The MOSDEF–LRIS Survey: the interplay between massive stars and ionized gas in high-redshift star-forming galaxies

Michael W. Topping,^{1★} Alice E. Shapley,¹ Naveen A. Reddy,² Ryan L. Sanders[®],³ Alison L. Coil,⁴ Mariska Kriek,⁵ Bahram Mobasher² and Brian Siana²

¹Department of Physics and Astronomy, University of California, Los Angeles, 430 Portola Plaza, Los Angeles, CA 90095, USA

²Department of Physics and Astronomy, University of California, Riverside, 900 University Avenue, Riverside, CA 92521, USA

³Department of Physics, University of California, Davis, 1 Shields Avenue, Davis, CA 95616, USA

⁴ Center for Astrophysics and Space Sciences, Department of Physics, University of California, San Diego, 9500 Gilman Drive., La Jolla, CA 92093, USA ⁵ Astronomy Department, University of California at Berkeley, Berkeley, CA 94720, USA

Accepted 2020 May 18. Received 2020 April 24; in original form 2019 December 21

ABSTRACT

We present a joint analysis of rest-ultraviolet (UV) and rest-optical spectra obtained using Keck/LRIS and Keck/MOSFIRE for a sample of 62 star-forming galaxies at $z \sim 2.3$. We divide our sample into two bins based on their location in the $[O III]5007/H\beta$ versus [N II] 6584/H α BPT diagram, and perform the first differential study of the rest-UV properties of massive ionizing stars as a function of rest-optical emission-line ratios. Fitting Binary Population and Spectral Synthesis (BPASS) stellar population synthesis models, including nebular continuum emission, to our rest-UV composite spectra, we find that high-redshift galaxies offset towards higher $[O III]\lambda 5007/H\beta$ and $[N II]\lambda 6584/H\alpha$ have younger ages $(\log(\text{Age/yr}) = 7.20^{+0.57}_{-0.20})$ and lower stellar metallicities $(Z_* = 0.0010^{+0.0011}_{-0.0003})$ resulting in a harder ionizing spectrum, compared to the galaxies in our sample that lie on the local BPT star-forming sequence (log(Age/yr) = $8.57^{+0.88}_{-0.84}$, $Z_* = 0.0019^{+0.0006}_{-0.0006}$). Additionally, we find that the offset galaxies have an ionization parameter of $log(U) = -3.04^{+0.06}_{-0.11}$ and nebular metallicity of $(12 + \log(\text{O/H}) = 8.40^{+0.06}_{-0.07})$, and the non-offset galaxies have an ionization parameter of $\log(U) = -3.11^{+0.08}_{-0.08}$ and nebular metallicity of $12 + \log(\text{O/H}) = 8.30^{+0.05}_{-0.06}$. The stellar and nebular metallicities derived for our sample imply that the galaxies offset from the local BPT relation are more α -enhanced $(7.28^{+2.52}_{-2.82} \text{ O/Fe}_{\odot})$ compared to those consistent with the local sequence $(3.04^{+0.95}_{-0.54} \text{ O/Fe}_{\odot})$. However, even galaxies that are entirely consistent with the local nebular excitation sequence appear to be α -enhanced – in contrast with typical local systems. Such differences must be considered when estimating gas-phase oxygen abundances at high redshift based on strong emission-line ratios. Specifically, 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.

Key words: galaxies: evolution – galaxies: high-redshift – galaxies: ISM.

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 $[N II]\lambda 6584/H\alpha$ and decreasing $[O III]\lambda 5007/H\beta$ 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 H II-region excitation of its constituent galaxies (Masters, Faisst & Capak 2016). Early observations with Keck/NIRSPEC suggested possible differences in the emission-line properties of high-redshift galaxies in the [O III] λ 5007/H β versus [N II] λ 6584/H α 'BPT' diagram (Baldwin, Phillips & Terlevich 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 [O III] λ 5007/H β and/or [N II] λ 6584/H α on average relative to local galaxies.

^{*} E-mail: mtopping@astro.ucla.edu

^{© 2020} The Author(s) Published by Oxford University Press on behalf of the Royal Astronomical Society

There are many possible causes for this observed difference between local and $z \sim 2$ galaxies, including higher ionization parameters, harder ionizing spectra at fixed nebular metallicity, higher densities, variations in gas-phase abundance patterns, and enhanced contributions from active galactic nuclei (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. 2016). 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 H II 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 restoptical 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 on to 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 time-scales 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 C IV $\lambda\lambda$ 1548, 1550 (Crowther et al. 2006; Leitherer et al. 2001) and He IIA1640 (Brinchmann, Pettini & Charlot 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 Fe III-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 $z \ge 2.5 M_* - Z_*$ 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 highredshift galaxies (Steidel et al. 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 $(Z_{neb}/Z_{\odot} \sim 0.5)$. 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_{neb} is tracing O/H. By analysing 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 2 describes our observations, data reduction, and methods. Section 3 presents the results of our analysis, and Section 4 provides a summary and discussion 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).

2 METHODS AND OBSERVATIONS

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 (IR) spectra of ~1500 H-band selected galaxies observed over 48.5 nights during 2012-2016 and targeted to lie within three distinct redshift intervals $(1.37 \le z \le 1.70, 2.09 \le z \le 2.61, and 2.95)$ $\leq z \leq 3.80$) 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 $1.40 \le z \le 1.90$, $1.90 \le z \le 2.65$, and $2.95 \le z \le 3.80$. 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 multiwavelength data sets. MOSDEF spectra were used to measure fluxes and redshifts of all rest-optical emission lines detected within



Figure 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 $2.09 \le z \le 2.61$ ($z_{\text{med}} = 2.28$) with all four BPT lines (H β , [O III] λ 5007, H α , [N II] λ 6584) detected at $\ge 3\sigma$ in MOSFIRE spectra from the MOSDEF survey, totalling 62 galaxies. Right: SFR calculated from the dust-corrected Balmer lines versus M_* for all objects with LRIS redshifts at $2.0 \le z \le 2.7$ (large circles). Blue and red points indicate galaxies included, respectively, in the *high* and *low* composite spectra described in Section 3. Galaxies from the MOSDEF survey within $2.0 \le z \le 2.7$ that have both H α and H β detected with $\ge 3\sigma$ are depicted by small grey points. The median error bar 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).

the *Y*, *J*, *H*, and *K* bands, the strongest of which are: $[O II]\lambda 3727$, $H\beta$, $[O III]\lambda\lambda 4959$, 5007, $H\alpha$, $[N II]\lambda 6584$, and $[S II]\lambda\lambda 6717$, 6731.

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 (H β , [O III], H α , [N II]) were detected with $\geq 3\sigma$. Next highest priority was given to objects where $H\alpha$, $H\beta$, and [O III] were detected, and an upper limit on [N II] 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 $1.90 \le z \le 2.65$ than those at $1.40 \le z \le 1.90$ or $2.95 \le 1.90$ $z \leq 3.80$), 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 catalogue (Momcheva et al. 2016) and lying within the MOSDEF target photometric redshift and apparent magnitude range. These targets comprise ~ 260 observed galaxies with redshifts $1.4 \le z \le 3.8^{1}$ which is large and diverse enough to create bins across multiple galaxy properties [e.g. location in the BPT diagram, stellar mass, and star formation rate (SFR)]. For this analysis, we do not include the small fraction of objects identified as AGN based on their Xray and rest-IR properties. Fig. 1 displays the redshift histogram and distributions of H α -based SFR and M_* derived from Spectral

¹Of the 260 observed galaxies, 214 galaxies had a redshift from the MOSDEF survey, with 32, 162, and 20 in the redshift intervals $1.40 \le z \le 1.90, 1.90 \le z \le 2.65$, and $2.95 \le z \le 3.80$ 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 $1.40 \le z \le 1.90, 1.90 \le z \le 2.65$, and $2.95 \le z \le 3.80$ respectively.

Energy Distribution (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 3.

A detailed description of the LRIS (Low-Resolution Imaging Spectrograph) data acquisition and data reduction procedures will be presented elsewhere, however a brief summary is provided here. The data were obtained using the LRIS (Oke et al. 1995) during five observing runs totalling ten nights between 2017 January and 2018 June. We observed nine multi-object slit masks with 1.2 arcsec 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 3400 Å on the blue side, and the 600 lines mm⁻¹ grating blazed at 5000 Å 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 resolution of $R \sim 800$, and the red side yielded a resolution of $R \sim 1300$. The median exposure time was 7.5 h, but ranged from 6 to 11 h 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 arcsec with typical values of 0.8 arcsec. Details of the observations are listed in Table 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 flatfielded 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 2D spectrum. We then performed a second-pass background subtraction on the stacked

Field	Mask name	RA	Dec.	$t_{\exp}^{\text{Blue}}(s)$	$t_{\exp}^{\text{Red}}(s)$	N objects
COSMOS	co_11	10:00:22.142	+02:14:25.623	25 200	24 080	33
COSMOS	co_12	10:00:22.886	+02:24:45.096	24 300	22716	31
COSMOS	co_15	10:00:29.608	+02:14:33.037	21 492	20736	27
COSMOS	co_16	10:00:39.965	+02:17:28.409	25 0 20	24 264	26
GOODS-S	gs_11	03:32:23.178	-27:43:08.900	39312	38 664	30
GOODS-N	gn_l1	12:37:13.178	+62:15:09.647	27 000	22 968	30
GOODS-N	gn_13	12:36:54.841	+62:15:32.920	32 400	31 500	27
AEGIS	ae_11	14:19:14.858	+52:48:02.128	28188	26964	31
AEGIS	ae_13	14:19:35.219	+52:54:52.570	34 056	33 1 2 0	25

Table 1. Summary of LRIS observations.

2D spectrum of each object while excluding the traces of objects in the slits in order to avoid oversubtraction 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 2D spectra to measure the traces of objects in each slit. The abundance of sky lines in the redside images prevented us from achieving an accurate second-pass background subtraction on the stacked 2D 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 fourth-order polynomial to the red and blue arc lamp spectra, resulting in typical residuals of ~ 0.035 and ~ 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 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. Fig. 2 shows some examples of reduced high-S/N (signal-to-noise ratio) continuum normalized spectra. Several strong absorption features are commonly visible, including Si IIA1260, Ο Ιλ1302 + Si ΙΙλ1304, C ΙΙλ1334, Si ΙVλλ1393, 1402, C ΙVλλ1548, 1550, and Al IIλ1670.

2.3 Redshift measurements

We measured a redshift for each object based on the Ly α emission line, as well as low-ionization interstellar (LIS) absorption lines, namely, Si II λ 1260, O I λ 1302 + Si II λ 1304, C II λ 1334, Si II λ 1526, Fe II λ 1608, and Al II λ 1670, where available. Due to the presence of galaxy-scale outflows, the Ly α emission and interstellar absorption lines are commonly Doppler shifted away from the systemic redshift, z_{sys} . Therefore, we defined two different redshift measurements, $z_{Lv\alpha}$, and z_{LIS} . We used the systemic redshift measured from nebular emission lines as part of the MOSDEF survey, when available, as an initial guess for $z_{Ly\alpha}$, and z_{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 Ly α 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_{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 Si IIA1260, C IIA1334, and Si IIA1526 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_{LIS} (162 objects). If an object did not have a redshift measurement for any of these three lines, we defined z_{LIS} by using the Al II λ 1670 line (1 object). If this line was also not available we used the blended $O_{I\lambda}1302 + Si_{II\lambda}1304$ 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_{svs} as:

$$\begin{cases} z_{\text{sys}} = z_{\text{LIS}} + 32.0 \left(\frac{1+z_{\text{LIS}}}{c}\right) & z_{\text{LIS}} \text{ only} \\ z_{\text{sys}} = z_{\text{LIS}} + 89.0 \left(\frac{1+z_{\text{LIS}}}{c}\right) & z_{\text{LIS}} \text{ and } z_{\text{Ly}\alpha} \\ z_{\text{sys}} = z_{\text{Ly}\alpha} - 153.0 \left(\frac{1+z_{\text{Ly}\alpha}}{c}\right) & z_{\text{Ly}\alpha} \text{ only}; \ z_{\text{Ly}\alpha} \le 2.7 \\ z_{\text{sys}} = z_{\text{Ly}\alpha} - 317.0 \left(\frac{1+z_{\text{Ly}\alpha}}{c}\right) & z_{\text{Ly}\alpha} \text{ only}; \ z_{\text{Ly}\alpha} \ge 2.7. \end{cases}$$

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_{sys} and $z_{Ly\alpha}$ or z_{LIS} , and for the remaining 24 objects we were not able to measure a redshift.



Figure 2. Ten continuum-normalized individual rest-UV spectra from our sample ordered by redshift. These spectra have the highest continuum S/N from our sample with median S/N/pixel = 7 ($4.5 \le S/N$ /pixel ≤ 12), measured over the wavelength range $1425 \text{ Å} \le \lambda \le 1500 \text{ Å}$. 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 cut-off 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.



Figure 3. Stacked spectrum composed of all galaxies with a redshift measured from LRIS and in the interval $2.09 \le z \le 2.55$, totalling 62 galaxies. The magnitude of the 1σ error spectrum is depicted by the thin black line. Spectral features are identified using the same labelling scheme as in Fig. 2, with the addition of the Lyman series marked by dashed-dotted purple lines.

2.4 The LRIS-BPT sample

The full MOSDEF–LRIS sample consists of 260 galaxies across three distinct redshift intervals $(1.40 \le z \le 1.90, 1.90 \le z \le 2.65,$ and $2.95 \le z \le 3.80$). 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 (H β , [O III] λ 5007, H α , [N II] λ 6584) at the $\ge 3\sigma$ 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. Fig. 3 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. Fig. 1 displays the SFR calculated from dust-corrected Balmer lines versus M_* for both galaxies in the LRIS-BPT and the full MOSDEF samples. The LRIS-BPT sample is characterized by a median SFR of $\log (SFR/(M_{\odot}/yr)) = 1.53 \pm 0.44$, and a median stellar mass of $\log (M_*/M_{\odot}) = 10.02 \pm 0.52$. 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(SFR/(M_{\odot}/yr)) = 1.36 \pm 0.50 and median mass of $\log (M_*/M_{\odot}) = 9.93 \pm 0.60$. 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.

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, resulting in longer main-sequence lifetimes. We considered BPASS stellar population models with all available stellar metallicities $(10^{-5} \le Z_* \le 0.04)$, which primarily trace Fe/H (Steidel et al. 2016; Strom et al. 2018), and ages between 10^7 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 cut-off of 100 M_{\odot} . By default, BPASS provides models of an instantaneous burst of star 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 $-2.0 \le \log (Z_{neb}/Z_{\odot}) \le 0.4$ in 0.2 dex steps, and ionization parameters of $-4.0 \le \log (U) \le -1.0$ in 0.4 dex steps. All models were run assuming an electron density typical of galaxies at this redshift of $n_e = 250 \text{ cm}^{-3}$ (Sanders et al. 2016; Strom et al. 2017). We set the abundance of nitrogen in the models using the log(N/O) versus log(O/H) relation from Pilyugin et al. (2012):

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

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_{neb}/Z_{\odot}) = -0.2; Sanders et al. 2016). Adjusting these parameters does not affect the nebular continuum significantly enough to alter the results of our model fitting. Fig. 4 shows the differences in the C IV $\lambda\lambda$ 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 C IV $\lambda\lambda$ 1548, 1550 profile



Figure 4. Zoomed-in regions around the $C IV\lambda\lambda 1548$, 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 $C IV\lambda\lambda 1548$, 1550 profile are highlighted in blue and orange respectively, showing how the overall $C IV\lambda\lambda 1548$, 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.

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.

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 Ly α 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 - 2000 Å. To determine the best-fitting age and metallicity, we first interpolated the model on to the wavelength

scale of our observed spectrum, and then calculated the χ^2 for each model in our grid:

$$\chi^2 = \sum_i \frac{(f_{\text{spec},i} - f_{\text{model},i})^2}{\sigma_i^2},\tag{1}$$

where $f_{\text{spec}, i}$, $f_{\text{model}, i}$, 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 ~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-fitting parameters. To calculate the uncertainties in these parameters, we perturbed the spectrum and repeated this process 1000 times to produce a distribution of bestfitting values. We then defined the boundaries of the 1 σ 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 ~0.6 Å pixel⁻¹ 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



Figure 5. [O III]/H β versus [N II]/H α 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]/H β and [N II]/H α (blue shaded region; *high* sample). The median [O III]/H β versus [N II]/H α values of galaxies in each bin are depicted as diamond-shaped symbols, with values of ([N II]/H α , [O III]/H β) = (-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.

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.

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 [N II]/H α values, but characterized by different rest-optical line ratios relative to the z = 0 BPT excitation sequence. Fig. 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 $[NII]/H\alpha$ and $[OIII]/H\beta$ 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. Fig. 6 shows the stacked, continuum-normalized spectra of galaxies in the high and low stacks in blue (top) and red (bottom), respectively. For completeness, Fig. 7 shows the positions of galaxies in our *high* and *low* stacks on the $[O III]\lambda 5007/H\beta$ versus $[S II]\lambda\lambda 6717$, 6731/H α and O_{32} versus R_{23} emission-line diagrams, where O_{32} = $[O III]\lambda\lambda 4959, 5007/[O II]\lambda\lambda 3726, 3729 \text{ and } R_{23} = ([O III]\lambda\lambda 4959,$ $5007 + [O II]\lambda\lambda 3726, 3729)/H\beta$. On both of these additional BPT



Figure 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 labelled as in Fig. 2. Fig. 9 shows a zoomed-in comparison of the CIV $\lambda\lambda$ 1548, 1550 and HeII λ 1640 lines. The 1 σ composite error spectrum is indicated by the shaded region surrounding each spectrum.



Figure 7. Left: [S II] 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 [O III] λ 5007/H β , however there is overlap between the two samples. A median error bar is shown in the bottom left. Right: log (O_{32}) versus log (R_{23}) diagram. Symbols are the same as in the left-hand panel. Galaxies in the *high* stack are offset on average toward higher log (R_{23}), though there is overlap with the *low* stack.

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. 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. Fig. 8 shows the best-fitting 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 S/N spectra in each bin. We find a stellar metallicity of $Z_* = 0.0010^{+0.001}_{-0.0003}$ 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-fitting stellar age for the *low* stack of $log(Age/yr) = 8.57^{+0.88}_{-0.84}$. 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 broad-band 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, Gunn & White (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 $\propto t \exp(-t/\tau)$, where *t* is the time since the onset of star formation, and τ is the characteristic star formation time-scale. 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 (M/M_{\odot}) = 10.05 ± 0.43 and 10.12 ± 0.32

for galaxies in the high and low stacks, respectively. Also, we find median SFR of log (SFR_{SED}/(M_{\odot}/yr)) = 1.33 ± 0.42 and 1.38 ± 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 ± 0.4) is younger than the median SED-based age in the *low* stack (log (Age/yr) = 8.6 ± 0.3). This result from the broad-band 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-S/N 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 and stellar metallicity) consistent with stacked spectrum results, two of which were upper limits on the age. The remaining galaxy had a bestfitting 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-fitting parameters of the stacked spectra have larger uncertainties compared to the individual spectra, which suggests that our bootstrap resampling method is capturing galaxyto-galaxy variations of age and Z_* in our sample.



Figure 8. Best-fitting Z_* and log (Age/yr) for our two stacked spectra, and the five highest continuum S/N individual spectra from each stack, ranging from $4.5 \le S/N/pixel \le 12$, measured in the wavelength range, $1425 \le \lambda \le 1500$. The remaining galaxies do not have high enough S/N (S/N/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 coloured 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.

In addition to our global rest-UV fitting procedure, which covers the full FUV spectrum at 1270 Å $\leq \lambda \leq 2000$ Å, evidence for a difference in age between our two stacks is visible in the wind lines produced by massive stars: C IV $\lambda\lambda$ 1548, 1550 and He II λ 1640. Fig. 9 shows these features for both of our stacks. The *high* stack has stronger CIV $\lambda\lambda$ 1548, 1550 emission (1552–1555Å), as well as stronger stellar wind absorption (1536–1545Å) 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 (Fig. 4). In addition, the *high* stack shows a significant He II λ 1640 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-fitting stellar population parameters, we can examine the ionizing spectrum predicted by the BPASS models. Fig. 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 (~10 Myr), and remains constant through most of our parameter space. As a result, the ionizing spectrum is similar between the two best-fitting 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 900 Å, and integrated over the range 200 Å $\leq \lambda \leq 912$ Å, is ~7 per cent 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 [O III]/H β versus [N II]/H α BPT diagram for the best-fitting stellar spectrum of each of our stacks (Fig. 11). We linearly interpolate



Figure 9. Zoomed-in regions around the C IV $\lambda\lambda$ 1548, 1550 (left) and He II λ 1640 (right) features for our *high* (blue) and *low* (red) stacked spectra. The *high* stacked spectrum has stronger He II λ 1640 and C IV $\lambda\lambda$ 1548, 1550 emission. Both features are signatures of massive stars, and are more prominent in younger populations. The 1 σ composite error spectrum is depicted by the shaded region surrounding each spectrum. Fig. 4 shows the age dependence of the C IV $\lambda\lambda$ 1548, 1550 line in the BPASS models.



Figure 10. Top: ionizing spectra of the best-fitting stellar population models for the *high* (blue) and *low* (red) stacks. The ionizing spectra are normalized at 900 Å. The normalized ionizing flux integrated in the range 200 Å $\leq \lambda \leq 912$ Å is ~7 per cent 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-fitting stellar population models for the *high* and *low* stacks (i.e. $\Delta F_{\nu}/F_{\nu}^{\text{low}} \equiv (F_{\nu}^{\text{high}} - F_{\nu}^{\text{low}})/F_{\nu}^{\text{low}})$.

the grid of [N II]/H α and [O III]/H β values produced by the CLOUDY models to determine which Z_{neb} and log(U) best match the median observed line ratios of each stack. To estimate the uncertainty, we perturb the median observed [N II]/H α and [O III]/H β of the stacks by their uncertainties, and repeat the process 1000 times to create a distribution of values. Fig. 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 $12 + \log(\text{O/H}) = 8.40^{+0.06}_{-0.07}$ 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 $12 + \log(O/H) =$ $8.30^{+0.05}_{-0.06}$. 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 restoptical 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 $n_e^{\text{high}} = 350 \pm 161 \text{ cm}^{-3}$ and $n_e^{\text{low}} = 334 \pm 282 \,\text{cm}^{-3}$ respectively. Both medians are consistent with the value assumed in the CLOUDY models ($n_e = 250 \text{ cm}^{-3}$), and a Kolmogorov-Smirnov test determines a 47 per cent probability that both samples are drawn from the same parent distribution. Table 2 summarizes the best-fitting physical parameters we find for the *high* and *low* stacks.

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 $[O III]/H\beta$ and $[N_{II}]/H\alpha$ compared to local galaxies. Several contributing factors have been proposed as the source of this offset, including varying 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 modelling 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 restoptical analyses have focused on average properties of the highredshift 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 time-scales ~ 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_{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_{neb} (nebular O/H), and using rest-UV spectral fitting to measure Z_* (stellar Fe/H), we find that both stacks have supersolar 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_o, 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 $[O III]\lambda 5007/H\beta$ versus [N II] λ 6584/H α 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 $12 + \log (O/H)$. 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 cut-off 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 an 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 [N II]/H α while keeping all other parameters



Figure 11. Predicted $[O III]/H\beta$ and $[N II]/H\alpha$ emission-line ratios corresponding to the CLOUDY+BPASS model grid of $12 + \log (O/H)$ and $\log (U)$ for a given ionizing spectrum. The centre of each point is colour coded by $\log (U)$, increasing from yellow to orange, while the border of each point is colour coded by $12 + \log (O/H)$, increasing from light to dark green. The scale for each parameter is indicated by the colour bars to the right of the panels. The median value and uncertainty in the observed $[O III]/H\beta$ and $[N II]/H\alpha$ 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-fitting stellar population for the high stack ($\log (Age/yr) = 7.2, Z_* = 0.001$). Right: line ratios assuming an ionizing spectrum inferred from the best-fitting stellar population model for the low stack ($\log (Age/yr) = 8.6, Z_{i*} = 0.002$).



Figure 12. Distributions of best-fitting 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]/H β and [N II]/H α 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-fitting ages. We find the *high* stack is younger (log(Age/yr) = $7.20^{+0.57}_{-0.20}$) compared to the *low* stack (log(Age/yr) = $8.57^{+0.88}_{-0.84}$). Top right: distribution of best-fitting stellar metallicities. We find that the *high* stack has an overall lower stellar metallicity ($Z_* = 0.0010^{+0.0010}_{-0.0003}$) compared to the *low* stack ($Z_* = 0.0019^{+0.0010}_{-0.0006}$). Bottom left: distributions of best-fitting nebular metallicities. We find best-fitting ionization parameters. We obtain values of $\log(U) = -3.04^{+0.06}_{-0.11}$ for the *high* stack, and $\log(U) = -3.11^{+0.08}_{-0.00}$ for the *high* stack. Bottom right: distributions of best-fitting nebular metallicities. We find best-fitting values of $12 + \log(O/H) = 8.40^{+0.06}_{-0.07}$ for the *high* stack, and $12 + \log(O/H) = 8.30^{+0.06}_{-0.06}$ for the *low* stack.

Table 2. Best-fitting physical parameters for the high and low stacks.

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

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 time-scale 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 and ~ 0.5 dex, respectively, at fixed O/H would be required. Given the age of both stacks, and the time-scale 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 versus O/H relations. For O/Fe to match between the high and low stacks, we would need to assume an N/O higher by ~ 0.6 dex for the high stack. For consistency at the 1σ level, the assumed N/O would need to be ~ 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 cut-off 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, [N II]/[O II], 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 [N II]/[O II] = -0.79 ± 0.25 and -0.99 ± 0.31 , respectively. Based on the calibration of N/O as a function of [N II]/[O II] 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 [N II]/[O II], suggesting that our spectra are well described by the models.

We check the predicted O_{32} distribution for the best-fitting 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-fitting 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 time-scale differences between core collapse and Type Ia supernovae predict that significant evolution of O/Fe will only occur on time-scales of \sim 1 Gyr, assuming smooth star formation histories (Weinberg, Andrews & Freudenburg 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.

5 SUMMARY AND CONCLUSIONS

We have obtained rest-UV spectra for a sample of 259 galaxies at 1.4 < z < 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 (H β , [O III] λ 5007, H α , [N II] λ 6584) 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 emissionline 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 [O III] λ 5007/H β and [N II] λ 6584/H α . 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 ($Z_* = 0.0010^{+0.0010}_{-0.0003}$) compared to the non-offset galaxies ($Z_* = 0.0019^{+0.0006}_{-0.0006}$). These results are displayed in Fig. 8.

(ii) We investigated how the ionizing spectrum of the best-fitting 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 (Fig. 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, Zneb, and ionization parameter, U (Fig. 11). Accordingly, our rest-UV spectral analysis enabled us to fix one of the input free parameters for photo-ionization modelling - that is 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_{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 nonoffset (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 $(12 + \log(\text{O/H}) = 8.40^{+0.06}_{-0.07})$ compared to the non-offset galaxy stack $(12 + \log(\text{O/H}) = 8.30^{+0.05}_{-0.06})$. The stellar and nebular metallicities we derived for our high and low stacks imply that the galaxies that are offset from the local BPT relation are more α -enhanced (7.28^{+2.52}_{-2.82} O/Fe_{\odot}) compared to those on the local sequence (3.04^{+0.95}_{-0.54} O/Fe_{\odot}).

Understanding the observed differences between local and highredshift 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 Zneb 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-S/N 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.

ACKNOWLEDGEMENTS

We thank the anonymous referee for their helpful comments. We acknowledge support from National Science Foundation (NSF) Astronomy and Astrophysics Research Grants (AAG) grants AST1312780, 1312547, 1312764, and 1313171, grant AR13907 from the Space Telescope Science Institute, and grant NNX16AF54G from the NASA Astrophysics Data Analysis Program (ADAP) program. We also acknowledge a NASA contract supporting the 'WFIRST Extragalactic Potential Observations (EXPO) Science Investigation Team' (15-WFIRST15-0004), administered by Goddard Space Flight Center (GSFC). This work made use of v2.2.1 of the BPASS models as described in Eldridge & Stanway et al. (2017); Stanway & Eldridge (2018). We 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. Based on data obtained at the W.M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and the National Aeronautics and Space Administration, and was made possible by the generous financial support of the W.M. Keck Foundation.

REFERENCES

- Abazajian K. N. et al., 2009, ApJS, 182, 543
- Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
- Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
- Berg D. A., Erb D. K., Henry R. B. C., Skillman E. D., McQuinn K. B. W., 2019, ApJ, 874, 93
- Brinchmann J., Pettini M., Charlot S., 2008, MNRAS, 385, 769
- Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
- Chabrier G., 2003, PASP, 115, 763
- Chisholm J., Rigby J. R., Bayliss M., Berg D. A., Dahle H., Gladders M., Sharon K., 2019, ApJ, 882, 182
- Conroy C., Gunn J. E., White M., 2009, ApJ, 699, 486
- Crowther P. A., Prinja R. K., Pettini M., Steidel C. C., 2006, MNRAS, 368, 895
- Cullen F. et al., 2019, MNRAS, 487, 2038
- Eldridge J. J., Stanway E. R., 2012, MNRAS, 419, 479
- Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, PASA, 34, e058
- Erb D. K., Shapley A. E., Pettini M., Steidel C. C., Reddy N. A., Adelberger K. L., 2006, ApJ, 644, 813
- Ferland G. J. et al., 2017, Rev. Mex. Astron. Astrofis., 53, 385
- Grogin N. A. et al., 2011, ApJS, 197, 35
- Halliday C. et al., 2008, A&A, 479, 417
- Kauffmann G. et al., 2003, MNRAS, 346, 1055
- Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, ApJ, 556, 121
- Kewley L. J., Dopita M. A., Leitherer C., Davé R., Yuan T., Allen M., Groves B., Sutherland R., 2013, ApJ, 774, 100
- Kobayashi C., Umeda H., Nomoto K., Tominaga N., Ohkubo T., 2006, ApJ, 653, 1145
- Kriek M., van Dokkum P. G., Labbé I., Franx M., Illingworth G. D., Marchesini D., Quadri R. F., 2009, ApJ, 700, 221
- Kriek M. et al., 2015, ApJS, 218, 15
- Kriek M. et al., 2016, Nature, 540, 248
- Leitherer C., Leão J. R. S., Heckman T. M., Lennon D. J., Pettini M., Robert C., 2001, ApJ, 550, 724
- Liu X., Shapley A. E., Coil A. L., Brinchmann J., Ma C.-P., 2008, ApJ, 678, 758
- Masters D. et al., 2014, ApJ, 785, 153
- Masters D., Faisst A., Capak P., 2016, ApJ, 828, 18
- Matteucci F., 2014, in Moore Ben, ed., Chemical Evolution of the Milky Way and Its Satellites. Springer, Berlin, Heidelberg, p. 145
- McLean I. S., et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. SPIE, Bellingham, p. 84460J
- Momcheva I. G. et al., 2016, ApJS, 225, 27

- Nomoto K., Tominaga N., Umeda H., Kobayashi C., Maeda K., 2006, Nucl. Phys. A, 777, 424
- Oke J. B. et al., 1995, PASP, 107, 375
- Pilyugin L. S., Vílchez J. M., Mattsson L., Thuan T. X., 2012, MNRAS, 421, 1624
- Rix S. A., Pettini M., Leitherer C., Bresolin F., Kudritzki R.-P., Steidel C. C., 2004, ApJ, 615, 98
- Sanders R. L. et al., 2016, ApJ, 816, 23
- Sanders R. L. et al., 2018, ApJ, 858, 99
- Sanders R. L. et al., 2020, MNRAS, 491, 1427
- Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, ApJ, 588, 65
- Shapley A. E., Coil A. L., Ma C.-P., Bundy K., 2005, ApJ, 635, 1006
- Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., Erb D. K., 2006, ApJ, 651, 688
- Shapley A. E. et al., 2015, ApJ, 801, 88
- Shapley A. E. et al., 2019, ApJ, 881, L35
- Shivaei I. et al., 2016, ApJ, 820, L23

- Sommariva V., Mannucci F., Cresci G., Maiolino R., Marconi A., Nagao T., Baroni A., Grazian A., 2012, A&A, 539, A136
- Stanway E. R., Eldridge J. J., 2018, MNRAS, 479, 75
- Steidel C. C. et al., 2014, ApJ, 795, 165
- Steidel C. C., Strom A. L., Pettini M., Rudie G. C., Reddy N. A., Trainor R. F, 2016, ApJ, 826, 159
- Strom A. L., Steidel C. C., Rudie G. C., Trainor R. F., Pettini M., Reddy N. A., 2017, ApJ, 836, 164
- Strom A. L., Steidel C. C., Rudie G. C., Trainor R. F., Pettini M., 2018, ApJ, 868, 117
- Veilleux S., Osterbrock D. E., 1987, ApJS, 63, 295
- Weinberg D. H., Andrews B. H., Freudenburg J., 2017, ApJ, 837, 183 Zahid H. J., Kudritzki R.-P., Conroy C., Andrews B., Ho I. T., 2017, ApJ,
- 847, 18

This paper has been typeset from a TeX/LATeX file prepared by the author.