The MOSDEF survey: a comprehensive analysis of the rest-optical emission-line properties of $z \sim 2.3$ star-forming galaxies

Jordan N. Runco [1]★ Alice E. Shapley,[1] Ryan L. Sanders [2]† Michael W. Topping,[1] Mariska Kriek,[3]
Naveen A. Reddy,[4] Alison L. Coil,[5] Bahram Mobasher,[4] Brian Siana,[4] William R. Freeman,[4]
Irene Shivaei,[6]† Mojegan Azadi,[7] Sedona H. Price,[8] Gene C. K. Leung,[5] Tara Fetherolf,[4] Laura de Groot,[9]
Tom Zick,[3] Francesca M. Fornasini [7] and Guillermo Barro[10]

[1]*Physics & Astronomy Department, University of California: Los Angeles, 430 Portola Plaza, Los Angeles, CA 90095, USA*
[2]*Department of Physics, University of California, Davis, One Shields Ave, Davis, CA 95616, USA*
[3]*Astronomy Department, University of California, Berkeley, CA 94720, USA*
[4]*Department of Physics & Astronomy, University of California, Riverside, 900 University Avenue, Riverside, CA 92521, USA*
[5]*Center for Astrophysics and Space Sciences, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0424, USA*
[6]*Department of Astronomy/Steward Observatory, 933 North Cherry Ave, Rm N204, Tucson, AZ 85721-0065, USA*
[7]*Harvard–Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*
[8]*Max-Planck-Institut für Extraterrestrische Physik, Postfach 1312, D-85741 Garching, Germany*
[9]*Department of Physics, The College of Wooster, 1189 Beall Avenue, Wooster, OH 44691, USA*
[10]*Department of Phyics, University of the Pacific, 3601 Pacific Ave, Stockton, CA 95211, USA*

ABSTRACT

We analyse the rest-optical emission-line spectra of $z \sim 2.3$ star-forming galaxies in the complete MOSFIRE Deep Evolution Field (MOSDEF) survey. In investigating the origin of the well-known offset between the sequences of high-redshift and local galaxies in the [O III]λ5008/Hβ versus [N II]λ6585/Hα ('[N II] BPT') diagram, we define two populations of $z \sim 2.3$ MOSDEF galaxies. These include the high population that is offset towards higher [O III]λ5008/Hβ and/or [N II]λ6585/Hα with respect to the local SDSS sequence and the low population that overlaps the SDSS sequence. These two groups are also segregated within the [O III]λ5008/Hβ versus [S II]$\lambda\lambda$6718,6733/Hα and the [O III]$\lambda\lambda$4960,5008/[O II]$\lambda\lambda$3727,3730 (O_{32}) versus ([O III]$\lambda\lambda$4960,5008+[O II]$\lambda\lambda$3727,3730)/Hβ (R_{23}) diagrams, which suggests qualitatively that star-forming regions in the more offset galaxies are characterized by harder ionizing spectra at fixed nebular oxygen abundance. We also investigate many galaxy properties of the split sample and find that the high sample is on average smaller in size and less massive, but has higher specific star formation rate (SFR) and SFR surface density values and is slightly younger compared to the low population. From Cloudy+BPASS photoionization models, we estimate that the high population has a lower stellar metallicity (i.e. harder ionizing spectrum) but slightly higher nebular metallicity and higher ionization parameter compared to the low population. While the high population is more α-enhanced (i.e. higher α/Fe) than the low population, both samples are significantly more α-enhanced compared to local star-forming galaxies with similar rest-optical line ratios. These differences must be accounted for in *all* high-redshift star-forming galaxies – not only those 'offset' from local excitation sequences.

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

1 INTRODUCTION

Rest-frame optical emission-line spectroscopy is one of the most important observational tools for understanding galaxy formation and evolution. Such spectra provide a wealth of information about a galaxy, including its star formation rate (SFR), dust extinction, active galactic nucleus (AGN) activity, virial and non-virial dynamics (e.g. outflows) and properties of the ionized interstellar medium (ISM) such as the metallicity, electron density (n_e), and ionization parameter

(U; i.e. the ratio of ionizing photon density to hydrogen, and therefore electron, density). Accordingly, knowledge about the emission-line properties of star-forming galaxies across cosmic time is essential for understanding the evolution of the stellar and gaseous content in galaxies. It is of particular interest to study rest-optical emission-line spectra of galaxies at $z \sim 2$, which represents the epoch of peak star formation in the universe (Madau & Dickinson 2014) and a time before the modern Hubble sequence was fully in place.

In early work, Baldwin, Phillips & Terlevich (1981) showed how diagnostic diagrams measuring the intensity ratios of [O III]λ5008/Hβ versus [N II]λ6585/Hα, i.e. the '[N II] BPT diagram,' can be used to distinguish between star formation and AGN activity as the ionizing source for a galaxy. Other authors have

★ E-mail: jrunco@astro.ucla.edu
† Hubble Fellow.

also considered the [O III]λ5008/Hβ versus [S II]λλ6718,6733/Hα diagram (originally introduced in Veilleux & Osterbrock 1987 and referred to hereafter as the '[S II] BPT diagram') when investigating the ionization mechanism of a galaxy. Indeed, the shape of the ionizing spectrum as well as the typical ionization parameter are different in gas excited by an AGN as opposed to by hot stars. Accordingly, AGNs and star-forming galaxies occupy distinct regions within rest-optical emission-line diagrams. These diagnostic diagrams also reveal information about the physical properties of the galaxies. For example, stellar mass (M_*) and metallicity have been found to increase with decreasing [O II]λ5008/Hβ and increasing [N II]λ6585/Hα along the local star-forming sequence in the [N II] BPT diagram (e.g. Masters, Faisst & Capak 2016).

Another diagram commonly used to describe star-forming galaxies is the [O III]λλ4960,5008/[O II]λλ3727,3730 (O_{32}) versus ([O III]λλ4960,5008+[O II]λλ3727,3730)/Hβ (R_{23}) diagram. O_{32} and R_{23} are rough tracers for ionization parameter and metallicity, respectively. This diagram allows us to probe such physical quantities in star-forming galaxies (e.g. Lilly, Carollo & Stockton 2003; Nakajima et al. 2013). For local galaxies, there is an increase in metallicity from the high-excitation end (high O_{32} & R_{23}) to the low-excitation tail (low O_{32} and R_{23}) on this diagram (Andrews & Martini 2013; Shapley et al. 2015).

Early studies with Keck/NIRSPEC found that the location of galaxies on the [N II] BPT diagram is redshift-dependent, as galaxies with $z > 1$ are found to be offset from local Sloan Digital Sky Survey (SDSS; York et al. 2000) galaxies, showing elevated [O III]λ5008/Hβ at fixed [N II]λ6585/Hα (or vice versa; e.g. Shapley et al. 2005; Erb et al. 2006; Liu et al. 2008). Based on new observations with multiobject near-infrared spectrographs on 8–10-m class telescopes, the sample of high-redshift galaxies with measurements of the BPT diagram emission lines now numbers in the hundreds. Of note, two large surveys of the high-redshift BPT diagram include the MOSFIRE Deep Evolution Field (MOSDEF; Kriek et al. 2015) survey and the Keck Baryonic Structure Survey (KBSS: Steidel et al. 2014), which find, based on much more robust statistical evidence, that high-redshift galaxies tend to have elevated [O III]λ5008/Hβ and/or [N II]λ6585/Hα compared to local galaxies (Steidel et al. 2014; Shapley et al. 2015, 2019; Sanders et al. 2016; Strom et al. 2017).

Understanding this offset in emission-line ratios is vital because we use strong rest-optical emission lines as empirical tracers for many physical properties (e.g. gas-phase oxygen abundance Pettini & Pagel 2004) of the ISM. Due to the observed offset for $z > 1$ galaxies, it is unclear if local metallicity calibrations, e.g. Pettini & Pagel (2004), yield accurate metallicities when applied in the high-redshift universe. Therefore, it is essential to gain a complete understanding of why high-redshift galaxies have elevated [O III]λ5008/Hβ and/or [N II]λ6585/Hα compared to the local SDSS sample.

There have been many proposed explanations of this systematic offset. Possible explanations include variations in physical properties of galaxies such as H II region electron densities (or proportionally pressures), density structure, H II region ionization parameter, H II region ionizing spectra at fixed metallicities, gas-phase N/O abundance ratio differences, unresolved AGN activity, and shocks (e.g. Brinchmann, Pettini & Charlot 2008; Liu et al. 2008; Wright et al. 2010; Kewley et al. 2013; Yeh et al. 2013; Masters et al. 2014; Coil et al. 2015; Sanders et al. 2016; Steidel et al. 2016; Strom et al. 2017; Freeman et al. 2019; Kashino et al. 2019; Shapley et al. 2019; Topping et al. 2020b). Galaxy selection effects could also be potential factors in this offset (Juneau et al. 2014).

Preliminary results from the MOSDEF survey (Shapley et al. 2015; Sanders et al. 2016) suggested that the offset in emission-line ratios is primarily due to elevated N/O at fixed O/H abundance patterns in $z ∼ 2.3$ galaxies relative to local ones. Studies using KBSS data (Steidel et al. 2016; Strom et al. 2017, 2018) argued for a harder ionizing spectrum at fixed nebular metallicity as the main cause of the offset while other works using the Fiber Multi-Object Spectrograph (FMOS)-COSMOS survey (Kashino et al. 2017, 2019) or local analogues of high-redshift galaxies (Bian et al. 2020) attribute the observed [N II] BPT offset to a higher ionization parameter. More recent results from the MOSDEF survey using the complete MOSDEF data set now suggest a harder ionizing spectrum drives the BPT offset (Shapley et al. 2019; Sanders et al. 2020b, a; Topping et al. 2020b).

The completed MOSDEF survey provides an ideal data set of high-redshift galaxies to help explore the observed offset on the [N II] BPT diagram. This survey provides access to spectra of ∼1500 galaxies in the $z ∼ 1.4–3.8$ redshift range, with full spectroscopic coverage of all the emission lines needed to complete the [N II] BPT, [S II] BPT, and O_{32} versus R_{23} diagrams in the $z ∼ 1.4−2.6$ redshift range.

In this work, we improve upon the previous $z > 1$ BPT offset studies based on early MOSDEF data by Shapley et al. (2015) and Sanders et al. (2016). We now have the full MOSDEF sample in hand and apply a more careful spectral energy distribution (SED) fitting method that incorporates corrections to broad-band photometric measurements for rest-optical emission-line fluxes. Accordingly, we derive unbiased age and stellar mass measurements, and can conduct a more thorough and complete investigation of the emission-line ratio properties of the MOSDEF sample. As an example of the level of improvement enabled by the full MOSDEF sample, the initial work by Shapley et al. (2015) explored the location of 53 star-forming galaxies at $1.9 ≤ z ≤ 2.7$ on the [N II] BPT diagram, and investigated the M_*, specific SFR (sSFR), and SFR surface density (Σ_{SFR}) of the sample. In this study, the sample increases to 180 star-forming galaxies on the [N II] BPT diagram, and we also investigate additional galaxy parameters such as the galaxy effective radius (R_e), SFR, n_e, and stellar population age. Sanders et al. (2016) investigated how a sample of 53 star-forming galaxies at $1.9 ≤ z ≤ 2.7$ translated from the [N II] BPT diagram to the [S II] BPT diagram and O_{32} versus R_{23} diagram. In the current study, that number is more than doubled to 122. Our more comprehensive analysis enables a better understanding of the observed offset between the MOSDEF sample and local galaxies on the [N II] BPT diagram. Through the analysis outlined above, we aim to determine the underlying physical causes of this offset.

The paper is organized as follows. In Section 2, we briefly describe the MOSDEF survey, and review our survey sample selection and data reduction. Section 3 presents the results of this study. Section 4 discusses the results and how they relate to photoionization models. Section 5 presents a summary of key results and looks ahead to future analyses. We adopt the following abbreviations for emission-line ratios used frequently throughout the paper.

$$N2 = [N\ II]\lambda6585/H\alpha, \tag{1}$$

$$S2 = [S\ II]\lambda\lambda6718, 6733/H\alpha, \tag{2}$$

$$O3 = [O\ III]\lambda5008/H\beta, \tag{3}$$

$$R_{23} = ([O\ III]\lambda\lambda4960, 5008 + [O\ II]\lambda\lambda3727, 3730)/H\beta, \tag{4}$$

$$O_{32} = [O\ III]\lambda\lambda4960, 5008/[O\ II]\lambda\lambda3727, 3730, \tag{5}$$

O3N2 = O3/N2.   (6)

O_{32} and R_{23} have already been defined, but we include them here for completeness. All emission-line wavelengths are vacuum wavelengths. Throughout this paper, we adopt a Lambda cold dark matter cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$. Also, we assume the solar abundance pattern from Asplund et al. (2009).

2 METHODS

2.1 MOSDEF sample and ancillary measurements

The MOSDEF survey was a 48.5-night observing program spanning multiple years (2012–2016), using the MultiObject Spectrometer For Infra-Red Exploration (MOSFIRE; McLean et al. 2012) on the 10-m Keck I telescope . The survey targets three redshift ranges: $1.37 \leq z \leq 1.70$, $2.09 \leq z \leq 2.61$, and $2.95 \leq z \leq 3.80$. These ranges are chosen to optimize the detection of strong rest-optical emission lines within windows of atmospheric transmission. There are 354, 678, and 266 galaxies that have been spectroscopically confirmed in the MOSDEF low-, medium-, and high-redshift ranges, respectively. The sample is *H*-band magnitude limited, and is located in the well-studied CANDELS and 3D-*HST* legacy fields (Grogin et al. 2011; Koekemoer et al. 2011; Momcheva et al. 2016): AEGIS, COSMOS, GOODS-N, GOODS-S, and UDS. The ancillary data from these fields enables us to probe other properties of the sample (e.g. galaxy size and stellar mass).

In this study, we restrict the sample to the central redshift bin, spanning the redshift range to $1.9 \leq z \leq 2.7$ to reflect the scatter between target photometric and measured spectroscopic redshifts (i.e. the fact that not all spectroscopic redshifts, when actually measured, fell precisely within the target $2.09 \leq z \leq 2.61$ window). For galaxies in this redshift range, [O II]$\lambda\lambda$3727,3730, Hβ, [O III]$\lambda\lambda$4960,5008, Hα, [N II]λ6585, and [S II]$\lambda\lambda$67186733 features are captured within the *J*, *H*, and *K* near-IR filter wavelength ranges. For MOSFIRE, the *J*, *H*, and *K* filters have typical respective wavelength coverages of 1.142–1.365, 1.450–1.826, and 1.897–2.427 μm and spectral resolutions of $R = 3000$, 3650, and 3600.

We imposed further restrictions to reach our final sample. Galaxies containing an AGN were identified and removed from the sample based on their IR colours, X-ray luminosity, or if N2 > 0.5 (Coil et al. 2015; Azadi et al. 2017; Leung et al. 2017). We also require an S/N ≥ 3 for each emission-line in our analysis. These criteria yielded a sample of 180 galaxies for initial classification in the [N II] BPT diagram, of which 122 galaxies are additionally detected in [S II]$\lambda\lambda$6718,6733 and [O II]$\lambda\lambda$3727,3730 and have size measurements from the *Hubble Space Telescope* (*HST*). The redshift distribution of the 180 MOSDEF star-forming galaxies at $z \sim 2.3$ with [N II] BPT diagram classifications is shown in Fig. 1.

We correct Hα and Hβ line fluxes for stellar Balmer absorption as described in Kriek et al. (2015) and Reddy et al. (2015). We also dust correct emission-line ratios for which the member features differ significantly in wavelength, including O_{32} and R_{23}. For such corrections, we assumed the Cardelli, Clayton & Mathis (1989) dust attenuation curve and an unreddened Hα/Hβ ratio of 2.86. In the case of emission-line ratios for which the lines are close in wavelength (O3, N2, and S2) dust correction was not applied.

In this study, we investigate additional galaxy photometric, spectroscopic, and structural properties for the MOSDEF sample in order to try to better understand the observed rest-optical emission-line properties. The Hao et al. (2011) calibration for a Chabrier (2003)
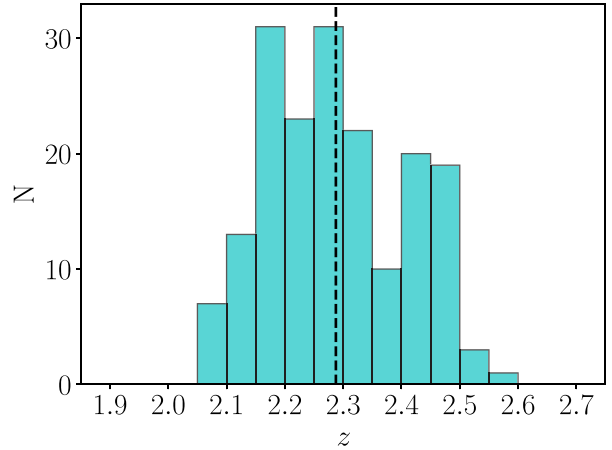
Figure 1. Redshift distribution of the 180 star-forming MOSDEF galaxies at $z \sim 2.3$ with [N II] BPT diagram classifications (i.e. S/N ≥ 3 detections in Hβ, [O III]$\lambda\lambda$4960,5008, Hα, and [N II]λ6585). The dashed vertical line displays the median redshift of the sample, $z_{med} = 2.29$.

initial mass function (IMF) and solar metallicity is used to estimate SFR(Hα) from stellar-Balmer-absorption-corrected, dust-corrected, and slit-loss-corrected Hα luminosities (Kriek et al. 2015; Reddy et al. 2015; Shivaei et al. 2015). We used the SED-fitting code, FAST (Kriek et al. 2009) to obtain key stellar population parameters including stellar mass and age. For this modelling, as a default, we assumed star formation histories of the delayed-τ form, where SFR $\propto t \times e^{-t/\tau}$. Here, t is the time since the onset of star formation (i.e. age), and τ is the characteristic star formation time-scale. Given the range of best-fitting τ values, the meaning of absolute ages is not necessarily clear. In order to obtain a better gauge of the relative maturities of the galaxy stellar populations in our sample, we used both normalized ages (i.e. t/τ), and also the age (t) obtained from constant star formation (CSF) models. Galaxy sizes, R_e, are taken as the F160W galaxy half-light radii from the van der Wel et al. (2014) catalogue, which were estimated using single-component Sérsic profile fits to the two-dimensional light distribution of galaxies in the CANDELS and 3D-*HST* fields. We combine SFR(Hα) and R_e to estimate the SFR surface density, Σ_{SFR} as

$$\Sigma_{SFR} = \frac{SFR(H\alpha)}{2\pi R_e^2} \tag{7}$$

We estimate the sSFR using SFR(Hα) and M_*, both of which are described above.

The electron density, n_e, is estimated with the [O II] and [S II] emission-line doublets using the method described in Sanders et al. (2016). In that study, the [O II] and [S II] electron densities agree within the uncertainties; both reliable tracers of the density within H II regions. [O II] is the preferred choice because the doublet typically has a higher S/N. We only use [S II] measurements when we do not have adequate [O II] data. This scenario arises for the following reasons:

(i) Low S/N. An S/N ≥ 3 is required for each emission-line in the doublet to estimate n_e; however, a less restrictive requirement of S/N ≥ 3 for the overall doublet is sufficient for a galaxy to be included in the [S II] BPT and O_{32} versus R_{23} diagrams. We only require S/N ≥ 3 for the individual doublet members when investigating the n_e of the sample.
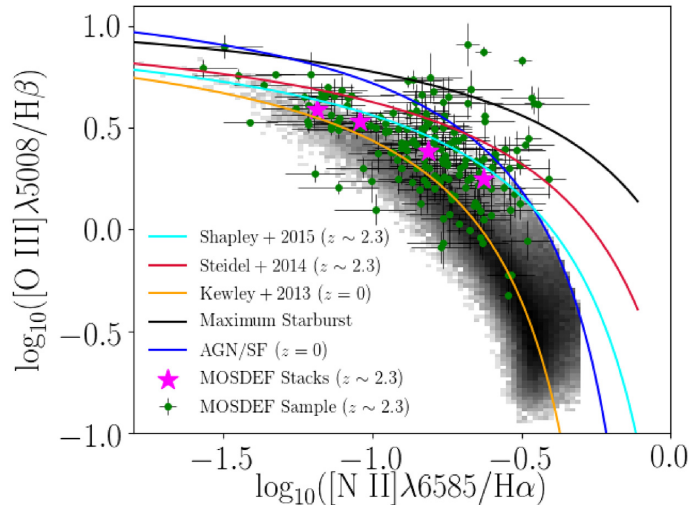
Figure 2. [N II] BPT diagram. Green points indicate $1.9 \leq z \leq 2.7$ MOSDEF galaxies with S/N ≥ 3 for all four plotted emission-lines. Large magenta stars represent measurements of stacks from composite spectra, binned by stellar mass, for $1.9 \leq z \leq 2.7$ MOSDEF galaxies. The grayscale 2D histogram indicates local SDSS galaxies. The cyan curve is the fit to the MOSDEF sample from Shapley et al. (2015). The orange curve is a fit to the $z \sim 0$ star-forming locus (Kewley et al. 2013). The red curve is the best fit to the $z \sim 2.3$ galaxies from KBSS (Steidel et al. 2014). The black curve is the maximum starburst line from Kewley et al. (2001). The blue curve is the empirical AGN-/star-forming galaxy dividing line from Kauffmann et al. (2003).

(ii) Rejection of the [O II] doublet based on visual inspection. Such cases occur when skylines affect the [O II] doublet but not the [S II] doublet, spurious detections, and poor fits.

We set the lower limit for n_e to be 1 cm^{-3}.

2.2 SDSS Comparison Sample

Throughout this study, we compare our high-redshift MOSDEF sample to local galaxies. For this comparison, we use archival data from the SDSS Data Release 7 (DR7; Abazajian et al. 2009). We obtain galaxy properties and emission-line measurements from the MPA-JHU DR7 release of spectrum measurements [1]. SDSS galaxies are selected within the $0.04 \leq z \leq 0.10$ redshift range. We impose similar restrictions on our SDSS sample to those applied to the MOSDEF sample by requiring that each emission-line used in the analysis has an S/N ≥ 3. For the SDSS sample, we remove AGN using equation (1) from Kauffmann et al. (2003). Galaxies are also identified as having an AGN component if N2 > 0.5. These criteria result in a comparison sample of 103 422 SDSS galaxies when considering the [N II] BPT diagram alone, and 74 726 SDSS galaxies when considering galaxies with simultaneous detections across all three emission-line diagrams analyzed in this work ([N II] BPT, [S II] BPT, and O$_{32}$ versus R$_{23}$).

3 RESULTS

3.1 The [N II] BPT diagram

We start by investigating the locations of $z \sim 2.3$ star-forming galaxies in the [N II] BPT diagram based on the complete MOSDEF sample (Fig. 2). We include the 180 galaxies with $\geq 3\sigma$ detections for all four emission-lines (Hβ, [O III]$\lambda\lambda$4960,5008, Hα, and [N II]λ6585) and the corresponding SDSS sample with the same four emission lines

detected. Similar to previous MOSDEF studies (e.g. Shapley et al. 2015; Sanders et al. 2016; Shapley et al. 2019; Topping et al. 2020b) and other studies from the literature (e.g. Shapley et al. 2005; Erb et al. 2006; Steidel et al. 2014), there is a systematic offset observed for the high-redshift MOSDEF sample from the local sequence (Fig. 2). The MOSDEF galaxies appear on average to be shifted towards the AGN region of the diagram with elevated N2 and/or O3 values, with some galaxies on the AGN side of the Kauffmann et al. (2003) AGN/SF boundary. There are even a small number of galaxies past the maximum starburst line from Kewley et al. (2001) as well.

To search for any biases in the sample based on the $\geq 3\sigma$ detection requirement in all four [N II] BPT lines, we construct spectral stacks for all MOSDEF galaxies at $1.9 \leq z \leq 2.7$ with Hα emission detected at S/N ≥ 3. There are four spectral stacks, divided into bins of stellar mass (see Sanders et al. 2018 for a full description of the method for constructing composite spectra). The emission-line ratios measured from stacked spectra follow the distribution of data points measured from MOSDEF galaxies with individual detections in all [N II] BPT lines. Therefore, the sample of individual detections seems to represent the parent $z \sim 2.3$ MOSDEF data set with minimal bias. The four stacks, similarly to the individually detected galaxies, are offset from the local SDSS sample with elevated N2 and/or O3 values.

3.2 Dividing the $z \sim 2.3$ sample in the [N II] BPT diagram

Fig. 2 shows that on the [N II] BPT diagram, part of the MOSDEF sample sits on or near the local SDSS sequence, while the remainder of the galaxies lie off the SDSS sequence and are shifted towards the local AGN region. We now investigate if the location of a high-redshift galaxy − either on or offset from the local SDSS sequence − is connected with other physical properties of the galaxy. For this analysis, we examine how galaxies on the [N II] BPT diagram populate the [S II] BPT and O$_{32}$ versus R$_{23}$ diagrams, and therefore require the sample to have $\geq 3\sigma$ detections for all emission-lines on these diagrams: [O II]$\lambda\lambda$3727,3730, Hβ, [O III]$\lambda\lambda$4960,5008, Hα, [N II]λ6585, and [S II]$\lambda\lambda$6718,6733. There are 123 MOSDEF galaxies that meet this criterion. One galaxy was removed because
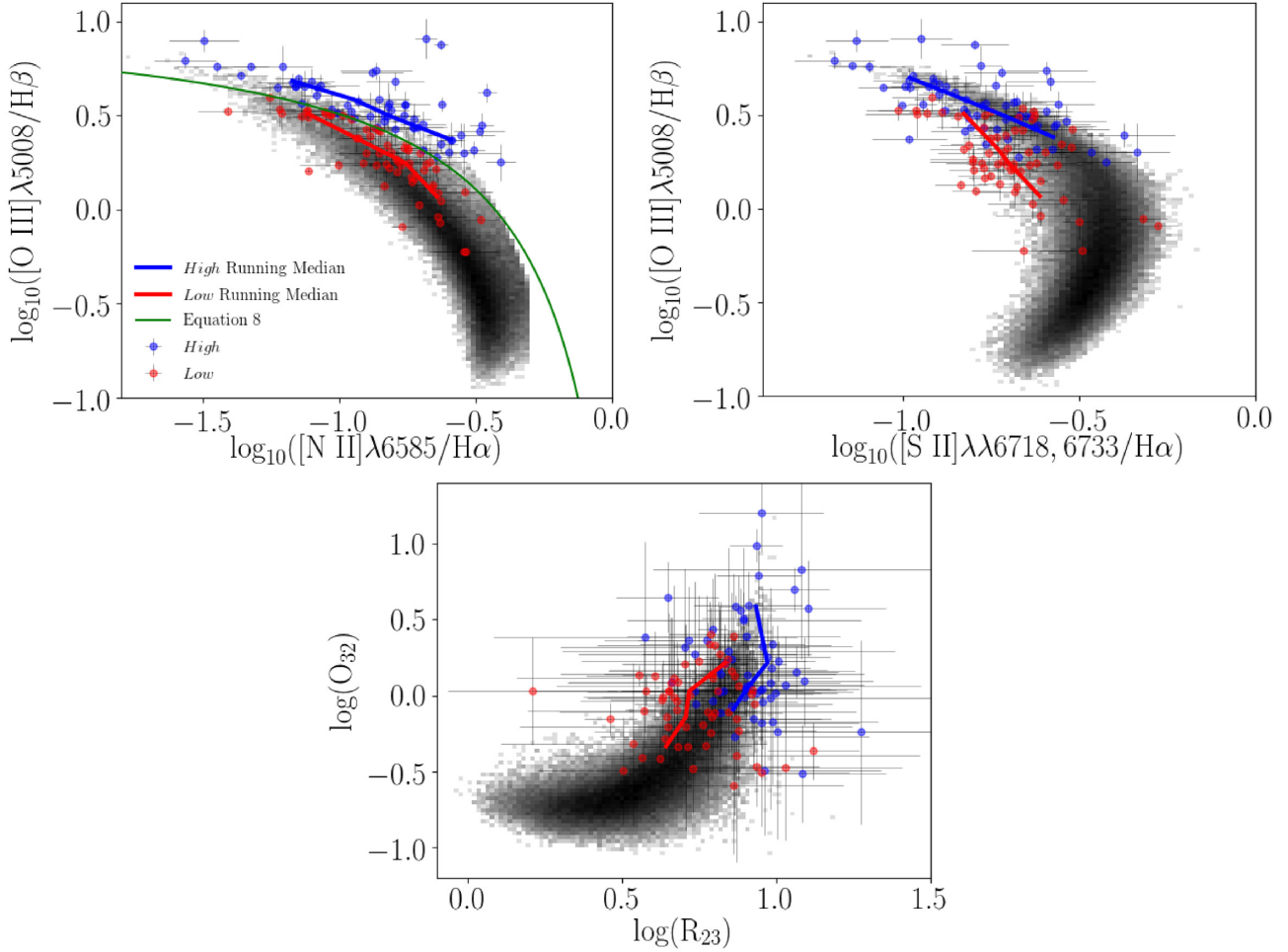
Figure 3. The sample of 122 MOSDEF galaxies that have $\geq 3\sigma$ detections in [O II]$\lambda\lambda$3727,3730, Hβ, [O III]$\lambda\lambda$4960,5008, Hα, [N II]λ6585, and [S II]$\lambda\lambda$6718,6733. Top left-hand panel: we modify the MOSDEF fit from Shapley et al. (2015) by altering the y-intercept, so a curve with the same functional form splits the data set evenly into two populations. The red galaxies (low sample) fall close to the local SDSS sequence, while the blue galaxies (high sample) are shifted off the local sequence. Top right-hand panel: the same two groups of data points as in the left-hand panel. The red and blue lines indicate binned medians as described in Section 3.2. These lines show that the two groups are also separated in the [S II] BPT diagram, with the high sample higher in S2 and/or O3 than the low sample. Bottom centre panel: the same two groups of data points as in the top panels. The red and blue lines again indicate binned medians as described in Section 3.2. The two groups are segregated in the O_{32} versus R_{23} diagram as well, with the high sample having a higher median R_{23} values at fixed O_{32} where there is overlap in O_{32}.

it does not have a R_e measurement, which brings the final sample to 122 MOSDEF galaxies.

For this analysis, we split the final sample of 122 galaxies into two groups using the functional form presented in Shapley et al. (2015), and adjusting the y-intercept so that our sample is divided in half. The equation used to split the MOSDEF sample into two groups of 61 galaxies each is

$$\log\big([\text{O III}]/\text{H}\beta\big) = \frac{0.67}{\log([\text{N II}]/\text{H}\alpha) - 0.20} + 1.065 \qquad (8)$$

The new curve is 0.055 dex lower in [O III]λ5008/Hβ compared to equation (1) in Shapley et al. (2015). This splitting of the sample on the [N II] BPT diagram is shown in the upper left-hand panel of Fig. 3, where galaxies above the curve are indicated with blue symbols and those below with red symbols. Hereafter, we refer to the galaxies above the curve on the [N II] BPT diagram as the high sample, and the group below the curve as the low sample, following the nomenclature in Topping et al. (2020b). We plot the two samples on both the [S II] BPT diagram (upper right-hand panel of Fig. 3)

and the O_{32} versus R_{23} diagram (bottom panel of Fig. 3). We also include binned median lines for both the high (blue) and low (red) populations. These binned medians are binned by $\log_{10}$(O3N2) for the [N II] BPT diagram, $\log_{10}$(O3) $-$ $\log_{10}$(S2) for the [S II] BPT diagram, and $\log_{10}(O_{32})$ + $\log_{10}(R_{23})$ for the O_{32} versus R_{23} diagram. The binning schemes above were adopted because they divide our sample into subgroups segregated roughly along the local star-forming sequence. For both the high and low samples, each of which contains 61 galaxies, there are four equally sized bins (three bins of 15 galaxies and one bin of 16).

For completeness, we checked for potential bias between the smaller sample of 122 galaxies with detections across all three emission-line diagrams and the parent sample of 180 galaxies defined based on detections in the [N II] BPT diagram alone. Of the 58 galaxies that were removed, 32 are above the curve dividing our sample (equation 8) and 26 are below it on the [N II] BPT diagram. Because these two groups are approximately equal in size, we can expect that the smaller subset with [O II] and [S II] also detected (122 galaxies) has approximately the same average BPT offset as the larger

[N II] BPT only sample (180 galaxies). Therefore, we can conclude that the cuts to create the sample of 122 galaxies with detections across all three emission-line diagrams do not introduce biases when compared to the larger parent sample of 180 galaxies.

3.2.1 Division in the other diagrams

We find separation between the high and low samples on both the [S II] BPT and O_{32} versus R_{23} diagrams. For the [S II] BPT diagram, the high sample has a systematically higher S2 at fixed O3 (or vice versa). For the O_{32} versus R_{23} diagram, the high sample has a systematically higher R_{23} value at fixed O_{32}.

Our new results on the [S II] BPT and O_{32} versus R_{23} diagrams update the early MOSDEF results from Sanders et al. (2016). In this earlier work, based on a significantly smaller sample of only 53 galaxies, subsamples split in the [N II] BPT diagram were found to be well mixed in the [S II] BPT and O_{32} versus R_{23} diagrams. In contrast, our results are consistent with those of Strom et al. (2017), based on the KBSS survey. In that study, the sample was split into two groups based on [N II] BPT diagram location − large and small offset from the local SDSS sequence. These subsamples were found to remain separated on the [S II] BPT diagram. Strom et al. (2017) did not investigate if this segregation remained on the O_{32} versus R_{23} diagram. Section 4 contains a more in-depth comparison with the results from Strom et al. (2017) and a discussion of the implications for the underlying causes of the observed shift in the high-redshift [N II] BPT diagram.

3.2.2 Physical properties of the galaxies

The fact that there is a segregation in all three plots in Fig. 3 implies that there are some key differences between galaxies offset from the local SDSS sequence and galaxies that overlap with it. We now investigate several galaxy physical properties to uncover any differences between the galaxies that fall into these two categories. In this study, as described in Section 2.1, we focus on: M_*, $\log(t/\tau)$ of the stellar population inferred from a delayed-τ star formation model, t of the stellar population assuming a CSF history, SFR(Hα), sSFR, Σ_{SFR}, R_e, and n_e. We have measurements of these properties for all 122 galaxies in the sample except for n_e. As discussed in Section 2.1, a more stringent requirement of an S/N $\geq$ 3 for each component of the [O II] or [S II] doublets is needed to estimate n_e while an S/N $\geq$ 3 for the combined doublet is needed to be plotted on the [S II] BPT and O_{32} versus R_{23} diagrams. Accordingly, while our fiducial sample for analysis contains 122 galaxies, only 90 of them (44 high and 46 low) have reliable n_e estimates.

We show how the high and low samples divide in the space of each of these parameters using histograms, marking the median value of the parameter with solid lines for each high and low population (Fig. 4). Showing the data in this format clearly highlights how the high and low samples separately distribute in each of the galaxy parameters and, accordingly, which of the galaxy parameters are correlated with the location of galaxies on the [N II] BPT diagram. The median values for the high and low populations are listed with 1σ uncertainties derived from bootstrap resampling in Table 1. We also list the probabilities, based on the Kolmogorov–Smirnov (K-S) test, of the null hypothesis that the high and low samples are drawn from the same parent distribution.

While the distributions of the high and low samples significantly overlap for some galaxy parameters, there are several parameters for which the high and low distributions are measurably offset in

the median and spread. These results indicate that high sample of MOSDEF galaxies tend to have a smaller R_e and M_*, but a larger sSFR and Σ_{SFR} compared to the low sample. We also note that even though R_e, sSFR, and Σ_{SFR} are correlated with M_*, at fixed M_* the high sample is still found to be smaller with larger sSFR and Σ_{SFR} values compared to the low sample. The variation in median Σ_{SFR} between the two samples can be attributed to the difference in median R_e. At fixed size, the median SFR(Hα) values for the high and low populations agree within the uncertainties. With smaller significance, the high sample is also younger (smaller median t/τ and CSF t). We find no mutual correlation for either n_e or SFR with the location of galaxies on the [N II] BPT diagram.

4 DISCUSSION

4.1 Comparison with previous work

We have divided the $z \sim 2.3$ MOSDEF sample according to location in the [N II] BPT diagram. When separated in this manner, our sample also shows segregation in the [S II] BPT and O_{32} versus R_{23} diagrams. These results update earlier MOSDEF work (Shapley et al. 2015; Sanders et al. 2016), and are in agreement with more recent MOSDEF studies (e.g. Shapley et al. 2019; Topping et al. 2020b). Our findings for the population segregation on the [S II] BPT diagram are also in agreement with those of Strom et al. (2017), in which a similar segregation in the [S II] BPT diagram is found for $z \sim$ 2.3 star-forming galaxies from the KBSS survey when separated by [N II] BPT location. Strom et al. (2017) did not consider the relative positions of the KBSS equivalent of our high and low galaxies on the O_{32} versus R_{23} diagram. However, these authors do show how photoionization models with different input parameters (i.e. varying ionizing spectrum at fixed nebular metallicity) vary across the O_{32} versus R_{23} parameter space. Specifically, it is shown that models with harder ionizing spectra at fixed nebular metallicity are offset towards higher R_{23} at fixed O_{32}. Strom et al. (2017) additionally show that models with harder ionizing spectra produce higher O3 at both fixed S2 and N2. These separations are at least qualitatively similar to that observed between our high and low samples.

The suggested implication of these segregations on both [S II] BPT and O_{32} versus R_{23} diagrams is a harder ionizing spectrum at fixed O/H, relative to local galaxies (e.g. Strom et al. 2017; Shapley et al. 2019; Topping et al. 2020b). A harder ionizing spectrum at fixed nebular metallicity in $z \sim 2.3$ galaxies may arise due to α-enhancement (i.e. supersolar O/Fe values) in the massive stars exciting the ionized gas in star-forming regions. Such abundance patterns may arise naturally in high-redshift galaxies, given the young median ages of their stellar populations (e.g. Steidel et al. 2016; Sanders et al. 2020a; Topping et al. 2020b).

We have also shown in Fig. 4 that the high sample is associated with more compact (i.e. smaller R_e and higher Σ_{SFR}), and intense (i.e. higher sSFR) star formation than the low sample. These differences between the two populations exist at fixed stellar mass as well. It will be important to explore the links between these global galaxy properties and the abundance patterns of massive stars using realistic galaxy formation simulations (e.g. FIRE-2; Hopkins et al. 2018).

4.2 Comparison with photoionization models

4.2.1 Modelling methodology

We use photoionization models to explain simultaneously the joint distributions of high and low galaxies in the [N II] and [S II] BPT and
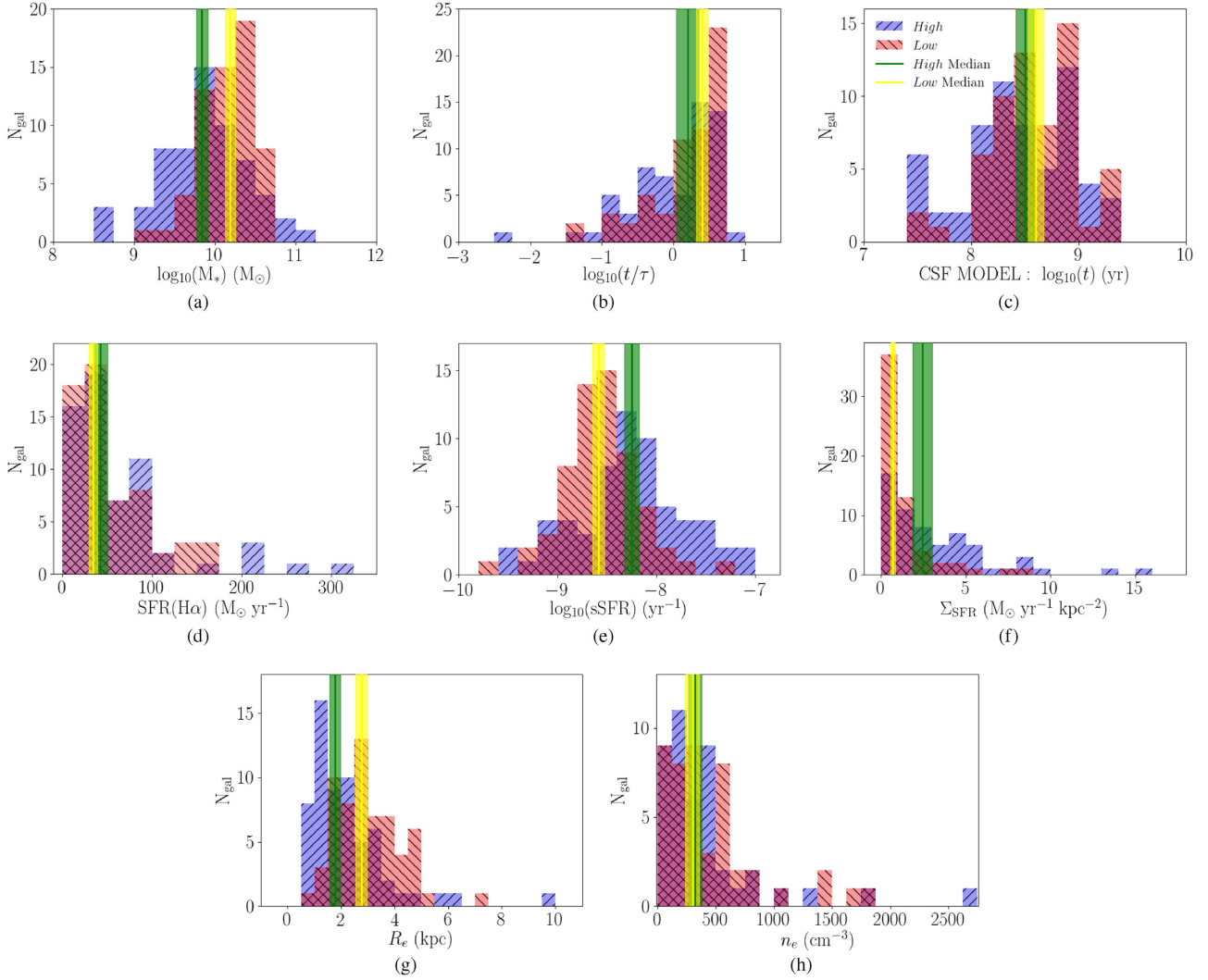
Figure 4. Distribution of physical properties for the 122 galaxy MOSDEF sample shown in Fig. 3 where the blue and red bins correspond to the high and low samples, respectively. The green lines are the median values for the high population, while the yellow lines are the median values for the low population. The median values with uncertainties for both populations are given in Table 1. The following galaxy properties are shown: (a) M_*, (b) $\log(t/\tau)$ of the stellar population using a delayed-τ star formation model, (c) t of the stellar population assuming a CSF history, (d) SFR(Hα), (e) sSFR, (f) Σ_{SFR}, (g) R_e, and (h) n_e estimated from the [O II] or [S II] doublet emission lines. We find that the median M_*, sSFR, Σ_{SFR}, and R_e differ significantly based on their location in the [N II] BPT Diagram, while the other properties are the within the uncertainties of each other.

Table 1. Column (1): physical property of the galaxies in the sample shown in Fig. 4. Column (2): median value with uncertainty of the high sample. Column (3): median value with uncertainty of the low sample. Column (4): Two-tailed *p*-value, based on the K-S test, estimating the probability that the null hypothesis can be rejected. Column (5): statistical significance (i.e. the σ value) that the *p*-value corresponds to.

Median values for physical properties of the *high* and *low* samples				
Physical property (1)	*High* median (2)	*Low* median (3)	*p*-value (4)	Statistical significance (σ) (5)
$\log_{10}(M_*/M_\odot)$	9.84 ± 0.07	10.19 ± 0.06	0.0040	2.9σ
$\log_{10}(t/\tau)$	0.20 ± 0.16	0.40 ± 0.08	0.2480	1.2σ
CSF $\log_{10}(t/\text{yr})$	8.50 ± 0.08	8.60 ± 0.07	0.1671	1.4σ
SFR(Hα) (M$_\odot$ yr^{-1})	42.7 ± 7.5	34.3 ± 4.9	0.7923	0.3σ
sSFR (yr^{-1})	-8.25 ± 0.07	-8.58 ± 0.06	0.0002	3.7σ
Σ_{SFR} (M$_\odot$ yr^{-1} kpc^{-2})	2.45 ± 0.56	0.70 ± 0.15	0.0002	3.7σ
R_e (kpc)	1.77 ± 0.20	2.75 ± 0.21	0.000 04	4.1σ
n_e (cm^{-3})	322 ± 58	300 ± 59	0.5341	0.6σ

O_{32} versus R_{23} diagrams. For this analysis, we use a combination of the code Cloudy (v17.01; Ferland et al. 2017) and the Binary Population And Spectral Synthesis (BPASS) v2.2.1 models (Eldridge et al. 2017; Stanway & Eldridge 2018). BPASS generates the spectra of model stellar populations, which we used as the input ionizing spectrum to Cloudy. We assume a CSF history in the BPASS models, as constructed in Topping et al. (2020b). Additionally, we use stellar population models that follow a Chabrier (2003) IMF and set 100 $M_\odot$ as the high-mass cutoff.

Using the BPASS input ionizing spectrum, Cloudy predicts the rest-optical emission-line strengths for different combinations of physical properties. For this analysis, we set the N/O abundance ratio using equation (2) in Pilyugin et al. (2012), use the Asplund et al. (2009) [S II] abundance pattern, and assume that $n_e = 250$ cm^{-3}, which is characteristic of $z \sim 2.3$ galaxies (Sanders et al. 2016; Strom et al. 2017; Harshan et al. 2020). This value for n_e is slightly lower than the median values of our high and low samples; however, it is within the uncertainties of the low population and not significantly less than the lower limit of the high population median. The best-fitting n_e values from this study also agree within the uncertainties with the those reported in Sanders et al. (2016) and Strom et al. (2017). In addition, we use Cloudy to estimate the contribution from the nebular continuum and added this into the BPASS stellar population models. To estimate this nebular contribution, we assume that $\log(U) = -2.5$ and $\log(Z_{neb}/Z_\odot) = -0.2$, which are typical of $z \sim 2.3$ galaxies (Sanders et al. 2016). Changing these conditions does not significantly affect the nebular continuum, which is small contribution to the overall spectrum. Therefore, adjustments to the initial parameters for the nebular continuum have minimal impact on the final model fitting (Topping et al. 2020b). As a sanity check on our models, we are able to reproduce the results from Steidel et al. (2016) using their initial input parameters.

For our Cloudy+BPASS model grids, we varied the following parameters: stellar metallicity (Z_*), t (i.e. age) of the stellar population assuming a CSF history, gas-phase oxygen abundance (Z_{neb}), and ionization parameter (U). We find that our choice of t over the range 10^7–$10^{9.8}$ yr (Topping et al. 2020b) has a negligible effect of the position of models on all three diagrams, so we chose a value of $10^{8.6}$ yr because it is approximately the median t (assuming a CSF history) of the $z \sim 2.3$ MOSDEF sample in this study.

In the photoionization model grids, ionization parameter and nebular metallicity range between $-3.60 \leq \log(U) \leq -1.40$ and $-1.3 \leq \log(Z_{neb}/Z_\odot) \leq 0.20$, respectively. Within the ranges of interest ($-3.40 \leq \log(U) \leq -2.40$ and $-1.00 \leq \log(Z_{neb}/Z_\odot) \leq 0.00$), these grids are finely sampled in steps of 0.02 dex. Outside of these regions, $\log(U)$ changes in steps of 0.20 dex. We also include three additional values for $\log(Z_{neb}/Z_\odot)$: 0.10 and 0.20, extending up towards higher metallicity, and -1.30 to round out the low-metallicity extreme. The stellar metallicity of the BPASS models includes discrete values of $Z_* = 0.00001, 0.0001, 0.001, 0.002, 0.003, 0.004, 0.006, 0.008, 0.01, 0.014, 0.02$, and 0.03. We assume Asplund et al. (2009) solar abundance (i.e. $Z_* = 0.014$).

4.2.2 Fixed Z_* models

In a recent MOSDEF study, Topping et al. (2020b) combined rest-UV and rest-optical spectra for a sample of 62 star-forming galaxies at $z \sim 2.3$ and constructed high and low samples using a similar

methodology as this study (i.e. based on offset from the local sequence on the [N II] BPT diagram). Fitting composite rest-frame UV spectra from their high and low stacked spectra resulted in best-fitting Z_* values of 0.001 and 0.002, respectively. Best-fitting Z_{neb} and $\log(U)$ values were then based on comparing the emission-line ratios measured from the stacked spectra with those generated by the same Cloudy photoionization model described above assuming the stellar metallicities obtained from the rest-UV spectra. We do not have rest-UV spectra for the full sample of 122 galaxies in our emission-line analysis, and therefore cannot obtain tight constraints on the median Z_* for the high and low samples of this work from our rest-optical data alone. Accordingly, we adopt the best-fitting Z_* values from Topping et al. (2020b) for our two MOSDEF populations. This is a reasonable assumption because 45/62 (73 per cent) of the galaxies in Topping et al. (2020b) are in this work, and the high and low samples are reasonably well matched to the corresponding samples in Topping et al. (2020b).

Fig. 5 shows Cloudy+BPASS photoionization models overplotted with our data, assuming Z_* values of 0.001 (left-hand column; high sample) and 0.002 (right-hand column; low sample). Using the solar value of $12+\log(O/H) = 8.69$ (Asplund et al. 2009), we convert $\log(Z_{neb}/Z_\odot)$ to $12+\log(O/H)$. Each curve corresponds to a different $12+\log(O/H)$ value, as indicated by colour. The triangular points along each nebular metallicity curve comprise the sequence of $\log(U)$ values. The points are smallest for $\log(U) = -3.60$ and largest for -1.40, therefore, increasing in size as $\log(U)$ increases. The Cloudy+BPASS model grids are overplotted on the MOSDEF and SDSS data. The green and yellow squares on the [N II] BPT, [S II] BPT, and O_{32} versus R_{23} diagrams indicate, respectively, the median values of the high and low samples. Uncertainties for these median values were obtained through bootstrap resampling and perturbing each resampled data point by its individual error according to a normal distribution. The sample median was calculated for each bootstrapped, perturbed sample, and then the 16th and 84th percentiles of the distribution of medians were taken as the lower and upper bounds, respectively, as the 1σ confidence interval on the plotted median value.

Using stellar metallicities of 0.001 and 0.002, respectively, for the high and low populations, we then estimate the best-fitting $12+\log(O/H)$ and $\log(U)$ values for the sample median data points using a χ^2 method that compares the N2, S2, O3, and O_{32} emission-line ratios between the models and data. R_{23} is not used because it is not independent of the other emission-line ratios (i.e. R_{23} can be found from a combination of O3 and O_{32}). While Fig. 5 displays $12+\log(O/H)$ and $\log(U)$ in large steps, we use the finer grids of 0.02 dex spacing for both physical properties when estimating the best-fitting sample values. The results for the best-fitting $12+\log(O/H)$ and $\log(U)$ values with 1σ uncertainties from the bootstrap resampling are shown in Table 2. The corner plot showing the distributions of these two properties obtained using the bootstrapped, perturbed N2, S2, O3, and O32 medians are shown in the top left-hand panel of Fig. 6.

We find that both $12+\log(O/H)$ and $\log(U)$ are larger for the median data point of the high population ($12+\log(O/H) = 8.37^{+0.04}_{-0.04}$, $\log(U) = -2.88^{+0.04}_{-0.04}$) compared to the median data point of the low population ($12+\log(O/H) = 8.31^{+0.04}_{-0.02}$, $\log(U) = -3.08^{+0.02}_{-0.04}$). The difference in $\log(U)$ between the two samples is more statistically significant than the difference in $12+\log(O/H)$, as median values of the latter agree within the uncertainties. Because the median electron densities of the high and low sample agree within the uncertainties (see Table 1), the difference in $\log(U)$ can be attributed to the high population having a higher number density of ionizing photons. This
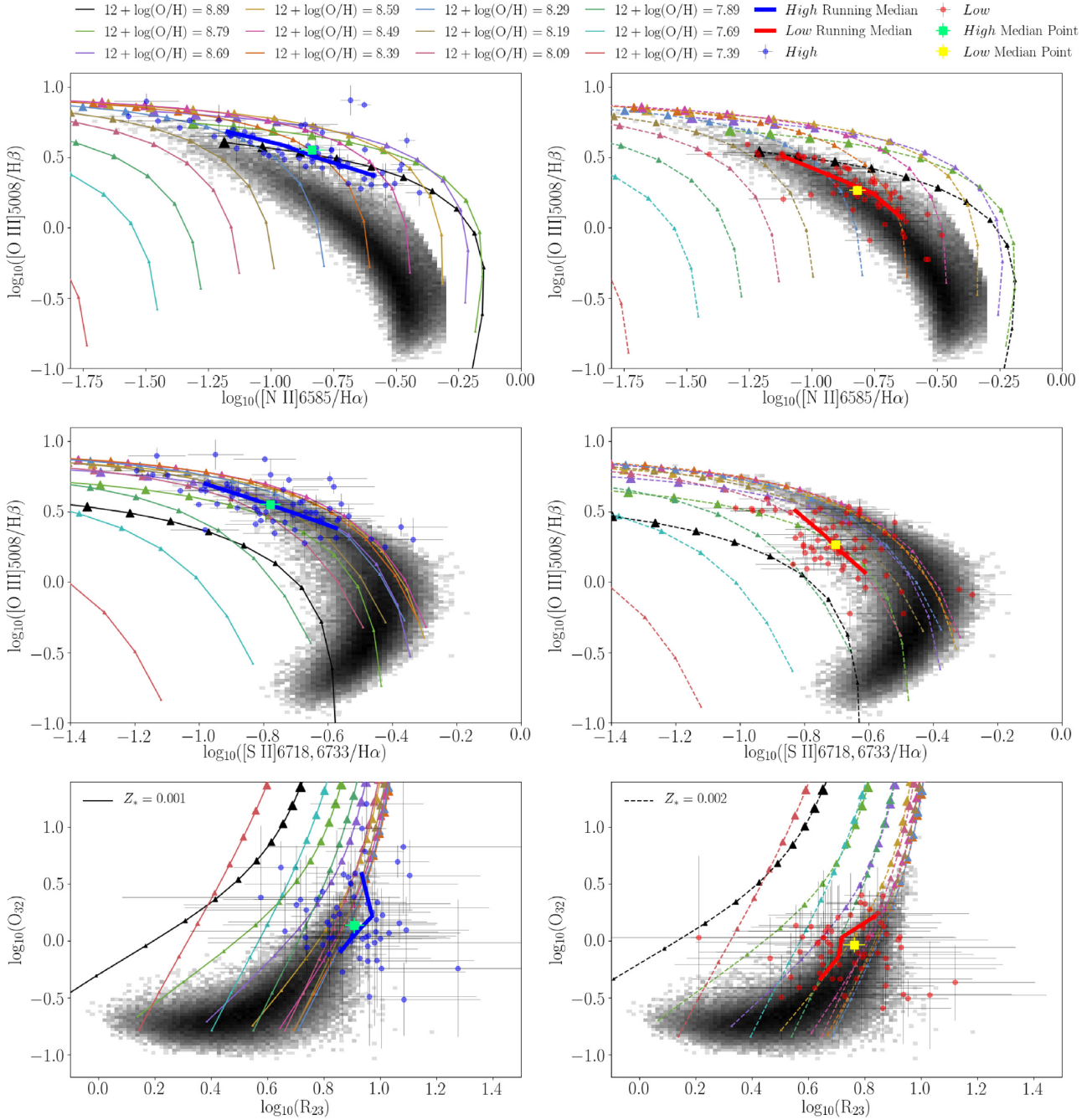
Figure 5. Left-hand column: $Z_* = 0.001$ (solid lines). Right-hand column: $Z_* = 0.002$ (dashed lines). Cloudy+BPASS model emission-line ratios shown on the [N II] BPT (top row), [S II] BPT (middle row), and O_{32} versus R_{23} diagrams (bottom row). The triangle data points on the curves increase in size as $\log(U)$ increases. To avoid overcrowding, we vary $\log(U)$ in steps of 0.20 dex and have nebular metallicity range between $-0.60 \leq 12 + \log(O/H) \leq 0.20$ in 0.10 dex steps, with additional values that extend deeper into the subsolar regime included as well (-0.80, -1.00, and -1.30). We adopt an age of $10^{8.6}$ yr because it is the approximate median age of the MOSDEF sample (given a CSF history), but the results are not sensitive to age. Also plotted are the local SDSS sample (grey 2D histogram) and the high (blue points) and low (red points) MOSDEF samples with binned medians as shown in Fig. 3. The green and yellow squares, with associated uncertainties, are the median values of the high and low populations.

is a reasonable assumption given that the high sample has a lower median stellar metallicity (i.e. a harder median ionizing spectrum and greater ionizing photon production efficiency) in addition to more concentrated star formation (i.e. higher Σ_{SFR}) compared to the low sample. The high population also appears to be more α-enhanced, due to it having a lower Z_* but a higher $12+\log(O/H)$ compared to the low population. Based on the variation of best-fitting $12+\log(O/H)$

and $\log(U)$ values between the high and low median data points, we conclude that quantifying both of these physical parameters, in addition to stellar metallicity (i.e. the hardness of the ionizing spectrum), is important when explaining the distribution of high-redshift galaxies in the [N II] BPT diagram. It is worth noting that we have shown the high population to be less massive (i.e. smaller M_*) but more metal-rich (i.e. larger $12+\log(O/H)$) compared to the low

Table 2. Column (1): physical property of the sample. Column (2): median value with uncertainty of the high sample. Column (3): median value with uncertainty of the low sample. For Columns (2) and (3), the Z_* values were selected and were not obtained from fitting the data. Also, the more finely spaced grids that vary 12+log(O/H) and log(U) in steps of 0.02 dex were used.

Single Z_* Cloudy+BPASS median values for physical properties of the *high* and *low* samples		
Physical property (1)	*High* median (2)	*Low* median (3)
Median points		
Z_*	0.001	0.002
12+log(O/H)	$8.37^{+0.04}_{-0.04}$	$8.31^{+0.04}_{-0.02}$
log(U)	$-2.88^{+0.04}_{-0.04}$	$-3.08^{+0.02}_{-0.04}$
Large-N2		
Z_*	0.002	0.003
12+log(O/H)	$8.49^{+0.02}_{-0.04}$	$8.41^{+0.02}_{-0.02}$
log(U)	$-3.04^{+0.04}_{-0.04}$	$-3.26^{+0.04}_{-0.04}$
Small-N2		
Z_*	0.001	0.001
12+log(O/H)	$8.31^{+0.04}_{-0.04}$	$8.23^{+0.02}_{-0.02}$
log(U)	$-2.60^{+0.06}_{-0.06}$	$-2.98^{+0.02}_{-0.02}$

population. This combination of physical properties is the opposite of what is expected based on results from previous studies investigating the mass–metallicity relationship (e.g. Tremonti et al. 2004; Steidel et al. 2014; Sanders et al. 2018, 2020a), which show that M_* and 12+log(O/H) have a positive relationship. However, the differences in mass and metallicity between the high and low samples reflect the scatter in the mass–metallicity relationship.

To check the validity of our assumption that the N/O ratio is consistent between the high and low populations, we estimate log(N/O) for the sample median points based on the [N II]λ6585/[O II]$\lambda\lambda$3727,3730 tracer. We use the calibration from Strom et al. (2018) to convert from log ([N II] λ6585 / [O II] $\lambda\lambda$3727, 3730) to log(N/O), and find that log(N/O) = -1.13 ± 0.05 and -1.08 ± 0.04 for the high and low samples, respectively. The consistency in log (N/O) between the high and low samples validates our Cloudy+BPASS model assumptions and rules out N/O variations as the primary driver of the [N II] BPT offset. This conclusion agrees with results from other current MOSDEF studies (e.g. Shapley et al. 2019; Sanders et al. 2020a, b; Topping et al. 2020a) as well as other z ∼ 2 studies (e.g. KBSS; Steidel et al. 2014; Strom et al. 2017, 2018).

For the most part, these results are in agreement with Topping et al. (2020b). The best-fitting 12+log(O/H) and log(U) values from both studies are in agreement within 1σ, except for the log(U) of the two high populations, which agree within 2σ. However, this study finds a larger difference in log(U) between the high and low samples, 0.20 dex, compared to Topping et al. (2020b), 0.07 dex. On the other hand, Topping et al. (2020b) find a larger variance in 12+log(O/H) between the two populations, 0.10 dex, compared to this study, 0.06 dex. These differences are due to the slightly different median rest-optical line ratios for the high and low populations in Topping et al. (2020b) and in the current work.

To characterize the variation of physical properties within each of our high and low populations, we apply the same methodology of finding the best-fitting 12+log(O/H) and log(U) values, but in this case for the two endpoints of the binned medians for the high and low samples in the [N II] BPT diagram. We use the same binning method for the [N II] BPT diagram (i.e. four equally sized bins based on O3N2 strength; see Section 3.2 above for complete details on binning) when calculating the median O3, N2, S2, and O_{32}

emission-line ratios for the endpoints of both binned median lines. Uncertainties for the line ratios of the median endpoints are estimated using the same methodology of bootstrap resampling coupled with perturbation of the data points of individual galaxies according to their error bars, as described above. On the [N II] BPT diagram, the binned medians move primarily in the N2 direction, therefore, we refer to the endpoints of both binned median lines as 'large-N2' and 'small-N2'. It is important to note that the large-N2 endpoints of both high and low binned median lines have larger S2, and smaller O_{32} compared to the small-N2 endpoints. In other words, while the binning method was not the same on all three diagrams in Fig. 3 (i.e. the four bins do not necessarily contain the same galaxies), the large-N2 bins on the [N II] BPT diagram have significant overlap with (and therefore roughly correspond to) the larger-S2, and smaller-O_{32} endpoints of the binned median lines on the [S II] BPT and O_{32} versus R_{23} diagrams.

The results of Topping et al. (2020a) suggest that a positive correlation exists between N2 and Z_* (i.e. stellar metallicity tends to be higher at larger N2 values). Therefore, we vary the Z_* values that we assume for the large-N2 and small-N2 endpoints of the binned median lines accordingly. For the low population, we assume stellar metallicities of 0.003 and 0.001 (i.e. ±0.001 from the median Z_*) for the large-N2 and small-N2 points, respectively. Similarly, for the high population, we assume stellar metallicity values of 0.002 and 0.001 for the large-N2 and small-N2 points, respectively. We do not lower the stellar metallicity of the small-N2 point of the high population to our next available Z_* model value (0.0001, a factor of 10 lower than the next highest Z_* value) because the results from Topping et al. (2020a) do not suggest that such a significant variation in stellar metallicity is observed. It is important to note that while we do vary Z_* in accordance with the results from Topping et al. (2020b), the best-fitting 12+log(O/H) and log(U) for the large-N2 and small-N2 endpoints of the high and low populations do not change significantly if we use $Z_* = 0.001$ and 0.002 (i.e. the Z_* values used for the median high and low population data points). We will discuss this systematic effect of how our choice in Z_* influences the best-fitting 12+log(O/H) and log(U) in more detail below in Sections 4.2.3 and 4.2.4.

Median 12+log(O/H) and log(U) values for the large-N2 and small-N2 endpoints of the high and low populations are included in Table 2. The corner plots showing the distributions of 12+log(O/H) and log(U) for the large-N2 and small-N2 endpoints are shown in the top right- and bottom-hand panels, respectively, of Fig. 6. In both samples, we find an anticorrelation between 12+log(O/H) and log(U) along the binned median lines. Moving from large-N2 to small-N2 (therefore also large-S2 to small-S2 and small-O_{32} to large-O_{32}) leads to a lower 12+log(O/H) but a higher log(U). This anticorrelation between log(U) and 12+log(O/H) has been observed in local star-forming galaxies and H II regions (Pérez-Montero 2014). In addition, this trend is consistent with the idea that the high O_{32} tail of the local sequence on the O_{32} versus R_{23} diagram includes galaxies with low metallicity, and high ionization parameter (Shapley et al. 2015). Also, on the [N II] BPT diagram both stellar and nebular metallicity decrease along the local sequence (i.e. from low O3 and high N2 to high O3 and low N2).

4.2.3 Variable Z_* models

As described above, thus far we have assumed Z_* values of 0.001 and 0.002 for the high and low population median points to match the best-fitting values from modelling the rest-frame UV spectra
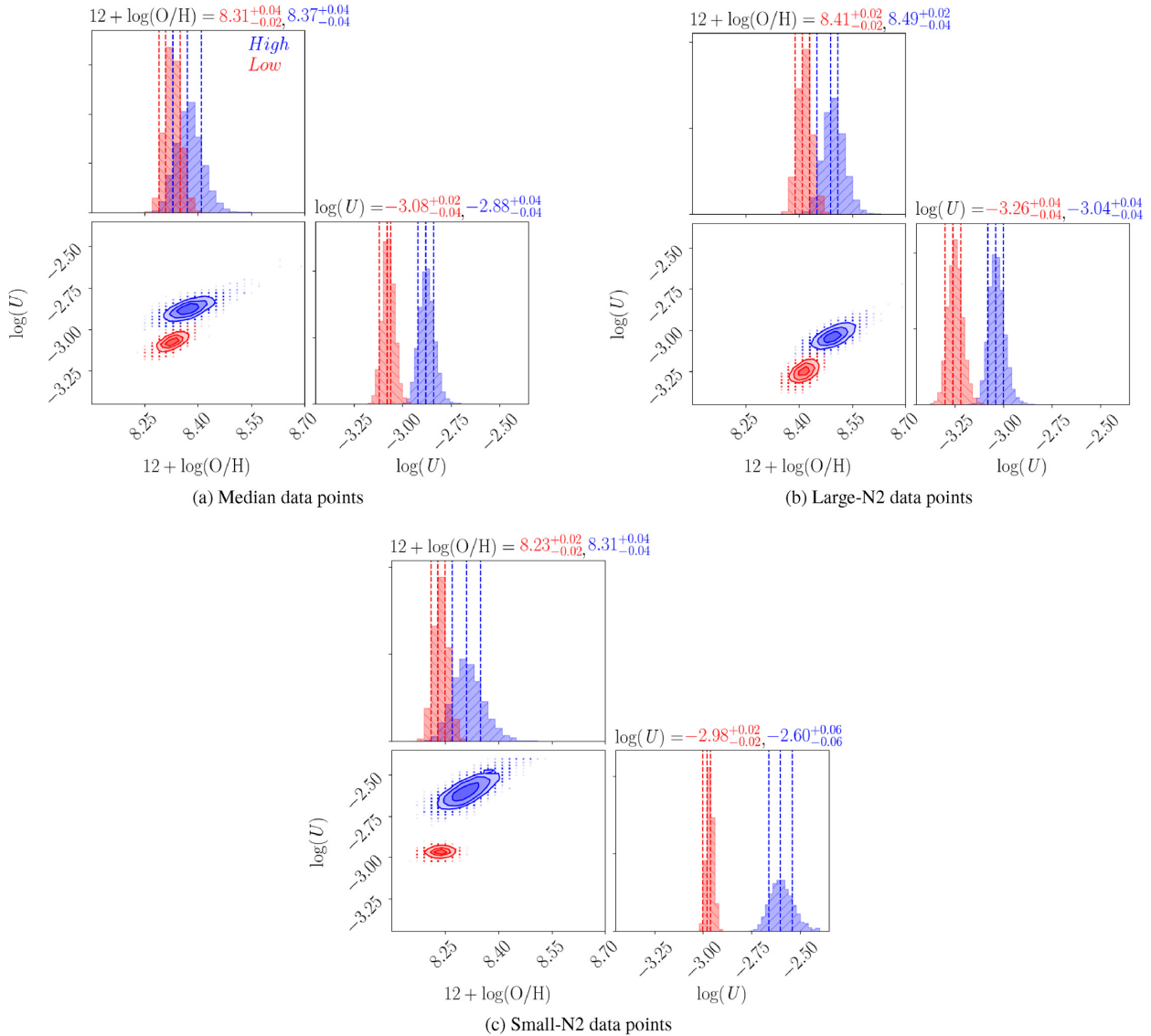
Figure 6. Corner plots comparing the distributions of 12+log(O/H) and log(U) for the high (blue) and low (red) MOSDEF populations. The distributions for the sample median points, i.e. green and yellow squares in Fig. 5, (upper left-hand panel), large-N2 endpoint of the sample binned medians (top right-hand panel) and small-N2 endpoint of the sample binned medians (bottom centre) are included. We assume Z_* values of 0.001 and 0.002 (median points; top left-hand panel), 0.002 and 0.003 (large-N2 points; top right-hand panel), and 0.001 and 0.001 (low-N2 points; bottom panel) for the high and low populations, respectively. The dashed lines on the histograms mark the median value and 1σ uncertainties (16th and 84th percentiles) for the distributions of 12+log(O/H) and log(U), and the three contours on the diagonal mark the 1σ, 1.5σ, and 2σ regions in the 2D 12+log(O/H)–log(U) parameter space. It is shown that moving from large- to small-N2 corresponds to decreasing nebular metallicity and increasing ionization parameter. In each of the three distributions, the high population has a higher 12+log(O/H) and log(U).

of $z \sim 2.3$ MOSDEF galaxies (Topping et al. 2020b). While this is a reasonable assumption, we also investigate how relaxing our requirement on stellar metallicity affects the best-fitting nebular metallicity and ionization parameter values of the two populations. For this additional analysis, we use the same χ^2 method described above (i.e. fitting the N2, S2, O3, and O$_{32}$ emission-line ratios). However, we not only fit for 12+log(O/H) and log(U), but also treat Z_* as a free parameter. We investigate two different model grids: one that allows a small range of Z_* for the high and low samples, including the best-fitting stellar metallicities from Topping et al. (2020b) for each population and those that are adjacent to the best-fitting values, and a fully unrestricted method that allows all Z_* values used by the BPASS models ranging from 0.00001 to 0.03.

This analysis shows that when we relax the constraints provided by rest-UV spectra (both with limited freedom and in a fully unconstrained manner), the rest-optical emission lines tend to favour higher stellar metallicities. When we allow for limited freedom in Z_* the stellar metallicities for the high and low samples prefer the largest allowed Z_* – i.e. 0.002 and 0.003, respectively. When all stellar metallicity values are allowed, the high and low populations favour Z_* values about four to five times greater than those found by Topping et al. (2020b) based on fitting rest-UV spectra. These shifts demonstrate the importance of imposing external constraints from a more direct probe of massive stars and stellar metallicity, i.e. the rest-UV continuum, as opposed to relying only on a joint fit of rest-optical emission lines. We also note that, while the best-fitting

values of Z_* increase when stellar metallicity is allowed to vary, the low population still always favours a higher Z_* than the high population.

The best-fitting 12+log(O/H) and log(U) values for the single-Z_* and limited freedom Z_* models are significantly consistent (i.e. within 1σ) for both the high and low populations. For the high sample, the best-fitting values for the single-Z_* and fully unconstrained Z_* models are significantly consistent as well. For the low sample, only some of the best-fitting values for the single-Z_* and fully unconstrained Z_* models are consistent within 1σ; however, all values are consistent within 2σ. In addition, similar to what we find for the single-Z_* models, we find that O/H increases and log(U) decreases with increasing N2 in both free-Z_* models.

The results from this section provide guidance on which wavelength regimes best constrain different galaxy properties. Supported by other studies (e.g. Topping et al. 2020b), rest-UV spectra are essential for accurately estimating Z_*. Rest-optical spectra, even with the combination of multiple emission-line ratios, cannot constrain the stellar metallicity. At high redshift, we have shown that modelling rest-optical spectra alone will lead to overestimates of stellar metallicity relative to more accurate methods that incorporate rest-UV spectral information. However, rest-optical spectra can be used to constrain 12+log(O/H) and log(U). Because the best-fitting values of nebular metallicity and ionization parameter vary based on which rest-optical emission-line diagram is used (see Fig. 5), the combination of line ratios from multiple diagrams will increase the accuracy of the models. These parameters can even be estimated with reasonable precision without detailed knowledge of the ionizing spectrum (i.e. the metallicity of the stellar population).

4.2.4 The Importance of α-enhancement

In summary, using rest-optical emission-line ratios alone without the constraints from fitting rest-UV spectra, we infer systematically higher Z_* values for both high and low samples. At the same time, when multiple rest-optical emission lines ratios are measured (N2, S2, O3, and O_{32}), the inferred median nebular parameters (12+log(O/H) and log(U)) and their variation across the BPT diagram, do not depend strongly on the allowed range of Z_*. However, the inclusion of constraints on Z_* from rest-UV fitting is essential for our understanding of the abundance patterns (i.e. α-enhancement) of $z \sim 2.3$ galaxies and we adopt the constrained values of Topping et al. (2020b) for our fiducial modelling procedure (Table 2). Because we see variation in Z_*, 12+log(O/H), and log(U) at the median and large- and small-N2 endpoints for the high and low samples, we conclude that constraints on all three of these physical properties are required for fully understanding the observed distribution of $z \sim 2.3$ star-forming galaxies in the [N II] BPT diagram. The importance of α-enhancement has also been highlighted by Steidel et al. (2016) and Cullen et al. (2019).

It is also important to note that while the high sample is more α-enhanced with a harder ionizing spectrum compared to the low sample, both populations are α-enhanced and both are characterized by harder ionizing spectra when compared with their equivalents (i.e. galaxies with similar rest-optical line ratios) at low redshift. Notably, such physical differences apply to the low population, even though it overlaps with the local [N II] BPT sequence. Specifically, an overlap in emission-line ratios does not correspond to the same properties for the ionizing radiation field and ionized gas. Therefore, we recommend determining the location of galaxies in multiple rest-optical emission-line diagrams, as well as having some knowledge

of the ionizing spectrum, when trying to identify local systems that can be used as true $z \sim 2$ analogues.

In addition, the differences between the high and low populations found in both this section and in Section 3.2.2 suggest that $z \sim 2.3$ galaxies in the high sample that generally have a lower Z_* with higher 12+log(O/H) and log(U) values are also typically smaller with higher sSFR and Σ_{SFR} values. Looking ahead, we must understand why there exists a connection between the local properties of ionized star-forming regions and the overall compactness/intensity of star formation. However, exploring this phenomenon is outside the scope of the current work.

4.2.5 Discrepancies between data and models

It is worth noting that in all three diagrams in Fig. 5, there are galaxies that do not lie within the Cloudy+BPASS grid space. On the [N II] BPT and [S II] BPT diagrams, there are galaxies with O3 values that exceed the extremes of the model grids. Similarly, on the O_{32} versus R_{23} diagram, we observe MOSDEF galaxies with higher R_{23} values than what the Cloudy+BPASS models predict. These offsets are primarily found among the high population. Such extreme galaxies are not specific to this study, as Strom et al. (2017) find similar discrepancies between data and models in the corresponding emission-line diagnostic diagrams.

For the Cloudy+BPASS model grids in this study, two possible solutions could be to either elevate n_{e} or lower Z_*. Increasing n_{e} has been shown to cause elevated O3 and N2 values (Kewley et al. 2013). Since the high population (defined by its elevated O3 and/or N2 values) also has elevated R_{23} values on average, increasing n_{e} in our Cloudy+PASS models theoretically could push the curves to higher R_{23}. The models in this study, which were constructed by Topping et al. (2020b), set $n_{\mathrm{e}} = 250$ cm^{-3}, which is typical of $z \sim 2.3$ galaxies in both the KBSS sample (Strom et al. 2017) and early MOSDEF work (Sanders et al. 2016). It is also comparable with the median electron densities reported in this work. The models employed in Strom et al. (2017) use $n_{\mathrm{e}} = 300$ cm^{-3}, and similarly cannot reach the highest observed O3 and R_{23} values in the KBSS survey.

Using $Z_* = 0.001$, we explore raising n_{e} to 1000 cm^{-3} in the Cloudy+BPASS models (right-hand column of Fig. 7). Comparing the elevated electron density models with the $Z_* = 0.001$, $n_{\mathrm{e}} = 250$ cm^{-3} scenario (left-hand column of Fig. 5) shows the differences in the predicted emission-line ratios caused by isolating variations in n_{e} on the [N II] BPT, [S II] BPT, and O_{32} versus R_{23} diagrams. We find that elevating n_{e} to such an extreme value has only a small effect on the emission-line ratios, and the models still do not encompass the galaxies observed at the highest N2, S2, O3, and R_{23} values.

A more promising approach is to lower the stellar metallicity. The left-hand column of Fig. 7 gives the model grids for $Z_* = 0.0001$, which reach the majority of the extreme data points. Therefore, a very low Z_* (i.e. a very hard ionizing spectrum) is the more likely than variations in n_{e} to be cause for these outliers.

In addition to a low Z_*, other effects may also lead to a harder ionizing spectrum (e.g. variations in the IMF slope and high-mass cutoff or AGN partially contributing to the integrated emission lines). Systematic uncertainties such as the BPASS models underpredicting the hardness of the ionizing spectrum at a given Z_* (particularly at lower values of Z_*), different star formation histories, or uncertainties in the dust corrections applied to O_{32} and R_{23} could play a role in the observed discrepancies between models and a minority of the sample as well.
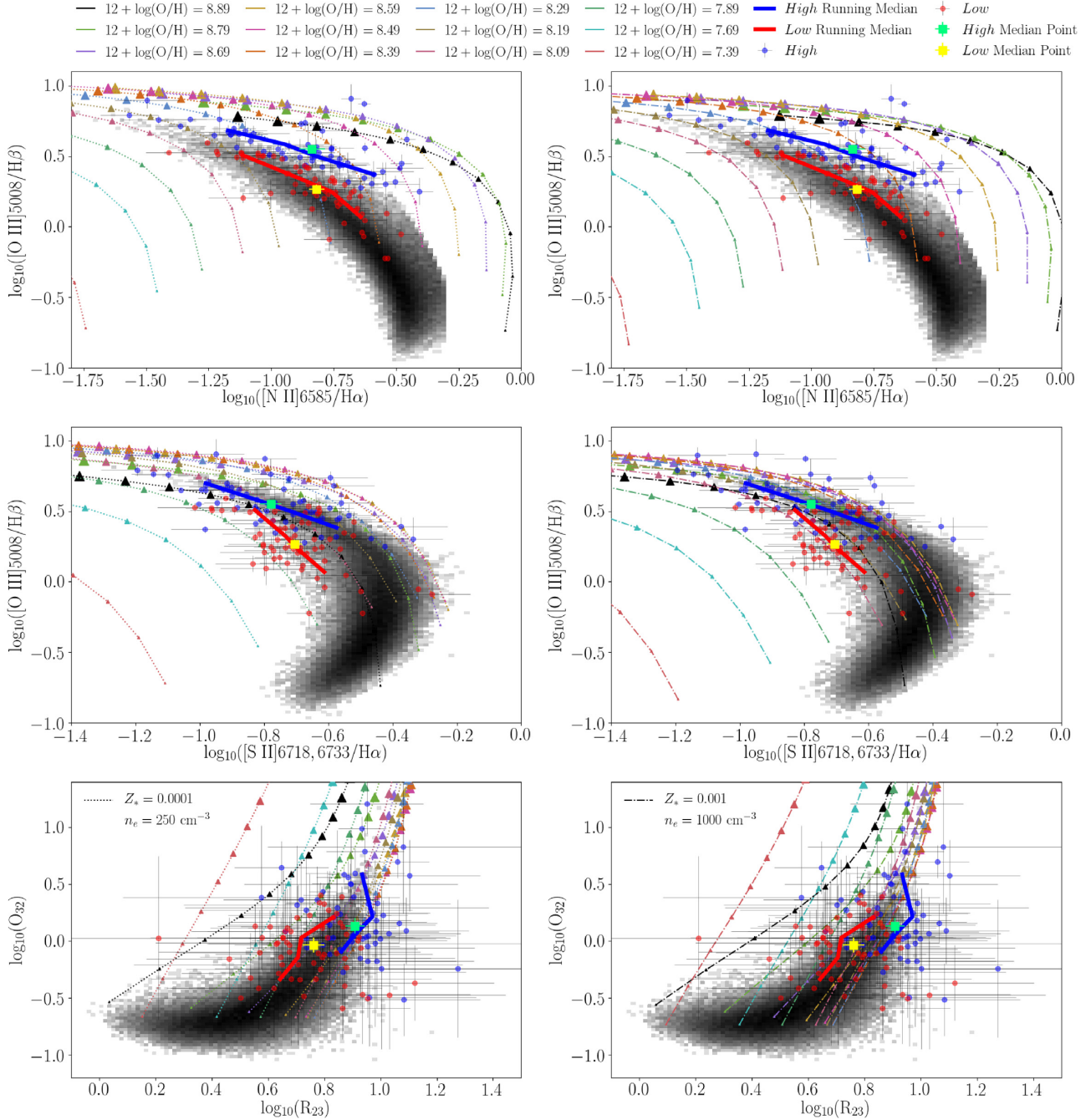
Figure 7. Left-hand column: $Z_* = 0.0001$, $n_e = 250\,\text{cm}^{-3}$ (dotted lines). Right-hand column: $Z_* = 0.001$, $n_e = 1000\,\text{cm}^{-3}$ (dot-dashed lines). Cloudy+BPASS model emission-line ratios shown on the [N II] BPT (top row), [S II] BPT (middle row), and O_{32} versus R_{23} diagrams (bottom row). A comparison of these models with the corresponding panels in Fig. 5 shows how raising the electron density and lowering stellar metallicity affects the predicted emission-line ratios. The correspondence between colours and symbol size and, respectively, 12+log(O/H) and log(U), matches that in Fig. 5. Once again, we adopt an age of $10^{8.6}$ yr because it is the approximate median age of the MOSDEF sample (given a CSF history) and include the local SDSS sample (grey 2D histogram), the high (blue) sample, and low (red) sample with binned medians as shown in Fig. 3. The green and yellow squares, with associated uncertainties, are the median values of the high and low populations as shown in Fig. 5. Aside from at roughly solar and supersolar nebular metallicities, a higher electron density has a minimal affect on the Cloudy+BPASS model grids and cannot reach the elevated O3 and O_{32} galaxies. A lower stellar metallicity is able to reach these values, suggesting that the most offset galaxies have the hardest ionizing spectrum.

5 SUMMARY

We present results on the emission-line properties of a sample of 122 star-forming galaxies at $1.9 \leq z \leq 2.7$ from the complete MOSDEF survey with $\geq 3\sigma$ detections for the [O II]$\lambda\lambda3727,3730$, Hβ, [O III]$\lambda\lambda4960,5008$, Hα, [N II]$\lambda6585$, and [S II]$\lambda\lambda6718,6733$

emission lines. To investigate the observed systematic offset of $z > 1$ star-forming galaxies on the [N II] BPT diagram relative to local systems, the MOSDEF sample is split into the high (offset with elevated O3 and/or N2 values) and low (overlapping with the local SDSS sequence) samples. We compare the location of galaxies in

both populations on the [S II] BPT and O_{32} versus R_{23} diagrams. Additionally, we compare physical properties − SFR, M_*, R_e, sSFR, Σ_{SFR}, n_e, and t (i.e. stellar population age) − of the galaxies in both groups using additional CANDELS and 3D-*HST* ancillary data. Finally, we use Cloudy+BPASS photoionization models to investigate potential differences in the stellar metallicity, nebular metallicity, and ionization parameter of the high and low groups.

The main results are as follows:

(i) On the [S II] BPT and O_{32} versus R_{23} diagrams, the high sample is offset on average towards higher S2 at fixed O3 (or vice versa) and higher R_{23} value at fixed O_{32}, respectively, relative to the low sample. These results update earlier work from MOSDEF (Sanders et al. 2016) and are consistent with results from Strom et al. (2017), based on the KBSS survey.

(ii) The high sample has a smaller median R_e and M_*, but higher median sSFR and Σ_{SFR} compared to the low sample. The observed differences in R_e, sSFR, and Σ_{SFR} are present at fixed stellar mass. The high sample is also slightly younger in median age (t); however, this age difference is not significant within the uncertainties. There is not a significant variation in SFR or n_e for the two populations. These results imply that the high population is associated with more concentrated and intense star formation than the low population.

(iii) Using photoionization models, we find that the high population has larger 12+log(O/H) and log(U) values compared to the low population. These conclusions hold both when we use the results from Topping et al. (2020b) and assume Z_* values of 0.001 and 0.002 for the high and low samples, respectively, as well as when we treat Z_* as a free parameter. Also, while the lower Z_*/Z_{neb} ratio of the high population implies that this sample is more α-enhanced than the low population, both samples are significantly α-enhanced relative to local star-forming galaxies with similar emission-line ratios assuming typical local galaxies have roughly solar α/Fe. These differences in the properties of the ionizing radiation field are critical to include when describing high-redshift galaxies – even those that overlap the local emission-line sequences.

(iv) Combining the results from the median galaxy properties of the high and low samples with the Cloudy+BPASS model grids leads to the conclusion that galaxies with a harder ionizing spectrum are associated with smaller sizes and higher sSFR and Σ_{SFR}.

(v) A harder ionizing spectrum at fixed nebular metallicity is favoured as one of the key drivers of the [N II] BPT offset. In this study, we find that variation in multiple physical parameters drives the variation in emission-lines in the MOSDEF sample, including Z_*, 12+log(O/H), and log(U). However, even in regions where the $z \sim 2.3$ galaxies overlap local ones in the space of emission-line ratios, the inferred underlying physical parameters for $z \sim 2.3$ galaxies are distinct from those of local galaxies. The full set of these parameters must be understood through modelling both rest-UV and rest-optical spectra, in order to understand the translation between empirical emission-line ratios and key physical quantities such as the gas-phase oxygen abundance.

Understanding the global properties of high-redshift galaxies using large samples at $z \sim 2.3$ is essential when comparing them with local galaxies. The use of large high-redshift samples has enabled us to discover more subtle trends in the MOSDEF data that were not found with more limited data sets (e.g. Shapley et al. 2015; Sanders et al. 2016). The work from this study has shown that galaxies with a harder ionizing spectrum also tend to be smaller in size with more intense and compact star formation. Why these factors are linked is not yet clear, and combining large $z \sim 2.3$ samples of star-forming

galaxies with realistic galaxy formation simulations (e.g. FIRE-2; Hopkins et al. 2018) will be important in finding this connection.

It has also been shown in both this work and in Strom et al. (2017) that models are not yet able to reach the highest observed O3 and R_{23} values for $z \sim 2.3$ star-forming galaxies on the [N II] BPT, [S II] BPT, and O_{32} versus R_{23} diagrams. The local sequence does not reach such elevated O3 and R_{23} values, and unlike their high-redshift counterparts, do not greatly exceed the data space covered by the model grids. This discrepancy indicates a need to develop both stellar population and emission-line models that better match the now large data sets of $z \sim 2.3$ star-forming galaxies.

DATA AVAILABILITY

The data underlying this paper will be shared on a reasonable request to the corresponding author.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Andrews B. H., Martini P., 2013, ApJ, 765, 140
Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Astropy Collaboration et al., 2013, A&A, 558, A33
Astropy Collaboration et al., 2018, AJ, 156, 123
Azadi M. et al., 2017, ApJ, 835, 27
Baldwin J. A., Phillips M. M., Terlevich R., 1981, Publ. Astron. Soc. Japan, 93, 5
Bian F., Kewley L. J., Groves B., Dopita M. A., 2020, MNRAS, 493, 580
Brinchmann J., Pettini M., Charlot S., 2008, MNRAS, 385, 769
Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245

[2]http://www.astropy.org

Chabrier G., 2003, Publ. Astron. Soc. Pac., 115, 763
Coil A. L. et al., 2015, ApJ, 801, 35
Cullen F. et al., 2019, MNRAS, 487, 2038
Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, Publ. Astron. Soc. Pac. , 34, e058
Erb D. K., Steidel C. C., Shapley A. E., Pettini M., Reddy N. A., Adelberger K. L., 2006, ApJ, 647, 128
Ferland G. J. et al., 2017, Rev. Mex. Astron. Astrofis., 53, 385
Foreman-Mackey D., 2016, J. Open Source Softw., 1, 24
Freeman W. R. et al., 2019, ApJ, 873, 102
Grogin N. A. et al., 2011, ApJS, 197, 35
Hao C.-N., Kennicutt R. C., Johnson B. D., Calzetti D., Dale D. A., Moustakas J., 2011, ApJ, 741, 124
Harshan A. et al., 2020, ApJ, 892, 77
Hopkins P. F. et al., 2018, MNRAS, 480, 800
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Juneau S. et al., 2014, ApJ, 788, 88
Kashino D. et al., 2017, ApJ, 835, 88
Kashino D. et al., 2019, ApJS, 241, 10
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
Koekemoer A. M. et al., 2011, ApJS, 197, 36
Kriek M. et al., 2015, ApJS, 218, 15
Kriek M., van Dokkum P. G., Labbé I., Franx M., Illingworth G. D., Marchesini D., Quadri R. F., 2009, ApJ, 700, 221
Leung G. C. K. et al., 2017, ApJ, 849, 48
Lilly S. J., Carollo C. M., Stockton A. N., 2003, ApJ, 597, 730
Liu X., Shapley A. E., Coil A. L., Brinchmann J., Ma C.-P., 2008, ApJ, 678, 758
McLean I. S. et al., 2012, in Ground-based and Airborne Instrumentation for Astronomy IV, Proc. SPIE, Vol. 8446. p. 84460J
Madau P., Dickinson M., 2014, ARA&A, 52, 415
Masters D. et al., 2014, ApJ, 785, 153
Masters D., Faisst A., Capak P., 2016, ApJ, 828, 18
Millman K. J., Aivazis M., 2011, Comput. Sci. Eng., 13, 9
Momcheva I. G. et al., 2016, ApJS, 225, 27
Nakajima K., Ouchi M., Shimasaku K., Hashimoto T., Ono Y., Lee J. C., 2013, ApJ, 769, 3

Oliphant T., 2006, NumPy: A Guide to NumPy. Trelgol Publishing, USA
Oliphant T. E., 2007, Comput. Sci. Eng., 9, 10
Perez F., Granger B. E., 2007, Comput. Sci. Eng., 9, 21
Pérez-Montero E., 2014, MNRAS, 441, 2663
Pettini M., Pagel B. E. J., 2004, MNRAS, 348, L59
Pilyugin L. S., Vílchez J. M., Mattsson L., Thuan T. X., 2012, MNRAS, 421, 1624
Reddy N. A. et al., 2015, ApJ, 806, 259
Sanders R. L. et al., 2016, ApJ, 816, 23
Sanders R. L. et al., 2018, ApJ, 858, 99
Sanders R. L. et al., 2020a, MNRAS, 491, 1427
Sanders R. L. et al., 2020b, ApJ, 888, L11
Shapley A. E. et al., 2015, ApJ, 801, 88
Shapley A. E. et al., 2019, ApJ, 881, L35
Shapley A. E., Coil A. L., Ma C.-P., Bundy K., 2005, ApJ, 635, 1006
Shivaei I. et al., 2015, ApJ, 815, 98
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
Topping M. W., Shapley A. E., Reddy N. A., Sanders R. L., Coil A. L., Kriek M., Mobasher B., Siana B., 2020a, MNRAS, 499, 1652
Topping M. W., Shapley A. E., Reddy N. A., Sanders R. L., Coil A. L., Kriek M., Mobasher B., Siana B., 2020b, MNRAS, 495, 4430
Tremonti C. A. et al., 2004, ApJ, 613, 898
van der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
van der Wel A. et al., 2014, ApJ, 788, 28
Veilleux S., Osterbrock D. E., 1987, ApJS, 63, 295
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Wright S. A., Larkin J. E., Graham J. R., Ma C.-P., 2010, ApJ, 711, 1291
Yeh S. C. C., Verdolini S., Krumholz M. R., Matzner C. D., Tielens A. G. G. M., 2013, ApJ, 769, 11
York D. G. et al., 2000, AJ, 120, 1579

This paper has been typeset from a TEX/LATEX file prepared by the author.