The Origin of the Observed Lyman alpha EW Distribution of Dwarf Galaxies at z 2

C. Snapp-Kolas, B. Siana, T. Gburek, A. Alavi, N. Emami, J. Richard, D.P. Stark, C. Scarlata

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **525,** 5500–5511 (2023) https://doi.org/10.1093/mnras/stad2412
Advance Access publication 2023 September 12

The origin of the observed Ly α EW distribution of dwarf galaxies at $z \sim 2$

Christopher Snapp-Kolas [1]★ Brian Siana,[1]★ Timothy Gburek,[1] Anahita Alavi,[2] Najmeh Emami,[3] Johan Richard [4], Daniel P. Stark,[5] Claudia Scarlata [6], Brian C. Lemaux [7,8] Giovanni Zamorani,[9] Sandro Bardelli,[9] Nimish Hathi [10] and Paolo Cassata[11,12]

[1]*Department of Physics & Astronomy,* University *of California, Riverside, CA 92521, USA*
[2]*IPAC, California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA*
[3]*Minnesota Institute for Astrophysics, University of Minnesota, Minneapolis, MN 55455, USA*
[4]*Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon, UMR5574, F-69230 Saint-Genis-Laval, France*
[5]*Steward Observatory, University of Arizona, 933 N Cherry Ave, Tucson, AZ 85721, USA*
[6]*Minnesota Institute for Astrophysics, University of Minnesota, Minneapolis, MN 55455, USA*
[7]*Gemini Observatory, NSF's NOIRLab, 670 N. A'ohoku Place, Hilo, HI 96720, USA*
[8]*Department of Physics and Astronomy, University of California, Davis, One Shields Ave., Davis, CA 95616, USA*
[9]*INAF – Astrophysics and Space Science Observatory Bologna, Via Piero Gobetti 93/3, I-40129 Bologna, Italy*
[10]*Space Telescope Science Institute, Baltimore, MD 21218, USA*
[11]*Dipartimento di Fisica e Astronomia Galileo Galilei, Università degli Studi di Padova Vicolo del L'Osservatorio 3, I-35122 Padova, Italy*
[12]*INAF Osservatorio Astronomico di Padova, vicolo dell'Osservatorio 5, I-35122 Padova, Italy*

Accepted 2023 July 11. Received 2023 June 20; in original form 2022 October 31

ABSTRACT

We present a rest-UV-selected sample of 32 lensed galaxies at $z \sim 2$ observed with joint Keck/LRIS rest-UV and Keck/MOSFIRE rest-optical spectra behind the clusters Abell 1689, MACS J0717, and MACS J1149. The sample pushes towards the faintest UV luminosities observed ($-19 \leq M_{\rm UV} \leq -17$) at this redshift. The fraction of dwarf galaxies identified as Ly α emitters (EW $\geq$ 20 Å) is $X_{\rm LAE} = 33^{+17}_{-12}$ per cent. We use the Balmer lines and UV continuum to estimate both the *intrinsic* EW and the Ly α escape fraction to determine their relative importance in dictating the *observed* EW distribution. Fainter galaxies ($M_{\rm UV} > -19$) show larger intrinsic EWs and escape fractions than brighter galaxies. Only galaxies with intrinsic EWs greater than 40 Å have escape fractions larger than 5 per cent. We find little to no anticorrelation between the escape fraction and dust in dwarf galaxies. The volumetric escape fraction of our sample is $f_{\rm esc}^{\rm Ly\alpha} = 4.6^{+2.0}_{-1.4}$ per cent in agreement with measurements found elsewhere in the literature. 60 per cent of the total integrated Ly α luminosity density comes from galaxies with EW$_{\rm obs} > 20$ Å.

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

1 INTRODUCTION

More than five decades ago, Ly α was proposed as a tracer of early stage star-forming galaxies (Partridge & Peebles 1967). Since then Ly α has been used to study star-forming galaxies locally and at the highest redshifts observed (e.g. Shapley et al. 2003; Jung et al. 2020). It has been used to study reionization (Bolton & Haehnelt 2013; Mesinger et al. 2015; Mason et al. 2018; Matthee et al. 2021), the interstellar medium (ISM, Du et al. 2021), the circumgalactic medium (Matsuda et al. 2012; Hayes et al. 2014), and has confirmed the redshifts of distant star-forming galaxies (Caruana et al. 2014; Hoag et al. 2019; Endsley et al. 2020; Jung et al. 2020).

Ly α is a resonant line and has a large cross-section. Therefore, it will be scattered even by low column density gas. This makes it a great tracer of gas distributions, but can result in difficulties in interpreting observations. Furthermore, Ly α can be easily absorbed by dust in the interstellar medium (ISM) reducing the strength of the line. Ly α observations are even more difficult when we try to observe galaxies during or before the epoch of reionization where there are large amounts of neutral hydrogen in the intergalactic medium (IGM) which can scatter whatever Ly α manages to escape. Each of these effects can drastically hamper the observability of Ly α and have strong effects on the distribution of observed Ly α equivalent widths (EWs, Stark et al. 2010; Stark, Ellis & Ouchi 2011). In fact, at high redshift the attenuation of Ly α can be a useful probe of the neutral fraction of hydrogen in the intergalactic medium. But in order to measure the attenuation, one must first know the *intrinsic* distribution of EW$_{\rm Ly\alpha}$ to compare with the observed distribution. The evolution of the EW$_{\rm Ly\alpha}$ distribution has been used to study the end of reionization and is well understood from $z = 3$–6 and for 'bright' galaxies with absolute UV magnitudes $M_{\rm UV} < -19$ (Stark et al. 2010, 2011; Pentericci et al. 2011; Curtis-Lake et al. 2012; Ono et al. 2012; Schenker et al. 2012; Cassata et al. 2015; De Barros et al. 2017; Caruana et al. 2018; Haro et al. 2018; De La Vieuville et al. 2020; Kusakabe et al. 2020; Zhang et al. 2021). At $z = 2$ there has been some study of the brighter star-forming galaxies ($M_{\rm UV} < -19$) (Reddy & Steidel 2009; Cassata et al. 2015; Hathi et al. 2016; Du et al. 2021; Zhang et al. 2021), but little work has been done to study the fainter population ($M_{\rm UV} > -19$). However, according to

★ E-mail: csnappkolas@gmail.com (CS-K); bsiana@ucr.edu (BS)

the luminosity functions of Alavi et al. (2016), Konno et al. (2016), and Bouwens et al. (2022) the faint galaxy population is larger at $z = 2$ and becomes more numerous with redshift. Therefore, if we want to understand reionization we need more information on the faint galaxy population, and $z = 2$ galaxies can help to constrain the $EW_{Ly\alpha}$ distribution at higher redshift where observation of Ly α is more difficult.

Ly α studies will often select galaxies via their rest-UV continuum luminosity density (via broad-band imaging), or their Ly α line emission (via narrow-band imaging, Berry et al. 2012; Konno et al. 2016; Hashimoto et al. 2017; Sobral et al. 2017; Sobral & Matthee 2019). Narrow-band-selected samples will be biased towards Ly α emitters (LAEs) as they typically require $EW^{rest}_{Ly\alpha} > 20$ Å for a galaxy to be detected (e.g. Gronwall et al. 2007; Guaita et al. 2010; Adams et al. 2011). Alternatively, narrow-band studies of Ly α often select on other emission lines in the optical (such as [O III]5007) (e.g. Ciardullo et al. 2014; Weiss et al. 2021). Erb et al. (2016) show that [O III]5007-selected galaxies are biased towards LAEs by comparing the fraction of LAEs (X_{LAE}) in their [O III]5007-selected sample with a sample of UV-selected galaxies. 80 per cent are found to be LAEs in their [O III]5007 sample while only 9 per cent are found to be LAEs in their UV-selected sample. This is further confirmed in Trainor et al. (2019) and Weiss et al. (2021), indicating that samples selected on nebular emission lines are biased toward Ly α emission. A sample chosen irrespective of the strength of nebular emission lines would be useful for understanding the full distribution of Ly α EWs. With an unbiased sample, it is possible to disentangle the effects of the stellar population (e.g. age and metallicity) and that of the gas and dust on the observed EW distribution. Broad-band surveys, which select galaxies independent of emission line strength, are able to probe down to $R_{AB} \sim 30$, but the follow-up spectroscopy necessary for accurate escape fraction and EW analyses limit the observations to $R_{AB} < 25.5$ (i.e. Reddy et al. 2008) in order to obtain sufficient S/N in the continuum.

The Ly α EW distribution is dependent on the escape fraction of Ly α photons. Past studies have measured the Ly α escape fraction by comparing the Ly α luminosity density with the Ha or UV continuum luminosity densities, obtained by integrating luminosity functions (Hayes et al. 2010; Ciardullo et al. 2014; Konno et al. 2016). However, comparing luminosity densities does not allow for analysis of individual galaxy escape fractions or correlations with other properties of the population. Others make use of narrow-band imaging of Ly α and one of the Balmer lines (e.g. Trainor et al. 2015) to make direct measurements of the Ly α escape fraction. Narrow-band imaging is biased towards high Ly α EW galaxies, and therefore does not represent the general population of star-forming galaxies as evidenced by the rather high average Ly α escape fractions of 30 per cent (Kornei et al. 2010; Blanc et al. 2011; Wardlow et al. 2014; Trainor et al. 2015) which differs greatly from other measures with more complete samples (e.g. 5.3 per cent, Hayes et al. 2010). Some try to model the escape fraction as a function of Ly α EW (e.g. Sobral & Matthee 2019), but this method must assume a star formation history and intrinsic ionizing continuum which are uncertain for low-mass galaxies. A direct measurement from a UV-selected sample with spectroscopic follow-up avoids these uncertainties.

What is needed is a measure of Ly α emission properties of a UV-selected sample of faint galaxies. We address this by selecting galaxies with photometric redshifts determined from HST photometry described in Alavi et al. (2014, 2016). We did not select on emission line flux and therefore the sample is unbiased towards Ly α emitters. Furthermore, by using gravitational lensing, we can study much fainter galaxies at $z = 2$. The goals of this paper are:

Table 1. List of data acquired with Keck/LRIS. In order from left to right we have the mask name used, the number of spectra obtained in each mask, the exposure time in seconds, the modified Julian date (MJD), and the seeing in arcseconds.

Mask (Name)	Spectra (Number)	Exposure time (s)	MJD	Seeing (arcsec)
A1689_1	15	18 000	55 325	0.9
A1689_3	16	12 286	55 981	0.9
Macsj0717	11	12 600	55 981	1.1
Macsj1149_1	16	5400	57 042	1.1
Macsj0717_2	8	9000	57 042	0.7
Macsj0717_1	8	9000	57 042	0.9
A1689_4	10	9000	57 042	1.4
Macsj1149_1_2	13	5400	57 043	0.8
A1689_z1_1	12	12 000	57 043	0.8
Macsj1149_2	5	4500	57 398	1.0
Macsj0717_3	8	6720	57 398	0.7
Macsj1149_3	11	4860	57 870	1.5
A1689_6	10	9600	57 870	1.0

(i) Measure the Ly α EW distribution and determine its dependence on the properties of the ionizing sources and the Ly α escape fraction.
(ii) Calculate the volumetric escape fraction of dwarf galaxies.
(iii) Identify trends in the EW distribution with absolute UV magnitude and redshift.

The remainder of the paper is organized as follows. In Section 2, we discuss the observations, data reduction, and sample selection. In Section 3, we discuss fits to continua and emission lines, SED fitting, magnification, and dust estimates. In Section 4, we present our measured values of the EW distribution and the Ly α escape fraction. In Section 5, we discuss the implications of our measurements. In Section 6, we summarize the work. We adopt a ΛCDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $h = 0.7$ throughout the paper and all magnitudes are in the AB system (Oke & Gunn 1983). All EWs are reported in the rest-frame with the convention that positive EWs indicate emission. f_{esc} refers to the Ly α escape fraction throughout this paper.

2 OBSERVATIONS, DATA REDUCTION, AND SAMPLE SELECTION

2.1 Observations and data reduction

Our sample consists of dwarf galaxies behind lensing clusters that have deep HST data and are accessible to the Keck Observatory (Abell 1689, MACS J1149, MACS J0717). Photometric redshifts were determined from HST photometry, as described in Alavi et al. (2014, 2016). Our selection criteria were photometric redshifts of $1.5 < z < 3.5$ and visual magnitudes brighter than $m_{F625W} < 26.3$. Slit allocation preference was given to galaxies with high magnification and photometric redshifts where the rest-optical lines are accessible in the YJHK atmospheric bands. In order to study physical properties (i.e. metallicity, f_{esc}, ionization parameter, etc.) of these dwarf galaxies we obtained Keck/MOSFIRE (McLean et al. 2010, 2012) spectra as described in Gburek et al. (2019). We also obtained Keck/LRIS optical spectra (Oke et al. 1995), for which the details are listed in Table 1. We observe eleven masks in three clusters with LRIS. Exposure times varied greatly (from 4500 to 18000 s) depending on the conditions and priority of objects in the mask, and seeing was typically ∼1 arcsec. We decreased read noise

by binning by 2 in the spectral direction of the CCD. We observed Macsj1149_1 on two separate nights and have named the second night Macsj1149_1_2 in Table 1. The A1689_1 mask was not targeted properly and we cannot confirm that the objects observed in the slits are those that were targeted and so we do not consider that mask in this work.

The data were reduced using a modified version of the PypeIt v1.x reduction pipeline (Prochaska et al. 2020b). This pipeline performs bias subtraction, flat fielding, cosmic ray rejection, wavelength calibration, sky subtraction, and extracts the 1D spectra. The 1D spectra are extracted using an optimal extraction b-spline fitting to the object profile along the spatial direction. The 1D spectra are then flux calibrated using the spectrum of a standard star. The 1D spectra from each frame are then combined using a weighted mean algorithm in PypeIt. We perform a final absolute flux normalization by scaling the spectra of compact bright continuum sources to the *Hubble* photometry. We take the median (per mask) of these and scale the remaining spectra according to this median. Most of our galaxies are faint and therefore may not have their continua well-detected in the 1D spectra, so the use of a median correction from bright continuum sources is necessary for the absolute calibration.

We correct for slit losses by convolving the Hubble images with a Gaussian such that point sources in the final images would have a Gaussian FWHM equal to the seeing. We then measure the total flux and the flux within the slit. We use the F435W and F475W bands of Hubble for the Abell 1689 and MACS J0717/MACS J1149 masks respectively. We apply this correction to our spectra, to obtain an estimate of the total flux, but refrain from applying the correction when displaying spectra. We do not account for possible offsets between the UV continuum and the Ly α emission line. By visual inspection of the 2D spectra we find only one galaxy in our sample with any offset and therefore expect insignificant correction to our measurements from spatial offsets (see Claeyssens et al. 2022, for a discussion of offsets). We also do not consider extended Ly α emission which is likely present at some level in our galaxies (e.g. Leclercq et al. 2017; Claeyssens et al. 2022). However, two qualities regarding extended Ly α emission suggest that consideration of this emission would have little effect on the qualitative trends found in this work. 1) The physical source of extended Ly α emission is uncertain. If the source is radiation from cold accreting gas, fluorescence, or emission from satellite galaxies, then it would seem unreasonable, physically, to include this flux in our measurements of the escape fraction and Ly α EW for this study which is concerned with sources of Ly α emission from the stellar populations of each galaxy. In fact, Lujan Niemeyer et al. (2022) suggest that the origin of extended Ly α emission originates outside of the galaxy itself. 2) If the extended emission originates from star-forming regions it should be included and could account for 2/3 of the total Ly α emission (Leclercq et al. 2017) if it is entirely produced from the star-forming regions. Nevertheless, our comparisons with other works which do not account for this correction will be the same given the same correction will be applied to their values as well.

When we have spectra of the same galaxy in different masks, we combine them using an inverse variance weighted mean. Some of our galaxies are multiply imaged, but combining multiple images involves demagnifying the spectra before combining and results in a spectrum of arbitrary magnification. Furthermore, the uncertainties in the magnification of high-magnification galaxies would result in large uncertainties in the flux with little gain in S/N. Therefore, we remove highly magnified images (according to the threshold described in Section 2.2) for the multiply imaged galaxies.

The LRIS slit masks use box slits to align the masks on bright stars in the field. Occasionally, the spectra of these alignment stars will contaminate nearby slits. As a result some of the slits near the alignment stars are rejected due to poor sky subtraction. After accounting for slits affected by nearby box slits, our final count of extracted 1D spectra is 127. After accounting for multiple images and spectra of individual galaxies our final parent sample comes to 89 galaxies. Each cluster takes up about one half of the LRIS detector. Because of this we have about half the number of spectra typically expected from an LRIS mask. We estimate the spectral resolution of our spectra given the slit width of our observations is $\sim$1 arcsec with a plate scale of 0.135 arcsec pix^{-1}, however, we have binned along the wavelength direction such that the effective plate scale is 0.27 arcsec pix^{-1}. The wavelength spacing is $\sim$2.18 Å pix^{-1}. So the resolution of our observations is $R \sim 500$ or $\Delta v \sim 600$ km s^{-1}. We show some example 1D spectra in Fig. 1.

2.2 Subsample selection

The remainder of the paper focuses on an analysis of the fraction of Ly α emitters ($X_{\rm LAE}$) and the escape fraction of Ly α photons ($f_{\rm esc}$). In order to measure $X_{\rm LAE}$ in an unbiased manner we need to select a subsample of galaxies for which we can identify galaxies with no Ly α emission. That is, a spectroscopic redshift can be measured independent of detection of the Ly α line. We call this subsample the $X_{\rm LAE}$ sample. In order to measure $f_{\rm esc}$ we must be able to measure H α. This is a more stringent requirement which eliminates a greater number of galaxies from the parent sample and would result in greater uncertainty in $X_{\rm LAE}$ were we to impose this condition on all of our data. Therefore, we choose to create an additional subsample which we call the H α sample.

2.2.1 X_{LAE} sample

We choose a subsample of galaxies from our sample of 89 in order to avoid biases in our Ly α equivalent width distribution. We account for biases from magnification, slit losses, and possible blending of sources. We mimic the subsample selection methods of Emami et al. (2020), but with values derived from our sample as follows:

(i) We require secure redshifts for all the galaxies in this sample so that galaxies with no Ly α or Ly α in absorption can be identified. Therefore, only galaxies with confirmed redshifts from our MOSFIRE data are kept. These redshifts are primarily determined by H α and [O III]$\lambda\lambda$5007.

(ii) We remove galaxies which do not have spectroscopic coverage of Ly α. Atmospheric absorption at $\lambda < 3200$ Å limits our analysis to galaxies at $z \gtrsim 1.6$.

(iii) We remove galaxies for which there is possible confusion/blending with nearby objects in the slit. These objects could have emission lines from one galaxy and continuum from another, affecting the equivalent width measurements.

(iv) We remove galaxies with large magnification, as they could suffer from differential magnification across the galaxy. This could introduce a bias into our sample which is selected on UV luminosity density. We therefore remove galaxies which have magnification $\mu > 30$ for Abell 1689 and $\mu > 15$ for MACS J0717 and MACS J1149.

(v) Finally, we remove galaxies with large slit loss correction in either LRIS or MOSFIRE. If a large fraction of the galaxy flux is outside of a slit, then it is not clear that a slitloss-corrected flux will reflect the true spectrum. Most of the galaxies have slit losses <2.2

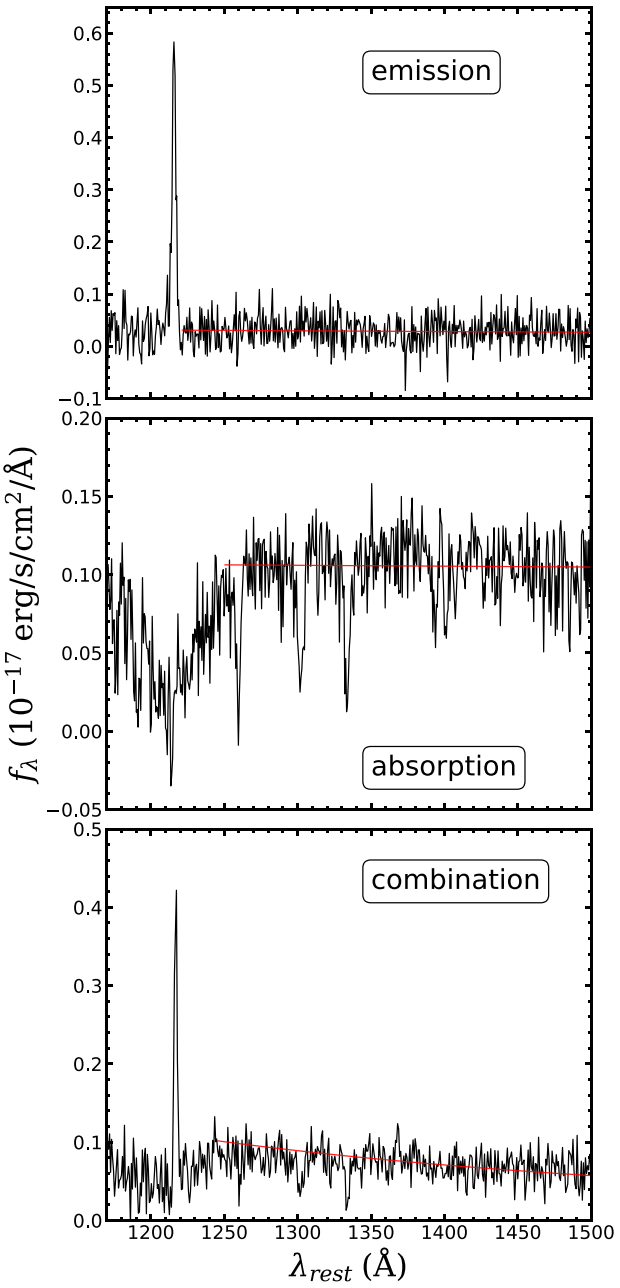

Figure 1. Three example spectra from our sample of galaxies. The red line is the continuum fit for each spectrum. The fitting method is described in Section 3. Top: An example of an 'emission' spectrum where we have Ly α in emission only. Middle: An example of an 'absorption' spectrum. Bottom: An example of a 'combination' spectrum where we see absorption and emission around Ly α.

and we thus set this as our upper limit. The slit loss distributions are shown in Fig. 2.

After this cleaning of the sample we obtain an $X_{\rm LAE}$ sample of 32 galaxies with a mean redshift $\langle z \rangle = 2.27$.

2.2.2 *H α sample*

In order to get a direct measurement of the Ly α escape fraction, H α must be observable. Because of this we create a second subsample

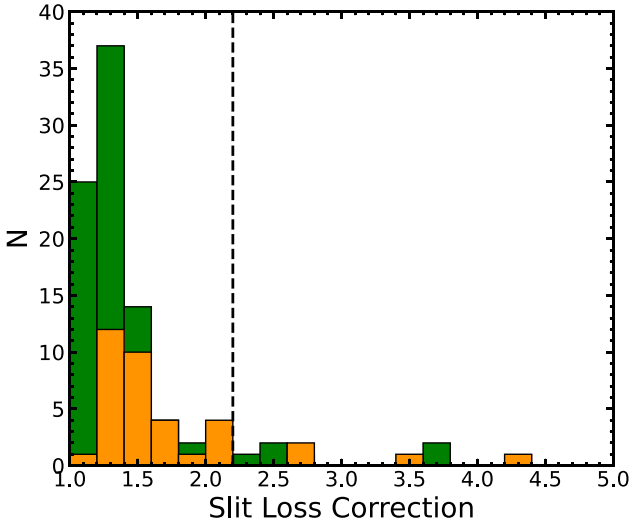

Figure 2. The distribution of slit loss corrections to our H α (yellow) and Ly α (green) spectra. The dashed line at 2.2 denotes the upper limit we set for including galaxies in our sample.

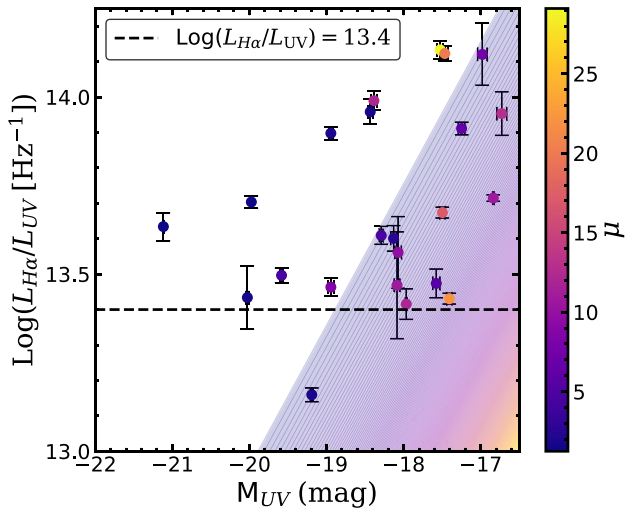

Figure 3. Plotted is the log($L_{\rm H\alpha}/L_{\rm UV}$) versus delensed absolute UV magnitude. Points are colour coded according to each galaxy's magnification. The black dashed line at log($L_{\rm H\alpha}/L_{\rm UV}$) = 13.4 denotes our completeness limit for the sample. The angled coloured space demonstrates the 3σ H α sensitivity limit of our detector colour coded according to magnification. Errors on $M_{\rm UV}$ do NOT include uncertainties on the lens model.

of galaxies from the $X_{\rm LAE}$ subsample for which H α could have been observed, which limits the redshift range to $z \lesssim 2.6$. Additionally, the rest-UV HST data are very deep and the KECK H α data are shallower which could bias us against galaxies that are faint in H α. To check whether we are excluding galaxies that are too faint in H α we calibrate our faint galaxy sample using our bright galaxies. We observe that our bright $M_{\rm UV}$ sample lies almost entirely above log($L_{\rm H\alpha}/L_{\rm UV}$) $\sim$ 13.4, as shown in Fig. 3, and therefore set this as our completeness limit for the sample. We take the median Hα = 4.3×10^{-18} erg s^{-1} cm^{-2} line flux error multiplied by 3 as our 3σ limit for detection of H α. We then check our galaxies magnification and absolute UV magnitude and ask whether the galaxy would have been observable at log($L_{\rm H\alpha}/L_{\rm UV}$) $\sim$ 13.4. If the galaxy would not have

been observable at 3σ at our completeness limit it was then removed. No galaxy fell below this criterion and therefore none were removed as shown in Fig. 3. We split our sample at $M_{\rm UV} = -19$ and measure the mean and error on the mean of the bright and faint samples to be $\log(L_{\rm H\alpha}/L_{\rm UV}) = 13.5 \pm 0.1$ and $\log(L_{\rm H\alpha}/L_{\rm UV}) = 13.7 \pm 0.1$ respectively. While all but one of our galaxies have $\log(L_{\rm H\alpha}/L_{\rm UV}) > 13.4$, 11 of the 18 galaxies fainter than $M_{\rm UV} = -19$ could have been observed below this line, given their magnifications. This suggests that there are indeed very few galaxies with $\log(L_{\rm H\alpha}/L_{\rm UV}) < 13.4$ and that our measured increase of $\log(L_{\rm H\alpha}/L_{\rm UV})$ for faint galaxies is real, and not due to a bias against galaxies with low $\log(L_{\rm H\alpha}/L_{\rm UV})$. Our final H α sample after all of these considerations is 23 galaxies.

3 METHODS

3.1 H α fits

The fits to the H α line were performed in a manner similar to Gburek et al. (2019) on the 1D extracted MOSFIRE spectra. We briefly review the procedure here. The spectra were fit using an MCMC sampler algorithm *emcee* (Foreman-Mackey et al. 2013). Each of the *J*, *H*, and *K* bands were fit separately and the emission lines were parametrized with a Gaussian and a linear continuum fit. The redshift was taken to be the weighted average of the values obtained between bands with emission lines present. The flux of the H α line was taken from the Gaussian parametrization.

3.2 UV continuum fits

We identified four types of spectra with respect to the Ly α line in our data set. Adopting the naming conventions of Kornei et al. (2010) we observe Ly α in 'emission', 'absorption', 'combination', and 'noise'. We measure the EWs of these spectra using the conventions of Du et al. (2018) with some variation. For each type of galaxy we set the bounds of numerical integration as follows:

(i) For emission-line galaxies we set the bounds to be where the emission line met the continuum.

(ii) For combination and absorption-line galaxies we set the region of integration to the space between 1208 and 1240 Å. We calculate the continuum level to be the midpoint flux value of a line passing through 1208 and 1240 Å, where we take the mean flux between 1206 and 1210 Å to be the flux at 1208 Å and the mean flux between 1225 and 1255 Å to be the flux at 1240 Å. We use this continuum level for calculating the EW.

(iii) For noise spectra we have no information on the Ly α profile. Therefore, we use boundaries derived by Kornei et al. (2010) at 1199.9 and 1228.8 Å rest frame for the short- and long-wavelength bounds, respectively, where the redshift is determined from the rest-optical emission lines.

In addition, we performed a parametrized fit of the Ly α profile (for emission and combination spectra) and continuum of each spectrum for our galaxy sample to measure the UV spectral slope β. The fitting procedure is performed using a Markov chain Monte Carlo code from pymc3. The wavelength range used for fitting varied based on redshift as the short-wavelength bound was set to just blueward of the Ly α emission line. The long-wavelength bound was set to be at 5450 Å in the observed frame, as the transmission decreases at longer wavelengths due to the dichroic at 5600Å. Typical emission and absorption lines (Si II$\lambda\lambda$1260 Å, O I+Si II$\lambda\lambda$1303 Å, C II$\lambda\lambda$1334 Å, Si IV$\lambda\lambda$1393 Å, C IV$\lambda\lambda$1549 Å, He II$\lambda\lambda$1640 Å, O III$\lambda\lambda$1666 Å, C III$\lambda\lambda$1909 Å, and O IV$\lambda\lambda$1343 Å) were masked when fitting each spectrum. We set the width for these masks to be 800 km s^{-1} except for the C IV$\lambda\lambda$1549 Å line which we set to have an asymmetric width of 1400 km s^{-1} with the blue side of the line extending to -1000 km s^{-1} because of the P-Cygni profile that C IV$\lambda\lambda$1549 Å tends to take. This process was repeated for all galaxies observed. The 'emission' spectra are modelled simultaneously with a power-law continuum and a Gaussian emission line to handle correlations between parameters,

$$f_\lambda = A_0 e^{-\frac{(\lambda-\lambda_0)^2}{2\sigma^2}} + B_0(\frac{\lambda}{\lambda_c})^\beta. \quad (1)$$

For the 'combination' spectra we add in a first-order approximation to the Ly α absorption profile,

$$f_\lambda = \begin{cases} A_0 e^{-\frac{(\lambda-\lambda_0)^2}{2\sigma^2}} + a\lambda + b, & \lambda < \lambda_s \\ B_0(\frac{\lambda}{\lambda_c})^\beta, & \lambda > \lambda_s \end{cases} \quad (2)$$

where λ_s is the wavelength where there exists a change in the UV spectral slope likely due to damped Ly α absorption. We fit this region with a linear function to account for the change in the UV spectral slope. The remaining galaxies are parametrized with a power-law continuum redward of Ly α (identified either from damped Ly α absorption or from the redshift estimate obtained with Keck/MOSFIRE),

$$f_\lambda = A_0(\frac{\lambda}{\lambda_0})^\beta. \quad (3)$$

We use an upper limit of FWHM = 4.9 Å rest frame to obtain upper limits on the flux from the 'absorption' and 'noise' spectra. The value of 4.9 Å is derived from the distribution of 49 FWHMs from the sample for which we have Ly α in emission. 4.9 Å is greater than all but two of 49 (4 per cent) of the FWHMs measured and therefore serves as an upper limit on the FWHM.

3.3 Magnification

Thanks to lensing, the signal from faint galaxies is magnified and we are able to extend spectroscopic studies to fainter magnitudes at $z \sim 2$ in just a few hours of Keck/LRIS and Keck/MOSFIRE time (see Table 1). Our galaxies are observed behind three lensing clusters. In the $X_{\rm LAE}$ and H α subsamples the magnifications span $1.25 \leq \mu \leq 29.12$. The median magnifications are $\langle\mu\rangle_{\rm median} = 7.11$ and $\langle\mu\rangle_{\rm median} = 7.45$, respectively. It is important to properly model the lensing in order to accurately measure the UV luminosity of the galaxies. Alavi et al. (2016) details the lens models used to measure the magnifications of the galaxies behind MACS J0717, MACS J1149, and Abell 1689. The models considered are constrained by the location and redshift of known multiply imaged sources. We use the models produced by the Clusters As Telescopes (CATS) collaboration.[1] As Alavi et al. (2016) states, we use the models of Limousin et al. (2007), Limousin, M. et al. (2016), and Jauzac et al. (2016) for Abell1689, MACS J0717, and MACS J1149, respectively. These models are derived using LENSTOOL[2] (Jullo et al. 2007). The basic assumptions of the model are that the mass distributions of the clusters are constructed from smooth parametrized large-scale potentials (see Priewe et al. 2017, and reference therein for more details). The uncertainties in the magnification are not considered throughout this work. For the Hubble frontier fields Priewe et al. (2017) show that the systematic uncertainty across lensing models will be on the order of 50 per cent

[1] https://archive.stsci.edu/prepds/frontier/lensmodels/
[2] https://projets.lam.fr/projects/lenstool/wiki

at our magnifications. This is small relative to our 'faint' to 'bright' delineation which we use for comparisons in M_{UV}. In addition, many of the quantities we consider are relative measures, and since magnification is achromatic the majority of quantities studied in this paper will not be affected by the magnification.

3.4 SED fits

We fit spectral energy distributions (SEDs) to our Hubble photometry in the near-UV, optical, and near-IR. Before SED-fitting we subtract off contributions from nebular emission lines using our slit-loss corrected spectra. We add an additional 3 per cent flux error to all of our bands to account for systematic errors in our photometry (Alavi et al. 2016). Using the code FAST[3] (Kriek et al. 2009), we fit Bruzual & Charlot (2003) stellar population synthesis models to the emission-line subtracted photometry under the following assumptions:

(i) Constant star formation histories (SFHs)
(ii) A Chabrier (2003) initial mass function (IMF)
(iii) Stellar metallicities of either 0.2 $Z_\odot$ or 0.4 $Z_\odot$
(iv) A Gordon et al. (2003, SMC) dust attenuation curve

Galaxy redshifts are set by the fit curves to the rest-optical spectroscopy. M_{UV} is calculated from the closest filter to the rest-frame 1700 Å that does not include Ly α, and uncertainties on M_{UV} are statistical.

3.5 Dust correction

In our analysis of the H α sample we need to correct the H α line and the UV continuum at 1700 Å for dust attenuation. We use the SMC curve of Gordon et al. (2003) and A_v estimates from our SED fitting to estimate the attenuation of the UV continuum. We make use of the SMC attenuation law for the UV continuum because it has been shown to be more accurate for high-redshift and low-metallicity galaxies than the Cardelli, Clayton & Mathis (1989) curve (e.g. Reddy et al. 2022). The correction for H α is estimated by measurement of the Balmer decrement (assuming Case-B recombination). However, many of our galaxies have Balmer lines that are either too faint or obscured by sky emission and we are unable to directly measure the Balmer decrement of each galaxy. We split our H α sample into bright ($M_{UV} < -19$) and faint ($M_{UV} > -19$) subsamples for which we estimate A_v. For the faint subsample we normalize the spectra by H α and perform a median stack of the galaxies. We use a Cardelli et al. (1989) curve and measure $\langle A_v \rangle = 0.54$ and apply this to all of our dwarf galaxies. Most of our reported results will be about the dwarf galaxies for which we use this average attenuation value. We recognize that there will be significant scatter around this value. Nonetheless, the correction is small, a factor of 1.5 for the H α luminosity, and does not affect the average trends reported in this paper. For our more massive galaxies we measure A_v for three of the galaxies ($A_v = 1.14, 0.88$, and 0.80). We then use the mean ($\langle A_v \rangle = 0.94$) on the remaining two galaxies for which we did not have H_β $S/N > 5$ and therefore could not reliably calculate the dust correction.

4 RESULTS

As explained above, our larger X_{LAE} sample is used to measure the EW distribution and the smaller H α subsample is used to measure the escape fraction. We compare our two samples in order to see if the two are similar and therefore both can be used to determine properties of dwarf galaxies at this redshift. We find that the median EWs agree within errors on the median as shown in Fig. 4. We calculate the uncertainty on the median for each sample by randomly selecting galaxies within the given sample with replacement and recalculating the median from this new sample. We perform this operation 10 000 times and define the errors as the 16th and 84th percentiles of the distribution of the median. The difference in mean redshift is small ($\langle z_{Ly\alpha} \rangle = 2.28$ and $\langle z_{H\alpha} \rangle = 2.23$). Since the distributions are similar we take the H α sample to be representative of trends in the X_{LAE} sample as well.

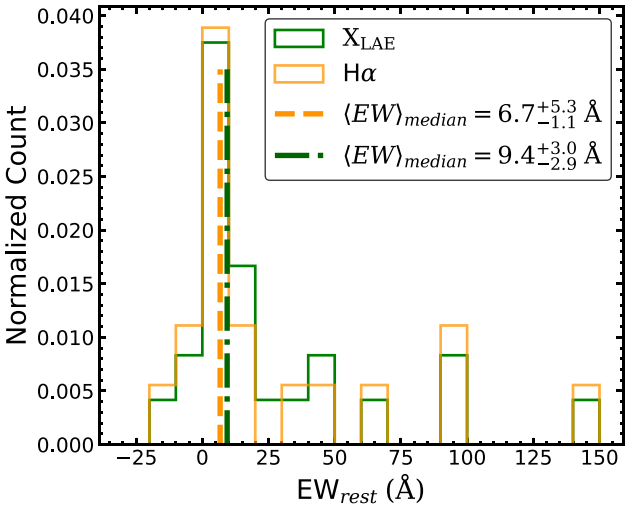

Figure 4. The green (orange) histogram is a normalized distribution of rest-frame EWs of our X_{LAE} (H α) sample. The coloured dashed vertical lines show the median EW of $9.4^{+3.0}_{-2.9}$ ($6.7^{+5.3}_{-1.1}$) for the X_{LAE} (H α) sample. The Ly α EW distributions of both samples are consistent.

4.1 Ly α EW distribution

We wish to investigate trends in the EW distribution with UV luminosity. Therefore we wish to compare with a sample of higher luminosity galaxies at similar redshifts where possible. Therefore, we compare our EW distribution with that of the higher luminosity ($M_{UV} \sim -20$) sample of Du et al. (2021) as shown in Fig. 5. We find that our sample is skewed towards larger Ly α EWs. The median EW for the Du et al. (2021) sample is $-6.0^{+1.9}_{-1.1}$ Å and the median for our sample is $9.4^{+3.0}_{-2.9}$ Å. The two distributions differ by $\sim 4\sigma$ showing that they are statistically different. We calculate the uncertainty for the Du et al. (2021) median in the same manner used for each of our subsamples.

The cut-off for defining an LAE in the literature varies significantly depending on the study (Stark et al. 2010, 2011; Caruana et al. 2018; Kusakabe et al. 2020, etc.). However, the most common cuts are EW > 20 Å, EW > 25 Å, and EW > 55 Å, so we will use these cuts as well. For these three cut-offs we measure $X_{LAE} = 33^{+17}_{-12}$ per cent (8/24), 33^{+17}_{-12} per cent (8/24), and 17^{+13}_{-8} per cent (4/24), respectively. We assume a Poisson distribution to estimate uncertainties on X_{LAE}. X_{LAE} measurements of brighter galaxies are ~ 11 per cent for EW > 20 Å at $z \sim 2$ [12 per cent from Reddy et al. (2008), 11.1 per cent from Hathi et al. (2016), and 10.7 per cent from Du et al. (2021)]. For brighter galaxies with EW > 25 Å and EW > 55 Å at $z \sim 2$ $X_{LAE} = 10$ per cent and 4 per cent respectively (Cassata et al.

[3] https://w.astro.berkeley.edu/~mariska/FAST.html

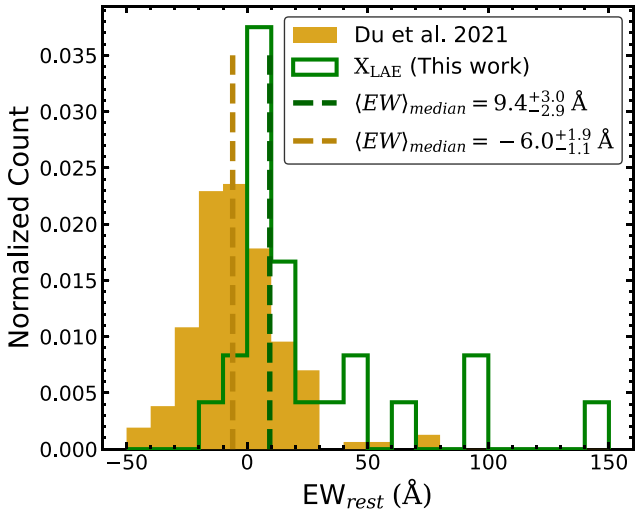

Figure 5. The green histogram is a normalized distribution of rest-frame EWs of our X_{LAE} sample. The gold histogram is a normalized distribution of rest-frame EWs of the brighter sample from Du et al. (2021). The vertical dashed lines of corresponding colour show the median Ly α EWs of each sample. The median EWs of the green and gold histograms are $\langle EW \rangle_{median} = 9.4^{+3.0}_{-2.9}$ Å and $\langle EW \rangle_{median} = -6.0^{+1.9}_{-1.1}$ Å, respectively.

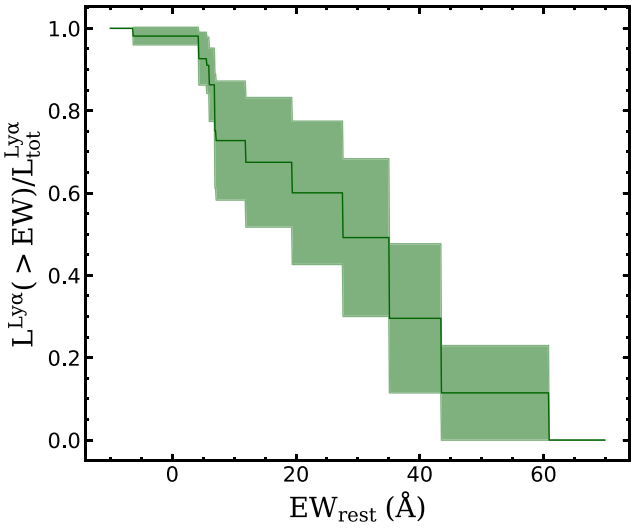

Figure 6. The integrated Ly α luminosity density as a function of EW. The uncertainty is given by the shaded green regions. We find that about 60 per cent of the integrated luminosity density comes from galaxies with $EW_{Ly\alpha} > 20$ Å.

2015). These data suggest that galaxies fainter than $M_{UV} = -19$ have greater numbers of LAEs in their population. We explore this in more detail in Section 5.3.

4.2 Escape fraction

In our H α sample for which we are able to measure the escape fraction of Ly α photons we assume a Case-B scenario for which the gas is taken to be optically thick for the Lyman series. The ratio of the flux of Ly α to the flux of H α under this assumption can vary depending on the temperature and electron density of the galaxy. We choose to use a ratio of 8.7 which is common in the literature (e.g. Matthee et al. 2016). We then measure the escape fractions of individual galaxies as

$$f_{esc} = \frac{F_{Ly\alpha}}{8.7\ F_{H\alpha,cor}} \quad (4)$$

where $F_{Ly\alpha}$ is the Ly α flux and $F_{H\alpha,cor}$ is the dust-corrected H α flux. The mean of the individual f_{esc} measurements is $4.3^{+1.6}_{-1.1}$ per cent. We estimate the uncertainty in the mean via a bootstrap method. We perturb each measurement of the sample according to a Gaussian distribution with standard deviation set by the uncertainty of the measurement and then resample by randomly selecting with replacement from our perturbed sample and then calculating the mean of the escape fractions. This process is performed 100 000 times. We then take the 16th and 84th percentiles of the 100 000 iterations to be the uncertainty on the mean escape fraction.

However, this value may not be representative of the actual number of Ly α photons escaping into the intergalactic medium. The intrinsic luminosity of galaxies varies greatly and may correlate with Ly α escape fraction. Thus, the mean escape fraction may be different from the net output from all galaxies, which we refer to as the *volumetric* escape fraction. To calculate the volumetric escape fraction, we first need to determine the galaxies' intrinsic luminosity and, thus, demagnify each of our galaxies. We then replace the fluxes in equation (4) with the sum of the respective luminosities. We measure a volumetric escape fraction of $4.6^{+2.0}_{-1.4}$ per cent. We estimate the uncertainty in a similar manner to the individual escape fraction estimate. We perturb each measurement of the sample according to a Gaussian distribution with standard deviation set by the uncertainty of the measurements, resample, and then calculate the volumetric escape fraction. We take the 16th and 84th percentiles of the 100 000 iterations to be the uncertainty on the volumetric escape fraction. This method includes contributions from uncertainties in the dust correction. This value agrees well with what is found in the literature (5.3 per cent $\pm$ 3.8 per cent at $M_{UV} \sim -19$ in Hayes et al. (2010), $4.4^{+2.1}_{-1.2}$ per cent at $M_{UV} \sim -19.5$ in Ciardullo et al. (2014), 5.1 per cent $\pm$ 0.2 per cent at $M_{UV} \sim -19.7$ in Sobral et al. (2017), and $5.8^{+0.7}_{-0.5}$ per cent at $M_{UV} \sim -19.5$ in Weiss et al. (2021)).

By the use of gravitational lensing we were able to directly measure a UV-selected sample of dwarf galaxies at $z = 2$. We are able to probe fainter H α and Ly α luminosities and have a greater sample of joint detections of H α and Ly α emitters than what was observed in Hayes et al. (2010).

4.3 Observed Ly α EW

We investigate which galaxies are emitting most of the Ly α luminosity. Fig. 6 shows that about 60 per cent of the integrated Ly α luminosity density comes from galaxies with $EW_{Ly\alpha} > 20$ Å. This suggests that LAEs are contributing a large fraction of the Ly α photons. This cumulative distribution can also be calculated from narrow-band surveys for Ly α. However, those surveys cannot determine the amount of total emissions from low EW galaxies. Here we show that ~10–20 per cent of the Ly α luminosity density comes from galaxies with $EW_{Ly\alpha} < 5$Å.

5 DISCUSSION

Here we seek to understand the origins of the observed EW distribution. Is the spread in observed EW primarily due to varying escape fractions, or are variations in stellar populations (in particular, starburst age), creating a large scatter in *intrinsic* EWs as well.

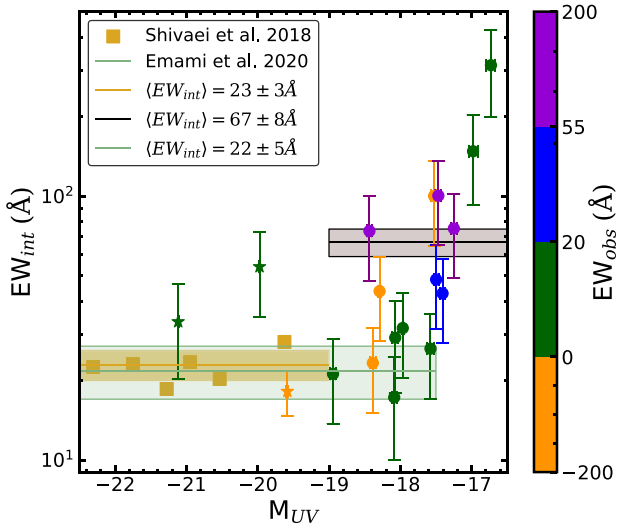

Figure 7. The *intrinsic* equivalent width as a function of absolute UV magnitude ($M_{\rm UV}$). Points are colour coded according to the *observed* equivalent width (see colour bar on the right) with blue and purple points representing galaxies observed as LAEs. The square gold points are average values of higher luminosity galaxies and are derived from the $\xi_{\rm ion}$ measurements of Shivaei et al. (2018). Galaxies brighter than −19 in absolute UV magnitude are plotted with a star, while fainter galaxies are plotted with filled circles. We show the mean intrinsic EW of Shivaei et al. (2018) and uncertainty in the mean with a gold shaded region, and we show the mean intrinsic EW and uncertainty in the mean of our dwarf galaxy sample in brown. We also show the mean intrinsic EW and uncertainty in the mean of Emami et al. (2020), which includes some of the same galaxies, with a green shaded region. Five galaxies did not have sufficient photometric coverage to perform SED fitting to obtain $A_{\rm V}$ values to correct the UV luminosity and are removed from any intrinsic EW analysis.

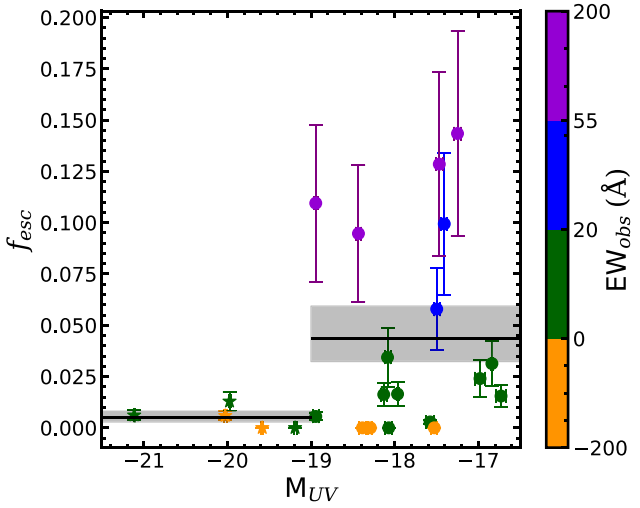

Figure 8. The Ly α escape fraction as a function of absolute UV magnitude ($M_{\rm UV}$). The colour coding and markers are the same as in Fig. 7. The mean escape fraction of the dwarf galaxies ($4.3^{+1.6}_{-1.1}$ per cent) and the more massive galaxies ($0.5^{+0.3}_{-0.2}$% per cent) are displayed with shaded black regions denoting the propagated uncertainty in the individual galaxies. However, the bright galaxy sample has only five galaxies and is therefore uncertain.

5.1 Intrinsic Ly α EW

To investigate influences from the stellar populations we make use of the intrinsic Ly α EW

$$\mathrm{EW_{int}} = \frac{8.7 L_{\mathrm{H}\alpha,\mathrm{cor}}}{L_{\lambda,1216,\mathrm{cor}}} \quad (5)$$

where $L_{\mathrm{H}\alpha,\,\mathrm{cor}}$ is the dust-corrected and demagnified H α luminosity, and $L_{\lambda,\,1216,\,\mathrm{cor}}$ is the dust-corrected and demagnified UV luminosity density extrapolated to 1216 Å. Fig. 7 shows the intrinsic EW as a function of absolute UV magnitude. We note that since we make use of $A_{\rm v}$ values from SED fitting to correct the UV luminosity we must remove five galaxies for which we do not have sufficient photometric coverage to perform SED fits from all analysis regarding intrinsic EWs.

We compare our measured $\mathrm{EW_{int}}$ values to those of more massive galaxies at the same redshift. We convert the ionizing photon production efficiency ($\xi_{\rm ion}$) of Shivaei et al. (2018) to $\mathrm{EW_{int}}$ by converting $\xi_{\rm ion}$ to $\mathrm{L_{Ly\alpha}/L_{UV}}$. In Shivaei et al. (2018) they calculate $\xi_{\rm ion}$ from weighted mean stacks of their galaxies in bins of absolute UV magnitude. They are shown as gold squares in Fig. 7. Shivaei et al. (2018) show a mean $\log(\xi_{\rm ion}) = 25.36 \pm 0.06$ for the SMC curve. This corresponds to an intrinsic EW of 23Å ± 3 Å. Emami et al. (2020) have $\xi_{\rm ion} = 25.22 \pm 0.10$ which corresponds to an intrinsic EW of 22 ± 5 Å for a similar sample. For our faint galaxies our mean intrinsic EW is 67 ± 8 Å. Therefore, we see a significant increase in the mean intrinsic EW for fainter galaxies. The faintest galaxies in our sample may be skewing the sample towards higher $\mathrm{EW_{int}}$, but if we remove the two faintest galaxies we still get an elevated mean intrinsic EW of 46 ± 5 Å. We also show in Fig. 7 the observed EWs via the colour bar where blue and purple points indicate LAEs. From the plot we observe large intrinsic EWs ($\mathrm{EW_{int}} > 40$ Å) facilitate observation of LAEs, but are not a guarantee that a galaxy will be observed as a LAE. This implies that the ionizing output of the stellar population in galaxies has significant bearing on whether a galaxy is observed to be a LAE. However, it is clear that the stellar population alone cannot account for the observability of LAEs.

5.2 Escape fraction $f_{\rm esc}$

The other factor driving the observed distribution in Ly α EWs is whether the photons actually escape from the galaxy. With our MOSFIRE spectra we are able to measure this value directly. For our absorption galaxies in the H α sample we set $f_{\rm esc} = 0$. We show the escape fraction as a function of absolute UV magnitude in Fig. 8.

We see large scatter in the escape fraction at faint absolute UV magnitudes. It is also the case that no escape fraction less than ∼ 5 per cent is observed as a Ly α emitter. Therefore, it is clear that the escape of ionizing photons plays a crucial role in the observed EW distribution and $X_{\rm LAE}$. To determine if the escape fraction and $\mathrm{EW_{int}}$ are correlated, we plot them against each other in Fig. 9.

We see that 5/9 of the galaxies with high intrinsic EWs (≥40 Å) have large escape fractions ($f_{\rm esc} > 5$ per cent). No galaxy (0/8) with $\mathrm{EW_{int}} < 40$ Å has an escape fraction larger than 5 per cent. We can set two conditions for LAEs; LAEs have escape fractions greater than 5 per cent and intrinsic EWs greater than 40 Å. We caution that our sample size is small, so this result could be made more robust by increasing the sample size. In order to understand this a little deeper we can look at possible drivers of the escape fraction.

Ly α can be heavily attenuated by dust. However, in our sample we have shown that our galaxies are typically not very dusty. Nevertheless we check to see if there is any relation between the dust attenuation estimated from SED fits and the escape fraction. Fig. 10 shows the escape fraction as a function of $A_{\rm v}$. We observe

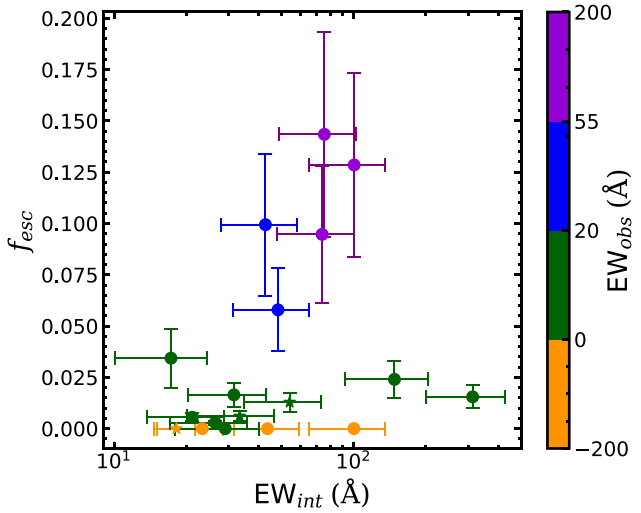

Figure 9. The Ly α escape fraction as a function of *intrinsic* equivalent width. The colour coding and markers are the same as in Fig. 7. Galaxies with no Ly α in emission are set to $f_{esc} = 0$. We observe that only galaxies with intrinsic equivalent widths greater than 40 Å have escape fractions greater than 5 per cent.

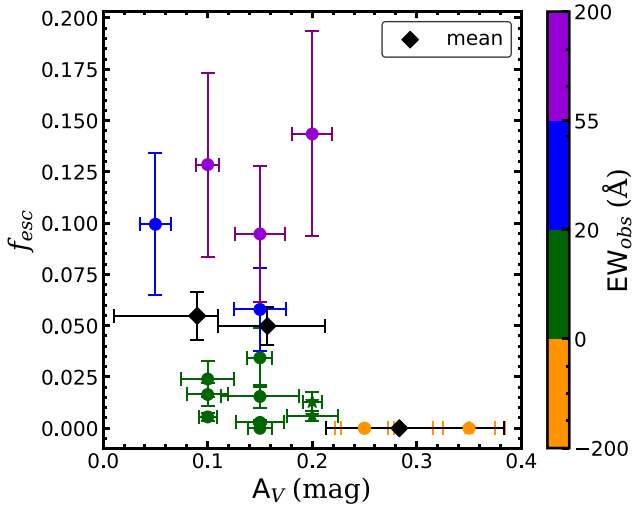

Figure 10. The Ly α escape fraction as a function of dust attenuation. The colour coding and markers are the same as in Fig. 7. Black diamonds show the mean escape fraction in three bins of A_V. The horizontal error bars denote the sizes of the bins and the vertical error bars denote the uncertainty in the mean. A_v is here derived from SED fitting of Hubble photometry and is in discreet steps.

an anticorrelation between f_{esc} and A_v. However, nearly all of our galaxies have $A_V \leq 0.2$. With little variance in dust content we are not able to test this robustly. Nevertheless, we do not observe escape fractions greater than 5 per cent for $A_V \geq 0.2$. To test the anti-correlation we perform an MCMC fit for a linear model and find the slope to be -0.015 ± 0.004. This establishes the anti-correlation at greater than 3σ. We also note that the two highest intrinsic EW objects also exhibit dust content similar to our LAEs. This suggests that it is primarily attenuation by neutral hydrogen dictating the escape of Ly α photons in dwarf galaxies. Cassata et al. (2015) show that there is larger variance in f_{esc} at a given reddening value (see their fig. 6) which is consistent with our results. Our highest dust bins

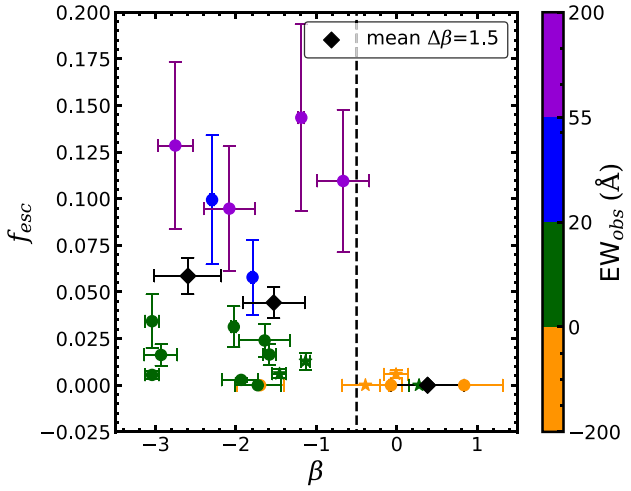

Figure 11. The Ly α escape fraction as a function of UV spectral slope (β). The colour coding and markers are the same as in Fig. 7. The black dashed vertical line is set as a limit for LAEs. All LAEs have slopes bluer than -0.5. There is large variance in beta for similar f_{esc}, perhaps suggesting bursty star formation. The black diamonds are mean values of the escape fraction in bins of $\Delta\beta = 1.5$. The error bars are calculated by propagating the statistical uncertainty of each datum. One object was excluded due to highly uncertain β from low S/N continuum.

have few data points and might not be representative. Our two lowest dust content bins may suggest a constant average escape fraction as a function of A_V which appears to be consistent with the lowest redshift bins of Matthee et al. (2016) with 3 arcsec apertures. This is in slight contrast to the measurements of Weiss et al. (2021) who show a steep increase in f_{esc} with decreasing reddening. However, the sample of Weiss et al. (2021) may be biased towards LAEs, particularly at the lowest mass/dust values ($\log(M^\star/M_\odot) < 9$), which is the regime of our study, since they are selecting on [O III]5007 emission (Erb et al. 2016). That they are biased at the low mass end is supported by the left-hand panel of their fig. 4 which indicates that their typical galaxy at low mass is an LAE.

The slope of the UV continuum can also be used as an indicator of dust and so we look at possible correlations with UV slope as well. This is shown in Fig. 11. Here we see a mild dependence in that the scatter of escape fractions increases for $\beta < -0.5$. We again test the anti-correlation using an MCMC routine and a linear model and find a slope of -0.0003 ± 0.0011. Therefore, our data are consistent with there being no correlation between f_{esc} and UV spectral slope. We suggest that the trend in the mean is marginal for our small data set. Matthee et al. (2016) show a change in the slope of the trend between f_{esc} and β which we do not see for our dwarf galaxies. However, we lack sufficient numbers of $\beta > 0$ galaxies to make any definitive statements regarding this change. We are inconsistent with Matthee et al. (2016) regarding the decrease in f_{esc} with β out to $\beta = 0$. Together, Figs 10 and 11 suggest that there is little to no anti-correlation between dust and f_{esc} for faint galaxies. This implies that attenuation of Ly α emission is due primarily to neutral hydrogen gas in our sample.

5.3 Trends in X_{LAE}

The decrement in X_{LAE} from the expected X_{LAE} at high redshift is often used to infer the IGM neutral fraction (Stark et al. 2010, 2011, e.g.). However, in order to investigate the number of faint galaxies

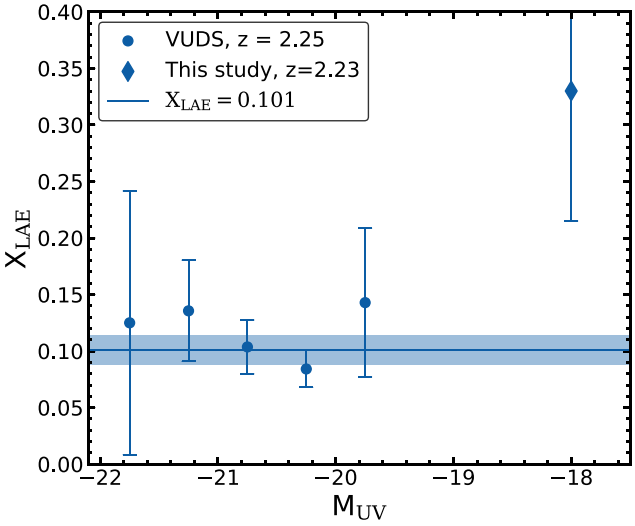

Figure 12. The Ly α emitter fraction, $X_{\rm LAE}$, as a function of $M_{\rm UV}$. Our data set is represented by a blue diamond and the VUDS data set by blue circles. We show a constant fit line, along with uncertainties in shaded blue, to the VUDS data at $X_{\rm LAE} = 0.101 \pm 0.012$ as there is no trend in the brighter $M_{\rm UV}$ sample.

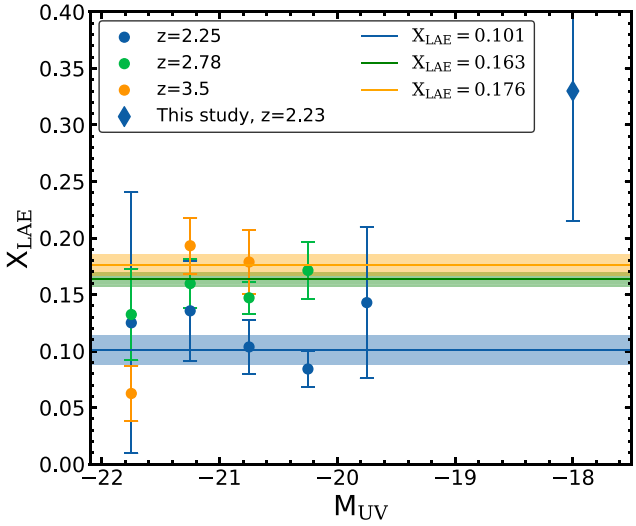

Figure 13. The Ly α emitter fraction, $X_{\rm LAE}$, as a function of absolute UV magnitude and redshift. The lines and shaded regions of corresponding colour to the data points are the constant fit curves to the data at each redshift. There are no trends in $M_{\rm UV}$ at any redshift in the VUDS data.

that contribute to reionization it is helpful to establish trends in $X_{\rm LAE}$ with $M_{\rm UV}$ at lower redshifts where Ly α is observable. To this end we compare with the VUDS sample (Cassata et al. 2015; Le Fèvre et al. 2015; Hathi et al. 2016) at higher UV luminosities. For this analysis we choose $EW_{Ly\alpha} > 20$ Å as our definition for a LAE. Fig. 12 shows the VUDS data set in bins of $M_{\rm UV}$ along with our data set. The uncertainty on the VUDS bins were calculated by resampling the distribution within each $M_{\rm UV}$ bin. The same number of galaxies within the bin were randomly selected with replacement and then the fraction of LAEs was calculated. This procedure was repeated 10 000 times on each bin and the standard deviation of $X_{\rm LAE}$ values from this procedure was taken to be the 1-σ uncertainty on the VUDS bins. For our $X_{\rm LAE}$ value we assume Poisson statistics (due to the small number of galaxies in our data set) to estimate the uncertainty. We find no trend in the VUDS data for brighter $M_{\rm UV}$, so we fit a constant line to the VUDS data set. We assume a binomial probability distribution for fitting $X_{\rm LAE}$ in the VUDS data of the form,

$$p(\text{data}|X_{\rm LAE}) \sim X_{\rm LAE}^{L}(1 - X_{\rm LAE})^{N-L}, \quad (6)$$

where N is the total number of galaxies and L is the number of LAEs. This gives a most likely value for the fit to the VUDS data of $X_{\rm LAE} = 0.101 \pm 0.012$. This is $\sim 2\sigma$ from our result of $X_{\rm LAE} = 0.33^{+0.17}_{-0.12}$. More data at $M_{\rm UV} \sim -19$ between the VUDS faintest $M_{\rm UV}$ and our sample is necessary to determine any trend in $X_{\rm LAE}$ with $M_{\rm UV}$.

We also investigate trends in redshift using the VUDS data set as shown in Fig. 13. We fit each redshift bin with a constant value, according to the same method used for the $z \sim 2$ sample, since we again see no trend with $M_{\rm UV}$ at any redshift. There is a clear trend towards larger $X_{\rm LAE}$ with redshift in agreement with other works in the literature (e.g. Stark et al. 2011; De La Vieuville et al. 2020; Kusakabe et al. 2020). The $X_{\rm LAE}$ values of low-luminosity galaxies at $z \sim 2$ in our sample are consistent with more massive galaxies at $z \geq 3$.

6 SUMMARY

In this paper we have investigated the distribution of Ly α EWs for dwarf galaxies at $z \sim 2$. We selected (via UV continuum) magnified galaxies behind lensing clusters, allowing us to obtain rest-UV spectroscopy for some of the faintest galaxies ($M_{\rm UV} < -17$, $\langle M^{\star} \rangle_{\rm median} = 10^{8.4}\,M_{\odot}$) observed at $z \sim 2$. We reduce our sample to avoid issues with differential magnification, large slit loss uncertainties, incompleteness due to faint Hα, and the observability of Ly α and H α within the bands of Keck/LRIS and Keck/MOSFIRE. We then analyse the EW distribution of the sample and draw the following conclusions:

(i) The observed $EW_{Ly\alpha}$ distribution of low UV luminosity galaxies is skewed towards larger $EW_{Ly\alpha}$ than higher luminosity galaxies such as those in the KBSS sample of Du et al. (2021). Our sample shows a median $EW_{Ly\alpha}$ of $9.4^{+3.0}_{-2.9}$ Å whereas the sample of Du et al. (2021) shows a median $EW_{Ly\alpha}$ of $-6.0^{+1.9}_{-1.1}$ Å.

(ii) The fraction of galaxies that are Ly α emitters, $X_{\rm LAE}$, is 33^{+17}_{-12}, 33^{+17}_{-12}, and 17^{+13}_{-8} per cent for $EW_{Ly\alpha} > 20$ Å, $EW_{Ly\alpha} > 25$ Å, and $EW_{Ly\alpha} > 55$ Å, respectively. These values are greater than $X_{\rm LAE}$ ($EW_{Ly\alpha} > 20$ Å) in higher luminosity samples such as Reddy et al. (2008, 12 per cent), Hathi et al. (2016, 11 per cent), and Du et al. (2021, 11 per cent).

(iii) We investigate possible trends in $X_{\rm LAE}$ with $M_{\rm UV}$ and redshift using the VUDS data set (Cassata et al. 2015; Le Fèvre et al. 2015; Hathi et al. 2016). We find no trend with $M_{\rm UV}$ between $-22 \leq M_{\rm UV} \leq -19.5$, but find our $X_{\rm LAE}$ at $M_{\rm UV} \sim -18$ to be about a factor of 3 greater at $\sim 2\sigma$ significance. There is a trend towards larger $X_{\rm LAE}$ with redshift in the VUDS data. Further investigation of dwarf galaxies at higher redshift could show increased $X_{\rm LAE}$ relative to higher mass galaxies as well.

(iv) We find that the total integrated Ly α luminosity is 60 per cent for galaxies with $EW_{Ly\alpha} > 20$ Å, suggesting that LAEs contribute a large fraction of the Ly α photons. We also find that $\sim 10 - 20$ per cent of the integrated luminosity comes from galaxies with $EW_{Ly\alpha} < 5$ Å.

(v) We estimate the intrinsic Ly α EW of galaxies, and show an increase in the mean $EW_{\rm int}$ for faint galaxies ($\langle EW_{\rm int} \rangle = 67 \pm 8$ Å)

compared to the brighter sample of Shivaei et al. (2018, $\langle EW_{int} \rangle = 23 \pm 3$ Å) and the intermediate luminosity sample of Emami et al. (2020, $\langle EW_{int} \rangle = 22 \pm 5$ Å). This suggests that younger ages and/or lower metallicities of the stellar populations of dwarf galaxies are increasing the intrinsic EW and contributing to the larger X_{LAE}.

(vi) We investigate the escape fraction of Ly α photons to further understand what is driving the EW distribution. We observe a marginal increase in the mean escape fraction with absolute UV magnitude. We also observe that only galaxies with $EW_{int} > 40$ Å have $f_{esc} > 5$ per cent. This suggests that only galaxies with higher EW_{int} will be classified as LAEs, not just because of their higher EW_{int}, but also because only those galaxies have large escape fractions.

(vii) We find an anti-correlation between f_{esc} and A_V, though for the lowest dust bins there is little to no correlation. Our data are consistent with no correlation between f_{esc} and β. This implies that it is primarily attenuation by neutral hydrogen dictating the escape of Ly α photons in dwarf galaxies.

(viii) We observe a global *volumetric* escape fraction of $f_{esc} = 4.6^{+2.0}_{-1.4}$ per cent in our sample, in good agreement with values inferred for other faint UV luminosity samples in the literature (e.g. Hayes et al. 2010; Weiss et al. 2021, etc.).

We were able to disentangle to some extent two primary drivers of the Ly α EW distribution, namely the ionizing sources (via the intrinsic EW), and the ISM gas/dust content (via the escape fraction and its dependencies). This sample can serve as a baseline with which to compare higher redshift and higher mass samples. Larger sample sizes of dwarf galaxies would serve to solidify some of the results of this work.

ACKNOWLEDGEMENTS

This research made use of `PypeIt`[4], a PYTHON package for semi-automated reduction of astronomical slit-based spectroscopy (Prochaska et al. 2020a, 2020c).

Some of the data presented herein were 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. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation.

The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

Based on observations made with the NASA/ESA *Hubble Space Telescope*, obtained from the Data Archive at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. These observations are associated with programs #9289, #11710, #11802, #12201, #12931, #13389, and #14209.

DATA AVAILABILITY

This paper is based on public data from the *Hubble Space Telescope* as well as from programs 12201, 12931, 13389, and 14209. The Spectroscopic data are from our own survey with the Keck Observatory. These data are available upon request from Christopher Snapp-Kolas or Dr Brian Siana.

[4]https://pypeit.readthedocs.io/en/latest/

REFERENCES

Adams J. J. et al., 2011, ApJ, 192, 5
Alavi A. et al., 2014, ApJ, 780, 143
Alavi A. et al., 2016, ApJ, 832, 56
Berry M. et al., 2012, ApJ, 749, 4
Blanc G. A. et al., 2011, ApJ, 736, 31
Bolton J. S., Haehnelt M. G., 2013, MNRAS, 429, 1695
Bouwens R. J., Illingworth G. D., Ellis R. S., Oesch P. A., Stefanon M., 2022, ApJ, 940, 55
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245
Caruana J., Bunker A. J., Wilkins S. M., Stanway E. R., Lorenzoni S., Jarvis M. J., Ebert H., 2014, MNRAS, 443, 2831
Caruana J. et al., 2018, MNRAS, 473, 30
Cassata P. et al., 2015, A&A, 573, 1
Chabrier G., 2003, PASP, 115, 763
Ciardullo R. et al., 2014, ApJ, 796, 64
Claeyssens A. et al., 2022, A&A, 666, A78
Curtis-Lake E. et al., 2012, MNRAS, 422, 1425
De Barros S. et al., 2017, A&A, 608, 123
De La Vieuville G. et al., 2020, A&A, 644, A39
Du X. et al., 2018, ApJ, 860, 75
Du X. et al., 2021, ApJ, 920, 95
Emami N., Siana B., Alavi A., Gburek T., Freeman W. R., Richard J., Weisz D. R., Stark D. P., 2020, ApJ, 895, 116
Endsley R., Stark D. P., Charlot S., Chevallard J., Robertson B., Bouwens R. J., Stefanon M., 2020, MNRAS, 000, 1
Erb D. K., Pettini M., Steidel C. C., Strom A. L., Rudie G. C., Trainor R. F., Shapley A. E., Reddy N. A., 2016, ApJ, 830, 52
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Gburek T. et al., 2019, ApJ, 887, 168
Gordon K. D., Clayton G. C., Misselt K. A., Landolt A. U., Wolff M. J., 2003, ApJ, 594, 279
Gronwall C. et al., 2007, ApJ, 667, 79
Guaita L. et al., 2010, ApJ, 714, 255
Haro P. A. et al., 2018, MNRAS, 478, 3740
Hashimoto T. et al., 2017, A&A, 608, A10
Hathi N. P. et al., 2016, A&A, 588, 26
Hayes M. et al., 2010, Nature, 464, 562
Hayes M. et al., 2014, ApJ, 782, 6
Hoag A. et al., 2019, ApJ, 878, 12
Jauzac M. et al., 2016, MNRAS, 457, 2029
Jullo E., Kneib J.-P., Limousin M., Elíasdóttir Á., Marshall P. J., Verdugo T., 2007, New J. Phys., 9, 447
Jung I. et al., 2020, ApJ, 904, 144
Konno A., Ouchi M., Nakajima K., Duval F., Kusakabe H., Ono Y., Shimasaku K., 2016, ApJ, 823, 20
Kornei K. A., Shapley A. E., Erb D. K., Steidel C. C., Reddy N. A., Pettini M., Bogosavljević M., 2010, ApJ, 711, 693
Kriek M., Van Dokkum P. G., Labbé I., Franx M., Illingworth G. D., Marchesini D., Quadri R. F., 2009, ApJ, 700, 221
Kusakabe H. et al., 2020, A&A, 638, A12
Le Fèvre O. et al., 2015, A&A, 576, A79
Leclercq F. et al., 2017, A&A, 608, A8
Limousin M. et al., 2007, ApJ, 668, 643
Limousin M. et al., 2016, A&A, 588, A99
Lujan Niemeyer M. et al., 2022, ApJ, 934, L26
Mason C. A., Treu T., Dijkstra M., Mesinger A., Trenti M., Pentericci L., de Barros S., Vanzella E., 2018, ApJ, 856, 2
Matsuda Y. et al., 2012, MNRAS, 425, 878
Matthee J., Sobral D., Oteo I., Best P., Smail I., Röttgering H., Paulino-Afonso A., 2016, MNRAS, 458, 449

Matthee J. et al., 2021, MNRAS, 505, 1382
McLean I. S. et al., 2010, in McLean I. S., Ramsay S. K., Takami H.eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference SeriesVol. 7735, Ground-based and Airborne Instrumentation for Astronomy III. p. 77351E
McLean I. S. et al., 2012, in McLean I. S., Ramsay S. K., Takami H.eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference SeriesVol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. p. 84460J
Mesinger A., Aykutalp A., Vanzella E., Pentericci L., Ferrara A., Dijkstra M., 2015, MNRAS, 446, 566
Oke J. B., Gunn J. E., 1983, ApJ, 266, 713
Oke J. B. et al., 1995, PASP, 107, 375
Ono Y. et al., 2012, ApJ, 744, 83
Partridge R. B., Peebles P. J. E., 1967, ApJ, 147, 868
Pentericci L. et al., 2011, ApJ, 743, 132
Priewe J., Williams L. L. R., Liesenborgs J., Coe D., Rodney S. A., 2017, MNRAS, 465, 1030
Prochaska J. X. et al., 2020a, pypeit/PypeIt: Release 1.0.0. Zenodo. Available at: https://doi.org/10.5281/zenodo.3743493
Prochaska J. X., Hennawi J. F., Westfall K. B., Cooke R. J., Wang F., Hsyu T., Davies F. B., Farina E. P., 2020b, preprint (arXiv:2005.06505)
Prochaska J. X. et al., 2020c, J. Open Source Softw., 5, 2308
Reddy N. A., Steidel C. C., 2009, ApJ, 692, 778
Reddy N. A., Steidel C. C., Pettini M., Adelberger K. L., Shapley A. E., Erb D. K., Dickinson M., 2008, ApJS, 175, 48
Reddy N. A. et al., 2022, ApJ, 926, 31
Schenker M. A., Stark D. P., Ellis R. S., Robertson B. E., Dunlop J. S., McLure R. J., Kneib J. P., Richard J., 2012, ApJ, 744, 179
Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, ApJ, 588, 65
Shivaei I. et al., 2018, ApJ, 855, 42
Sobral D., Matthee J., 2019, A&A, 623
Sobral D. et al., 2017, MNRAS, 466, 1242
Stark D. P., Ellis R. S., Chiu K., Ouchi M., Bunker A., 2010, MNRAS, 408, 1628
Stark D. P., Ellis R. S., Ouchi M., 2011, ApJ, 728, 2
Trainor R. F., Steidel C. C., Strom A. L., Rudie G. C., 2015, ApJ, 809, 89
Trainor R. F., Strom A. L., Steidel C. C., Rudie G. C., Chen Y., Theios R. L., 2019, ApJ, 887, 85
Wardlow J. L. et al., 2014, ApJ, 787, 9
Weiss L. H. et al., 2021, ApJ, 920, 95
Zhang Y. et al., 2021, ApJ, 922, 167

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