The Origin of the Observed Lyα EW Distribution of Dwarf Galaxies at $z \sim 2$

Christopher Snapp-Kolas,1,* Brian Siana,1, Timothy Gburek1, Anahita Alavi2, Najmeh Emami3,4
Johan Richard5,6 Daniel P. Stark5 Claudia Scarlata6,7

1Department of Physics & Astronomy, University of California, Riverside, CA 92521, US
2IPAC, California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA
3Minnesota Institute for Astrophysics, University of Minnesota, Minneapolis, MN, 55455, USA
4Univ Lyon, Univ Lyon1, ENS de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230, Saint-Genis-Laval, France
5Steward Observatory, University of Arizona, 933 N Cherry Ave, Tucson, AZ 85721, USA
6Minnesota Institute for Astrophysics, University of Minnesota, Minneapolis, MN, 55455, USA

Accepted XXX. Received YYY; in original form ZZZ

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 ($-19 \leq M_{UV} \leq -17$) at this redshift. The fraction of dwarf galaxies identified as Lyα emitters ($EW \geq 20$) is $X_{LAE} = 25^{+15}_{-10} \%$. We use the Balmer lines and UV continuum to estimate the intrinsic EW allowing us to distinguish the effects of the ionizing spectra and escape fraction on the observed EW distribution. Fainter galaxies ($M_{UV} > -19$) show larger intrinsic EWs and escape fractions than brighter galaxies. Only galaxies with intrinsic EWs greater than 40 have escape fractions larger than 0.05. We find an anti-correlation between the escape fraction and $A_V$ as well as UV spectral slope. The volumetric escape fraction of our sample is $\int_{esc}^{Ly\alpha} = 4.59^{+2.0}_{-1.4} \%$ in agreement with measurements found elsewhere in the literature. About half of the total integrated Lyα luminosity density comes from galaxies with $EW_{obs} > 20$ Å.

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

1 INTRODUCTION

More than five decades ago, Lyα was proposed as a tracer of early stage star forming galaxies (Partridge & Peebles 1967). Since then Lyα has been used to study star-forming galaxies locally and at the highest redshifts observed (i.e. 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 the most distant star-forming galaxies (Caruana et al. 2014; Hoag et al. 2019; Endsley et al. 2020; Jung et al. 2020).

Lyα 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, Lyα can be easily absorbed by dust in the interstellar medium (ISM) reducing the strength of the line. Lyα 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 Lyα manages to escape. Each of these effects can drastically hamper the observability of Lyα and have strong effects on the distribution of observed Lyα equivalent widths (EWs) (Stark et al. 2010, 2011). In fact, at high redshift the attenuation of Lyα 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$_{Lya}$ to compare with the observed distribution. The evolution of the EW$_{Lya}$ distribution has been used to study the end of reionization and is well understood from $z \sim 3$ and for “bright” galaxies with absolute UV magnitudes $M_{UV} < -19$ (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; Haro et al. 2018; Caruana et al. 2018; Fuller et al. 2020; 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 ($M_{UV} < -19$) (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 ($M_{UV} > -19$). However, according to 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$_{Lya}$ distribution at higher redshift where observation of Lyα is more difficult.

Lyα studies will either select galaxies via their rest-UV continuum luminosity density (via broadband imaging), or their Lyα line emission (via narrowband imaging, Berry et al. 2012; Konno et al.
2016; Sobral et al. 2017; Sobral & Matthee 2019; Hashimoto et al. 2017). Narrow-band selected samples will be biased towards Ly\(\alpha\) emitters (LAEs) as it requires that the line be significantly detected in emission. Alternatively, narrow-band studies of Ly\(\alpha\) often select on other emission lines in the optical (such as [OIII]5007) (i.e. Ciardullo et al. 2014; Trainor et al. 2015; Weiss et al. 2021). Erb et al. (2016) show that [OIII]5007-selected galaxies are biased towards LAEs by comparing the fraction of LAEs (\(X_{[\text{OIII}]5007}\)) in their [OIII]5007 selected sample with a sample of UV-selected galaxies. 80% are found to be LAEs in their [OIII]5007 sample while only 9% are found to be LAEs in their UV selected sample. This is further confirmed in Trainor et al. (2019); Weiss et al. (2021), indicating that samples selected on nebular emission lines are biased toward Ly\(\alpha\) emission. A sample chosen irrespective of the strength of nebular emission lines would be useful for understanding the full distribution of Ly\(\alpha\) 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 \(R \sim 30\), but the followup spectroscopy necessary for accurate escape fraction and EW analyses limit the observations to \(R < 25.5\) (i.e Reddy et al. 2008) in order to obtain sufficient S/N in the continuum.

The Ly\(\alpha\) EW distribution is dependent on the escape fraction of Ly\(\alpha\) photons. Past studies have measured the Ly\(\alpha\) escape fraction by comparing the Ly\(\alpha\) 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 narrowband imaging of Ly\(\alpha\) and one of the Balmer lines (Trainor et al. 2015) to make direct measurements of the Ly\(\alpha\) escape fraction. Narrow-band imaging is biased towards high Ly\(\alpha\) EW galaxies, and therefore does not represent the general population of star-forming galaxies as evidenced by the rather high average Ly\(\alpha\) escape fractions of 30% (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 (i.e. 5.3%, Hayes et al. 2010). Some try to model the escape fraction as a function of Ly\(\alpha\) EW (i.e. Sobral & Matthee 2019), but this method must assume a star formation history and intrinsic ionization continuum which are uncertain for low-mass galaxies. A direct measurement from a UV-selected sample with spectroscopic followup avoids these uncertainties.

What is needed is a measure of Ly\(\alpha\) emission properties of a UV-selected sample of faint galaxies. We address this by selecting galaxies with photometric redshifts 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, \(f_{\text{esc}}\), ionization parameter, etc.) of these dwarf galaxies we obtained Keck/MOSFIRE spectra as described in Gburek et al. (2019). We also obtained Keck/LRIS optical spectra, for which the details are listed in Table 1. We observe eleven masks in three clusters with LRIS. Exposure times varied greatly (from 4500s-18000s) depending on the conditions and priority of objects in the mask, and seeing was typically \(\sim 1''\). We decreased read noise by binning by 2 in the spectral direction of the CCD.

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 Ly$_\alpha$ to obtain an estimate of the total flux, but refrain from applying the correction when displaying spectra.

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 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 1D spectra is 127. The clusters are about 3'x3' which 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 0.135 pix$^{-1}$, however we have binned along the wavelength direction such that the effective plate scale is 0.27 pix$^{-1}$. The wavelength spacing is $\sim 2.18$ pix$^{-1}$. So the resolution of our observations is $R \sim 500$ or $\Delta \nu \sim 600$ km s$^{-1}$. We show some example 1D spectra in figure 1.

2.2 Sub-Sample Selection

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

2.2.1 $X_{LAE}$ sample

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

- We require secure redshifts for all the galaxies in this sample so that galaxies with no Ly$_\alpha$ or Ly$_\beta$ in absorption can be identified. Therefore, only galaxies with confirmed redshifts from our MOSFIRE data are kept. These redshifts are primarily determined by H$_\alpha$ and [OIII]$_\lambda$5007.
- We remove galaxies which do not have spectroscopic coverage of Ly$_\alpha$. Atmospheric absorption at $\lambda < 3200$ Å limits our analysis to galaxies at $z \geq 1.6$.
- 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.
- 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.

Figure 1. Three example spectra from our sample of galaxies. The red line is the fit for each spectrum. The fitting method is described in section 3. Top: An example of an "emission" spectrum where we have Ly$_\alpha$ 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$_\alpha$.
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 slit loss-corrected flux will reflect the true spectrum. Most of the galaxies have slit losses < 2.2 and we thus set this as our upper limit. The slit loss distributions are shown in figure 2.

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

2.2.2 H\textalpha sample

In order to get a measurement of the Ly\textalpha escape fraction, H\textalpha must be observable. Because of this we create a second sub-sample of galaxies from the X$_{\text{LAE}}$ sub-sample for which H\textalpha could have been observed, which limits the redshift range to $z \leq 2.6$. Additionally, the rest-UV HST data are very deep and the KECK H\textalpha data are shallower which could bias us against galaxies that are faint in H\textalpha. To check whether we are excluding galaxies that are too faint in H\textalpha we observe that our bright M$_{\text{UV}}$ sample lies almost entirely above log(L$_{\text{H\alpha}}$/L$_{\text{UV}}$) $\approx$ 13.4, as shown in figure 3, and therefore set this as our completeness limit for the sample. We take the median H\textalpha $= 4.3 \times 10^{-18}$ erg s$^{-1}$ cm$^{-2}$ line flux error multiplied by 3 as our 3$\sigma$ limit for detection of H\textalpha. We then check our galaxies magnification and absolute UV magnitude and ask whether the galaxy would have been observable at log(L$_{\text{H\alpha}}$/L$_{\text{UV}}$) $\approx$ 13.4. If the galaxy would not have been observable at 3$\sigma$ our completeness limit it was then removed. No galaxy fell below this criterion and therefore none were removed as shown in figure 3. We split our sample at M$_{\text{UV}}$ = −19 and measure the mean and error on the mean of the bright and faint samples to be log(L$_{\text{H\alpha}}$/L$_{\text{UV}}$) = 13.8 $\pm$ 0.003 and log(L$_{\text{H\alpha}}$/L$_{\text{UV}}$) = 13.5 $\pm$ 0.002 respectively. While all but one of our galaxies have log(L$_{\text{H\alpha}}$/L$_{\text{UV}}$) $>$ 13.4, 11 of the 18 galaxies fainter than M$_{\text{UV}}$ = −19 could have been observed below this line, given their magnifications. This suggests that there are indeed very few galaxies with log(L$_{\text{H\alpha}}$/L$_{\text{UV}}$) $<$ 13.4 and that our measured increase of log(L$_{\text{H\alpha}}$/L$_{\text{UV}}$) for faint galaxies is real, and not due to a bias against galaxies with low log(L$_{\text{H\alpha}}$/L$_{\text{UV}}$).

Our final H\textalpha sample after all of these considerations is 23 galaxies.
spectrum for our galaxy sample. The fitting procedure is performed using a Markov chain Monte Carlo code from pymc3. The wavelength range used for fitting varied based on the redshift as the short-wavelength bound was set to just blueward of the \( \text{Ly}\alpha \) 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 5600 Å. Typical emission and absorption lines (Si\(\text{II}\)λ11260 Å, O\(\text{I}\)λ1303 Å, CIIλ1334 Å, SiIVλ1393 Å, CIVλ1549 Å, HeIIλ1640 Å, OIIIλ1666 Å, CIIλ1909 Å, and OIVλ1343 Å) were masked when fitting each spectrum. We set the width for these masks to be 800 km s\(^{-1}\) except for the CIVλ1549 Å line which we set to have an asymmetric width of 1400 km s\(^{-1}\) with the blue side of the line extending to \( \approx 1000 \) km s\(^{-1}\) because of the P-Cygni profile that CIVλ1549 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,  

\[
f_1 = A_0 e^{-\frac{(\lambda - \lambda_0)^2}{2\sigma^2}} + B_0 \left( \frac{\lambda}{\lambda_s} \right)^\beta
\]

(1)

For the ‘combination’ spectra we add in a first order approximation to the Ly\(\alpha\) absorption profile,  

\[
f_1 = \begin{cases} 
    A_0 e^{-\frac{(\lambda - \lambda_0)^2}{2\sigma^2}} + a \lambda + b, & \lambda < \lambda_s \\
    B_0 \left( \frac{\lambda}{\lambda_s} \right)^\beta, & \lambda > \lambda_s 
\end{cases}
\]

(2)

Where \( \lambda_s \) is the wavelength where the linear approximation to the Voigt profile meets the UV continuum. The remaining parameters are powered with a power law continuum redward of Ly\(\alpha\) (identified either from damped Ly\(\alpha\) absorption or from the redshift estimate obtained with Keck/MOSFIRE),  

\[
f_1 = A_0 \left( \frac{\lambda}{\lambda_0} \right)^\beta
\]

(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\(\alpha\) in emission. 4.9 Å is greater than all but two of 49 (4%) of the FWHMs measured and therefore serves as an upper limit on the FWHM. We determine the Ly\(\alpha\) EW and flux by numerical integration around Ly\(\alpha\) using the methods of Du et al. (2021) and then divide by the continuum level extrapolated to rest-frame 1216 Å to determine the EW.

### 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 ??). Our galaxies are observed behind three lensing clusters. In the X\(\text{LAE}\) and H\(\alpha\) sub-samples the magnifications span 1.25 \( \leq \mu \leq 29.12 \). The median magnifications are \( \langle \mu \rangle_{\text{median}} = 7.11 \) and \( \langle \mu \rangle_{\text{median}} = 7.45 \) 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) collabo-

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).

### 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% flux error to all of our hands to account for systematic errors in our photometry (Alavi et al. 2016). Using the code FAST3 (Kriek et al. 2009), we fit Bruzual & Charlot (2003) stellar population synthesis models to the emission-line subtracted photometry under the following assumptions:

- Constant star formation histories (SFHs)
- A Chabrier (2003) initial mass function (IMF)
- Stellar metallicities of either 0.2 Z\(_{\odot}\) or 0.4 Z\(_{\odot}\)
- A Gordon et al. (SMC 2003) dust attenuation curve

Galaxy redshifts are set by the fit curves to the spectroscopy.

### 3.5 Dust Correction

In our analysis of the H\(\alpha\) sample we need to correct the H\(\alpha\) 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. The correction for H\(\alpha\) is estimated by direct 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 observe that the Balmer decrement is well determined for three of our five brightest galaxies, while our fainter galaxies are less certain. Therefore, We split our H\(\alpha\) sample into bright (M\(_{UV} < -19\)) and faint (M\(_{UV} > -19\)) sub-samples for which we estimate \( A_V \). For the faint sub-sample we normalize the spectra by H\(\alpha\) 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, 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\(\beta\) S/N > 5 and therefore could not reliably calculate the dust correction.

### 4 RESULTS

As explained above, our larger X\(\text{LAE}\) sample is used to measure the EW distribution and the smaller H\(\alpha\) sub-sample is used to measure the escape fraction. We compare our two samples in order to see if the two are similar and therefore both representative of dwarf galaxies at this redshift. We find that the median EWS are consistent with

---

1. [https://archive.stsci.edu/prepds/frontier/lensmodels/](https://archive.stsci.edu/prepds/frontier/lensmodels/)
2. [https://projets.lam.fr/projects/lenstool/wiki](https://projets.lam.fr/projects/lenstool/wiki)
3. [https://www.astro.berkeley.edu/~mariska/FAST.html](https://www.astro.berkeley.edu/~mariska/FAST.html)
The cutoff for defining a 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 cutoffs we measure $X_{LAE} = 25_{-10}^{+15} \%$ (6/24), $25_{-10}^{+15} \%$ (6/24), and $17_{-8}^{+13} \%$ (4/24), respectively. We assume a Poisson distribution to estimate uncertainties on $X_{LAE}$. $X_{LAE}$ measurements of brighter galaxies are $\sim 11%$ for $EW > 20$ Å at $z \sim 2$ (12% from Reddy et al. (2008), 11.1% from Hathi et al. (2016), and 10.7% Du et al. (2021)). For brighter galaxies with $EW > 25$ Å and $EW > 55$ Å at $z \sim 2 X_{LAE} = 10%$ and 4% respectively (Cassata et al. 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 Hα sample for which we are able to measure the escape fraction of Lyα 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 Lyα to the flux of Hα 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 (i.e. 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 Lyα flux and $F_{H\alpha, cor}$ is the dust corrected Hα flux. The mean of the individual $f_{esc}$ measurements is $4.3^{+1.6}_{-1.0} \%$. 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 100000 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 photons escaping into the intergalactic medium. The intrinsic luminosity of galaxies varies greatly and may correlate with Lyα 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, deredshift each of our galaxies. We then replace the fluxes in $f_{esc}$ with the sum of the respective luminosities. We measure a volumetric escape fraction of $4.59^{+2.0}_{-1.4} \%$. 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% ± 3.8% in Hayes et al. (2010), 4.4$^{+2.1}_{-1.2} %$ in Ciardullo et al. (2014), 5.1$^{+0.2}_{-0.1} %$ in Sobral et al. (2017), and 5.8$^{+0.7}_{-0.5} %$ 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 Hα and Lyα luminosities and have a greater sample of joint...
Figure 6. The intrinsic equivalent width as a function of absolute UV magnitude ($M_{UV}$). Points are color coded according to the observed equivalent width (see color 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 $\xi_{ion}$ measurements of Shivaei et al. (2018). Galaxies brighter than -19 in absolute UV magnitude are plotted with an x, 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 black. Five galaxies did not have sufficient photometric coverage to perform SED fitting and are removed from any intrinsic EW analysis.

Figure 7. The Ly$\alpha$ escape fraction as a function of absolute UV magnitude ($M_{UV}$). The color coding and markers are the same as in figure 6. The mean escape fraction of the dwarf galaxies (4.3$^{+1.7}_{-1.1}$%) and the more massive galaxies (0.5$^{+0.3}_{-0.2}$%) are displayed with shaded black regions denoting the propagated uncertainty in the individual galaxies. However, the bright galaxy sample has only 5 galaxies and is therefore uncertain.

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.

5.1 Intrinsic Ly$\alpha$ EW

To investigate influences from the stellar populations we make use of the intrinsic Ly$\alpha$ EW

$$EW_{int} = \frac{8.7L_{H\alpha, cor}}{L_{L,1216, cor}}$$

(5)

where $L_{H\alpha, cor}$ is the dust-corrected and de-magnified H$\alpha$ luminosity, and $L_{L,1216, cor}$ is the dust corrected and de-magnified UV luminosity density extrapolated to 1216 Å. Figure 6 shows the intrinsic EW as a function of absolute UV magnitude.

We compare our measured EW$_{int}$ values to those of more massive galaxies at the same redshift. We convert the ionizing photon production efficiency ($\xi_{ion}$) of Shivaei et al. (2018) to EW$_{int}$ by converting $\xi_{ion}$ to $L_{Ly\alpha}/L_{UV}$. In Shivaei et al. (2018) they calculate $\xi_{ion}$ from weighted mean stacks of their galaxies in bins of absolute UV magnitude. They are shown as gold squares in figure 6. Shivaei et al. (2018) show a mean log($\xi_{ion}$) = 25.36 ± 0.06 for the SMC curve. This corresponds to an intrinsic EW of 23 ± 3 Å. Emami et al. (2020) have $\xi_{ion} = 25.22 ± 0.10$ 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 toward higher EW$_{int}$. We also show in figure 6 the observed EWs via the colorbar where blue and purple points indicate LAEs. From the plot we observe large intrinsic EWs ($EW_{int} > 40$ 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 $f_{esc}$

The other factor driving the observed distribution in Ly$\alpha$ 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$\alpha$ sample we set $f_{esc} = 0$. We show the escape fraction as a function of absolute UV magnitude in figure 7.

We see large scatter in the escape fraction at faint absolute UV magnitudes. It is also the case that no escape fraction less than ~10% is observed as a Ly$\alpha$ emitter. Therefore, it is clear that the escape of ionizing photons plays a crucial role in the observed EW distribution and $X_{LAE}$.

To determine if the escape fraction and EW$_{int}$ are correlated, we plot them in figure 8.

We see that 4/9 of the galaxies with high intrinsic EWs (≥ 40 Å) have large escape fractions ($f_{esc} > 0.05$). No galaxy (0/8) with EW$_{int} < 40$ Å has an escape fraction larger than 0.05. We can set two conditions for LAEs; LAEs have escape fractions greater than 5% and intrinsic EWs greater than 40 Å. We caution that our sample detections of H$\alpha$ and Ly$\alpha$ emitters than what was observed in Hayes et al. (2010). Furthermore, our sample is not biased towards emitters as in Ciardullo et al. (2014), Sobral et al. (2017), and Weiss et al. (2021). Nevertheless, the volumetric escape fraction in each study is consistent with our measurement, and all within 4%-6%.
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.

Lya 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. Figure 9 shows the escape fraction as a function of AV. We observe an anti-correlation between fesc and AV. However, nearly all of our galaxies have AV ≤ 0.2. With little variance in dust content we aren’t able to test this robustly. Nevertheless, we don’t observe escape fractions greater than 0.05 for AV ≥ 0.2.

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 figure 10. Here we see a mild dependence in that the scatter of escape fractions increases for β < −0.5. Together, these plots suggest that there is a slight anti-correlation between dust and fesc for faint galaxies. This is consistent with figure 9 which shows an anti-correlation between fesc and AV.

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, i.e.). However, in order to investigate the number of faint galaxies that contribute to reionization it is helpful to establish trends in X_{LAE} with M_{UV} at lower redshifts where Lya is observable. To this end we compare with the VUDS sample (Le Fèvre et al. 2015; Hathi et al. 2016; Cassata et al. 2015) at higher UV luminosities. Figure 11 shows the VUDS data set in bins of M_{UV} along with our data set. We find no trend in the VUDS data for brighter M_{UV}, so we fit a constant line to the VUDS data set and find X_{LAE} = 0.106 ± 0.011. This is 1.4σ from our result of X_{LAE} = 0.25^{+0.15}_{-0.10}. More data between the VUDS faintest M_{UV} and our sample is necessary to determine any trend in X_{LAE} with M_{UV}.

We also investigate trends in redshift using the VUDS data set as shown in figure 12. We fit each redshift bin with a constant value since we again see no trend with M_{UV} at any redshift. There is large variance in beta for similar X_{LAE} values of M_{UV}. Together, these plots suggest that there is a slight anti-correlation between dust and fesc for faint galaxies. This is consistent with figure 9 which shows an anti-correlation between fesc and AV.
5.4 Observed Ly\textalpha{} EW

We investigate which galaxies are emitting most of the Ly\textalpha{} luminosity. Figure 13 shows that about 50\% of the integrated Ly\textalpha{} luminosity density comes from galaxies with EW\textsubscript{Ly\textalpha{}} > 20 \textmu{}m. This suggests that LAEs are contributing a large fraction of the Ly\textalpha{} photons. This cumulative distribution can also be calculated from narrow-band surveys for Ly\textalpha{}. However, those surveys can not determine the amount of total emissions from low EW galaxies. Here we show that ~10\% of the Ly\textalpha{} luminosity density EW\textsubscript{rest} < 5\textmu{}m, however, because of the size of our sample, this result is quite uncertain.

6 SUMMARY

In this paper we have investigated the distribution of Ly\textalpha{} 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\textsubscript{UV} < -17, $\langle M^*_{\text{stellar}} \rangle = 10^{8.4} M_{\odot}$) observed at $z \sim 2$. We reduce our sample to avoid issues with differential magnification, large slit loss uncertainties, incompleteness due to faint UV luminosity, and H\textalpha{} within the bands of Keck/LRIS and Keck/MOSFIRE. We then analyze the EW distribution of the sample and draw the following conclusions:

- The observed EW distribution of low UV luminosity galaxies is skewed towards larger EWs than higher luminosity galaxies such as those in the KBSS sample of Du et al. (2021). Our sample shows a median EW of 5.8 \textmu{}m whereas the sample of Du et al. (2021) shows a median EW of 6.0 \textmu{}m.

- The fraction of galaxies that are Ly\textalpha{} emitters X\textsubscript{LAE} is 25\% for EW > 20 \textmu{}m, 17\% for EW > 25 \textmu{}m, and 10\% for EW > 55 \textmu{}m respectively. These values are greater than X\textsubscript{LAE} in higher luminosity samples such as Reddy et al. (2008); Hathi et al. (2016); Cassata et al. (2015); Du et al. (2021).

- We investigate possible trends in X\textsubscript{LAE} with M\textsubscript{UV} and redshift using the VUDS data set (Le Fèvre et al. 2015; Hathi et al. 2016; Cassata et al. 2015). We find no trend with M\textsubscript{UV} between $-22 \leq M_{\text{UV}} \leq -19.5$, but find our X\textsubscript{LAE} at $M_{\text{UV}} \sim -18$ to be ~15\% greater at 1.4\sigma{} significance. There is a trend towards larger X\textsubscript{LAE} with redshift in the VUDS data. Further investigation of dwarf galaxies at higher redshift could show increased X\textsubscript{LAE} relative to higher mass galaxies as well.

- We find that the total integrated Ly\textalpha{} luminosity is about 50\% for galaxies with EW\textsubscript{Ly\textalpha{}} > 20 \textmu{}m, suggesting that LAEs contribute a large fraction of the Ly\textalpha{} photons. However, because of small numbers this value is uncertain.

- We estimate the intrinsic Ly\textalpha{} EW of galaxies, and show an increase in the mean EW\textsubscript{int} for faint galaxies (EW) = 67 $\pm$ 8 compared to the brighter sample of Shivaei et al. (2018). (EW) = 23 $\pm$ 3) and the intermediate luminosity sample of Emami et al.
(2020, ⟨EW⟩ = 22 ± 5). 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}.

- We investigate the escape fraction of Lyα 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 have f_{esc} > 0.05. 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.

- We find an anti-correlation between f_{esc} and both ΔV and β. This implies that low dust-content facilitates the escape of Lyα photons.

- We observe a global volumetric escape fraction of f_{esc} = 4.59^{+2.01}_{-1.44} % in our sample, in good agreement with values found for other faint UV luminosity samples in the literature (i.e. Hayes et al. 2010; Weiss et al. 2021, etc.).

We were able to disentangle to some extent two primary drivers of the Lyα 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.

ACKNOWLEDGMENTS

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

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.

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, #14209.

DATA AVAILABILITY

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

REFERENCES

Alavi A., et al., 2014, ApJ, 780, 143
Alavi A., et al., 2016, ApJ, 832, 56
Berry M., et al., 2012, ApJ, 749
Blanc G. A., et al., 2011, ApJ, 736, 31

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

Bolton J. S., Haehnelt M. G., 2013, MNRAS, 429, 1695
Bouwens R. J., Illingworth G. D., Ellis R. S., Oesch P. A., Stehanon M., 2022, arXiv e-prints, p. arXiv:2205.11526
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
Curtis-Lake E., et al., 2012, MNRAS, 422, 1425
De Barros S., et al., 2017, A&A, 608, 123
De La Veyrie 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
Fuller S., et al., 2020, ApJ, 896, 156
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
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
Julio E., Kneib J.-P., Limousin M., Eliasdóttir Á., Marshall P. J., Verdugo T., 2007, New Journal of Physics, 9, 447
Jung I., et al., 2020, ApJ, 904, 144
Kron 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., Marchesi D., Quadri R. F., 2009, ApJ, 700, 221
Kusakabe M., et al., 2020, A&A, 638, A12
Le Fèvre O., et al., 2015, A&A, 576, A79
Limousin M., et al., 2016, A&A, 588, A99
Limousin M., et al., 2007, ApJ, 668, 643
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
Mathee J., Sobral D., Oteo I., Best P., Smail I., Röttgering H., Paulino-Afonso A., 2016, MNRAS, 458, 449
Mathee J., et al., 2021, MNRAS, 000, 1
Mesinger A., AYKUTALP A., Vanzella E., Pentericci L., Ferrara A., Dijkstra M., 2015, MNRAS, 446, 566
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
Prochaska J. X., et al., 2020a, pypeit/PypeIt: Release 1.0.0, doi:10.5281/zenodo.3743493
Prochaska J. X., Hennawi J. F., Westfall K. B., Cooke R. J., Wang F., Hsu T., Davies F. B., Farina E. F., 2020b, arXiv e-prints, p. arXiv:2005.06505
Prochaska J. X., et al., 2020c, Journal of Open Source Software, 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, ApJ Supplement Series, 175, 48

MNras 000, 1–11 (2022)
This paper has been typeset from a \TeX/\LaTeX\ file prepared by the author.