, or a photometric redshift, with 9, 31, and 6 in the redshift intervals $1.40 \leq z \leq 1.90$, $1.90 \leq z \leq 2.65$, and $2.95 \leq z \leq 3.80$ respectively.

Table 4.1. Summary of LRIS observations.

Field	Mask Name	R.A.	decl.	$t_{\rm exp}^{\rm Blue}$ [s]	$t_{\rm exp}^{\rm Red}$ [s]	$N_{\rm objects}$
COSMOS	co_l1	10:00:22.142	+02:14:25.623	25200	24080	33
COSMOS	co_l2	10:00:22.886	+02:24:45.096	24300	22716	31
COSMOS	co_l5	10:00:29.608	+02:14:33.037	21492	20736	27
COSMOS	co_l6	10:00:39.965	+02:17:28.409	25020	24264	26
GOODS-S	gs_l1	03:32:23.178	−27:43:08.900	39312	38664	30
GOODS-N	gn_l1	12:37:13.178	+62:15:09.647	27000	22968	30
GOODS-N	gn_l3	12:36:54.841	+62:15:32.920	32400	31500	27
AEGIS	ae_l1	14:19:14.858	+52:48:02.128	28188	26964	31
AEGIS	ae_l3	14:19:35.219	+52:54:52.570	34056	33120	25

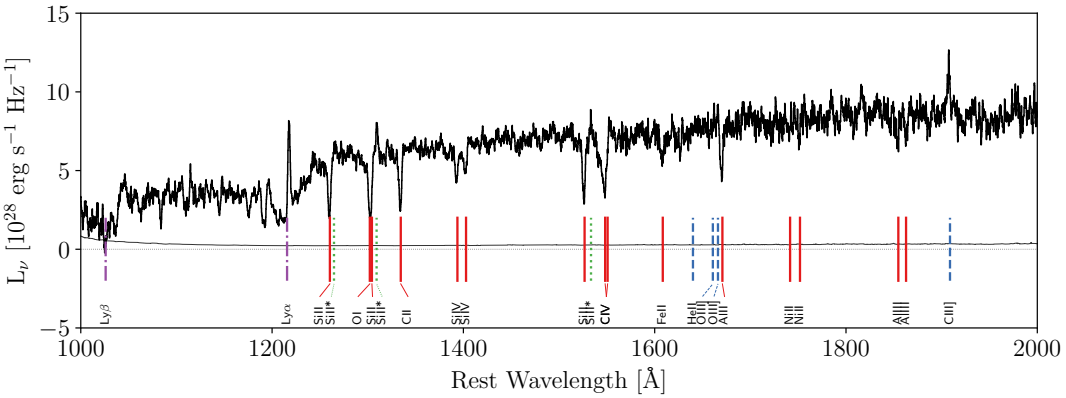

Figure 4.4: Stacked spectrum composed of all galaxies with a redshift measured from LRIS and in the interval $2.09 \leq z \leq 2.55$, totalling 62 galaxies. The magnitude of the 1σ error spectrum is depicted by the thin black line. Spectral features are identified using the same labeling scheme as in Figure 4.2, with the addition of the Lyman series marked by dashed-dotted purple lines.

resolution of $R \sim 800$, and the red side yielded a resolution of $R \sim 1300$. The median exposure time was 7.5 hours, but ranged from 6–11 hours on different masks. One night was lost completely due to weather. On 6/9 of the remaining nights the conditions were clear, and on 3/9 of the remaining nights there were some clouds, although we collected data on all three of those nights. The seeing ranged from 0.″6 to 1.″2 with typical values of 0.″8. Details of the observations are listed in Table 4.1.

We reduced the data from the LRIS red and blue detectors using custom `iraf`,

`idl`, and `python` scripts. We first fit polynomials to the traces of each slit edge, and rectified each slit accordingly, straightening the slit-edge traces. For blue-side images, we then flat fielded each frame using twilight sky flats, and dome flats for the red side. We cut out the slitlet for each object in all flat-fielded exposures. Following this step we used slightly different methods to reduce the red- and blue-side images. For each object, the blue-side slitlets were first cleaned of cosmic rays. Then, slitlets from each individual blue frame were background subtracted, registered and combined to create a stacked two-dimensional spectrum. We then performed a second-pass background subtraction on the stacked two-dimensional spectrum of each object while excluding the traces of objects in the slits in order to avoid over-subtraction of the background (Shapley et al. 2006). For the red-side images, we first background subtracted the individual frames, and cut out the slitlet for each object in all images. These individual slitlets were then registered and median combined using minmax rejection to remove cosmic rays, which more significantly contaminate the red-side slitlets. We then used the stacked two-dimensional spectra to measure the traces of objects in each slit. The abundance of sky lines in the red-side images prevented us from achieving an accurate second-pass background subtraction on the stacked two-dimensional spectra. Therefore, we masked out the spectral traces in the individual red side slitlets, and recalculated the background subtraction on the individual slitlets. These individual, background subtracted slitlets were re-registered and median combined with rejection again to create the final stacked image.

Following these steps, we extracted and wavelength calibrated the blue and red side 1D spectrum of each object. The wavelength solution was calculated by fitting a 4th-order polynomial to the red and blue arc lamp spectra, resulting in typical residuals of $\sim$ 0.035 Å and $\sim$ 0.3 Å for the red and blue side spectra, respectively. We repeated this reduction procedure a second time without background subtraction and measured the centroid of several known sky lines. We shifted the wavelength solution zeropoint so that the sky lines appear at their correct

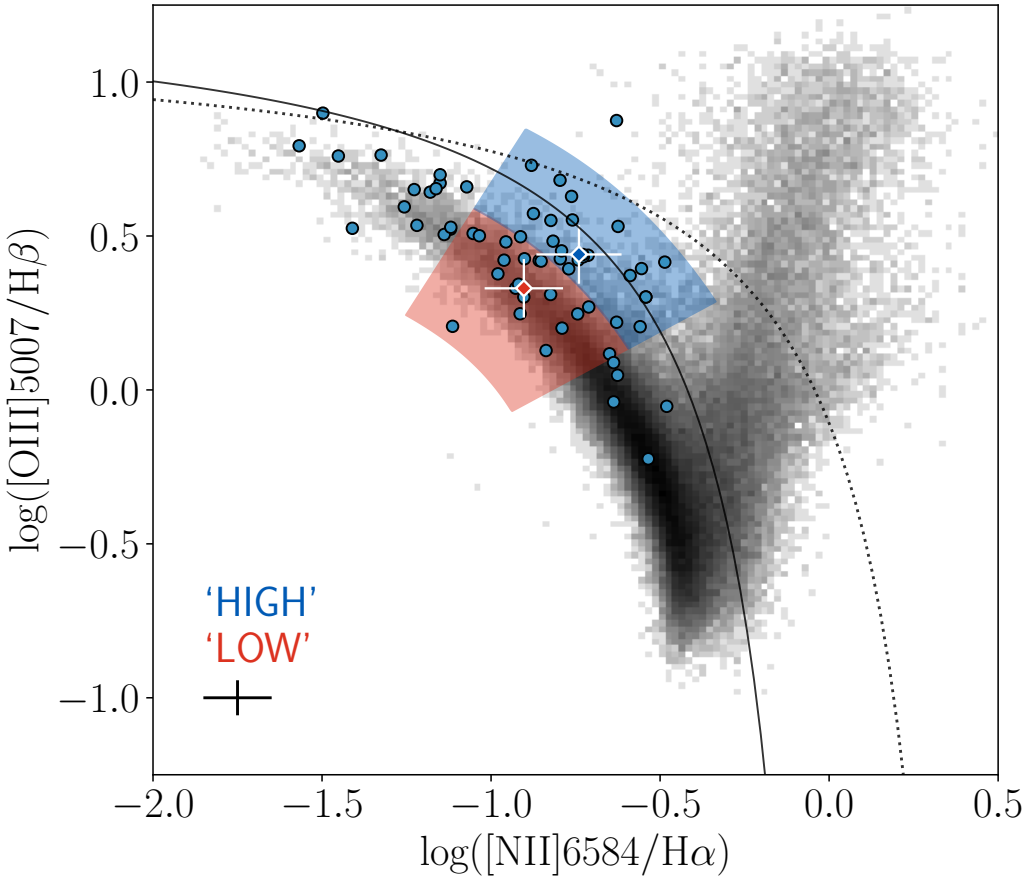

Figure 4.5: [O III]/Hβ vs. [N II]/Hα BPT diagram for $z \sim 2.3$ galaxies in our sample (blue points) as well as local SDSS galaxies (grey; Abazajian et al. 2009). We split the region of the BPT diagram most densely covered by our sample into two bins, one consisting of galaxies along the locus of $z = 0$ galaxies (red shaded region; *low* sample), and one bin of galaxies toward higher [O III]/Hβ and [N II]/Hα (blue shaded region; *high* sample). The median [OIII]/Hβ vs. [N II]/Hα values of galaxies in each bin are depicted as diamond-shaped symbols, with values of ([N II]/Hα, [OIII]/Hβ)=(-0.90 ± 0.12, 0.33 ± 0.09) for the *low* stack, and (-0.74 ± 0.13, 0.44 ± 0.09) for the *high* stack. For reference, the 'maximum starburst' model of Kewley et al. (2001) (dotted curve) and star-formation/AGN boundary from Kauffmann et al. (2003) (solid curve) are plotted. A median error bar for the $z \sim 2.3$ sample is shown in the bottom left.

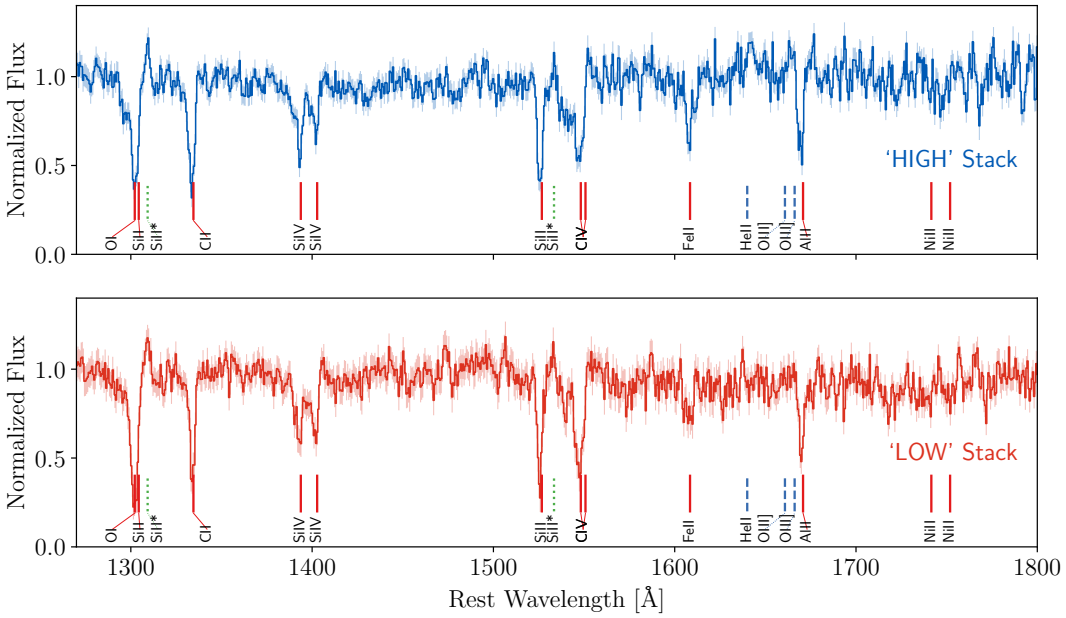

Figure 4.6: Top: A composite rest-UV spectrum of the 22 galaxies comprising our *high* stack. Bottom: A composite rest-UV spectrum of the 19 galaxies comprising our *low* stack. Spectral features are labeled as in Figure 4.2. Figure 4.9 shows a zoomed-in comparison of the CIV$\lambda\lambda$1548, 1550 and HeIIλ1640 lines. The 1σ composite error spectrum is indicated by the shaded region surrounding each spectrum.

wavelength values, and found the median required shift had a magnitude of $\sim$ 4Å in either direction. To apply the flux calibration, we used a first pass calibration based on spectrophotometric standard star observations obtained through a long slit during each observing run. We performed a final, absolute flux calibration for each galaxy by comparing 3D-HST photometric measurements with spectrophotometric measurements calculated from our spectra, and normalized our spectra so that our calculated magnitudes matched the 3D-HST values. After this absolute calibration, we checked that the continuum levels of the red and blue side spectra matched on either side of the dichroic cut-off at 5000 Å. Figure 4.2 shows some examples of reduced high-SNR continuum normalized spectra. Several strong absorption features are commonly visible, including SiIIλ1260, OIλ1302+SiIIλ1304, CIIλ1334, SiIV$\lambda\lambda$1393, 1402, CIV$\lambda\lambda$1548, 1550 , and AlIIλ1670.

4.2.3 Redshift Measurements

We measured a redshift for each object based on the Lyα emission line, as well as low-ionization interstellar (LIS) absorption lines, namely, SiIIλ1260, OIλ1302+SiIIλ1304, CIIλ1334, SiIIλ1526, FeIIλ1608, and AlIIλ1670, where available. Due to the presence of galaxy-scale outflows, the Lyα emission and interstellar absorption lines are commonly Doppler shifted away from the systemic redshift, $z_{\rm sys}$. Therefore, we defined two different redshift measurements, $z_{\rm Ly\alpha}$, and $z_{\rm LIS}$. We used the systemic redshift measured from nebular emission lines as part of the MOSDEF survey, when available, as an initial guess for $z_{\rm Ly\alpha}$, and $z_{\rm LIS}$. If no redshift was present for an object in the MOSDEF survey, we manually inspected the LRIS spectrum and measured the redshift based on any available features. This manually measured redshift was then used as an initial guess for our redshift measurement analysis. We measured the centroid of each line by simultaneously fitting the local continuum and spectral line with a quadratic function and a single Gaussian respectively. We restricted the amplitude of the Gaussian to be ≥ 0 for the Lyα emission line, and ≤ 0 for the absorption lines. We repeated this fitting process 100 times for each line, and with every iteration we perturbed the spectrum by its corresponding error spectrum. The average and standard deviation of the centroids from the 100 trials became the measured redshift and uncertainty for each spectral line. We manually inspected the fits to each line, and excluded that line if the fits were poor. We calculated the final $z_{\rm LIS}$ using the available interstellar absorption lines for each galaxy by giving priority to absorption lines that provide a more accurate measurement of the redshift. The SiIIλ1260, CIIλ1334, and SiIIλ1526 absorption lines provide the best options to use as a redshift measurement, as they are not contaminated by nearby features (Shapley et al. 2003). We averaged any successful redshift measurement from these three lines to obtain $z_{\rm LIS}$ (162 objects). If an object did not have a redshift measurement for any of these three lines, we defined $z_{\rm LIS}$ by using the AlIIλ1670 line (1 object). If this line was

also not available we used the blended OIλ1302+SiIIλ1304 line (6 objects). We established relations between systemic redshifts from the MOSDEF survey and redshift measurements from the rest-UV spectrum to infer the systemic redshift for galaxies without MOSDEF measurements. In particular, we set $z_{\,\mathrm{sys}}$ as:

$$\begin{cases} z_{\,\mathrm{sys}} = z_{\,\mathrm{LIS}} + 32.0(\frac{1+z_{\,\mathrm{LIS}}}{c}) & z_{\,\mathrm{LIS}} \text{ only} \\ z_{\,\mathrm{sys}} = z_{\,\mathrm{LIS}} + 89.0(\frac{1+z_{\,\mathrm{LIS}}}{c}) & z_{\,\mathrm{LIS}} \text{ and } z_{\,\mathrm{Ly}\alpha} \\ z_{\,\mathrm{sys}} = z_{\,\mathrm{Ly}\alpha} - 153.0(\frac{1+z_{\,\mathrm{Ly}\alpha}}{c}) & z_{\,\mathrm{Ly}\alpha} \text{ only; } z_{\,\mathrm{Ly}\alpha} \leq 2.7 \\ z_{\,\mathrm{sys}} = z_{\,\mathrm{Ly}\alpha} - 317.0(\frac{1+z_{\,\mathrm{Ly}\alpha}}{c}) & z_{\,\mathrm{Ly}\alpha} \text{ only; } z_{\,\mathrm{Ly}\alpha} \geq 2.7. \end{cases}$$

Finally, we used the systemic redshifts to shift each spectrum into the rest-frame. Out of the total 260 objects in our sample, 214 had systemic redshifts measured from the MOSDEF survey, 22 utilized our relations between $z_{\,\mathrm{sys}}$ and $z_{\,\mathrm{Ly}\alpha}$ or $z_{\,\mathrm{LIS}}$, and for the remaining 24 objects we were not able to measure a redshift.

4.2.4 The LRIS-BPT Sample

The full MOSDEF-LRIS sample consists of 260 galaxies across three distinct redshift intervals ($1.40 \leq z \leq 1.90$, $1.90 \leq z \leq 2.65$, and $2.95 \leq z \leq 3.80$). We define a subset of this sample, hereafter referred to as the LRIS-BPT sample, which is composed of galaxies in the central redshift range that have detections in the four primary BPT emission lines (Hβ, [OIII]λ5007, Hα, [NII]λ6584) at the $\geq 3\sigma$ level from the MOSDEF survey and a redshift measured from the LRIS spectrum. These criteria result in a sample of 62 galaxies that we define as "the LRIS-BPT sample." Due to the requirement of detections in the four rest-optical emission lines listed above, all 62 galaxies in this sample have a directly measured systemic redshift. Figure 4.4 displays the median-combined composite spectrum of the 62 galaxies in the LRIS-BPT sample.

We compared the population of galaxies in the LRIS-BPT sample with that

of the full MOSDEF survey. Figure 4.1 displays the SFR calculated from dust-corrected Balmer lines vs. M_* for both galaxies in the LRIS-BPT sample and the full MOSDEF sample. The LRIS-BPT sample is characterized by a median SFR of $\log(\mathrm{SFR}/(M_\odot/\ \mathrm{yr})) = 1.53 \pm 0.44$, and a median stellar mass of $\log(\mathrm{M_*/M_\odot}) = 10.02 \pm 0.52$. The median values are consistent with the properties of galaxies in the central redshift range ($1.90 \leq z \leq 2.65$) of the full MOSDEF survey, which has a median SFR of $\log(\ \mathrm{SFR}/(M_\odot/\mathrm{yr})) = 1.36 \pm 0.50$ and median mass of $\log(\ \mathrm{M_*/M_\odot}) = 9.93 \pm 0.60$. The similarity in median SFRs for the LRIS-BPT and total MOSDEF $z \sim 2$ samples also holds when using SFRs based on SED fitting, instead of from dust-corrected Balmer lines (Shivaei et al. 2016). These comparisons suggest that our LRIS-BPT sample is an unbiased subset of the full $z \sim 2$ MOSDEF sample.

4.2.5 Stellar Population Models

In order to determine the physical properties of the stars within our target galaxies we compared their observed spectra to a grid of stellar population models created with varying parameters. We used the Binary Population And Spectral Synthesis (BPASS) v2.2.1 models (Eldridge et al. 2017; Stanway & Eldridge 2018) because, relative to other recent models, they more accurately incorporate many key processes in the evolution of massive stars, including the addition of binary stars, rotational mixing, and Quasi-Homogeneous Evolution (QHE), resulting in longer main sequence lifetimes. We considered BPASS stellar population models with all available stellar metallicities ($10^{-5} \leq\ \mathrm{Z_*} \leq 0.04$), which primarily trace Fe/H (Steidel et al. 2016; Strom et al. 2018), and ages between 10^7yr and $10^{9.8}$yr in steps of 0.4 dex. The upper limit in age for this grid was chosen to include the age of the universe at the lowest redshift in our sample. We used the stellar population models that assume a Chabrier (2003) IMF, and have a high-mass cutoff of $100 M_\odot$. By default, BPASS provides models of an instantaneous burst of star

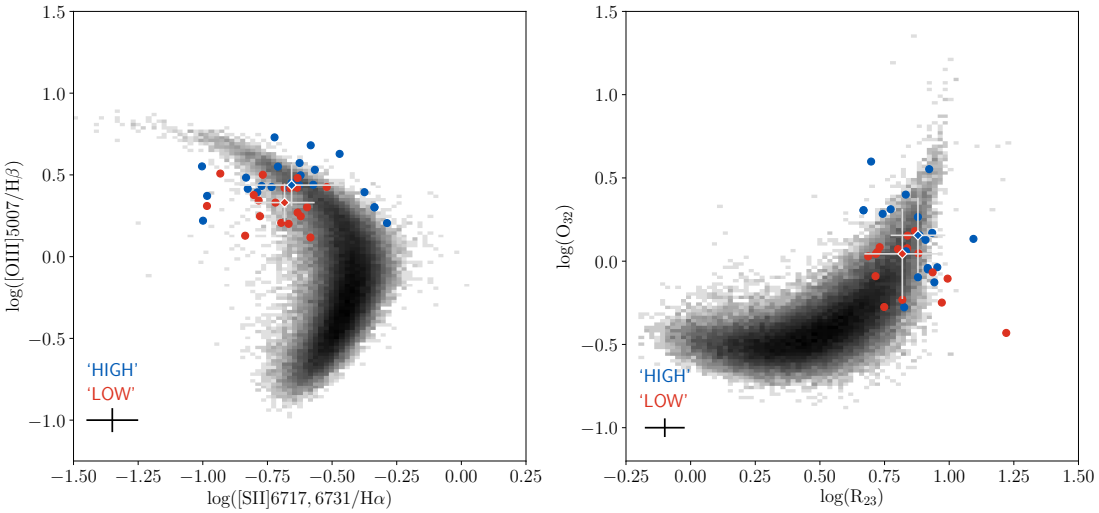

Figure 4.7: Left: [SII] BPT diagram. Galaxies in the *high* and *low* stacks are shown, respectively, using blue and red symbols. For comparison, the grey histogram shows the distribution of local SDSS galaxies. Galaxies in the *high* stack are offset on average toward higher [OIII]λ5007/Hβ, however there is overlap between the two samples. A median errorbar is shown in the bottom left. Right: log(O_{32}) vs. log(R_{23}) diagram. Symbols are the same as in the left panel. Galaxies in the *high* stack are offset on average toward higher log(O_{32}) and log(R_{23}), though there is overlap with the *low* stack.

formation. We constructed models assuming a constant star-formation history, by summing up the burst models, weighted by their ages.

In order to accurately compare our models with our observed spectra we must include contributions from the nebular continuum. To model the nebular continuum component of the UV spectrum we used the radiative transfer code Cloudy v17.01 (Ferland et al. 2017). For each individual BPASS stellar population of a given age and stellar metallicity, we ran a grid of Cloudy models with a range of nebular metallicities (i.e., gas-phase O/H) and ionization parameters. Our Cloudy grids include a range of nebular metallicities of $-2.0 \leq \log(Z_{\rm neb}/Z_{\odot}) \leq 0.4$ in 0.2 dex steps, and ionization parameters of $-4.0 \leq \log(U) \leq -1.0$ in 0.4 dex steps. All models were run assuming an electron density typical of galaxies at this redshift of $n_e = 250$ cm^{-3} (Sanders et al. 2016a; Strom et al. 2017). We set the abundance of nitrogen in the models using the log(N/O) vs. log(O/H) relation from Pilyugin et al. (2012):

$$\log(\text{ N/O}) = -1.493$$
$$\text{for } 12 + \log(\text{ O/H}) < 8.14$$
$$\log(\text{ N/O}) = 1.489 \times [12 + \log(\text{ O/H})] - 13.613$$
$$\text{for } 12 + \log(\text{ O/H}) \geq 8.14.$$

When using the stellar population models we added the contribution from the nebular continuum calculated assuming parameters typical of galaxies at this redshift ($\log(U) = -2.5$, $\log(Z_{\rm neb}/\ {\rm Z}_{\odot}) = -0.2$; Sanders et al. 2016a). Adjusting these parameters does not affect the nebular continuum significantly enough to alter the results of our model fitting. Figure 4.3 shows the differences in the CIV$\lambda\lambda$1548, 1550 profile for a subset of age and stellar metallicity models used in our analysis. Two key features are highlighted in blue and red, both of which are located within regions of the CIV$\lambda\lambda$1548, 1550 profile that are not strongly affected by contamination from interstellar absorption. Both of these features increase in strength towards higher stellar metallicity and younger ages. While the strengths of these features do not necessarily represent a unique combination of age and stellar metallicity, this degeneracy is broken by considering the full rest-UV spectrum.

4.2.6 Spectra fitting

We fit the combined stellar population plus nebular continuum models to our observed spectra in order to determine which stellar population parameters produce a spectrum that most closely matches our observed spectra. We first continuum normalized the observed and model spectra. In fitting the continuum level, we only considered the rest-frame spectral region at 1270 Å $\leq \lambda \leq$ 2000 Å to avoid the Lyα feature on the blue end, and a decrease in the quality of our spectra redwards of 2000 Å. To define an accurate continuum, we used spectral windows in regions of the spectrum relatively unaffected by stellar or nebular features, as defined by Rix et al. (2004). We averaged the flux in each of the windows and fit

a cubic spline through the windows to obtain the continuum level.

The models that we used consist of stellar and nebular continuum components only, so we masked out regions of the spectrum that contain other features, such as interstellar absorption. For this purpose, we adopted 'Mask 1' from Steidel et al. (2016) in the wavelength range 1270 Å − 2000 Å. To determine the best-fit age and metallicity, we first interpolated the model onto the wavelength scale of our observed spectrum, and then calculated the χ^2 for each model in our grid:

$$\chi^2 = \sum_i \frac{(f_{\mathrm{spec},i} - f_{\mathrm{model},i})^2}{\sigma_i^2}, \tag{4.1}$$

where $f_{\mathrm{spec},i}$, $f_{\mathrm{model},i}$, and σ_i^2 are the individual pixel values of the masked, continuum-normalized observed spectrum, masked, continuum-normalized model spectrum, and variance in the spectrum, respectively. We did not smooth either the models or the observed spectra as their resolutions were comparable with values of ~ 1Å in the rest-frame. This sum was typically carried out over $\sim$ 1000 wavelength elements, and resulted in a χ^2 surface in the log(Age/yr)-Z_* plane, which we interpolated using a 2D cubic spline and minimized to find the best-fit parameters. To calculate the uncertainties in these parameters, we perturbed the spectrum and repeated this process 1000 times to produce a distribution of best-fit values. We then defined the boundaries of the 1σ confidence interval at the 16th and 84th percentiles of this distribution.

In addition to fitting individual spectra, we applied our method to fit composite spectra. To construct a composite spectrum, we first interpolated continuum-normalized individual spectra to a common wavelength grid with the sampling of the typical blue-side spectra (i.e., the lower resolution side), resulting in a typical sampling of $\sim$ 0.6 Å/pixel in the rest frame. We then median combined the interpolated spectra to produce the final composite spectrum. We constructed the composite error spectrum using a bootstrap resampling method. For a composite spectrum composed of a given number of galaxy spectra, we first selected an equal

number of spectra from the composite sample with replacement. We perturbed the selected spectra by their corresponding error spectra, and median combined them to create a composite spectrum. This process was repeated 1000 times to assemble an array of composite spectra. Finally, the composite error spectrum was determined as the standard deviation of the distribution of flux values of the perturbed composite spectra at each wavelength element.

4.3 Results

To determine how galaxy properties vary across the BPT diagram, we create two stacks of galaxies with roughly comparable oxygen abundance based on their similar [NII]/Hα values, but characterized by different rest-optical line ratios relative to the $z = 0$ BPT excitation sequence. Figure 4.5 shows the regions we use to define our stacks. We label the stack of galaxies consistent with the $z = 0$ BPT locus as the *low* stack, and the stack of galaxies at higher [NII]/Hα and [OIII]/Hβ as the *high* stack. The two stacks contain a majority of the galaxies in our LRIS-BPT sample, with the *low* and *high* stacks comprising 19 and 22 galaxies respectively. Despite being composed of a large number of galaxies, each stack covers a small enough area on the BPT diagram to sample galaxies with similar emission line properties. Figure 4.6 shows the stacked, continuum-normalized spectra of galaxies in the *high* and *low* stacks in blue (top) and red (bottom) respectively. For completeness, Figure 4.7 shows the positions of galaxies in our *high* and *low* stacks on the [OIII]λ5007/Hβ vs. [SII]$\lambda\lambda$6717, 6731/ Hα and O_{32} vs. R_{23} emission-line diagrams, where $O_{32} =$ [OIII]$\lambda\lambda$4959, 5007/[OII]$\lambda\lambda$3726, 3729 and $R_{23} =$ ([OIII]$\lambda\lambda$4959, 5007 + [OII]$\lambda\lambda$3726, 3729)/$H\beta$. On both of these additional BPT diagrams, the median positions of the two stacks are offset, however there is overlap between the samples.

In order to estimate the average physical properties of galaxies in our two stacks, we fit models to our stacked spectra using the procedure described above.

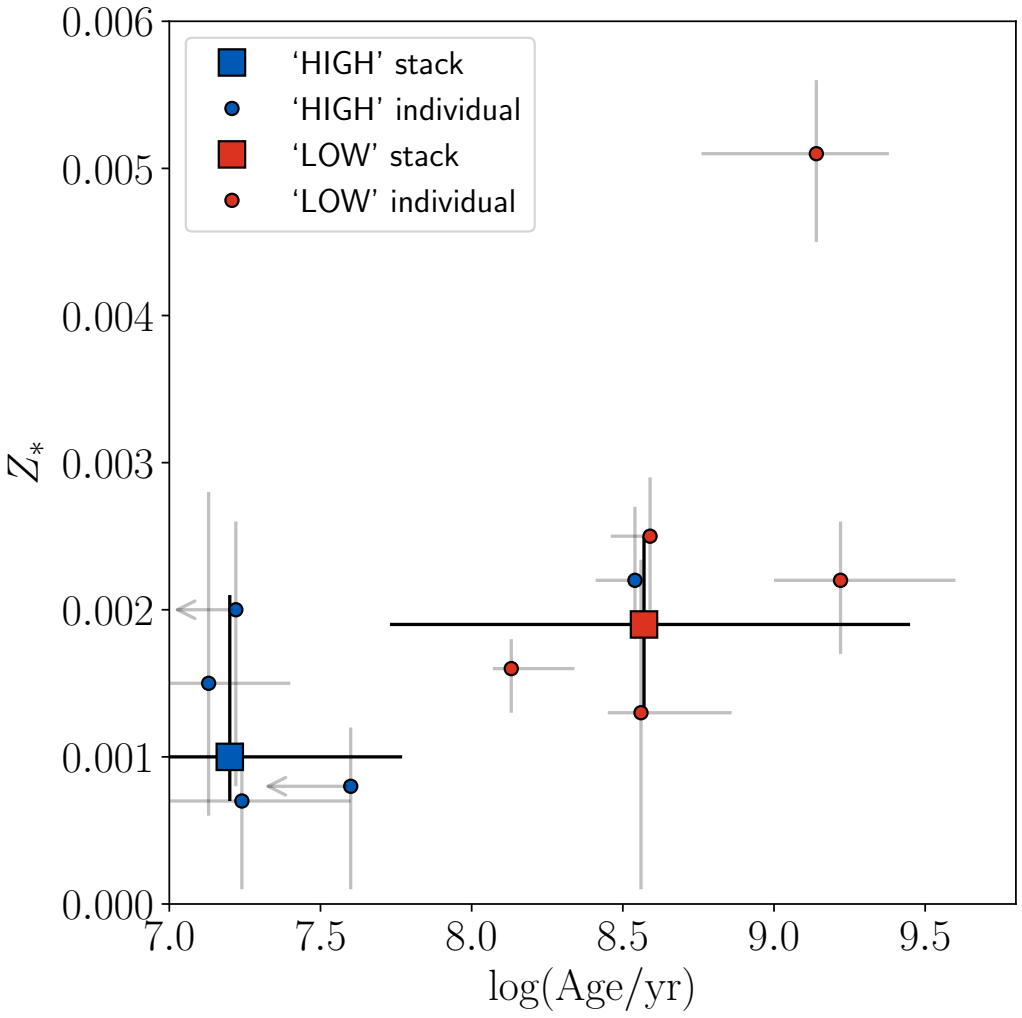

Figure 4.8: Best-fit Z_* and log(Age/yr) for our two stacked spectra, and the five highest continuum SNR individual spectra from each stack, ranging from $4.5 \leq$ SNR/pixel ≤ 12, measured in the wavelength range, $1425 \leq \lambda \leq 1500$. The remaining galaxies do not have high enough SNR (SNR/pixel $\lesssim 4$) for our fitting procedure to produce reliable results without stacking. The large square points show results from the two stacks, and the small individual points are for individual galaxy measurements. Points corresponding to the *high* (22 galaxies) stack are indicated in blue, those from the *low* (19 galaxies) stack are colored red. The results from fitting individual galaxy spectra are predominantly consistent with the stacked results, however one galaxy from the *high* stack has an age older than the stack. Two galaxies from the *high* stack had 1σ uncertainties at the edge of our grid, so they are represented as upper limits here. The *high* sample is characterized by a younger age and lower stellar metallicity compared to the *low* stack.

To measure uncertainties in these properties, we repeated the fitting process 1000 times, during each of which we recreated the stack using galaxy spectra randomly chosen from the original stack with replacement and perturbed by their corresponding error spectrum. Figure 4.8 shows the best-fit stellar parameters that we determined for our two stacks. Also shown are the results from applying our fitting procedure to the five individual galaxies with the highest SNR spectra in each bin. We find a stellar metallicity of $Z_* = 0.0010^{+0.0011}_{-0.0003}$ for the *high* stack, and a stellar metallicity of $Z_* = 0.0019^{+0.0006}_{-0.0006}$ for the *low* stack. Both of these metallicities are consistent with each other within 1σ. We find a best-fit stellar age for the *low* stack of $\log(\text{ Age/yr}) = 8.57^{+0.88}_{-0.84}$. At this age, the number of O-stars, and therefore the FUV spectrum, has largely equilibrated in a stellar population with a constant star-formation history, which results in the large error bars (Eldridge & Stanway 2012). We find $\log(\text{ Age/yr}) = 7.20^{+0.57}_{-0.20}$ for the *high* stack. This result suggests that the galaxies consistent with the *high* stack typically have younger stellar populations compared to those in the *low* stack.

We check the properties for the galaxies in each stack estimated by comparing their broadband SEDs to stellar population synthesis models. Briefly, this analysis uses the fitting code FAST (Kriek et al. 2009) to fit stellar population models from Conroy et al. (2009), assuming a Chabrier (2003) IMF and the Calzetti et al. (2000) dust reddening curve. The models also assume a "delayed-τ" star-formation history of the form: SFR $\propto t \times e^{(-t/\tau)}$, where t is the time since the onset of star formation, and τ is the characteristic star formation timescale. For a full description of the SED fitting procedure see Kriek et al. (2015). Based on the SED fitting, we find median stellar masses of $\log(\text{ M}/M_\odot) = 10.05 \pm 0.43$ and $\log(\text{ M}/M_\odot) = 10.12 \pm 0.32$ for galaxies in the *high* and *low* stacks respectively. Also, we find median SFR of $\log(\text{ SFR}_{\text{SED}}/(M_\odot/\text{yr})) = 1.33 \pm 0.42$ and $\log(\text{ SFR}_{\text{SED}}/(M_\odot/\text{yr})) = 1.38 \pm 0.34$ for the galaxies in the *high* and *low* stacks respectively. Therefore, both stacks comprise galaxies that are well matched in SFR and M_*. Additionally, the median SED-based age for galaxies in the *high*

stack (log(Age/yr) $= 8.5 \pm 0.4$) is younger than the median SED-based age in the *low* stack (log(Age/yr) $= 8.6 \pm 0.3$). This result from the broadband SED fitting agrees qualitatively with the younger age we find for the *high* stack based on the full rest-UV fitting. However, the SED-based ages for the two stacks are not significantly different. The differences between the ages inferred from the rest-UV spectra, and those reported from SED fitting likely arise for a couple of reasons. First, the rest-UV fitting only accounts for light from the most massive stars, while the SED-based results also include information from longer wavelengths. In addition, for the fitting in this work, we only consider a constant star-formation history, and the SED fitting employs a larger range of 'delayed-τ' star-formation histories of the form $\tau \times e^{-t/\tau}$, where both t and τ are fitted parameters. Incorporating more complex star-formation histories into our rest-UV fitting will be the subject of a future work.

In addition, the results from fitting model spectra to the high-SNR individual galaxy spectra are largely consistent with the results from using the stacked spectra. Four out of five individual galaxies from the *high* stack that we fit had stellar properties (age, stellar metallicity) consistent with stacked spectrum results, two of which were upper limits on the age. The remaining galaxy had a best fit age that was substantially older than the stack. All five individual galaxies in the *low* stack are consistent with an older population, and all but one object showed consistent metallicities with the stack. The best-fit parameters of the stacked spectra have larger uncertainties compared to the individual spectra, which suggests that our bootstrap resampling method is capturing galaxy-to-galaxy variations of age and Z_* in our sample.

In addition to our global rest-UV fitting procedure, which covers the full FUV spectrum at $1270\text{Å} \leq \lambda \leq 2000\text{Å}$, evidence for a difference in age between our two stacks is visible in the wind lines produced by massive stars: CIV$\lambda\lambda$1548, 1550 and HeIIλ1640 . Figure 4.9 shows these features for both of our stacks. The *high* stack has stronger CIV$\lambda\lambda$1548, 1550 emission ($1552\text{Å} - 1555\text{Å}$), as well as stronger stellar

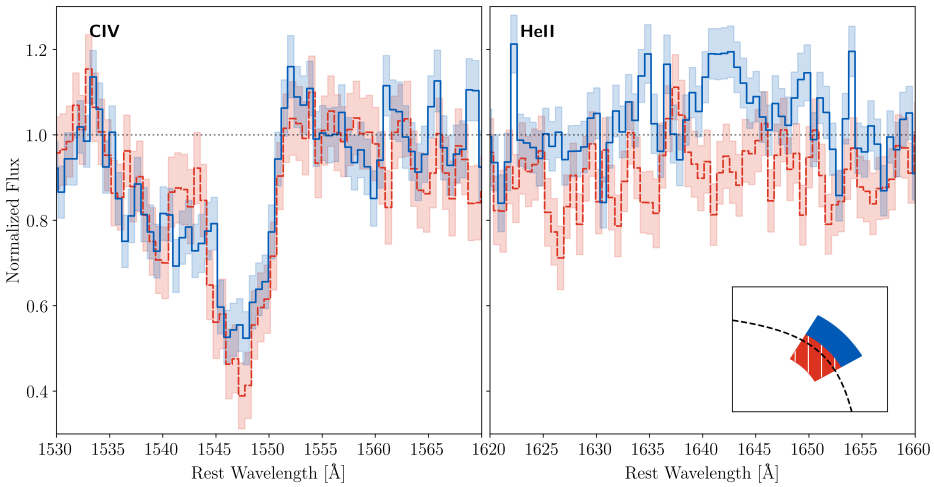

Figure 4.9: Zoomed-in regions around the CIV$\lambda\lambda$1548, 1550 (left) and HeIIλ1640 (right) features for our *high* (blue) and *low* (red) stacked spectra. The *high* stacked spectrum has stronger HeIIλ1640 and CIV$\lambda\lambda$1548, 1550 emission. Both features are signatures of massive stars, and are more prominent in younger populations. The 1σ composite error spectrum is depicted by the shaded region surrounding each spectrum. Figure 4.3 shows the age dependence of the CIV$\lambda\lambda$1548, 1550 line in the BPASS models.

wind absorption (1536Å − 1545Å) when compared to the *low* stack. This result is confirmed qualitatively by looking at the CIV$\lambda\lambda$1548, 1550 profiles produced by stellar population models, which predict stronger CIV$\lambda\lambda$1548, 1550 emission for younger stellar populations (Figure 4.3). In addition, the *high* stack shows a significant HeIIλ1640 emission line, whereas the *low* stack has none visible. Both of these features confirm the results of our fitting analysis suggesting that the stack of *high* galaxies shows evidence for stellar youth.

Using the best-fit stellar population parameters, we can examine the ionizing spectrum predicted by the BPASS models. Figure 4.10 shows the predicted ionizing spectrum for both the *high* and *low* stacks. The most massive stars, which are responsible for producing the ionizing radiation, have lifetimes much shorter than the ages of most of our models. Due to the assumed constant star-formation history in our models, the number of these massive stars equilibrates quickly ($\sim$ 10 Myr), and remains constant through most of our parameter space. As a result, the ionizing spectrum is similar between the two best-fit models to our observed

spectra, however the model corresponding to the *high* stack has a harder ionizing spectrum due to its lower stellar metallicity. Specifically, the ionizing flux normalized at 900Å, and integrated over the range 200 Å $\leq \lambda \leq$ 912 Å, is $\sim 7\%$ higher in the *high* stack compared to the *low* stack.

Using the predicted ionizing spectrum from our fitting analysis, we infer the nebular line fluxes expected for a given set of nebular parameters using Cloudy. We place our grid of Cloudy models on the [OIII]/Hβ vs. [NII]/Hα BPT diagram for the best-fit stellar spectrum of each of our stacks (Figure 4.11). We linearly interpolate the grid of [NII]/Hα and [OIII]/Hβ values produced by the Cloudy models to determine which $Z_{\rm neb}$ and $\log(U)$ best match the median observed line ratios of each stack. To estimate the uncertainty, we perturb the median observed [NII]/Hα and [OIII]/Hβ of the stacks by their uncertainties, and repeat the process 1000 times to create a distribution of values. Figure 4.12 (bottom row) displays the distributions of nebular metallicity and ionization parameter obtained from this analysis. We find an ionization parameter of $\log(U) = -3.04^{+0.06}_{-0.11}$ and nebular metallicity of $12 + \log(\ \mathrm{O/H}) = 8.40^{+0.06}_{-0.07}$ for the *high* stack. For the *low* stack, we find an ionization parameter of $\log(U) = -3.11^{+0.08}_{-0.08}$ and nebular metallicity of $12 + \log(\ \mathrm{O/H}) = 8.30^{+0.05}_{-0.06}$. While these differences are consistent to $\sim 1\sigma$, and are small given the dynamic range of ionization parameter in high-redshift star-forming galaxies, and systematic uncertainties in nebular metallicities, they have a measurable effect on the rest-optical emission ratios for the *high* and *low* stacks. We achieve similar results by instead fixing the ionizing spectrum for all galaxies in each stack, and inferring a distribution of nebular metallicities and ionization parameters of individual objects within the stack using the same method described above. Furthermore, we find the *high* and *low* stacks comprise samples with comparable electron density distributions, with median values of $n_e^{high} = 350 \pm 161$ cm^{-3} and $n_e^{low} = 334 \pm 282$ cm^{-3} respectively. Both medians are consistent with the value assumed in the Cloudy models ($n_e = 250$ cm^{-3}), and a Kolmogorov-Smirnov test determines a 47% probability that both samples

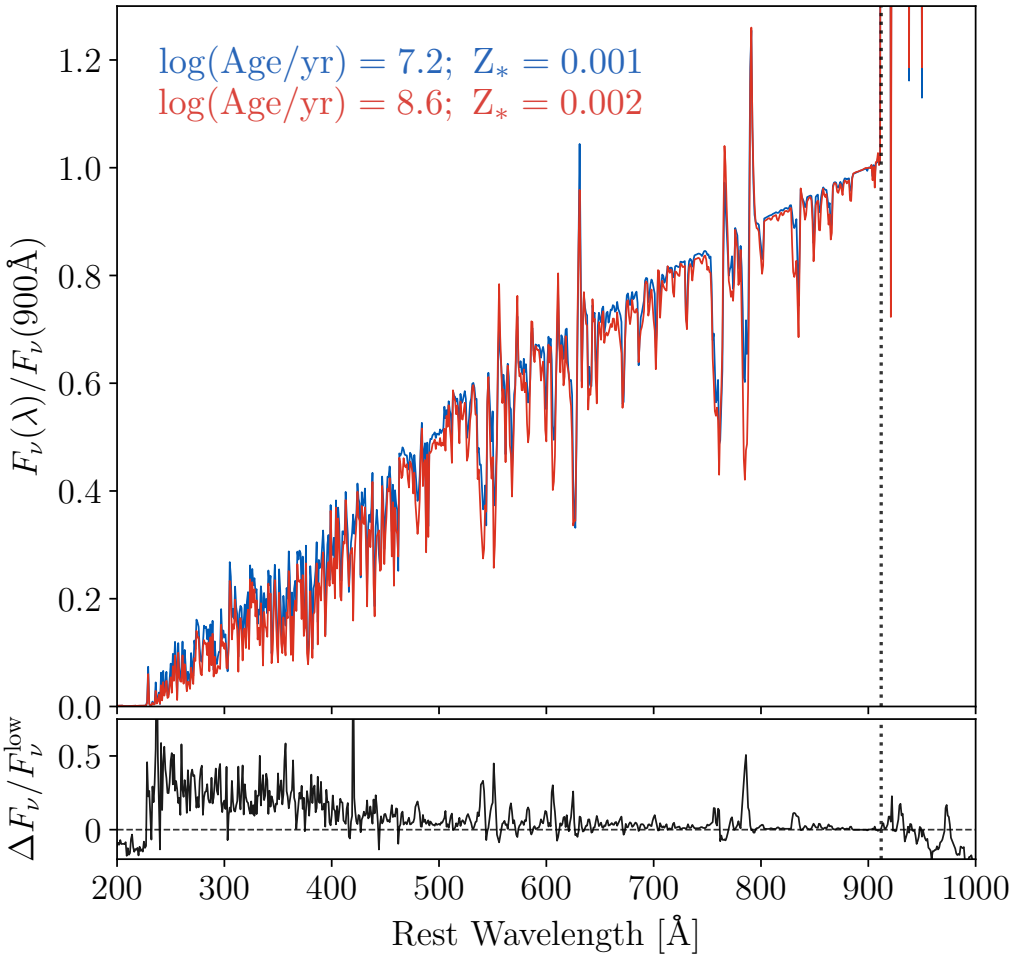

Figure 4.10: Top: Ionizing spectra of the best-fit stellar population models for the *high* (blue) and *low* (red) stacks. The ionizing spectra are normalized at 900Å. The normalized ionizing flux integrated in the range 200 Å $\leq \lambda \leq$ 912 Å is $\sim$ 7% higher in the *high* stack compared to the *low* stack. The vertical dotted line indicates the Lyman limit at 912 Å. Bottom: Fractional difference between the ionizing spectra of the best-fit stellar population models for the *high* and *low* stacks (i.e., $\Delta F_\nu/F_\nu{}^{\rm low} \equiv (F_\nu{}^{\rm high} - F_\nu{}^{\rm low})/F_\nu{}^{\rm low}$).

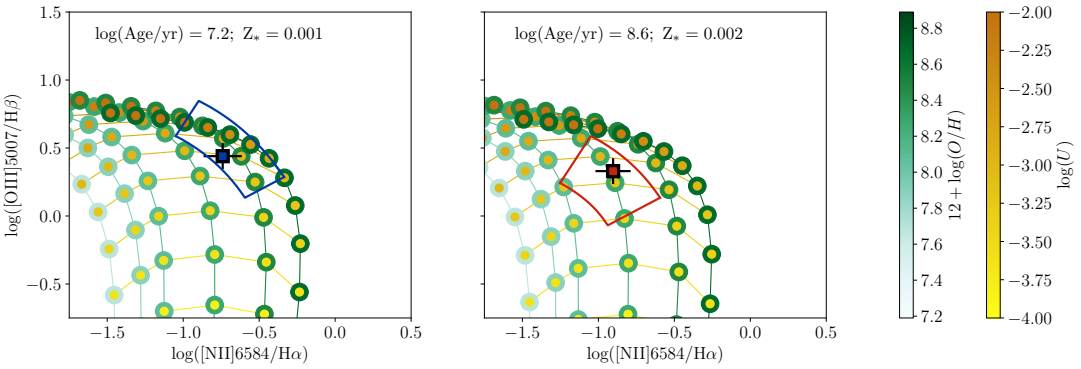

Figure 4.11: Predicted [OIII]/Hβ and [NII]/Hα emission-line ratios corresponding to the Cloudy+BPASS model grid of $12 + \log(\text{O/H})$ and $\log(U)$ for a given ionizing spectrum. The center of each point is color-coded by $\log(U)$, increasing from yellow to orange, while the border of each point is color-coded by $12 + \log(\text{O/H})$, increasing from light to dark green. The scale for each parameter is indicated by the color bars to the right of the panels. The median value and uncertainty in the observed [OIII]/Hβ and [NII]/Hα for galaxies in each bin are marked by the square points in each panel. The blue and red solid lines outline the regions of the BPT diagram we used to define our *high* and *low* stacks respectively. Left: Grid of line ratios assuming an ionizing spectrum corresponding to the best-fit stellar population for the high stack ($\log(\text{Age/yr}) = 7.2$, $Z_* = 0.001$). Right: Line ratios assuming an ionizing spectrum inferred from the best-fit stellar population model for the low stack ($\log(\text{Age/yr}) = 8.6$, $Z_* = 0.002$).

	$\log(\text{Age/yr})$	Z_*	$\log(U)$	$12 + \log(\text{O/H})$
high stack	$7.20^{+0.57}_{-0.20}$	$0.0010^{+0.0011}_{-0.0003}$	$-3.04^{+0.06}_{-0.11}$	$8.40^{+0.06}_{-0.07}$
low stack	$8.57^{+0.88}_{-0.84}$	$0.0019^{+0.0006}_{-0.0006}$	$-3.11^{+0.08}_{-0.08}$	$8.30^{+0.05}_{-0.06}$

Table 4.2: Best-fit physical parameters for the *high* and *low* stacks.

are drawn from the same parent distribution. Table 4.2 summarizes the best-fit physical parameters we find for the *high* and *low* stacks.

4.4 Discussion

Sensitive multiplexed spectroscopic instruments on large telescopes have enabled the study of rest-optical spectra for statistical samples of galaxies at high-redshift. These studies have established an offset of high-redshift star-forming galaxies towards higher [O III]/Hβ and [N II]/Hα compared to local galaxies. Several contributing factors have been proposed as the source of this offset, including varying

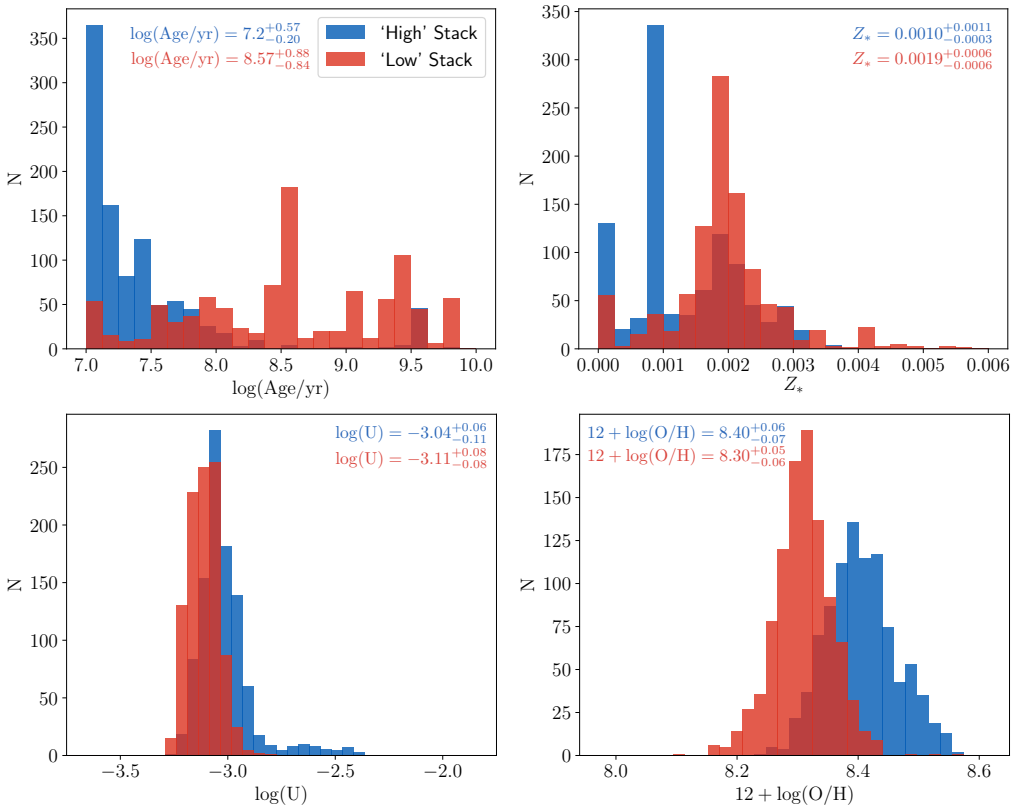

Figure 4.12: Distributions of best-fit age, stellar metallicity (Z_*), ionization parameter (U), and nebular metallicity (O/H) for our two stacks. The parameter distributions of age and stellar metallicity are produced by finding the minimum χ^2 from fitting the grid of Cloudy+BPASS stellar population models to a bootstrap resampled composite spectrum 1000 times. The parameter distributions of ionization parameter and nebular metallicity for each galaxy sample are produced by perturbing the sample median [O III]/Hβ and [N II]/Hα values by their uncertainties 1000 times and comparing the perturbed values to the inferred line ratios from our grid of Cloudy+BPASS models. All panels display parameter distributions for the *high* and *low* stacks in blue and red, respectively. Top Left: Distribution of best-fit ages. We find the *high* stack is younger ($\log(\text{ Age/yr}) = 7.20^{+0.57}_{-0.20}$) compared to the *low* stack ($\log(\text{ Age/yr}) = 8.57^{+0.88}_{-0.84}$). Top Right: Distribution of best-fit stellar metallicities. We find that the *high* stack has an overall lower stellar metallicity ($Z_* = 0.0010^{+0.0011}_{-0.0003}$) compared to the *low* stack ($Z_* = 0.0019^{+0.0006}_{-0.0006}$). Bottom Left: Distributions of best-fit ionization parameters. We obtain values of $\log(U) = -3.04^{+0.06}_{-0.11}$ for the *high* stack, and $\log(U) = -3.11^{+0.08}_{-0.08}$ for the *high* stack. Bottom Right: Distributions of best-fit nebular metallicities. We find best-fit values of $12 + \log(\text{ O/H}) = 8.40^{+0.06}_{-0.07}$ for the *high* stack, and $12 + \log(\text{ O/H}) = 8.30^{+0.05}_{-0.06}$ for the *low* stack.

abundance patterns, changes in the ionizing spectra, stellar and nebular metallicities, different ages of the stellar populations, and a different ionization parameter (Shapley et al. 2005; Erb et al. 2006; Liu et al. 2008; Kewley et al. 2013). One essential aspect of understanding these differences in high-redshift galaxies is a robust constraint on the ionizing spectrum produced by massive stars. Rest-UV spectroscopy of star-forming galaxies traces the properties of the massive star populations, and (given a set of stellar population synthesis modeling assumptions) provides constraints on the ionizing radiation field. In turn, photoionization modelling enables us to connect the ionizing spectrum and massive star population, with rest-optical nebular line ratios including those in the BPT diagram.

Currently, studies utilizing this combined rest-UV and rest-optical analysis have focused on average properties of the high-redshift population. By dividing our sample into two bins based on their location on the BPT diagram, we investigated how stellar population properties change as galaxies move away from the local BPT sequence. We found that the stack of galaxies above the local BPT sequence have younger ages, and lower stellar metallicities compared to galaxies along the local sequence at $z \sim 2$. Additionally, we find that galaxies above the local BPT sequence have harder ionizing spectra compared to their *low* stack counterparts. In our models, which assume a constant star-formation history, the most massive star population equilibrates on timescales ~ 10 Myr. Therefore, the difference in ionizing spectrum is not due to the age difference between our two stacks, but instead the lower stellar metallicity (i.e., Fe/H) in those galaxies that are offset.

While the most notable difference between our *low* and *high* stacks is a factor of ~ 2 lower stellar metallicity in the *high* stack, our photoionization modelling reveals small differences in the additional nebular parameters U and $Z_{\rm neb}$. All three of these parameters contribute to the observed rest-optical emission line ratios of the *high* and *low* stacks. Using photoionization modelling to measure $Z_{\rm neb}$ (nebular O/H), and using rest-UV spectral fitting to measure Z_* (stellar Fe/H), we find that both stacks have super-solar O/Fe, with our *low* and *high*

stacks having values of $3.04^{+0.95}_{-0.54}$ O/Fe$_\odot$ and $7.28^{+2.52}_{-2.82}$ O/Fe$_\odot$ respectively. While α-enhancement has previously been presented as an explanation for the offset of $z \sim 2$ galaxies in the BPT diagram, we stress that even galaxies that are entirely consistent with the local excitation sequence in the [OIII]λ5007/Hβ vs. [NII]λ6584/Hα diagram (i.e., the *low* stack) appear to be α-enhanced – in contrast with local systems. Such differences must be considered in order to accurately model the properties of these galaxies and to infer gas-phase oxygen abundances based on strong emission-line ratios. Without accounting for these differences, models will produce nebular metallicities biased toward higher $12 + \log(\text{O/H})$. The O/Fe value for the *high* stack is above the $\sim 5.5 \times$ O/Fe$_\odot$ theoretical limit assuming a Salpeter IMF and high-mass cutoff of 50 $M_\odot$ (Nomoto et al. 2006), but is still consistent within 1σ. However, the exact value of this theoretical limit is dependent on supernova yield models, which are not well constrained (Kobayashi et al. 2006). Kriek et al. (2016) found comparable α-enhancement in a massive quiescent galaxy at $z = 2.1$, reporting a Mg/Fe $= 3.9 \times$ Mg/Fe$_\odot$.

The assumed nitrogen abundance at fixed O/H affects where photoionization model grids fall in the BPT diagram, such that increasing N/O increases [NII]/Hα while keeping all other parameters fixed. Consequently, if our assumed N/O-O/H relation does not hold for typical $z \sim 2$ galaxies, then our inferred oxygen abundances will be systematically biased. An underestimate in N/O leads to an overestimate of O/H, and vice versa. Therefore, the high α-enhancement inferred in our offset galaxy stack could be due in part to differences in N/O at fixed O/H, perhaps due to the timescale of nitrogen enrichment in stellar populations (Berg et al. 2019). However, in order for the *high* and *low* stacks to each have solar O/Fe, an enhancement of N/O by ~ 1 dex and ~ 0.5 dex respectively at fixed O/H would be required. Given the age of both stacks, and the timescale of Fe enrichment from Type Ia supernovae (~ 1 Gyr), the absence of α-enhancement in either stack is unlikely. Another question is whether the difference in inferred α-enhancement for the two stacks can be explained by different N/O vs. O/H

relations. For O/Fe to match between the *high* and *low* stacks, we would need to assume an N/O higher by $\sim$ 0.6 dex for the *high* stack. For consistency at the 1σ level, the assumed N/O would need to be $\sim$0.2 dex higher for the *high* stack. Additionally, an O/Fe exceeding the theoretical limit of Nomoto et al. (2006) could be explained by a top-heavy IMF, or by increasing the high-mass cutoff of the stellar population. Investigating these possible differences in stellar populations is an avenue for future analysis.

To verify that our assumptions for the N/O ratio are reasonable, we compute the N/O ratio using the tracer, [NII]/[OII], for all objects in our stacks that have detections with $> 3\sigma$ in both lines. We find that the high and low stacks are characterized by a median [NII]/[OII]= -0.79 ± 0.25 and [NII]/[OII]= -0.99 ± 0.31 respectively. Based on the calibration of N/O as a function of [NII]/[OII] from Strom et al. (2018), these line ratios correspond to a log(N/O) = -1.05 ± 0.13 for the high stack, and log(N/O) = -1.15 ± 0.16 for the low stack. Using the N/O to O/H relation from Pilyugin et al. (2012), and the inferred nebular metallicity for out two stacks, we infer a nitrogen abundance of log(N/O) = -1.1 for the high stack, and log(N/O) = -1.25 for the low stack. These inferred values are both consistent with the nitrogen abundances computed based on [NII]/[OII], suggesting that our spectra are well described by the models.

We check the predicted O_{32} distribution for the best-fit nebular metallicity and ionization parameter inferred from our models, and compare it to the observed O_{32} distributions for our two stacks. We find that, on average, models for galaxies in the high stack have $O_{32} = 0.16 \pm 0.22$ while models for galaxies in the low stack have $O_{32} = 0.04 \pm 0.13$. These values are in agreement with the distributions of observed O_{32} measured from galaxies in our two stacks, for which we find $O_{32} = 0.15 \pm 0.22$ for the high stack, and $O_{32} = 0.03 \pm 0.16$ for the low stack. This agreement suggests that the best-fit models can self-consistently reproduce the observed O_{32} line ratio.

An intriguing question is if the high-redshift galaxies that lie along the local sequence (i.e., the *low* sample) be interpreted as descendants of the offset galaxies (i.e., the *high* sample). Qualitatively, it is suggestive that this may be the case based on the age dependence of α-enhancement seen in galactic bulge stars (Matteucci et al. 2016). However, chemical evolution models that incorporate realistic timescale differences between core collapse and Type Ia supernovae predict that significant evolution of O/Fe will only occur on timescales of $\sim$ 1 Gyr, assuming smooth star-formation histories (Weinberg et al. 2017), which is significantly longer than the age difference inferred between the high and low rest-UV composite spectra. In contrast, in the models of Weinberg et al. (2017), a sudden burst of star formation could temporarily boost O/Fe by $\sim$ 0.3 dex. Accordingly, galaxies in the high stack may show the evidence of recent bursts of star formation, and follow systematically different star-formation histories from those in the low stack. More detailed modelling will be required to see if this proposed explanation is applicable.

4.5 Summary & Conclusions

We have obtained rest-UV spectra for a sample of 259 galaxies at $1.4 \leq z \leq 3.8$ that were observed as part of the MOSDEF survey, enabling a combined analysis of rest-UV probes of massive stars and rest-optical probes of ionized gas. Of these galaxies, 62 are at $z \sim 2.3$ ($2.09 \leq z \leq 2.55$), and have all four BPT emission lines (Hβ, [OIII]λ5007, Hα, [NII]λ6584) detected at $\geq 3\sigma$. We constructed two composite rest-UV spectra of a subset of these 62 galaxy spectra based on their location on the BPT diagram. We tested how galaxy properties, including the age, stellar metallicity, nebular metallicity, and ionization parameter vary for galaxies on and off the local sequence. To derive these properties, we first fit a grid of Cloudy+BPASS stellar population synthesis models to constrain the age and stellar metallicity of the massive star population, therefore fixing the intrinsic

ionizing spectrum. With the ionizing spectrum established, we then computed optical emission line flux ratios using Cloudy for a grid of nebular metallicities and ionization parameters. Finally, we set the nebular metallicity and ionization parameter for our spectra based on the models that best reproduced the observed rest-optical emission line ratios. We summarize our main results and conclusions below.

(i) Using Cloudy+BPASS stellar population synthesis models we investigated how the age and stellar metallicity varies for high-redshift galaxies that lie on the local BPT sequence compared to those that are offset toward higher [OIII]λ5007/Hβ and [NII]λ6584/Hα. We found that the offset galaxies have younger ages (log(Age/yr) = $7.20^{+0.57}_{-0.20}$) compared to the galaxies in our sample that lie on the local sequence (log(Age/yr) = $8.57^{+0.88}_{-0.84}$). Additionally, we found that the offset galaxies had overall lower stellar metallicities ($Z_* = 0.0010^{+0.0011}_{-0.0003}$) compared to the non-offset galaxies ($Z_* = 0.0019^{+0.0006}_{-0.0006}$). These results are displayed in Figure 4.8.

(ii) We investigated how the ionizing spectrum of the best-fit stellar population synthesis models varies across the BPT diagram, and found that the galaxies that are offset from the local BPT sequence have a harder ionizing spectrum compared to those that are not offset (Figure 4.10). This difference is due to the lower stellar metallicity in the offset galaxies. Inferred ages for both composites are old enough such that in constant star-formation models, the number of O-stars has reached an equilibrium, and the age of the population no longer has a significant effect on the ionizing spectrum.

(iii) Using the ionizing spectrum inferred for each stack from the rest-UV spectral fitting, we computed the resulting emission line fluxes for a grid of nebular metallicity, $Z_{\rm neb}$, and ionization parameter, U (Figure 4.11). Accordingly, our rest-UV spectral analysis enabled us to fix one of the input free parameters for photo-ionization modeling – i.e., the form of the ionizing spectrum. We compared the resulting emission line flux ratios to the median observed ratios of our

stacks from the MOSDEF survey in order to infer $Z_{\rm neb}$ and U for our two galaxy stacks. We found that the offset (*high*) galaxies have an ionization parameter of $\log(U) = -3.04^{+0.06}_{-0.11}$ and the non-offset (*low*) galaxies have an ionization parameter of ($\log(U) = -3.11^{+0.08}_{-0.08}$). In addition, the offset galaxy stack has a slightly higher nebular metallicity ($12+\log(\ \mathrm{O/H}) = 8.40^{+0.06}_{-0.07}$) compared to the non-offset galaxy stack ($12 + \log(\ \mathrm{O/H}) = 8.30^{+0.05}_{-0.06}$). The stellar and nebular metallicities we derived for our *high* and *low* stack imply that the galaxies that are offset from the local BPT relation are more α-enhanced ($7.28^{+2.52}_{-2.82}$ $\mathrm{O/Fe}_{\odot}$) compared to those on the local sequence ($3.04^{+0.95}_{-0.54}$ $\mathrm{O/Fe}_{\odot}$).

Understanding the observed differences between local and high-redshift galaxies in terms of their physical properties is required for a complete galaxy evolution model. Thus far, these differences have mainly been probed in a sample-averaged sense, therefore variations across the high-redshift galaxy population cannot be determined. By stacking our sample based on BPT location we observed which differences were enhanced in high-redshift galaxies that are most offset from the local sequence. We found that high-redshift galaxies had several factors contributing to the offset, namely that the most offset galaxies have younger ages, lower stellar metallicities, higher ionization parameters, and higher nebular oxygen abundances. Notably, the offset galaxies are more α-enhanced compared to high-redshift galaxies that lie along the local sequence. Any photoionization modelling of $z \sim 2.3$ galaxies that do not take these differences into account, instead using local properties, will yield biased results. While α-enhancement was found to be heightened in the most offset galaxies, some level of enhancement is present throughout the high-redshift sample–even those coincident with the local sample. Therefore, interpreting the agreement between the location local galaxies and some high-redshift galaxies (i.e., our *low* sample) on the BPT diagram as a similarity of physical properties is an oversimplification. While our method of inferring Z_* from rest-UV spectral fitting, and $Z_{\rm neb}$ from photoionization modelling has not been applied to local galaxies, joint studies of the local stellar and gas-phase mass-

metallicity relations suggest that $\sim L*$ star-forming galaxies in the local universe are not α-enhanced (Zahid et al. 2017).

While we have refined the results of previous studies by measuring variations in high-redshift galaxy properties on and off the local sequence, a further refinement of composite spectra, or large numbers of high-SNR individual galaxies is still required. In addition, during this analysis we made several assumptions about the stellar populations of these galaxies, namely constant star-formation histories, and a single IMF. Future investigations will need to examine more general star-formation histories and variations in the IMF in order to more accurately constrain galaxy properties at high redshift.

CHAPTER 5

The MOSDEF-LRIS Survey: Individual Galaxy Analysis of Massive Stars and Ionized Gas at High Redshift

5.1 Introduction

Studies of large numbers of high-redshift galaxies in the rest-optical have revealed a wealth of information about the physical conditions of their interstellar medium (ISM). In the local universe, measurements of optical emission lines reveal that star-forming galaxies follow a tight sequence of simultaneously increasing [NII]λ6584/Hα and decreasing [OIII]λ5007/Hβ (e.g., Veilleux & Osterbrock 1987; Kauffmann et al. 2003). Analogous observations of galaxies at high-redshift expose a similar sequence, but offset toward higher [OIII]λ5007/Hβ and [NII]λ6584/Hα on the "BPT" diagram relative to local galaxies(Baldwin et al. 1981; Shapley et al. 2005; Erb et al. 2006; Liu et al. 2008).

Many properties of galaxies at high redshift may be responsible for this observed difference on the BPT diagram, including higher ionization parameters, harder ionizing spectra at fixed nebular metallicity, higher densities, variations in gas-phase abundance patterns, and enhanced contributions from AGNs and shocks at high redshift (see e.g., Kewley et al. 2013, for a review). Initial results from the MOSFIRE Deep Evolution Field (MOSDEF; Kriek et al. 2015) survey, suggested that the observed offset is primarily caused by an enhanced N/O ratio at fixed oxygen abundance, in addition to higher physical densities in high-redshift galax-

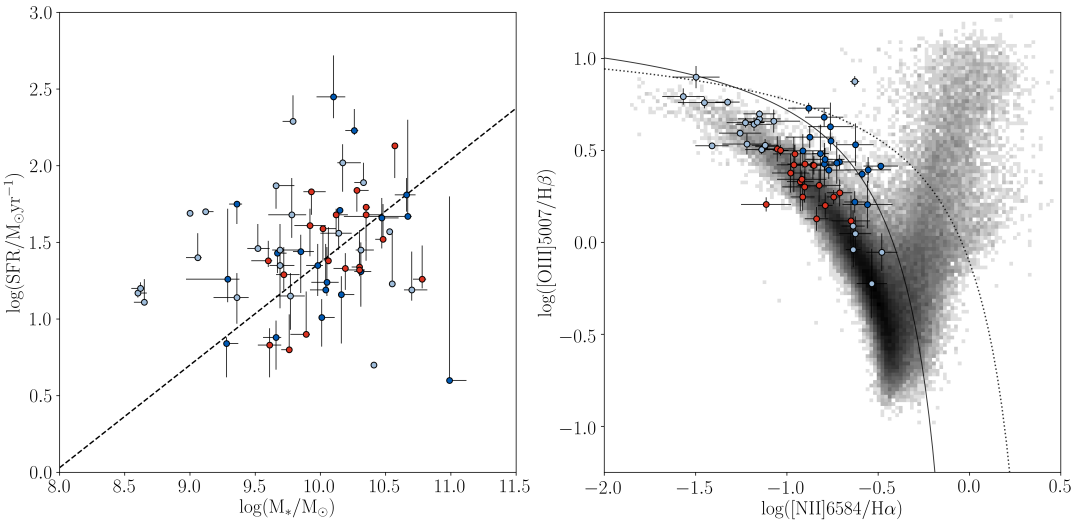

Figure 5.1: Properties of the LRIS-BPT sample. Left: SFR calculated from the dust-corrected Balmer lines vs. M_* for all objects with LRIS redshifts at $2.0 \leq z \leq 2.7$. Blue and red points indicate galaxies included, respectively, in the *high* and *low* composite spectra described in Topping et al. (2019). Right: Location of the LRIS-BPT sample on the [OIII]λ5007/Hβ vs, [NII]λ6584/Hα BPT diagram. The grey histogram shows the location of SDSS galaxies (grey; Abazajian et al. 2009). As in the left panel, galaxies comprising either the *high* or *low* stack are colored blue and red respectively. For reference, the 'maximum starburst' model of Kewley et al. (2001) (dotted curve) and star-formation/AGN boundary from Kauffmann et al. (2003) (solid curve) are plotted.

ies compared to local systems (Masters et al. 2014; Shapley et al. 2015; Sanders et al. 2016b). Results from the Keck Baryonic Structure Survey (KBSS; Steidel et al. 2014) instead suggested that the offset is driven primarily by a harder intrinsic ionizing spectrum at fixed oxygen abundance (Steidel et al. 2016; Strom et al. 2017). Updated results from the MOSDEF survey support the explanation of a harder ionizing spectrum at fixed oxygen abundance (Sanders et al. 2019; Shapley et al. 2019). Furthermore, a harder stellar ionizing spectrum at fixed oxygen abundance arises naturally due to lower stellar metallicity, which primarily traces Fe/H, reflecting an α-enhancement of the most massive stars that produce the bulk of the ionizing radiation in star-forming galaxies. This α-enhancement is expected in high-redshift galaxies due to their rapid formation timescale resulting in enrichment primarily from Type II supernovae.

The rest-optical emission lines observed in high-redshift star-forming galaxies are strongly affected be the intrinsic ionizing spectrum primarily produced by the most massive stars. Several properties of the massive stars affect the production of ionizing photons, including stellar metallicity, IMF, and stellar binarity (Topping & Shull 2015; Steidel et al. 2016). In addition to controlling the ionizing spectrum, the formation of massive stars is regulated by gas accretion onto galaxies, and in turn regulates the resulting chemical enrichment of the ISM through the deposition of metals by supernova explosions and stellar winds, and the ejection of metals through star-formation-driven galaxy-scale outflows. Strom et al. (2018) investigated the relationship between properties of the ionized ISM and factors contributing to the excitation state within galaxies, including the stellar metallicity. Also in an effort to connect factors affecting the ionizing spectrum with properties of the ISM, Sanders et al. (2019) used photoionization modelling to constrain the ionization parameter and stellar metallicity of massive stars. However, breaking the degeneracy between the ionization parameter and the intrinsic ionizing spectrum is challenging when only rest-optical emission line ratios are available. Therefore, direct constraints on the ionizing spectrum are imperative

to fully understand the physical conditions within high-redshift galaxies.

While directly observing the intrinsic ionizing spectrum within high-redshift galaxies is challenging, information about the massive star population can be determined based on the rest-UV stellar spectrum. Specifically, features including the CIV$\lambda\lambda$1548, 1550 and HeIIλ1640 stellar wind lines, and a multitude of stellar photospheric features are sensitive to the properties of massive stars (Leitherer et al. 2001; Crowther et al. 2006; Rix et al. 2004). For example, Halliday et al. (2008) utilized FeIII absorption lines to measure a sub-solar stellar metallicity for a composite spectrum of $z \sim 2$ galaxies. Sommariva et al. (2012) tested the ability of additional photospheric absorption line indicators to accurately determine the stellar metallicity of the massive star population of high-redshift galaxies. Recently, instead of using integrated regions of the rest-UV spectrum, Cullen et al. (2019) instead fit stellar population models to the full rest-UV spectrum of multiple composite spectra to investigate galaxy properties across $2.5 < z < 5.0$ and a stellar mass range of $8.5 < \log(M_*/M_\odot) < 10.2$.

Crucially, recent studies have concurrently studied the production of the ionizing spectrum with the rest-optical properties of high-redshift galaxies by utilizing simultaneous rest-UV and rest-optical spectroscopy (Steidel et al. 2016; Chisholm et al. 2019; Topping et al. 2019). Steidel et al. (2016) constructed a composite spectrum of 30 $z \sim 2.4$ star-forming galaxies from KBSS, and found that their rest-UV properties could only be reproduced by stellar population models that include binary stars, have a low stellar metallicity ($Z_*/Z_\odot \sim 0.1$), and a different, higher, nebular metallicity ($Z_{\rm neb}/Z_\odot \sim 0.5$). By analyzing a single composite rest-UV spectrum, Steidel et al. (2016) only probed average properties of their high-redshift galaxy sample. With single rest-UV and rest-optical composite spectra it is not possible to probe the average rest-UV spectral properties as a function of the location in the BPT diagram. In contrast, Chisholm et al. (2019) fit linear combinations of stellar population models to 19 individual galaxy rest-UV spectra at $z \sim 2$, and determined light-weighted properties. In Topping et al. (2019) we

compared the properties of two composite spectra one of which included galaxies lying along the local sequence and the other including galaxies offset towards high [NII]λ6584/Hα and [OIII]λ5007/Hβ. This analysis indicated that galaxies offset from the local sequence had younger ages and lower stellar metallicities, resulting in a harder ionizing spectrum in addition to a higher ionization parameter and a higher α-enhancement, all of which contributed to the different BPT diagram locations. Intriguingly, we found that even high-redshift galaxies coincident with local star-forming sequence on the BPT diagram were α-enhanced relative to their local counterparts, requiring consideration when modelling. In this paper, we improve the methods presented in Topping et al. (2019) by expanding the stellar population synthesis models to consider more complex star formation histories (SFHs), and including additional rest-optical emission lines to the photoionization modelling. Furthermore, we test the capability of the models to be fit to lower SNR spectra, and analyze a sample of $\sim$ 30 individual galaxies with combined rest-UV and rest-optical spectra.

The organization of this chapter is as follows: Section 5.2 describes our observations, data reduction, and methods. Section 5.3 presents the results of our analysis, Section 5.4 provides a discussion, and Section 5.5 gives summary of our key results. Throughout this paper we assume a cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $H_0 = 70$km s^{-1} Mpc^{-1}, and adopt solar abundances from Asplund et al. (2009) (i.e., $Z_\odot = 0.014$).

5.2 Methods and Observations

5.2.1 Rest-Optical Spectra and the MOSDEF survey

Our analysis utilizes rest-optical spectroscopy of galaxies from the MOSDEF survey (Kriek et al. 2015) at $z \sim 2.3$, observed using the Multi-Object Spectrometer for Infra-Red Exploration (MOSFIRE; McLean et al. 2012) over 48.5 nights during

2012–2016. This rest-optical spectroscopic sample is composed of $\sim$ 1500 near-infrared spectra at moderate resolution (R $\sim$ 3500) of H-band selected galaxies targeted to lie within three distinct redshift intervals ($1.37 \leq z \leq 1.70$, $2.09 \leq z \leq 2.61$, and $2.95 \leq z \leq 3.80$). Based on the scatter between photometric and spectroscopic redshifts of the MOSDEF targets, the actual redshift ranges slightly differ from the initial target ranges. Therefore, we define the true redshift ranges as $1.40 \leq z \leq 1.90$, $1.90 \leq z \leq 2.65$, and $2.95 \leq z \leq 3.80$. In addition to rest-optical spectra from the MOSDEF survey, MOSDEF targets have extensive ancillary datasets from the CANDELS (Grogin et al. 2011) and 3D-HST (Momcheva et al. 2016) surveys. MOSDEF spectra were used to measure fluxes and redshifts of all rest-optical emission lines detected within the Y, J, H, and K bands, the strongest of which are: [OII]λ3727, Hβ, [OIII]$\lambda\lambda$4959, 5007, Hα, [NII]λ6584, and [SII]$\lambda\lambda$6717, 6731.

5.2.2 Rest-UV Spectra and the MOSDEF-LRIS sample

A full description of the rest-UV data collection and reduction procedures will be described in a future work, however we provide a brief description here. We selected a subset of MOSDEF galaxies for rest-UV spectroscopic followup using the Low-Resolution Imaging Spectrograph (LRIS; Oke et al. 1995). Target priorities were set using the following prescription. Highest priority was given to objects from the MOSDEF survey that had detections in all four BPT emission lines (Hβ, [OIII], Hα, [NII]λ6584) with $\geq 3\sigma$. Then, objects were added to the sample with detections in Hβ, [OIII]λ5007, and Hα with $\geq 3\sigma$, and an upper limit in [NII]λ6584. With decreasing priority, the remaining targets were selected by having a spectroscopic redshift measurement from the MOSDEF survey, objects from the MOSDEF survey without a successfully measured redshift, and objects not targeted as part of the MOSDEF survey, but that were part of the 3D-HST survey catalog (Momcheva et al. 2016) with properties within the MOSDEF survey

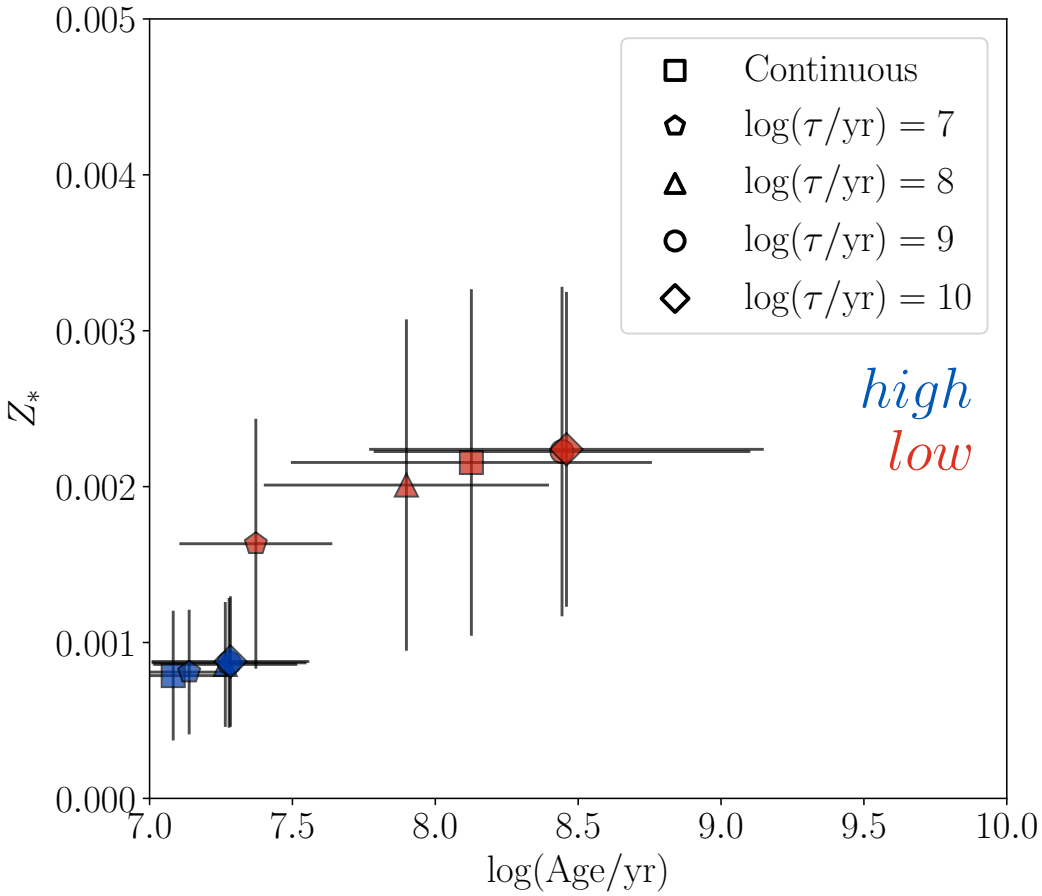

Figure 5.2: Best-fit stellar metallicity and age for the *high* (blue) and *low* (red) for five different star formation histories including a continuous SFH, and four realizations of the delayed-τ model, each depicted by a different shape. In all cases, galaxies in the *high* stack have younger ages and lower stellar metallicities compared to the *low* stack. The best-fit age and stellar metallicity increases with increasing τ when considering models with a 'delayed-τ SFH.

photometric redshift and apparent magnitude ranges. In total, these targets comprise a sample of 260 galaxies. Of those targets with spectroscopic redshifts from the MOSDEF survey, 32, 162, and 20 were in the redshift ranges $1.40 \leq z \leq 1.90$, $1.90 \leq z \leq 2.65$, and $2.95 \leq z \leq 3.80$ respectively. The remaining galaxies, with either a spectroscopic redshift not from MOSDEF, or a photometric redshift, made up 9, 31, and 6 galaxies in the redshift intervals $1.40 \leq z \leq 1.90$, $1.90 \leq z \leq 2.65$, and $2.95 \leq z \leq 3.80$ respectively. For this analysis, we excluded the few objects that have been identified to be AGN based on their X-ray and rest-frame near-IR properties.

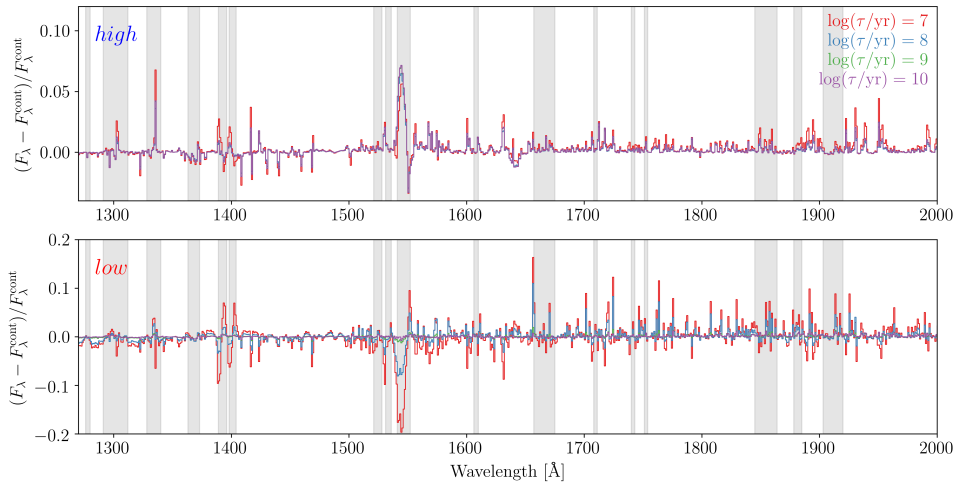

Figure 5.3: Fractional difference between the best-fit model spectra from models using a delayed-τ SFH compared to the model spectrum using a constant SFH. Best-fit models fit to the *high* and *low* stacks are displayed in the top and bottom panels respectively. The regions masked out using 'mask1' from Steidel et al. (2016) defined to include contamination from non-stellar sources is shown in grey. On average, the models assuming a delayed-τ SFH differ from those with a constant SFH at the few percent level.

Rest-UV spectra were obtained using Keck/LRIS during ten nights across five observing runs between January 2017 and June 2018. Our target sample totals 260 distinct galaxies on 9 multi-object slit masks with 1.″2 slits in the COSMOS, AEGIS, GOODS-S, and GOODS-N fields. In order to obtain continuous wavelength coverage from the atmospheric cut-off at 3100Å up to a median red wavelength limit of ∼7650Å, we observed all slit masks using the d500 dichroic, the 400 lines mm^{-1} grism blazed at 3400Å on the blue side, and the 600 lines mm^{-1} grating blazed at 5000Å on the red side. This setup yielded a resolution of $R \sim 800$ on the blue side, and a resolution of $R \sim 1300$ on the red side. The exposure times ranged from 6–11 hours on different masks, with a median exposure time of ~ 7.5 hours for the full sample. The data were collected with seeing ranging from 0.″6 to 1.″2 with typical values of 0.″8.

We reduced the red- and blue-side data from LRIS using custom `iraf`, `idl`, and `python` scripts. First, we fit polynomials to the edges of it, and transformed each slit to be rectilinear. The subsequent steps required slightly different treat-

ment for the red and blue images. We flat fielded each image using twilight sky flats for the blue side, and dome flats for the red side images. Then we cut out each slitlet for each object in every flat-fielded exposure. Following this step, the blue-side slitlets were cleaned of cosmic rays, and background subtracted. These images were registered and median combined to create a stacked two-dimensional spectrum. In order to prevent over-estimation of the background due to the target, we measured the trace of each object in the stacked two-dimensional spectrum, and masked it out for a second-pass background subtraction (Shapley et al. 2006). For the red-side slitlets, we constructed a stacked two-dimensional spectrum by first registering and median combining the images using minmax rejection to remove cosmic rays. We used this stacked image to measure the object traces in each slitlet. We then recomputed the background subtraction in the individual images with the object traces masked out, as the stacked image is too contaminated by sky lines to achieve a good background subtraction. After the second pass background subtraction, the individual red-side slitlets were combined to create the final stacked image.

We extracted the 1D spectrum of each object from the red and blue side stacked slitlets. We calculated the wavelength solution by fitting a 4th-order polynomial to the red and blue arc lamp spectra, which resulted in residuals of ~ 0.035Å and ~ 0.3Å respectively. We repeated this step on a set of frames that had not had sky lines removed. Using the resulting sky spectra, we measured the centroid of several sky lines and shifted the wavelength solution zeropoint until the sky lines aligned with their known wavelengths. This shift typically had a magnitude of ~ 4Å throughout the sample. We applied an initial flux calibration based on spectrophotometric standard star observations obtained during each observing run. We checked the flux calibration by comparing spectrophotometric measurements calculated from our objects to measurements in the 3D-HST photometric catalog, and applied a multiplicative factor to correct our calibration. Following the final flux calibration, we ensured that continuum levels on either side of the dichroic at

log(Age/yr)	7.0, 7.3, 7.4, 7.5, 7.7, 8.0, 8.5, 9.0, 10.0
Z_*	$10^{-5}, 10^{-4}, 10^{-3}, 0.002, 0.003, 0.004, 0.006, 0.008, 0.01, 0.014, 0.02, 0.03, 0.04$
$\log(Z_{\rm neb}/Z_\odot)$	-1.3, -1.0, -0.8, -0.6, -0.5, -0.4, -0.3, -0.2, -0.1, 0.0, 0.1, 0.2
$\log(U)$	-3.6, -3.4, -3.2, -3.0, -2.8, -2.6, -2.4, -2.2, -2.0, -1.8, -1.6, -1.4

Table 5.1: Summary of model grid parameters. The age and stellar metallicity values correspond to BPASS models we fit to our observed spectra. For each combination of age and stellar metallicity, we computed a set of photoionzation models with the listed nebular metallicity and ionization parameter values.

5000Å were consistent. Common features visible in the spectra include: SiIIλ1260, OIλ1302+SiIIλ1304, CIIλ1334, SiIV$\lambda\lambda$1393, 1402, CIV$\lambda\lambda$1548, 1550 , FeIIλ1608, and AlIIλ1670. While the full sample described above consists of 260 galaxies across three distinct redshift intervals ($1.40 \leq z \leq 1.90$, $1.90 \leq z \leq 2.65$, and $2.95 \leq z \leq 3.80$), we focus on a subset of this sample composed of galaxies in the central redshift window that have detections in four primary BPT lines (Hβ, [OIII]λ5007, Hα, [NII]λ6584) at $\geq 3\sigma$ from the MOSDEF survey. This 'LRIS-BPT' sample comprises 62 galaxies, each of which has a systemic redshift.

5.2.3 Stellar Population Synthesis and Photoionization Models

For this analysis, we used the version 2.2.1 Binary Population and Spectral Synthesis (BPASS) stellar population models to interpret our observed rest-UV galaxy spectra (Eldridge et al. 2017; Stanway & Eldridge 2018). Notably, these stellar population models incorporate the effects of stellar rotation, quasi-homogeneous evolution, stellar winds, and binary stars. These effects can have a substantial effect on the spectrum of a model stellar population, and in particular, the EUV spectrum produced by massive, short-lived stars. The BPASS models are computed with multiple Initial Mass Functions (IMFs), including the Chabrier (2003) IMF, and IMFs with high-mass ($M \geq 0.5{\rm M}_\odot$) slopes of $\alpha = -2.00$, -2.35, and -2.70. In addition, the models using each IMF were computed with a high-mass cutoff of $100{\rm M}_\odot$ and $300{\rm M}_\odot$. For this analysis, we only considered

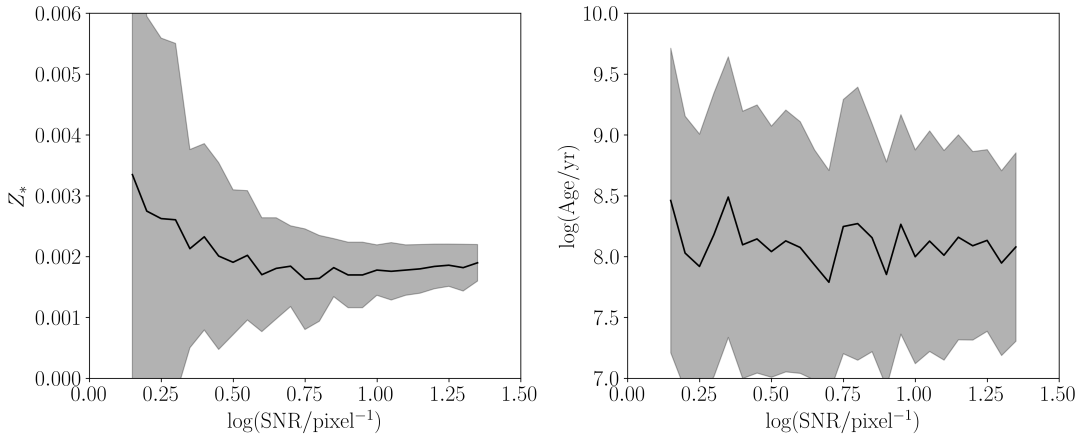

Figure 5.4: The best-fit stellar metallicity (left) and stellar age (right) as a function of spectra SNR created by artificially adding an increasing amount of noise to one of our composite spectra. The best-fit stellar metallicity remains consistent to the high-SNR value in the range 4.0 ≤SNR/pixel. The best-fit age remains consistent at all SNR values, however the 1σ uncertainties increase to the size of the parameter space ($7.0 \leq \log(\mathrm{Age/yr}) \leq 9.6$) at low SNR.

models computed assuming a Chabrier (2003) IMF, and a high-mass cutoff of $100\mathrm{M}_{\odot}$. Finally, the BPASS models have been computed with ages between $\log(\mathrm{Age/yr}) = 6.0 - 11.0$ in increments of 0.1 dex, and stellar metallicities of $Z_* = 10^{-5} - 0.04$. While we considered all available stellar metallicities in our analysis, we restricted the ages to $\log(\mathrm{Age/yr}) = 7.0 - 9.6$. At ages younger than $\log(\mathrm{Age/yr}) = 7.0$ we would be probing timescales shorter than the dynamical timescale of the galaxies, and therefore could not accurately attribute physical properties to the entire galaxy simultaneously. Also, at the lowest redshift galaxy in our sample, the age of the universe was $\sim$ 4Gyr, so including older stellar populations is not necessary.

We constructed stellar population models that assume different star formation histories (SFH) by combining the BPASS models, which describe a coeval stellar population, using:

$$F(\lambda) = \Psi_{t_0} f(\lambda)_{t_0} \Delta t_0 + \sum_{i=1}^{t_{\max}} \Psi_{t_i} f(\lambda)_{t_{\max} - t_i} (t_i - t_{i-1}), \tag{5.1}$$

where $t_{\max}$ is the age of the population, Ψ_{t_i} is the star-formation rate of the

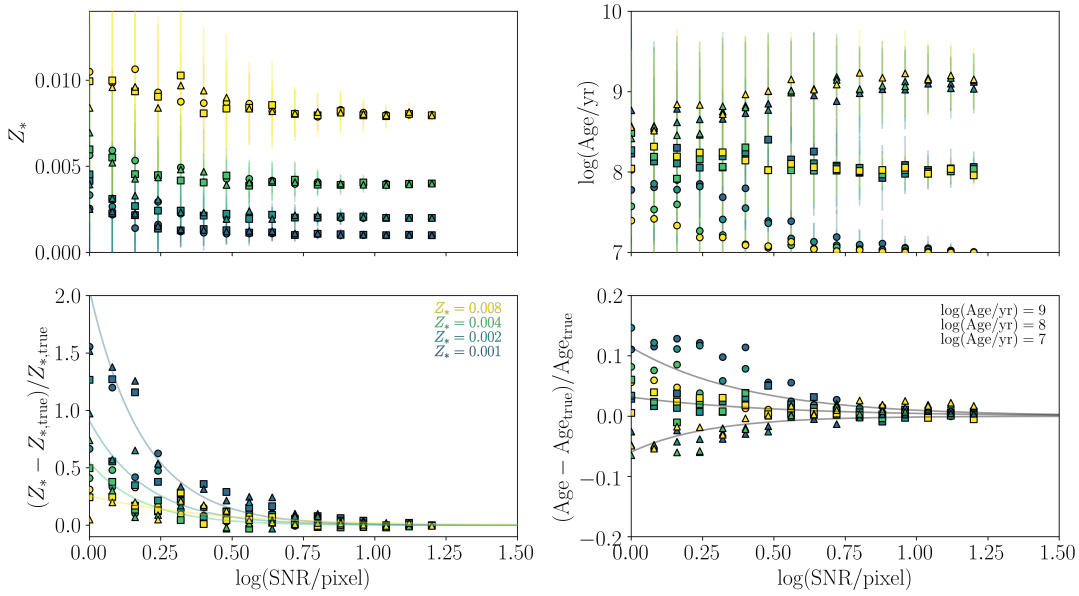

Figure 5.5: Best-fit stellar parameters as a function of SNR computed by adding varying amounts of noise to an array of BPASS model spectra for which the log(Age/yr) and Z_* are known. For all panels the color corresponds to the model input stellar metallicity, and the symbol depicts the input stellar age. Top Right: Best-fit stellar metallicities as a function of SNR computed for four different input values (0.001, 0.002, 0.004, 0.008) each of which has been computed at three different ages. At high SNR/pixel all of the models asymptote to their input values, however at SNR/pixel ≤ 4 the best-fit values are biased high. Bottom Right: Same as top right but for the fractional difference between the best-fit Z_* and the input Z_*. At the lowest SNR, the best-fit values can be biased high by $\sim 50\% - 150\%$. Top Left: Best-fit stellar ages as a function of SNR for three different input ages at a range of stellar metallicities. At the lowest SNR, the uncertainties expand to fill the parameter search range, and the best-fit values are biased toward log(Age/yr) ~ 8.5. Bottom Right: Fractional difference between the best-fit age and the input age for each of our models. The best-fit results begin to significantly diverge from their input values at SNR/pixel ~ 4.

population at time t_i, $f(\lambda)_{t_{\mathrm{max}}-t_i}$ is the model spectrum with age $t_{\mathrm{max}} - t_i$ (i.e., the model that began t_i years prior to the final age, t_{max}), and $(t_i - t_{i-1})$ is the time between subsequent model spectra. For the case of a constant SFH, all of the SFR weightings, Ψ_{t_i}, are set to unity. In addition to a constant SFH, we considered several models with a 'delayed-τ' SFH of the form SFR $\propto t \times e^{-t/\tau}$, with $\log(\tau/\mathrm{yr}) = 7, 8, 9, 10$. With this set of models we covered three schematically different SFHs. These SFHs allowed us to explore different regions of the SFH, including regions where it is rising, falling, and peaked.

We processed the stellar population model spectra using the photoionization

code Cloudy (Ferland et al. 2017). Using this code, we input an ionizing spectrum from BPASS and, given a set of ISM properties, calculated expected emission line fluxes. We compared the simulated line fluxes to the observed rest-optical emission lines of galaxies in our sample to infer properties of the ISM. We assumed a fixed electron density of $n_e = 250$ cm^{-3}, which is representative of the galaxies in our sample (Sanders et al. 2016a). In addition, while we vary the nebular oxygen abundance, we assume solar abundance ratios for most elements. However, we adopt the log(N/O) vs. log(O/H) relation from Pilyugin et al. (2012):

$$\log(\mathrm{N/O}) = -1.493 \quad \text{for } 12 + \log(\mathrm{O/H}) < 8.14$$
$$\log(\mathrm{N/O}) = 1.489 \times [12 + \log(\mathrm{O/H})] - 13.613 \quad \text{for } 12 + \log(\mathrm{O/H}) \geq 8.14.$$

For each BPASS model, we ran a grid of Cloudy models for a range of nebular metallicity ($-1.6 \leq \log(Z_{\mathrm{neb}}/Z_{\odot}) \leq 0.3$) and ionization parameter ($-3.6 \leq \log(U) \leq -1.4$). We have made several updates to the parameters of the model grids described in Topping et al. (2019), in order to more finely sample the parameter space in regions of interest. Table 5.1 summarizes the parameters, and lists each value for which we compute a model. An additional component of the photoionization models is the nebular continuum. The nebular continuum contributes a relatively small amount of flux to the UV spectrum, compared to the stellar component. We explicitly compute the nebular continuum for BPASS models listed in Table 5.1, however, it changes smoothly with age, so we interpolate the nebular contribution for the remaining BPASS models.

5.2.4 Composite Spectra and Fitting

To compute a composite spectrum, we first interpolated each of the individual galaxy spectra onto a common wavelength grid. We chose the sampling of this

common wavelength grid to be 0.8Å, which corresponds to the rest-frame sampling of our spectra at the median redshift of our sample. Then, at each wavelength, we median combined all spectra that had coverage at that wavelength. We defined the error spectrum as the standard deviation of all contributing spectra at each wavelength. Our fitting analysis utilized continuum normalized spectra for comparison with the models. Because of this approach, we did not need to consider effects that smoothly affect the continuum (e.g., reddening). In addition, using continuum normalized spectra simplifies the process of stacking the spectra in a cohesive manner. We first extracted regions of the spectra that are not contaminated by absorption lines, in the windows defined by Rix et al. (2004). We then fit a cubic spline to the median flux values within each window to define the continuum level.

To fit the BPASS stellar population synthesis models to our individual galaxy and composite spectra, we masked out regions of the observed spectra that include components not present in the models. Then, we continuum normalized both the observed and BPASS model spectra. We then interpolated the BPASS models onto the wavelength grid of the galaxy spectra. Following this step, we calculated the χ^2 for each BPASS model in the grid, and determined which age and stellar metallicity produced the minimum χ^2 value. We determined the uncertainties in these parameters by perturbing the observed spectrum and calculating which age and stellar metallicity best-fit the observed spectrum. In the case of an individual galaxy spectrum, this perturbation is simply adding in noise to each wavelength element pulled from a normal distribution with a standard deviation defined by the magnitude of the error spectrum at that wavelength element. For a simulated composite spectrum, we selected a new sample of galaxies from the initial composite spectrum sample with replacement. Then, each galaxy was perturbed using the method described above before being combined. After repeating this process 1000 times, we defined the best-fit value and upper and lower 1σ uncertainties as the median, $16th$ and $84th$ percentile of the distribution.

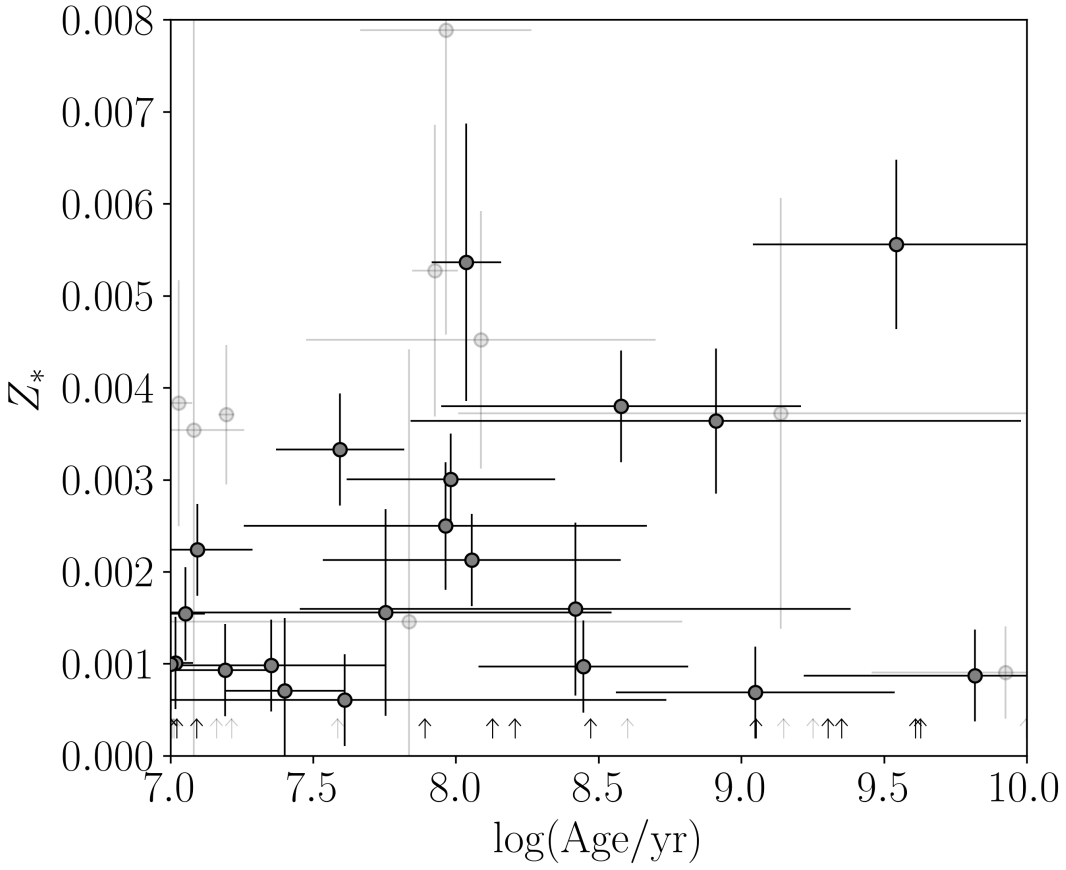

Figure 5.6: Best-fit age and stellar metallicity for all galaxies in the LRIS-BPT sample. For completeness, galaxies with SNR/pixel$\leq$ 4 are displayed as faint grey symbols. The sample comprises galaxies with ages in the range $7.0 \leq \log(\mathrm{Age/yr}) \leq 9.6$, with the majority of galaxies having stellar metallicities of $0.0005 \leq Z_* \leq 0.004$.

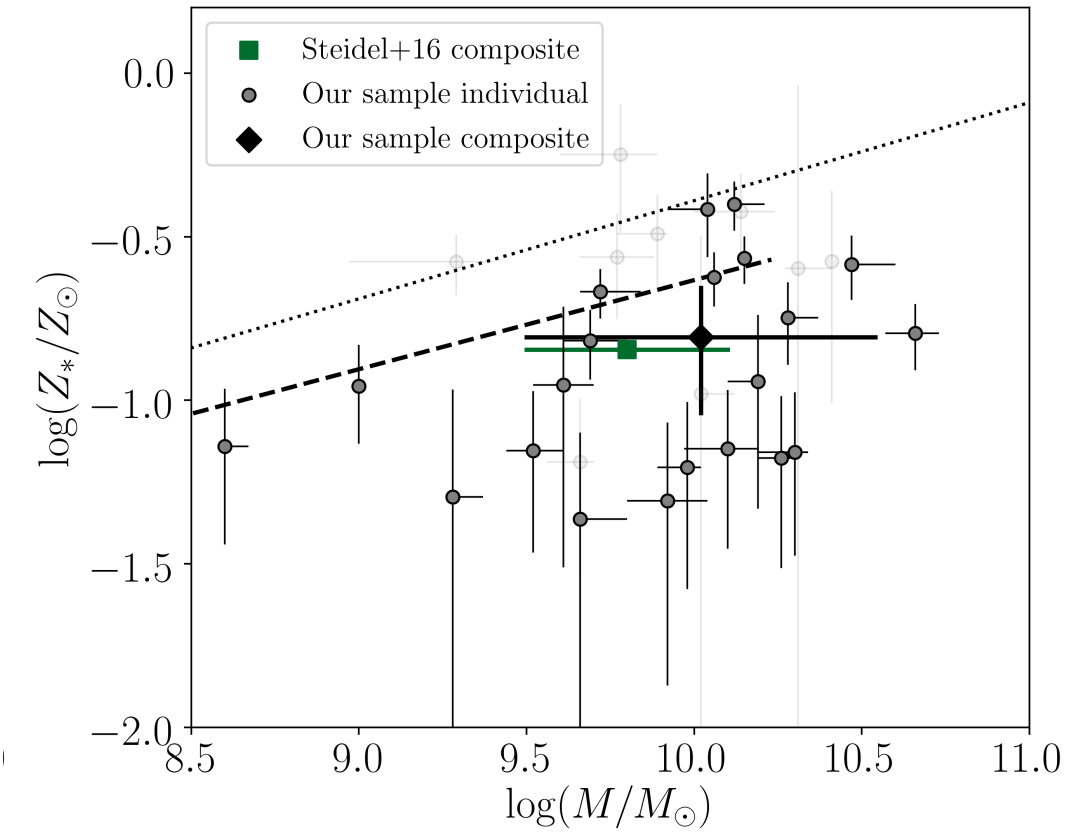

Figure 5.7: Stellar metallicity measured from BPASS fitting plotted against SED-based stellar mass. In the individual measurements we see a correlation between these two parameters such that the galaxies with the highest measured stellar metallicity are at the massive end of our sample. Measurements displayed in light grey are galaxies with rest-UV spectra with ≤ 4 SNR/pixel. The median M_* and Z_* measured for a composite spectrum composed of galaxies in the LRIS-BPT sample is depicted by the black diamond. The Z_*-M_* relation from Cullen et al. (2019) measured for galaxies at $2.5 \leq z \leq 5.0$ is displayed as the dashed line. The green square depicts the stellar metallicity and age found for the KBSS-LM1 composite from Steidel et al. (2016). For reference, the dotted line shows the best-fit $12 + \log(\mathrm{O/H})$-M_* relation for the full $z \sim 2.3$ MOSDEF sample calculated using the O3N2 line ratio from Sanders et al. (2018).

5.2.5 Testing models with additional SFHs

We expanded the model grid used in Topping et al. (2019), which only considered stellar population models that assume a constant SFH. We repeated fitting the model grids to the *high* and *low* stacks using our updated models that assume different SFHs. For each SFH, we found the results are consistent with those of Topping et al. (2019). In particular, we fit models that assume a 'delayed-τ' SFH, with $\log(\tau/\mathrm{yr}) = 7, 8, 9,$ and 10. Figure 5.2 shows the best-fit age and stellar metallicity of the *high* and *low* stacks for each model grid. For each SFH, we find that the *high* stack has lower stellar metallicity, and a younger stellar age compared to the *low* stack. While this trend between the properties of the *high* and *low* stacks persists for each SFH we considered, the exact values of the stellar metallicity and age differ between the assumed models. In particular, the delayed-τ model with $\log(\tau/\mathrm{yr}) = 7$ has the youngest best-fit age and stellar metallicity, and both the age and stellar metallicity increase when assuming an increasing τ.

For each different assumed SFH, we recovered the same qualitative trend found in Topping et al. (2019), according to which the *high* stack had a younger age and lower stellar metallicity relative to the *low* stack. We also investigated if the stellar population models yielded any constraint on the form of the SFH for each stack. We tested this question by measuring the minimum χ^2 value for the best-fit model of each SFH. For the *high* and *low* stacks, none of the SFHs were preferred, suggesting that a given UV stellar spectrum is not unique to a particular SFH. Figure 5.3 compares the best-fit spectrum of a constant SFH model to models with a delayed-τ SFH. The best-fit models for the *high* stack are nearly identical, and at some wavelengths are different at the few percent level. The *low* stack models vary more, in particular for the $\log(\tau/\mathrm{yr}) = 7$ model, which has some signatures of a young population not seen in the other best-fit models. In particular, this model has slightly enhanced CIV$\lambda\lambda$1548, 1550 and HeIIλ1640 compared to the other models. However, the majority of the models are

in agreement, with differences of only up to $\sim 10\%$ in a few wavelength elements.

5.2.6 The low SNR boundary to avoid biased results

While the composite spectra achieve good fits due to their high SNR, the SNR of individual galaxy spectra can be much lower. We tested how the SNR of a a spectrum affects the best-fit stellar properties by manually introducing noise to one of our composite spectra, refitting the models, and checking if biases arise as the spectrum drops in quality. Figure 5.4 displays how the best-fit stellar metallicity and age change as a function of the amount of added noise. For this composite, the best-fit stellar metallicity retains an unbiased estimate of the value obtained in the high-SNR limit down to a SNR/pixel$\sim$ 4. However even above this limit, the stellar metallicity uncertainty increases with decreasing SNR. The best-fit age remains consistent throughout the range of SNR/pixel, yet as the SNR decreases, the uncertainty grows to ≥ 2 dex, leaving the age unconstrained.

This test showcases how biases in the best-fit stellar parameters may occur in lower SNR spectra. In order to quantify this effect, we repeated the process used on the composite spectrum, except on BPASS models for which the 'true' parameters are known. We added noise selected from a normal distribution to the BPASS models at each wavelength element. We repeated this for a combination of ages and stellar metallicities to determine if these biases exist throughout the range of parameters. Figure 5.5 shows how the best-fit age and stellar metallicity changes as error in introduced into the model spectra. At all stellar metallicities, a low SNR/pixel introduces a positive bias, which at the lowest SNR, can be up to $\sim 150\%$. The best-fit stellar age also changes at low SNR/pixel, however in contrast to the bias of the stellar metallicity, the trend of the bias depends on the 'true' age. At low SNR/pixel, models with an old input age are biased younger, and models with a young age are biased toward older values. While these biases exist at low SNR/pixel for both age and stellar metallicity, parameters can be

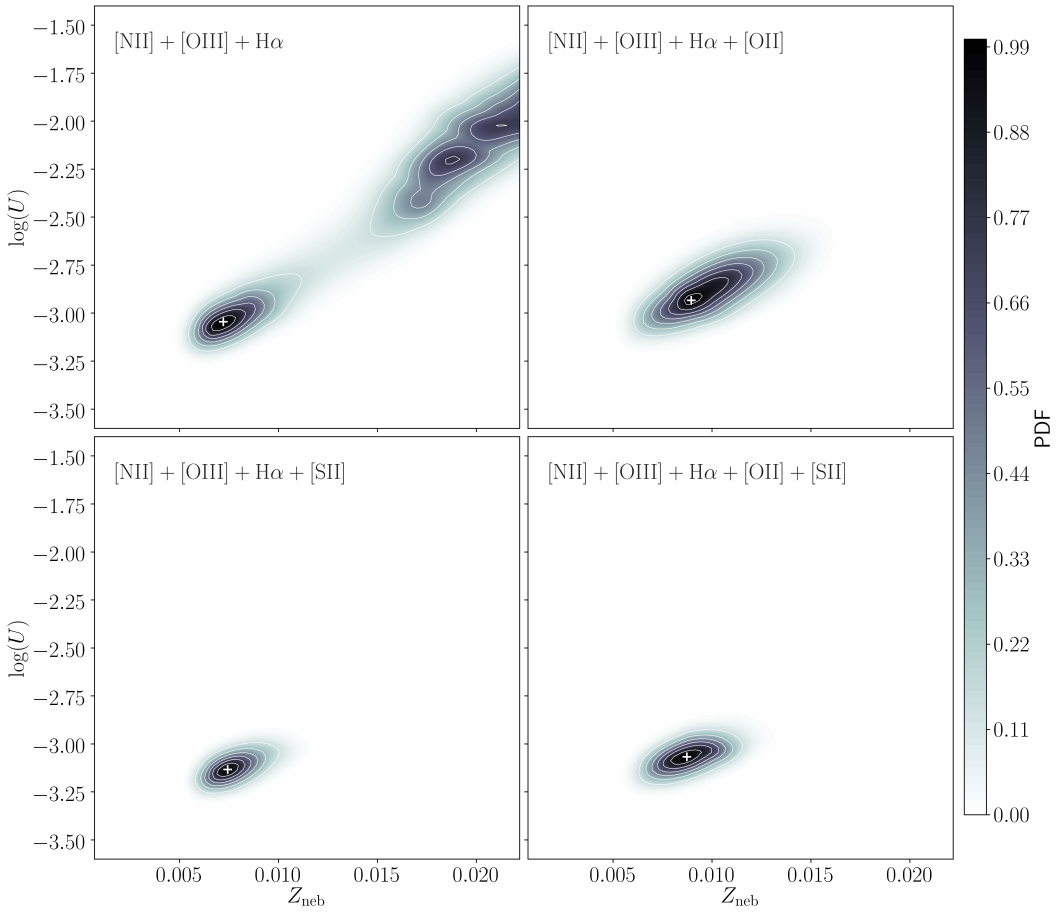

Figure 5.8: Probability density functions for inferring the ionization parameter ($\log(U)$) and nebular metallicity ($Z_{\rm neb}$) when different sets of rest-optical emission lines are used. The text in the top left of each panel displays which lines correspond to each PDF. All of the emission line fluxes are scaled to the observed Hβ flux. These panels demonstrate that the ionization parameter and nebular metallicity are better constrained when lines beyond [NII]λ6584, [OIII]λ5007, and Hα are included in the fitting procedure.

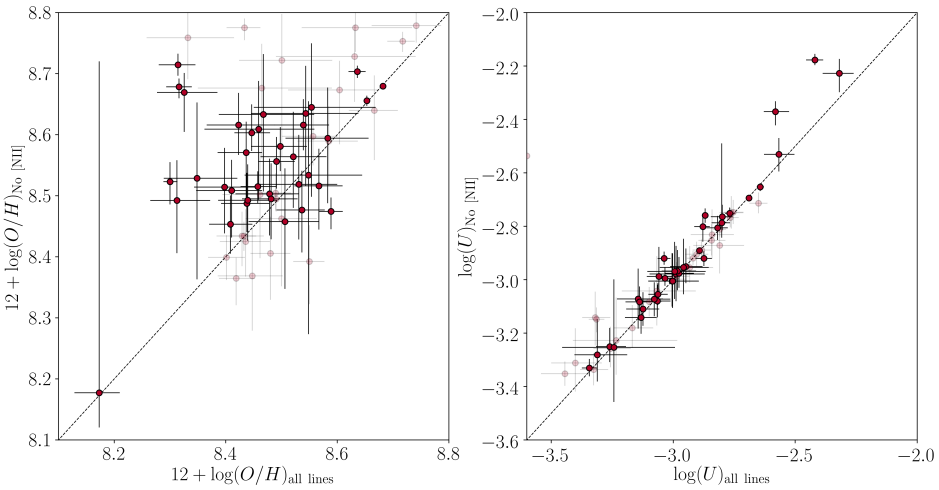

Figure 5.9: Comparison between inferred nebular parameters when [NII]λ6584 is excluded from the fitting procedure. The dotted line displays the one-to-one relation in each panel. Left: Best-fit $12 + \log(\mathrm{O/H})$ inferred without [NII]λ6584 vs. $12 + \log(\mathrm{O/H})$ inferred with it included. The majority of galaxies are consistent within their uncertainties using both methods. Right: Same as the left panel except for the $\log(U)$. The values inferred with and without [NII]λ6584 agree remarkably well for nearly all galaxies.

accurately determined for spectra with SNR/pixel $\geq\sim 4$.

5.3 Results

5.3.1 Stellar Metallicities and Masses

Based on our tests on the BPASS models, we can achieve accurate age and stellar metallicity measurements for individual spectra with SNR/pixel $\geq\sim 4$. Figure 5.6 displays these values for all galaxies in the LRIS-BPT sample, highlighting those with high enough SNR. The stellar metallicity for the individual galaxies ranges between $0.001 \leq Z_* \leq 0.006$, consistent with the best-fit metallicities found for our composite spectra. In some cases, it was not possible to constrain the stellar metallicity to lie within our grid. In such cases, stellar metallicity is therefore displayed as a lower limit. Importantly, based on the stellar metallicity and age we are able to constrain the shape of the ionizing spectrum for each of these individual galaxies.

Using the best-fit stellar metallicities from the BPASS stellar population fitting, we see how the stellar metallicity changes as a function of mass. Figure 5.7 shows Z_* vs. M_* for all of the individual galaxies in our sample. There appears to be a positive correlation between these two quantities, such that galaxies with the highest stellar metallicities are the most massive. This distribution has a Spearman correlation coefficient $r_s = 0.37$ and probability of being drawn from an uncorrelated distribution of 0.09. In addition, we note that at all masses we find stellar metallicities that are sub-solar. For reference, Figure 5.7 displays the $12 + \log(\mathrm{O/H})$-M_* for the $z \sim 2.3$ MOSDEF sample from Sanders et al. (2018). It is worth noting that the oxygen abundances displayed here from Sanders et al. (2018) were calculated using the O3N2 indicator that is calibrated using local samples, which may introduce systematics when applied to high-redshift samples. The Z_*-M_* relation found for a sample of $2.5 \leq z \leq 5.0$ galaxies from Cullen et al. (2019) passes through our sample, however the majority of data points in our sample fall below that relation. Furthermore, the results of our individual galaxies are consistent with the stellar metallicity and median stellar mass for a sample of 30 $z \sim 2.4$ star-forming galaxies comprising the KBSS-LM1 sample from (Steidel et al. 2016) ($\log(M_*/M_\odot = 9.8 \pm 0.3$, $Z_* = 0.002$. We find results consistent to those of the Steidel et al. (2016) composite when fitting BPASS models to a stack of galaxies in our LRIS-BPT sample. This sample constitutes a median mass of $\log(M_*/M_\odot) = 10.02 \pm 0.52$ and stellar metallicity of $Z_* = 0.0022 \pm 0.0009$.

5.3.2 Ionized Gas Properties

The ionizing spectrum emitted from the stellar population drives the production of the emergent rest-optical emission line ratios. Therefore, constraining the ionizing spectrum within galaxies is crucial in order use photoionization models to extract nebular properties from the observed nebular emission lines. In particular, using Cloudy we vary the nebular metallicity and ionization parameter, and then

catalog the emergent line fluxes. We then compare the resulting catalog of nebular emission line fluxes to those observed form an individual or composite spectrum. The inferred nebular metallicity and ionization parameter is set by which model best reproduces the observed emission lines. To understand the uncertainty in these quantities, we perturb the observed emission line fluxes by their corresponding uncertainties, and recompute the best-fit nebular parameters. Topping et al. (2019) used an approach that compared the locations of the models and observed galaxies on the [NII]λ6584/Hα vs. [OIII]λ5007/Hβ BPT diagram. In this analysis, we use a slightly different approach that simultaneously fits the [NII]λ6584, Hα, and [OIII]λ5007 fluxes, scaled to Hβ. Furthermore, we include additional strong lines, [OII]λ3727 and [SII]$\lambda\lambda$6717, 6731, in order to better constrain the nebular parameters. We investigate the effect that including these additional has on the inferred nebular parameters. Figure 5.8 shows an example of how the constraint on nebular metallicity and ionization parameter changes for different sets of nebular emission lines. In this example, the inferred nebular properties are consistent when considering different sets of lines, however the ionization parameter and nebular metallicity are better constrained when additional emission lines are included.

One assumption made in the photoionization modelling is the form of the N/O vs. O/H relation. The median nitrogen abundance of HII regions in the local universe has been measured to vary by $\sim$ 0.5dex for $8.2 \leq 12 + \log(O/H) \leq 8.6$, with scatter of $\sim$ 0.2dex at fixed O/H (Pilyugin et al. 2012). This assumption strongly affects the output [NII]λ6584 flux in our photoionization models. These Nitrogen abundance variations can result in a disparity of the [NII]λ6584/Hβ ratio $\sim$ 0.5 dex, resulting in a biased inference of $Z_{\rm neb}$ and $\log(U)$. Figure 5.9 shows the effect of removing [NII]λ6584 from our fitting procedure, eliminating the uncertainty surrounding the N/O relation. Without [NII]λ6584, the inferred ionization parameters are well matched to those inferred when using [NII]λ6584, with nearly all galaxies falling on the one-to-one relation. In addition, the majority

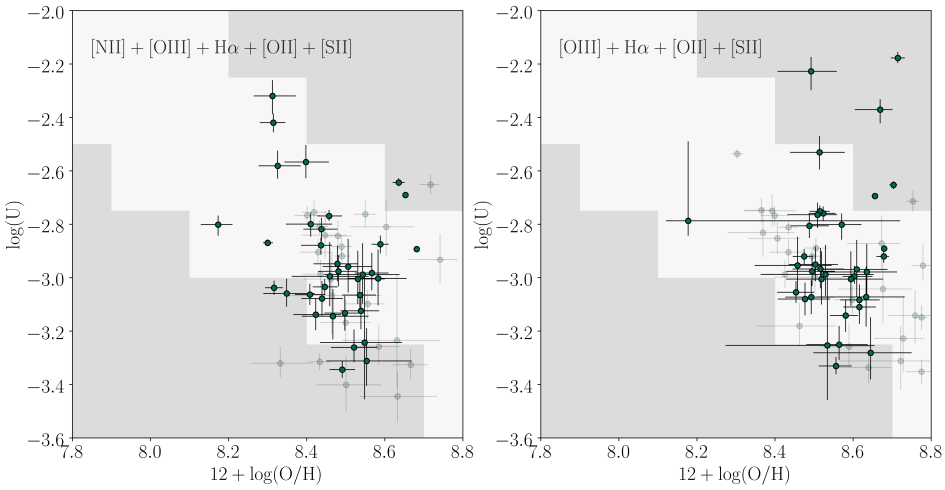

Figure 5.10: Inferred $\log(U)$ vs. $12 + \log(\mathrm{O/H})$ for each galaxy in the LRIS-BPT sample. Galaxies with low SNR/pixel (≤ 4) are displayed as the faint points. The majority of galaxies lie within the area that hosts local HII regions defined by Pérez-Montero (2014).

of the galaxies scatter around the one-to-one relation for $12 + \log(O/H)$, however some outliers are present at the lowest metallicity end of our sample. Based on this result, we conclude that the N/O vs. O/H relation we assumed in our photoionization modelling is appropriate for the galaxies in our sample, and will not significantly bias our inferred nebular parameters.

Figure 5.10 shows $\log(U)$ against $12 + \log(O/H)$ for individual galaxies in the LRIS-BPT sample. The data show a trend decreasing ionization parameter with increasing nebular metallicity. A majority of the galaxies in our sample fall within the region populated by local HII regions (Pérez-Montero 2014). Our sample has median values of our sample are $12+\log(\mathrm{O/H}) = 8.48 \pm 0.11$ and $\log(U) = -2.98 \pm 0.25$. Furthermore, this result remains largely the same when considering $\log(U)$ and $12 + \log(O/H)$ inferred without [NII]λ6584. However, the different methods used to measure the oxygen abundances between our sample, and those from Pérez-Montero (2014), which used the 'direct' method could introduce systematics. In particular, (Esteban et al. 2014) demonstrated that metallicities measured using the direct method are ~ 0.24 lower than those which use nebular recombination

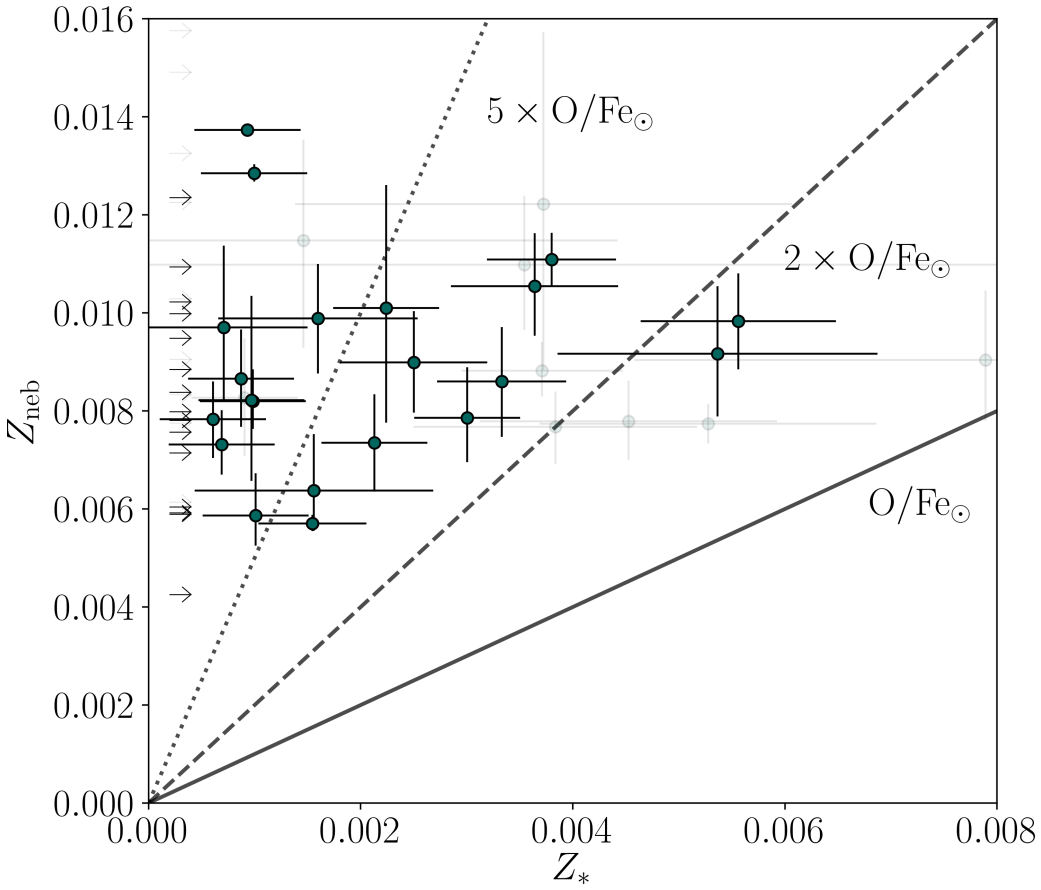

Figure 5.11: Nebular metallicity inferred from photoionization modelling plotted against stellar metallicity measured for each galaxy in our sample. Lines of constant α-enhancement (i.e., O/Fe) are displayed as solid, dashed, and dotted lines, respectively, for O/Fe$_\odot$, 2×O/Fe$_\odot$, and 5×O/Fe$_\odot$. All galaxies in our sample show evidence for α-enhancement. Additionally, some galaxies are in the regime above 5O/Fe$_\odot$, which has been suggested as the theoretical limit based on supernova yield models. The galaxies for which the stellar metallicity could not be determined are displayed as lower limits.

lines on average.

5.3.3 Combined Stellar and Nebular Properties

To connect the stellar and nebular properties of the individual galaxies in our sample, we combine the nebular O/H abundance inferred from photoionization modelling with the Fe/H measured from the BPASS model fitting to look at the α-enhancement of individual galaxies. Figure 5.11 compares the nebular metallicity and the stellar metallicity for each galaxy in our sample. Noticeably, all of the galaxies in our sample show evidence for α-enhancement. These values range from $\sim 1.75\mathrm{O/Fe}_{\odot}$ to $\geq 5\mathrm{O/Fe}_{\odot}$. A number of objects fall above the expected theoretical limit (Nomoto et al. 2006; Kobayashi et al. 2006). However, this limit depends on the details of the stellar population and expected Type II SNe yields. For example, the theoretical O/Fe limit increases when calculated assuming a top-heavy IMF. Therefore, different assumptions of the IMF or supernova yields could remove the tension between the theoretical limit and some of our observed galaxies.

5.4 Discussion

Since the first evidence that suggested that high-redshift galaxies are offset on the BPT diagram, several hypotheses have been proposed to explain the underlying cause. Among the proposed sources for this offset between local and high-redshift galaxies are harder ionizing spectra at fixed nebular metallicity, higher electron densities, contributions from AGNs and shocks at high redshift, and variations in gas-phase abundance patterns. Recently, two prevailing theories suggest that the offset is primarily driven by higher ioniziation parameters at fixed gas-phase metallicity (Kewley et al. 2015; Kashino et al. 2017; Cullen et al. 2018; Bian et al. 2018), or that high-redshift galaxies exhibit a harder intrinsic ionizing spectrum at fixed nebular metallicity driven by α-enhancement at high redshift (Steidel et al.

2016; Sanders et al. 2019).

To answer this question of the origin of the BPT offset, Sanders et al. (2019) used the 'direct' method to estimate oxygen abundances for a sample of 18 high-redshift galaxies at low nebular metallicities, and found they lie along $\log(U)$ vs. $12 + \log(O/H)$ relation of local HII regions (Pérez-Montero 2014). This result suggests that the high ionization parameter measured in their sample is due to their low nebular metallicity, and that their sample has consistent ionization parameter with local HII regions that share the same O/H. Furthermore, Shapley et al. (2019) demonstrated that high-redshift galaxies are also offset toward higher [SII]$\lambda\lambda 6717, 6731/\mathrm{H}\alpha$ and [OIII]$\lambda 5007/\mathrm{H}\beta$ when using the appropriate comparison to local galaxies with low contribution from diffuse ionized gas (DIG) within the ISM. Using photoionization models from Sanders et al. (2016a), Shapley et al. (2019) concluded that both the offset on the [NII] and [SII] BPT diagrams is best explained by a harder ionizing spectrum at fixed nebular metallicity.

Finally, the results described in this paper suggest that $z \sim 2$ galaxies do not have an elevated ionization parameter compared to local HII regions that share the same $12 + \log(O/H)$. Our analysis illustrates the importance of an independently constrained ionizing spectrum. Without such a constraint, the degeneracy between ionization parameter and the intrinsic ionizing spectrum can bias inferences of the ionization parameter. It is important to note that the method used to infer oxygen abundances of our sample is different than the method of our local HII region comparison sample (Pérez-Montero 2014), which could introduce systematics. However, the offset between these two methods results in ~ 0.24 dex lower oxygen abundance when using the direct method relative to the nebular recombination lines. Therefore, after correcting for this offset, the galaxies in our sample remain at or below the nebular metallicity of local HII regions. Based on our stellar and nebular results, we find that the the offset on the BPT diagram is primarily due to a harder ionizing spectrum resulting from super-solar O/Fe values relative to local galaxies.

5.5 Summary & Conclusions

We used combined rest-UV and rest-optical spectra for a sample of 62 galaxies to investigate the physical conditions within galaxies at $z \sim 2.3$. We expanded upon the results of Topping et al. (2019) which constructed composite spectra based on location in the [NII]λ6584 BPT diagram, and found that galaxies offset from the local sequence typically had younger ages, lower stellar metallicities, higher ionization parameters, and were more α-enhanced. We expanded the fitting analysis to include additional SFHs and rest-optical emission line fluxes. In addition, we quantitatively determined the SNR limit above which we can determine the physical properties of individual galaxies in the rest-UV We summarize our main results and conclusions below.

(i) We constructed additional BPASS stellar population models for a variety of SFHs. We repeated the fitting analysis of the two stacked spectra defined by Topping et al. (2019) and found that for each SFH, the stack composed of galaxies offset from the local sequence on the BPT diagram had a younger age and lower stellar metallicity. Additionally, when fitting across all SFHs for a single stack, we do not find any preference for one SFH over another. Therefore, we cannot determine which SFH best characterizes the rest-UV spectra.

(ii) We tested which individual galaxy spectra are suitable to be fit using this analysis. Based on the test of perturbing model spectra with known stellar parameters with increasing amounts of noise, we find that the best-fit stellar metallicity is biased high when the spectrum reaches a SNR/pixel < 4, and the best-fit age is biased toward the middle of the grid (log(age) $\sim$8.5), with an uncertainty that fills the parameter space. The best-fit age and stellar metallicity remain consistent with the true value for spectra with SNR/pixel > 4.

(iii) Based on the SNR requirements described above, we found that 30 galaxies in our sample satisfied the criteria to be fit on an individual basis. We find that galaxies in this sample have ages that span our parameter space, and that most

galaxies have stellar metallicities in the range $\sim 0.001 < Z_* < 0.004$. In addition, we see evidence for a correlation between stellar metallicity and mass, and find that our individual stellar masses and stellar metallicities are consistent with those found for the KBSS-LM1 composite in Steidel et al. (2016).

(iv) We examined how different rest-optical emission lines affect the inferred ionization parameter and nebular metallicity. Previously, we inferred nebular parameters by comparing observed [NII]λ6584/Hα and[OIII]λ5007/Hβ with a suite of photoionization models. In this analysis, we tested how adding [SII]$\lambda\lambda$6717, 6731 and [OII]λ3727 to the fitting procedure affects the resultant parameters. In general, adding the additional lines produces results with smaller uncertainties. Additionally, because one assumption we made is the N/O vs O/H relation as an input to our models, we tested fitting the rest-optical lines that are not affected by this assumption, namely [NII]λ6584. We find that by excluding [NII]λ6584 from fitting, the nebular metallicity and ionization parameter are consistent for the majority of galaxies.

(v) With the constrained ionizing spectrum for each individual galaxy, we used photoionization models to we infer ionization parameters and nebular metallicities for each galaxy in our sample, and found that the inferred ionization parameters ($\log(U)_{\rm med} = -2.98 \pm 0.25$) are consistent with those measured in local HII regions that share the same oxygen abundance ($12 + \log({\rm O/H})_{\rm med} = 8.48 \pm 0.11$). This results suggests that the offset of high-redshift galaxies on the BPT diagram relative to local galaxies is not due to elevated ionization parameters, but instead a harder ionizing spectrum resulting from elevated O/Fe in high-redshift galaxies.

(vi) Combining the best-fit stellar metallicities from fitting BPASS model spectra to the nebular metallicities inferred from photoionization modelling we find that all of our individual galaxies are α-enhanced compared to local galaxies. The range of O/Fe values found ranges from $\sim 1.75\times{\rm O/Fe}_\odot$ to above the theoretical limit ($\sim 5\times{\rm O/Fe}_\odot$ Nomoto et al. 2006). This limit could be affected by details

of the IMF used to calculate it. Furthermore, the stellar metallicities may change as stellar modelling of the most metal-poor massive stars are better understood. In particular, if the lowest Z_* stars actually produce harder ionizing spectra compared to current models, we would infer a higher stellar metallicity for our rest-UV spectra.

Directly observing the sources of ionizing radiation exciting HII regions with the physical properties of the ISM itself is a crucial step toward a complete model of high-redshift galaxy evolution. In order to fully understand high-redshift galaxies we must explore how their properties differ from local galaxies, but also how the population of high-redshift galaxies varies itself. Ultimately, detailed modelling of large numbers of individual galaxies will be required to expand our understanding of galaxies beyond the level of a sample average. The rest-optical emission lines of galaxies can be best interpreted using photoionzation models when the ionizing spectrum can be constrained. Rest-UV spectrum is the ideal tool to gain insight into the massive star population, and therefore the ionizing spectrum, for a sample of individual galaxies at high redshift. This type of analysis is key in order to compare the internal properties of high-redshift galaxies to those of local HII regions and galaxies.

CHAPTER 6

Summary and Future Work

6.1 Summary

This dissertation presents results surrounding two important aspects of the high-redshift universe: the largest-bound structures, and internal properties of individual galaxies. The main results are listed below.

In Chapter 2 I presented evidence for substructure within a protocluster at $z \sim 3$ in the SSA22 field. While the existence of a large overdensity in the SSA22 field, I identified anisotropy present within the redshift distribution in the form of a double-peaked structure. Further investigation reveals that the sample of galaxies within each peak of the redshift distribution are also segregated on the sky. This suggests that the two structures are physically distinct.

Chapter 3 further extends the work on the protocluster in the SSA22 field. In particular, I quantified the properties of each overdensity peak in the redshift distribution and found the two peaks have overdensities of $\delta_{r,gal} = 9.5 \pm 2.0$ and $\delta_{b,gal} = 4.8 \pm 1.8$, and masses of $M_r = (2.5 \pm 0.32) \times 10^{15} h^{-1} M_\odot$ and $M_r = (0.76 \pm 0.17) \times 10^{15} h^{-1} M_\odot$ for the red and blue peak respectively. Furthermore I identified similar structures in the Small MultiDark Planck comsological N-body simulation. These structures could only be identified in configurations of multiple large ($\leq 10^{14} h^{-1} M_\odot$) overdensities in close proximity. I traced these structures to the present time and found that the multiple structures remain distinct at $z = 0$.

Shifting focus away from the largest bound structures, to the properties of individual galaxies at high redshift, Chapter 4 utilizes combined rest-UV and

rest-optical spectra of 260 $z \sim 2.3$ star-forming galaxies. I constructed two composite spectra comprising galaxies in different locations of the BPT diagram in order to investigate the cause for high-redshift galaxies' elevated [OIII]λ5007/Hβ and [NII]λ6584/Hα relative to local galaxies. I found that offset high-redshift galaxies have a younger ages and lower stellar metallicities, resulting in a harder ionizing spectrum. In addition, the offset galaxies have higher ionization parameters and are more α-enhanced compared to their counterparts that lie along the local sequence. Importantly, I found that all high-redshift galaxies are α-enhanced compared to local galaxies, including those with consistent [OIII]λ5007/Hβ and [NII]λ6584/Hα. These differences are crucial to take into account when modelling high-redshift galaxies to avoid getting biased results.

Finally, Chapter 5 first improved the stellar population models to include models that assume different star-formation histories, in addition to the continuous star-formation history models presented in Chapter 4. When considering additional star formation histories, I found that the stellar properties of the two composite spectra defined in Chapter 4 remained distinct in terms of their stellar age and stellar metallicity. While the additional SFHs make the model fitting more robust, in the case of the two composite spectra, no single SFH was preferred. Following these improvements, I tested how well the models are able to fit low SNR spectra, and in particular, the spectra of individual galaxies. I found that at low SNR (≤ 4/pixel) the fitting produced stellar metallicities that were biased high, and similarly, the ages of spectra below this SNR cutoff could not be constrained. I then fit the galaxies that fit these criteria to get best-fit stellar ages and stellar metallicities for a sample of ~ 30 galaxies and found that most galaxies have stellar metallicities in the range $\sim 0.001 < Z_* < 0.004$. Furthermore, there is a positive correlation between stellar mass and stellar metallicity for the galaxies fit in our sample. Using the ionizing spectrum which has been constrained from the rest-UV stellar fitting, I used photoionization models to extract nebular properties from the observed rest-optical emission line fluxes. I tested how including

additional rest-optical nebular emission lines to the fitting procedure affected the inferred nebular metallicity and ionization parameter. The addition of the [SII] and [OII] lines provided a tighter constraint of both nebular parameter lines. The ionization parameters of galaxies in our sample are consistent with those measured in local HII regions that share the same oxygen abundance. This results suggests that high-redshift galaxies are offset on the BPT diagram not due to elevated ionization parameters, but instead a result of a harder ionizing spectrum relative to local galaxies. Finally, combining stellar and nebular properties reveals that all individual galaxies in our sample are α-enhanced compared to local galaxies, ranging from O/Fe of $\sim 1.75\times$O/Fe$_\odot$ to above the theoretical limit of $\sim 5.0\times$O/Fe$_\odot$. Easing this tension of α-enhancement will require different assumptions about the IMF of high-redshift stellar populations, or updated supernova yield models that allow for higher O/Fe values.

6.2 Future Work

Understanding the connection between gas and stars, in and around galaxies, is vital in order to form a complete picture of galaxy formation and evolution. The importance of these effects is enhances at $z \sim 2$, when the universe hosts much more activity compared to the current epoch. At these redshifts, supermassive black hole accretion rates are highest, the cosmic star-formation rate density is at its peak, and galaxy outflows are common in star-forming galaxies. Due to this high activity, this is the epoch when many of the local galaxy population's properties are established. The rest-optical spectrum provides a wealth of information about the physical conditions within these galaxies. At these redshifts, the rest-optical emission lines are redshifted into the near-IR, restricting the redshift ranges at which these galaxies can be observed in order to accommodate atmospheric absorption windows. The recent advent of sensitive multiplex NIR spectrographs on 10m class telescopes, such as Keck/MOSFIRE, has allowed de-

tailed studies of large ($\sim$ 1000s) of galaxies. In the local universe, rest-optical studies show that star-forming galaxies follow a tight relation on the [O III]/Hβ vs. [N II]/Hα emission line, or, "BPT" diagram. One important result found from these high-redshift studies, is an offset toward higher [O III]/Hβ and [N II]/Hα relative to local galaxies. Understanding the cause of this offset is a key step in obtaining an accurate model of galaxy evolution.

While rest-optical spectroscopy informs on properties of the ISM, rest-UV spectroscopy can provide insights to the massive star populations within galaxies. Massive stars provide an impetus for nearly all aspects of a galaxy's, and its surroundings' evolution. These stars are responsible for producing strong winds, creating black holes, regulating star-formation, dispersing metals throughout the ISM, and one more thing. In addition to shaping the evolution of galaxies, massive stars provide several useful tools to aid in our understanding of galactic processes. Due to the short-lived nature of these stars, their existence allows for a relatively high frequency look into the current star-forming status of galaxies. Their FUV light is reprocessed by the ISM, allowing the determination of physical properties internal to the galaxies.

Despite the increased understanding of high-redshift galaxies, there are still many outstanding questions. These include, what processes are driving star formation, what are the connections between the ISM and CGM, and how is stellar mass assembled in galaxies? The addition of rest-UV spectroscopy is a natural extension to previous studies using only the rest-optical as it allows a direct investigation of the most massive stars. Understanding massive star populations in high redshift galaxies has several complications. First, the modelling of such stars has several unknowns: massive stars are thought to be more likely found in binaries, the interaction between stars in multiples can affect their evolution in complicated ways, and difficulties in observing them means the FUV spectrum of an O-star has never been measured. Features in the FUV stellar spectrum are often near, or overlapping ISM absorption lines, requiring these regions to be

excluded in our analysis.

Complementary to combined rest-UV and rest-optical studies of high-redshift galaxies, we require spatially resolved information provided by adaptive optics-assisted IFU spectroscopy. With Keck/OSIRIS, we can obtain maps of strong emission 'BPT' lines, Hβ, [OIII]$\lambda\lambda$4959, 5007, Hα, [NII]λ6584, and [SII]$\lambda\lambda$6717, 6731. Using these lines, I will construct spatially resolved BPT diagrams in order to detect regions with significant AGN or LINER contributions. Such contributions have been proposed as an effect responsible for the observed BPT offset in high redshift galaxies. I will more broadly investigate the spatial variations of emission line ratios in these galaxies, and see if regions that are more 'offset' in the BPT diagram trace regions of high star-formation surface density, or position within the galaxies. Such a correlation could help reveal the source of the BPT offset as either the ionizing conditions with HII regions, or more related to non-solar abundance ratios. As we gain insight into the cause of the BPT offset in high-redshift galaxies, spatial sampling as described above is the next step toward a complete understanding of high-redshift star-forming galaxies, and their differences to local galaxies.

Bibliography

Abazajian, K. N., J. K. Adelman-McCarthy, M. A. Agüeros, S. S. Allam, C. Allende Prieto, D. An, K. S. J. Anderson, S. F. Anderson, J. Annis, N. A. Bahcall, & et al. 2009, ApJS, 182, 543

Adelberger, K. L., C. C. Steidel, A. E. Shapley, & M. Pettini. 2003, ApJ, 584, 45

Asplund, M., N. Grevesse, A. J. Sauval, & P. Scott. 2009, ARA&A, 47, 481

Baldwin, J. A., M. M. Phillips, & R. Terlevich. 1981, PASP, 93, 5

Barkana, R. & A. Loeb. 2004, ApJ, 609, 474

Behroozi, P. S., R. H. Wechsler, & H.-Y. Wu. 2013a, ApJ, 762, 109

Behroozi, P. S., R. H. Wechsler, H.-Y. Wu, M. T. Busha, A. A. Klypin, & J. R. Primack. 2013b, ApJ, 763, 18

Berg, D. A., D. K. Erb, R. B. C. Henry, E. D. Skillman, & K. B. W. McQuinn. 2019, ApJ, 874, 93

Bian, F., L. J. Kewley, & M. A. Dopita. 2018, ApJ, 859, 175

Brinchmann, J., M. Pettini, & S. Charlot. 2008, MNRAS, 385, 769

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

Chabrier, G. 2003, PASP, 115, 763

Chapman, S. C., D. Scott, R. A. Windhorst, D. T. Frayer, C. Borys, G. F. Lewis, & R. J. Ivison. 2004, ApJ, 606, 85

Chiang, Y.-K., R. A. Overzier, K. Gebhardt, S. L. Finkelstein, C.-T. Chiang, G. J. Hill, G. A. Blanc, N. Drory, T. S. Chonis, G. R. Zeimann, A. Hagen, D. P. Schneider, S. Jogee, R. Ciardullo, & C. Gronwall. 2015, ApJ, 808, 37

Chisholm, J., J. R. Rigby, M. Bayliss, D. A. Berg, H. Dahle, M. Gladders, & K. Sharon. 2019, ApJ, 882, 182

Conroy, C., J. E. Gunn, & M. White. 2009, ApJ, 699, 486

Conroy, C., A. E. Shapley, J. L. Tinker, M. R. Santos, & G. Lemson. 2008, ApJ, 679, 1192

Crowther, P. A., R. K. Prinja, M. Pettini, & C. C. Steidel. 2006, MNRAS, 368, 895

Cullen, F., R. J. McLure, S. Khochfar, J. S. Dunlop, C. Dalla Vecchia, A. C. Carnall, N. Bourne, M. Castellano, A. Cimatti, M. Cirasuolo, D. Elbaz, J. P. U. Fynbo, B. Garilli, A. Koekemoer, F. Marchi, L. Pentericci, M. Talia, & G. Zamorani. 2018, MNRAS, 476, 3218

Cullen, F., R. J. McLure, J. S. Dunlop, S. Khochfar, R. Davé, R. Amorín, M. Bolzonella, A. C. Carnall, M. Castellano, A. Cimatti, M. Cirasuolo, G. Cresci, J. P. U. Fynbo, F. Fontanot, A. Gargiulo, B. Garilli, L. Guaita, N. Hathi, P. Hibon, F. Mannucci, F. Marchi, D. J. McLeod, L. Pentericci, L. Pozzetti, A. E. Shapley, M. Talia, & G. Zamorani. 2019, MNRAS, 487, 2038

Eldridge, J. J. & E. R. Stanway. 2012, MNRAS, 419, 479

Eldridge, J. J., E. R. Stanway, L. Xiao, L. A. S. McClelland , G. Taylor, M. Ng, S. M. L. Greis, & J. C. Bray. 2017, , 34, e058

Erb, D. K., A. E. Shapley, M. Pettini, C. C. Steidel, N. A. Reddy, & K. L. Adelberger. 2006, ApJ, 644, 813

Esteban, C., J. García-Rojas, L. Carigi, M. Peimbert, F. Bresolin, A. R. López-Sánchez, & A. Mesa-Delgado. 2014, MNRAS, 443, 624

Ferland, G. J., M. Chatzikos, F. Guzmán, M. L. Lykins, P. A. M. van Hoof, R. J. R. Williams, N. P. Abel, N. R. Badnell, F. P. Keenan, R. L. Porter, & P. C. Stancil. 2017, Revista Mexicana de Astronomia y Astrofisica, 53, 385

Gawiser, E., H. Francke, K. Lai, K. Schawinski, C. Gronwall, R. Ciardullo, R. Quadri, A. Orsi, L. F. Barrientos, G. A. Blanc, G. Fazio, J. J. Feldmeier, J.-s. Huang, L. Infante, P. Lira, N. Padilla, E. N. Taylor, E. Treister, C. M. Urry, P. G. van Dokkum, & S. N. Virani. 2007, ApJ, 671, 278

Geach, J. E., Y. Matsuda, I. Smail, S. C. Chapman, T. Yamada, R. J. Ivison, T. Hayashino, K. Ohta, Y. Shioya, & Y. Taniguchi. 2005, MNRAS, 363, 1398

Geach, J. E., D. M. Alexander, B. D. Lehmer, I. Smail, Y. Matsuda, S. C. Chapman, C. A. Scharf, R. J. Ivison, M. Volonteri, T. Yamada, A. W. Blain, R. G. Bower, F. E. Bauer, & A. Basu-Zych. 2009, ApJ, 700, 1

Geach, J. E., D. Narayanan, Y. Matsuda, M. Hayes, L. Mas-Ribas, M. Dijkstra, C. C. Steidel, S. C. Chapman, R. Feldmann, A. Avison, O. Agertz, Y. Ao, M. Birkinshaw, M. N. Bremer, D. L. Clements, H. Dannerbauer, D. Farrah, C. M. Harrison, M. Kubo, M. J. Michałowski, D. Scott, D. J. B. Smith, M. Spaans, J. M. Simpson, A. M. Swinbank, Y. Taniguchi, P. van der Werf, A. Verma, & T. Yamada. 2016, ApJ, 832, 37

Grogin, N. A., D. D. Kocevski, S. M. Faber, H. C. Ferguson, A. M. Koekemoer, A. G. Riess, V. Acquaviva, D. M. Alexander, O. Almaini, M. L. N. Ashby, M. Barden, E. F. Bell, F. Bournaud, T. M. Brown, K. I. Caputi, S. Casertano, P. Cassata, M. Castellano, P. Challis, R.-R. Chary, E. Cheung, M. Cirasuolo, C. J. Conselice, A. Roshan Cooray, D. J. Croton, E. Daddi, T. Dahlen, R. Davé, D. F. de Mello, A. Dekel, M. Dickinson, T. Dolch, J. L. Donley, J. S. Dunlop, A. A. Dutton, D. Elbaz, G. G. Fazio, A. V. Filippenko, S. L. Finkelstein, A. Fontana, J. P. Gardner, P. M. Garnavich, E. Gawiser, M. Giavalisco, A. Grazian, Y. Guo, N. P. Hathi, B. Häussler, P. F. Hopkins, J.-S. Huang, K.-H.

Huang, S. W. Jha, J. S. Kartaltepe, R. P. Kirshner, D. C. Koo, K. Lai, K.-S. Lee, W. Li, J. M. Lotz, R. A. Lucas, P. Madau, P. J. McCarthy, E. J. McGrath, D. H. McIntosh, R. J. McLure, B. Mobasher, L. A. Moustakas, M. Mozena, K. Nandra, J. A. Newman, S.-M. Niemi, K. G. Noeske, C. J. Papovich, L. Pentericci, A. Pope, J. R. Primack, A. Rajan, S. Ravindranath, N. A. Reddy, A. Renzini, H.-W. Rix, A. R. Robaina, S. A. Rodney, D. J. Rosario, P. Rosati, S. Salimbeni, C. Scarlata, B. Siana, L. Simard, J. Smidt, R. S. Somerville, H. Spinrad, A. N. Straughn, L.-G. Strolger, O. Telford, H. I. Teplitz, J. R. Trump, A. van der Wel, C. Villforth, R. H. Wechsler, B. J. Weiner, T. Wiklind, V. Wild, G. Wilson, S. Wuyts, H.-J. Yan, & M. S. Yun. 2011, ApJS, 197, 35

Halliday, C., E. Daddi, A. Cimatti, J. Kurk, A. Renzini, M. Mignoli, M. Bolzonella, L. Pozzetti, M. Dickinson, G. Zamorani, S. Berta, A. Franceschini, P. Cassata, G. Rodighiero, & P. Rosati. 2008, A&A, 479, 417

Harikane, Y., M. Ouchi, Y. Ono, S. Saito, P. Behroozi, S. More, K. Shimasaku, J. Toshikawa, Y.-T. Lin, M. Akiyama, J. Coupon, Y. Komiyama, A. Konno, S.-C. Lin, S. Miyazaki, A. J. Nishizawa, T. Shibuya, & J. Silverman. 2017, ArXiv e-prints, arXiv:1704.06535

Hayashino, T., Y. Matsuda, H. Tamura, R. Yamauchi, T. Yamada, M. Ajiki, S. S. Fujita, T. Murayama, T. Nagao, K. Ohta, S. Okamura, M. Ouchi, K. Shimasaku, Y. Shioya, & Y. Taniguchi. 2004, AJ, 128, 2073

Hayatsu, N. H., Y. Matsuda, H. Umehata, N. Yoshida, I. Smail, A. M. Swinbank, R. Ivison, K. Kohno, Y. Tamura, M. Kubo, D. Iono, B. Hatsukade, K. Nakanishi, R. Kawabe, T. Nagao, A. K. Inoue, T. T. Takeuchi, M. Lee, Y. Ao, S. Fujimoto, T. Izumi, Y. Yamaguchi, S. Ikarashi, & T. Yamada. 2017, PASJ, 69, 45

Hogg, D. W. 1999, ArXiv Astrophysics e-prints, astro-ph/9905116

Kaiser, N. 1984, ApJ, 284, L9

—. 1987, MNRAS, 227, 1

Kashino, D., J. D. Silverman, D. Sanders, J. S. Kartaltepe, E. Daddi, A. Renzini, F. Valentino, G. Rodighiero, S. Juneau, L. J. Kewley, H. J. Zahid, N. Arimoto, T. Nagao, J. Chu, N. Sugiyama, F. Civano, O. Ilbert, M. Kajisawa, O. Le Fèvre, C. Maier, D. Masters, T. Miyaji, M. Onodera, A. Puglisi, & Y. Taniguchi. 2017, ApJ, 835, 88

Kauffmann, G., J. M. Colberg, A. Diaferio, & S. D. M. White. 1999, MNRAS, 307, 529

Kauffmann, G., S. D. M. White, & B. Guiderdoni. 1993, MNRAS, 264, 201

Kauffmann, G., T. M. Heckman, C. Tremonti, J. Brinchmann, S. Charlot, S. D. M. White, S. E. Ridgway, J. Brinkmann, M. Fukugita, P. B. Hall, Ž. Ivezić, G. T. Richards, & D. P. Schneider. 2003, MNRAS, 346, 1055

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

Kewley, L. J., M. A. Dopita, R. S. Sutherland, C. A. Heisler, & J. Trevena. 2001, ApJ, 556, 121

Kewley, L. J., H. J. Zahid, M. J. Geller, M. A. Dopita, H. S. Hwang, & D. Fabricant. 2015, ApJ, 812, L20

Klypin, A., G. Yepes, S. Gottlöber, F. Prada, & S. Heß. 2016, MNRAS, 457, 4340

Kobayashi, C., H. Umeda, K. Nomoto, N. Tominaga, & T. Ohkubo. 2006, ApJ, 653, 1145

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

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

Kriek, M., C. Conroy, P. G. van Dokkum, A. E. Shapley, J. Choi, N. A. Reddy, B. Siana, F. van de Voort, A. L. Coil, & B. Mobasher. 2016, Nature, 540, 248

Kubo, M., T. Yamada, T. Ichikawa, M. Kajisawa, Y. Matsuda, & I. Tanaka. 2015, ApJ, 799, 38

Kuiper, E., N. A. Hatch, G. K. Miley, N. P. H. Nesvadba, H. J. A. Röttgering, J. D. Kurk, M. D. Lehnert, R. A. Overzier, L. Pentericci, J. Schaye, & B. P. Venemans. 2011, MNRAS, 415, 2245

Lacey, C. & S. Cole. 1993, MNRAS, 262, 627

Lee, K.-G., J. F. Hennawi, M. White, J. X. Prochaska, A. Font-Ribera, D. J. Schlegel, R. M. Rich, N. Suzuki, C. W. Stark, O. Le Fèvre, P. E. Nugent, M. Salvato, & G. Zamorani. 2016, ApJ, 817, 160

Lee, K.-S., M. Giavalisco, O. Y. Gnedin, R. S. Somerville, H. C. Ferguson, M. Dickinson, & M. Ouchi. 2006, ApJ, 642, 63

Lehmer, B. D., D. M. Alexander, J. E. Geach, I. Smail, A. Basu-Zych, F. E. Bauer, S. C. Chapman, Y. Matsuda, C. A. Scharf, M. Volonteri, & T. Yamada. 2009a, ApJ, 691, 687

—. 2009b, ApJ, 691, 687

Lehmer, B. D., D. M. Alexander, S. C. Chapman, I. Smail, F. E. Bauer, W. N. Brandt, J. E. Geach, Y. Matsuda, J. R. Mullaney, & A. M. Swinbank. 2009c, MNRAS, 400, 299

Leitherer, C., J. R. S. Leão, T. M. Heckman, D. J. Lennon, M. Pettini, & C. Robert. 2001, ApJ, 550, 724

Lemaux, B. C., O. Cucciati, L. A. M. Tasca, O. Le Fèvre, G. Zamorani, P. Cassata, B. Garilli, V. Le Brun, D. Maccagni, L. Pentericci, R. Thomas, E. Vanzella, E. Zucca, R. Amorín, S. Bardelli, P. Capak, L. P. Cassarà, M. Castellano, A. Cimatti, J. G. Cuby, S. de la Torre, A. Durkalec, A. Fontana, M. Giavalisco, A. Grazian, N. P. Hathi, O. Ilbert, C. Moreau, S. Paltani, B. Ribeiro, M. Salvato, D. Schaerer, M. Scodeggio, V. Sommariva, M. Talia, Y. Taniguchi, L. Tresse, D. Vergani, P. W. Wang, S. Charlot, T. Contini, S. Fotopoulou, R. R. Gal, D. D. Kocevski, C. López-Sanjuan, L. M. Lubin, Y. Mellier, T. Sadibekova, & N. Scoville. 2014, A&A, 572, A41

Liu, X., A. E. Shapley, A. L. Coil, J. Brinchmann, & C.-P. Ma. 2008, ApJ, 678, 758

Madau, P. & M. Dickinson. 2014, ARA&A, 52, 415

Mandelbaum, R., A. Slosar, T. Baldauf, U. Seljak, C. M. Hirata, R. Nakajima, R. Reyes, & R. E. Smith. 2013, MNRAS, 432, 1544

Masters, D., A. Faisst, & P. Capak. 2016, ApJ, 828, 18

Masters, D., P. McCarthy, B. Siana, M. Malkan, B. Mobasher, H. Atek, A. Henry, C. L. Martin, M. Rafelski, N. P. Hathi, C. Scarlata, N. R. Ross, A. J. Bunker, G. Blanc, A. r. G. Bedregal, A. Domínguez, J. Colbert, H. Teplitz, & A. Dressler. 2014, ApJ, 785, 153

Matsuda, Y., T. Yamada, T. Hayashino, H. Tamura, R. Yamauchi, M. Ajiki, S. S. Fujita, T. Murayama, T. Nagao, K. Ohta, S. Okamura, M. Ouchi, K. Shimasaku, Y. Shioya, & Y. Taniguchi. 2004, AJ, 128, 569

Matsuda, Y., T. Yamada, T. Hayashino, H. Tamura, R. Yamauchi, T. Murayama,

T. Nagao, K. Ohta, S. Okamura, M. Ouchi, K. Shimasaku, Y. Shioya, & Y. Taniguchi. 2005, ApJ, 634, L125

Matsuda, Y., T. Yamada, T. Hayashino, R. Yamauchi, Y. Nakamura, N. Morimoto, M. Ouchi, Y. Ono, K. Kousai, E. Nakamura, M. Horie, T. Fujii, M. Umemura, & M. Mori. 2011, MNRAS, 410, L13

Matteucci, F., E. Spitoni, D. Romano, & A. Rojas Arriagada. 2016, in Frontier Research in Astrophysics II (FRAPWS2016), 27

McLean, I. S., C. C. Steidel, H. W. Epps, N. Konidaris, K. Y. Matthews, S. Adkins, T. Aliado, G. Brims, J. M. Canfield, J. L. Cromer, J. Fucik, K. Kulas, G. Mace, K. Magnone, H. Rodriguez, G. Rudie, R. Trainor, E. Wang, B. Weber, & J. Weiss. 2012, in Proc. SPIE, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 84460J

Miyazaki, S., Y. Komiyama, M. Sekiguchi, S. Okamura, M. Doi, H. Furusawa, M. Hamabe, K. Imi, M. Kimura, F. Nakata, N. Okada, M. Ouchi, K. Shimasaku, M. Yagi, & N. Yasuda. 2002, PASJ, 54, 833

Momcheva, I. G., G. B. Brammer, P. G. van Dokkum, R. E. Skelton, K. E. Whitaker, E. J. Nelson, M. Fumagalli, M. V. Maseda, J. Leja, M. Franx, H.-W. Rix, R. Bezanson, E. Da Cunha, C. Dickey, N. M. Förster Schreiber, G. Illingworth, M. Kriek, I. Labbé, J. Ulf Lange, B. F. Lundgren, D. Magee, D. Marchesini, P. Oesch, C. Pacifici, S. G. Patel, S. Price, T. Tal, D. A. Wake, A. van der Wel, & S. Wuyts. 2016, ApJS, 225, 27

Muldrew, S. I., N. A. Hatch, & E. A. Cooke. 2015, MNRAS, 452, 2528

Naoz, S. & R. Barkana. 2005, MNRAS, 362, 1047

Nestor, D. B., A. E. Shapley, K. A. Kornei, C. C. Steidel, & B. Siana. 2013, ApJ, 765, 47

Nestor, D. B., A. E. Shapley, C. C. Steidel, & B. Siana. 2011, ApJ, 736, 18

Nomoto, K., N. Tominaga, H. Umeda, C. Kobayashi, & K. Maeda. 2006, Nuclear Physics A, 777, 424

Oke, J. B., J. G. Cohen, M. Carr, J. Cromer, A. Dingizian, F. H. Harris, S. Labrecque, R. Lucinio, W. Schaal, H. Epps, & J. Miller. 1995, PASP, 107, 375

Ouchi, M., K. Shimasaku, S. Okamura, H. Furusawa, N. Kashikawa, K. Ota, M. Doi, M. Hamabe, M. Kimura, Y. Komiyama, M. Miyazaki, S. Miyazaki, F. Nakata, M. Sekiguchi, M. Yagi, & N. Yasuda. 2004, ApJ, 611, 660

Padmanabhan, T. 1993, Structure Formation in the Universe, 499

Peebles, P. J. E. 1980, The large-scale structure of the universe

Pérez-Montero, E. 2014, MNRAS, 441, 2663

Pilyugin, L. S., J. M. Vílchez, L. Mattsson, & T. X. Thuan. 2012, MNRAS, 421, 1624

Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, & et al. 2014, A&A, 571, A16

Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, & et al. 2015, ArXiv e-prints, arXiv:1502.01589

Quadri, R., P. van Dokkum, E. Gawiser, M. Franx, D. Marchesini, P. Lira, G. Rudnick, D. Herrera, J. Maza, M. Kriek, I. Labbé, & H. Francke. 2007, ApJ, 654, 138

Rix, S. A., M. Pettini, C. Leitherer, F. Bresolin, R.-P. Kudritzki, & C. C. Steidel. 2004, ApJ, 615, 98

Rodríguez-Puebla, A., P. Behroozi, J. Primack, A. Klypin, C. Lee, & D. Hellinger. 2016, MNRAS, 462, 893

Sanders, R. L., A. E. Shapley, M. Kriek, N. A. Reddy, W. R. Freeman, A. L. Coil, B. Siana, B. Mobasher, I. Shivaei, S. H. Price, & L. de Groot. 2016a, ApJ, 816, 23

—. 2016b, ApJ, 816, 23

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

Sanders, R. L., A. E. Shapley, N. A. Reddy, M. Kriek, B. Siana, A. L. Coil, B. Mobasher, I. Shivaei, W. R. Freeman, M. Azadi, S. H. Price, G. Leung, T. Fetherolf, L. de Groot, T. Zick, & F. M. F. G. Barro. 2019, MNRAS, 2653

Shapley, A. E., A. L. Coil, C.-P. Ma, & K. Bundy. 2005, ApJ, 635, 1006

Shapley, A. E., C. C. Steidel, M. Pettini, & K. L. Adelberger. 2003, ApJ, 588, 65

Shapley, A. E., C. C. Steidel, M. Pettini, K. L. Adelberger, & D. K. Erb. 2006, ApJ, 651, 688

Shapley, A. E., N. A. Reddy, M. Kriek, W. R. Freeman, R. L. Sanders, B. Siana, A. L. Coil, B. Mobasher, I. Shivaei, S. H. Price, & L. de Groot. 2015, ApJ, 801, 88

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

Sheth, R. K. & G. Tormen. 1999, MNRAS, 308, 119

Shivaei, I., M. Kriek, N. A. Reddy, A. E. Shapley, G. Barro, C. Conroy, A. L. Coil, W. R. Freeman, B. Mobasher, B. Siana, R. Sanders, S. H. Price, M. Azadi, I. Pasha, & H. Inami. 2016, ApJ, 820, L23

Sommariva, V., F. Mannucci, G. Cresci, R. Maiolino, A. Marconi, T. Nagao, A. Baroni, & A. Grazian. 2012, A&A, 539, A136

Springel, V., S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas, H. Couchman, A. Evrard, J. Colberg, & F. Pearce. 2005, Nature, 435, 629

Stanway, E. R. & J. J. Eldridge. 2018, MNRAS, 479, 75

Steidel, C. C., K. L. Adelberger, M. Dickinson, M. Giavalisco, M. Pettini, & M. Kellogg. 1998, ApJ, 492, 428

Steidel, C. C., K. L. Adelberger, A. E. Shapley, D. K. Erb, N. A. Reddy, & M. Pettini. 2005, ApJ, 626, 44

Steidel, C. C., K. L. Adelberger, A. E. Shapley, M. Pettini, M. Dickinson, & M. Giavalisco. 2000, ApJ, 532, 170

—. 2003, ApJ, 592, 728

Steidel, C. C., M. Bogosavljević, A. E. Shapley, J. A. Kollmeier, N. A. Reddy, D. K. Erb, & M. Pettini. 2011, ApJ, 736, 160

Steidel, C. C., A. L. Strom, M. Pettini, G. C. Rudie, N. A. Reddy, & R. F. Trainor. 2016, ApJ, 826, 159

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

Strom, A. L., C. C. Steidel, G. C. Rudie, R. F. Trainor, & M. Pettini. 2018, ApJ, 868, 117

Strom, A. L., C. C. Steidel, G. C. Rudie, R. F. Trainor, M. Pettini, & N. A. Reddy. 2017, ApJ, 836, 164

Tinker, J., A. V. Kravtsov, A. Klypin, K. Abazajian, M. Warren, G. Yepes, S. Gottlöber, & D. E. Holz. 2008, ApJ, 688, 709

Topping, M. W., A. E. Shapley, N. A. Reddy, R. L. Sanders, A. L. Coil, M. Kriek, B. Mobasher, & B. Siana. 2019, arXiv e-prints, arXiv:1912.10243

Topping, M. W., A. E. Shapley, & C. C. Steidel. 2016, ApJ, 824, L11

Topping, M. W., A. E. Shapley, C. C. Steidel, S. Naoz, & J. R. Primack. 2018, ApJ, 852, 134

Topping, M. W. & J. M. Shull. 2015, ApJ, 800, 97

Toshikawa, J., N. Kashikawa, R. Overzier, T. Shibuya, S. Ishikawa, K. Ota, K. Shimasaku, M. Tanaka, M. Hayashi, Y. Niino, & M. Onoue. 2014, ApJ, 792, 15

Trainor, R. F. & C. C. Steidel. 2012, ApJ, 752, 39

Trainor, R. F., C. C. Steidel, A. L. Strom, & G. C. Rudie. 2015, ApJ, 809, 89

Umehata, H., Y. Tamura, K. Kohno, R. J. Ivison, D. M. Alexander, J. E. Geach, B. Hatsukade, D. H. Hughes, S. Ikarashi, Y. Kato, T. Izumi, R. Kawabe, M. Kubo, M. Lee, B. Lehmer, R. Makiya, Y. Matsuda, K. Nakanishi, T. Saito, I. Smail, T. Yamada, Y. Yamaguchi, & M. Yun. 2015, ApJ, 815, L8

Veilleux, S. & D. E. Osterbrock. 1987, ApJS, 63, 295

Venemans, B. P., H. J. A. Röttgering, G. K. Miley, W. J. M. van Breugel, C. de Breuck, J. D. Kurk, L. Pentericci, S. A. Stanford, R. A. Overzier, S. Croft, & H. Ford. 2007, A&A, 461, 823

Wang, T., D. Elbaz, E. Daddi, A. Finoguenov, D. Liu, C. Schreiber, S. Martín, V. Strazzullo, F. Valentino, R. van der Burg, A. Zanella, L. Ciesla, R. Gobat,

A. Le Brun, M. Pannella, M. Sargent, X. Shu, Q. Tan, N. Cappelluti, & Y. Li. 2016, ApJ, 828, 56

Weinberg, D. H., B. H. Andrews, & J. Freudenburg. 2017, ApJ, 837, 183

White, S. D. M. & C. S. Frenk. 1991, ApJ, 379, 52

White, S. D. M. & M. J. Rees. 1978, MNRAS, 183, 341

Yamada, T., Y. Nakamura, Y. Matsuda, T. Hayashino, R. Yamauchi, N. Morimoto, K. Kousai, & M. Umemura. 2012, AJ, 143, 79

Zahid, H. J., R.-P. Kudritzki, C. Conroy, B. Andrews, & I. T. Ho. 2017, ApJ, 847, 18