Predicting Lyα escape fractions with a simple observable[*]

Lyα in emission as an empirically calibrated star formation rate indicator

David Sobral[1,2][**] and Jorryt Matthee[2]

[1] Department of Physics, Lancaster University, Lancaster, LA1 4YB, UK
[2] Leiden Observatory, Leiden University, P.O. Box 9513, NL-2300 RA Leiden, The Netherlands

March 28, 2018

ABSTRACT

Lyman-α (Lyα) is intrinsically the brightest line emitted from active galaxies. While it originates from many physical processes, for star-forming galaxies the intrinsic Lyα luminosity is a direct tracer of the Lyman-continuum (LyC) radiation produced by the most massive O- and early-type B-stars ($M_\star \gtrsim 10 \, M_\odot$) with lifetimes of a few Myrs. As such, Lyα luminosity should be an excellent instantaneous star formation rate (SFR) indicator. However, its resonant nature and susceptibility to dust as a rest-frame UV photon makes Lyα very hard to interpret due to the uncertain Lyα escape fraction, $f_{esc,Ly\alpha}$. Here we explore results from the CAlibrating LYMan-α with Hα (CALYMHA) survey at $z = 2.2$, follow-up of Lyα emitters (LAEs) at $z = 2.2 - 2.6$ and a $z \sim 0 - 0.3$ compilation of LAEs to directly measure $f_{esc,Ly\alpha}$ with Hα. We derive a simple empirical relation that robustly retrieves $f_{esc,Ly\alpha}$ as a function of Lyα rest-frame EW (EW$_0$): $f_{esc,Ly\alpha} = 0.0048 \, \mathrm{EW}_0[\text{Å}] \pm 0.05$ and we show that the relation is driven by a tight sequence between high ionisation efficiencies and low dust extinction in LAEs. Observed Lyα luminosities and EW$_0$ are easy measurable quantities at high redshift, thus making our relation a practical tool to estimate intrinsic Lyα and LyC luminosities under well controlled and simple assumptions. Our results allow observed Lyα luminosities to be used to compute SFRs for LAEs at $z \sim 0 - 2.6$ within ± 0.2 dex of the Hα dust corrected SFRs. We apply our empirical SFR(Lyα,EW$_0$) calibration to several sources at $z \geq 2.6$ to find that star-forming LAEs have SFRs typically ranging from 0.1 to 20 $M_\odot$ yr^{-1} and that our calibration might be even applicable for luminous LAEs within the epoch of re-ionisation. Our results imply higher than canonical ionisation efficiencies and low dust content in LAEs across cosmic time, and will be easily tested with future observations with *JWST* which can obtain Hα and Hβ measurements for high-redshift LAEs.

Key words. Galaxies: star formation, starburst, evolution, statistics, general, high-redshift; Ultraviolet: galaxies.

1. Introduction

With a vacuum rest-frame wavelength of 1215.67 Å, the Lyman-α (Lyα) recombination line ($n = 2 \rightarrow n = 1$) plays a key role in the energy release from ionised hydrogen gas, being intrinsically the strongest emission line in the rest-frame UV and optical (e.g. Partridge & Peebles 1967; Pritchet 1994). Lyα is emitted from ionised gas around star-forming regions (e.g. Charlot & Fall 1993; Pritchet 1994) and AGN (e.g. Miley & De Breuck 2008) and it is routinely used as a way to find high redshift sources ($z \sim 2 - 7$; see e.g. Malhotra & Rhoads 2004).

Several searches for Lyα-emitting sources (Lyα emitters; LAEs) have led to samples of thousands of star-forming galaxies (SFGs) and AGN (e.g. Sobral et al. 2018b, and references therein). LAEs are typically faint in the rest-frame UV, including many that are too faint to be detected by continuum based searches even with the *Hubble Space Telescope* (e.g. Bacon et al. 2015). The techniques used to detect LAEs include narrow-band surveys (e.g. Rhoads et al. 2000; Ouchi et al. 2008; Hu et al. 2010; Matthee et al. 2015), Integral Field Unit (IFU) surveys (e.g. van Breukelen et al. 2005; Drake et al. 2017) and blind slit spectroscopy (e.g. Martin & Sawicki 2004; Rauch et al. 2008; Cassata et al. 2011). Galaxies selected through their Lyα emission allow for easy spectroscopic follow-up due to their high

EWs (e.g. Hashimoto et al. 2017) and typically probe low stellar masses (see e.g. Gawiser et al. 2007; Hagen et al. 2016).

The intrinsic Lyα luminosity is a direct tracer of the ionising Lyman-continuum (LyC) luminosity and thus a tracer of instantaneous star formation rate (SFR), in the same way as Hα is (e.g. Kennicutt 1998). Unfortunately, inferring intrinsic properties of galaxies from Lyα observations is extremely challenging. This is due to the complex resonant nature and sensitivity to dust of Lyα (see e.g. Dijkstra 2017, for a detailed review on Lyα), which contrasts with Hα. For example, a significant fraction of Lyα photons is scattered in the Inter-Stellar Medium (ISM) and in the Circum-Galactic Medium (CGM) as evidenced by the presence of extended Lyα halos in LAEs (e.g. Momose et al. 2014; Wisotzki et al. 2016), but also in the more general population of $z \sim 2$ SFGs sampled by Hα emitters (Matthee et al. 2016), and the bluer component of such population traced by UV-continuum selected galaxies (e.g. Steidel et al. 2011). Such scattering leads to kpc-long random-walks which take millions of years and that significantly increase the probability of Lyα photons being absorbed by dust particles. The complex scattering and consequent higher susceptibility to dust absorption typically leads to low and uncertain Lyα escape fractions ($f_{esc,Ly\alpha}$; the ratio between observed and intrinsic Lyα luminosity; see e.g. Atek et al. 2008).

"Typical" star-forming galaxies at $z \sim 2$ have low $f_{esc,Ly\alpha}$ ($\sim 1 - 5\%$; e.g. Oteo et al. 2015; Cassata et al. 2015), likely because significant amounts of dust present in their ISM easily absorb Lyα photons (e.g. Ciardullo et al. 2014; Oteo et al. 2015;

Oyarzún et al. 2017). However, sources selected through their Lyα emission typically have ~ 10 times higher $f_{esc,Ly\alpha}$ (e.g. Song et al. 2014; Sobral et al. 2017), with Lyα escaping over ≈ 2× larger radii than Hα (Sobral et al. 2017).

Furthermore, one expects $f_{esc,Ly\alpha}$ to depend on several physical properties which could be used as predictors of $f_{esc,Ly\alpha}$. For example, $f_{esc,Ly\alpha}$ anti-correlates with stellar mass (e.g. Oyarzún et al. 2017), dust attenuation (e.g. Verhamme et al. 2008; Hayes et al. 2011; Matthee et al. 2016; An et al. 2017) and SFR (e.g. Matthee et al. 2016). However, most of these relations require derived properties (e.g. Yang et al. 2017), show a large scatter, may evolve with redshift and sometimes reveal complicated trends (e.g. dust dependence; see Matthee et al. 2016).

Interestingly, the Lyα rest-frame equivalent width (EW_0), a simple observable, seems to be the simplest direct predictor of $f_{esc,Ly\alpha}$ in LAEs (Sobral et al. 2017; Verhamme et al. 2017) with a relation that shows no strong evolution from $z \sim 0$ to $z \sim 2$ (Sobral et al. 2017) and that might be applicable at least up to $z \sim 5$ (Harikane et al. 2017). Such empirical relation may hold the key for a simple but useful calibration of Lyα as a direct tracer of the intrinsic LyC luminosity by providing a way to estimate $f_{esc,Ly\alpha}$, and thus as a good SFR indicator for LAEs (see also Dijkstra & Westra 2010). We fully explore such possibility and its implications in this work. In §2 we present the samples at different redshifts and methods used to compute $f_{esc,Ly\alpha}$. In §3 we present and discuss the results, their physical interpretation and our proposed empirical calibration of Lyα as a SFR indicator. Finally, we present the conclusions in §4. We adopt a flat cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70\,km\,s^{-1}\,Mpc^{-1}$.

2. Sample and Methods

2.1. LAEs at low redshift ($z \leq 0.3$)

For our lower redshift sample, we explore a compilation of 30 sources presented in Verhamme et al. (2017) which have accurate (Hα derived) $f_{esc,Ly\alpha}$ measurements and sample a range of galaxy properties. The sample includes high EW Hα emitters (HAEs) from the Lyman Alpha Reference Sample at $z = 0.02 - 0.2$ (LARS, e.g. Hayes et al. 2013, 2014), a sample of LyC leakers (LyCLs) investigated in Verhamme et al. (2017) at $z \sim 0.3$ (Izotov et al. 2016a,b) and a more general 'green pea' (GPs) sample (e.g. Cardamone et al. 2009; Henry et al. 2015; Yang et al. 2016, 2017). These are all LAEs at low redshift with available Lyα, Hα and dust extinction information required to estimate $f_{esc,Ly\alpha}$ (see §2.4) and for which Lyα EW_0s are available. For more details on the sample, see Verhamme et al. (2017) and references therein.

2.2. LAEs at cosmic noon ($z = 2.2 - 2.6$)

For our sample at the peak of star formation history we use 188 narrow-band selected LAEs with Hα measurements from the CALYMHA survey at $z = 2.2$ (Matthee et al. 2016; Sobral et al. 2017) presented in Sobral et al. (2017), for which $f_{esc,Ly\alpha}$ measurements are provided as a function of EW_0. In addition, we explore spectroscopic follow-up of CALYMHA sources with X-SHOOTER on the VLT (Sobral et al. 2018a) and individual measurements for four sources (CALYMHA-67, -93, -147 and -373; see Sobral et al. 2018a). For those sources we measure Lyα, Hα and Hβ. Furthermore, we also use a sample of 29 narrow-band selected LAEs at $z \sim 2.6$ presented by Trainor et al. (2016), for which Lyα and Hα measurements are available.

2.3. Higher redshift LAEs ($2.6 \leq z \leq 6$)

As an application of our results, we explore the publicly available sample of 3,908 LAEs in the COSMOS field (SC4K survey; Sobral et al. 2018b) which provides Lyα luminosities and rest-frame EWs for all LAEs. We also explore published median or average values for the latest MUSE samples, containing 417 LAEs (e.g. Hashimoto et al. 2017). Note that for all these higher redshift samples, Hα is not directly available, thus $f_{esc,Ly\alpha}$ cannot be directly measured (but see Harikane et al. 2017).

2.4. Measuring the Lyα escape fraction ($f_{esc,Ly\alpha}$) with Hα

We use dust corrected Hα luminosity to predict the intrinsic Lyα luminosity. We then compare the latter to the observed Lyα luminosity to obtain the Lyα escape fraction ($f_{esc,Ly\alpha}$). Assuming case B recombination[1], a temperature of 10^4 K and an electron density of $350\,cm^{-3}$, we can use the observed Lyα luminosity ($L_{Ly\alpha}$), the observed Hα luminosity ($L_{H\alpha}$) and the dust extinction affecting $L_{H\alpha}$ ($A_{H\alpha}$[2], in mag) to compute $f_{esc,Ly\alpha}$ as:

$$f_{esc,Ly\alpha} = \frac{L_{Ly\alpha}}{8.7\,L_{H\alpha} \times 10^{0.4 \times A_{H\alpha}}}. \tag{1}$$

This means that with our assumptions so far, and provided that we know $f_{esc,Ly\alpha}$, we can use the observed $L_{Ly\alpha}$ to obtain the intrinsic Hα luminosity. Therefore, one can use Lyα as a star formation rate (SFR) indicator[3] following Kennicutt (1998) for a Salpeter (Chabrier) IMF ($0.1 - 100\,M_\odot$):

$$SFR_{Ly\alpha}\,[M_\odot\,yr^{-1}] = \frac{7.9(4.4) \times 10^{-42}}{(1 - f_{esc,LyC})}\,\frac{L_{Ly\alpha}}{8.7\,f_{esc,Ly\alpha}} \tag{2}$$

where $f_{esc,LyC}$ is the escape fraction of ionising LyC photons (see e.g. Sobral et al. 2018b). In practice, $f_{esc,LyC}$ is typically assumed to be ≈ 0, but it may be ≈ 0.1 − 0.15 for LAEs (see discussions in e.g. Matthee et al. 2017a; Verhamme et al. 2017).

2.5. Statistical fits and errors

For all fits and relations in this work (e.g. $f_{esc,Ly\alpha}$ vs. EW_0), we vary each data-point within its full Gaussian probability distribution function independently (both in EW_0 and $f_{esc,Ly\alpha}$), and re-fit 10,000 times. We present the best-fit relation as the median of all fits, and the uncertainties (lower and upper) are the 16 and 84 percentiles. For bootstrapped quantities (e.g. for fitting the low redshift sample) we obtain 10,000 samples randomly picking half of the total number of sources and computing that specific quantity. We fit relations in the form $y = Ax + B$.

3. Results and Discussion

3.1. The observed $f_{esc,Ly\alpha}$-EW_0 relation at $z \sim 0.1 - 2.6$

Figure 1 shows that $f_{esc,Ly\alpha}$ correlates with Lyα EW_0 with apparently no redshift evolution between $z = 0 - 2.6$ (see also Verhamme et al. 2017; Sobral et al. 2017). We find that $f_{esc,Ly\alpha}$ varies continuously from ≈ 0.2 to ≈ 0.7 for LAEs from the lowest (≈ 30 Å) to the highest (≈ 120 − 160 Å) Lyα rest-frame EWs.

---

[1] We use Lyα/Hα = 8.7, but vary the Lyα/Hα case B ratio between 8.0 and 9.0 to test for its effect; see §3.5.
[2] With our case B assumptions the intrinsic Balmer decrement is: Hα/Hβ = 2.86. Using a Calzetti et al. (2000) dust law we use $A_{H\alpha}$ = $6.531 \log_{10}(H\alpha/H\beta) - 2.981$ (see details in e.g. Sobral et al. 2012).
[3] For continuous star formation over 10 Myr timescales.
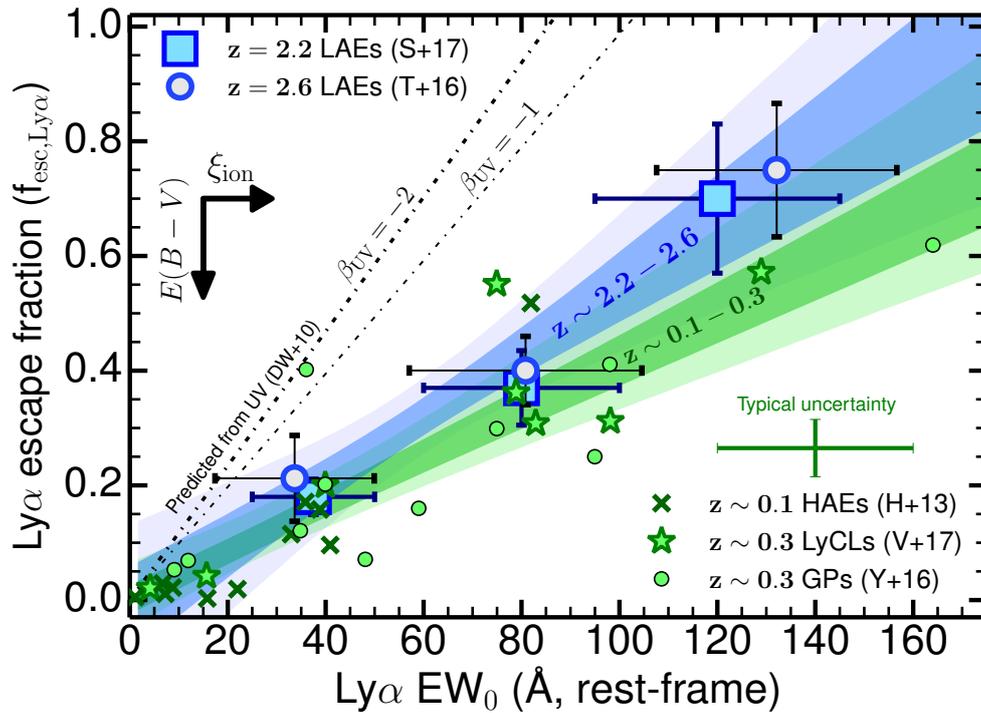
Fig. 1. The relation between $f_{esc,Ly\alpha}$ and Lyα EW_0 for $z \sim 2.2$ (stacks; see Sobral et al. 2017), $z \sim 2.6$ (binning; Trainor et al. 2015) and comparison with $z \sim 0 - 0.3$ samples (e.g. Cardamone et al. 2009; Hayes et al. 2013; Henry et al. 2015; Yang et al. 2016, 2017; Verhamme et al. 2017), estimated from dust-corrected Hα luminosities (Equation 1). We show the 1σ and 2σ range for the fits at $z \sim 2.2 - 2.6$ and $z \sim 0 - 0.3$ separately, and find them to be consistent within those uncertainties, albeit with a potential steeper relation at higher redshift. We find a combined best fitting relation given by $f_{esc,Ly\alpha} = 0.0048\,EW_0 \pm 0.05$. The observed relation is significantly away from what would be predicted (DW+10) based on the UV (see Dijkstra & Westra 2010), and implies not only a higher ξ_{ion} than the canonical value, but also an increasing ξ_{ion} as a function of EW_0.

We use our samples at $z \sim 0 - 0.3$ and $z \sim 2.2 - 2.6$, separately and together, to obtain linear fits to the relation between $f_{esc,Ly\alpha}$ and Lyα EW_0 (see §2.5). These fits allow us to provide a more quantitative view on the empirical relation and evaluate any subtle redshift evolution; see Table 1.

The relation between $f_{esc,Ly\alpha}$ and Lyα EW_0 is statistically significant at 5 to 10σ for all redshifts. We note that all linear fits are consistent with a zero escape fraction for a null EW_0 (Table 1), suggesting that the trend is well extrapolated for weak LAEs with $EW_0 \approx 0 - 20$ Å. Furthermore, as Table 1 shows, the fits to the individual (perturbed) samples at different redshifts result in relatively similar slopes and normalisations within the uncertainties, and thus are consistent with the same relation from $z \sim 0$ to $z \sim 2.6$. Nevertheless, we note that there is minor evidence for a shallower relation at lower redshift for the highest EW_0 (Figure 1), but this could be driven by current samples selecting sources with more extreme properties (including LyC leakers). Given our findings, we decide to combine the samples and obtain joint fits, with the results shown in Table 1. The slope of the relation is consistent with being ≈ 0.005 with a null $f_{esc,Ly\alpha}$ for $EW_0 = 0$ Å.

3.2. The $f_{esc,Ly\alpha}$-EW_0 relation: expectation vs. reality

The existence of a relation between observed Lyα luminosity and EW_0 (Figure 1) is not surprising. This is because Lyα EW_0 is sensitive to the ratio between Lyα and the UV, which can be used as a proxy of the $f_{esc,Ly\alpha}$ (see Sobral et al. 2018b). However, the slope, normalisation and scatter of such relation depend on complex physical conditions such as dust obscuration, differential dust geometry, scattering of Lyα photons and the production

Table 1. The results from fitting the relation between $f_{esc,Ly\alpha}$ and Lyα EW_0 as $f_{esc,Ly\alpha} = A \times EW_0 + B$, with EW_0 in Å (see §2.5). [i: individual sources used for fitting; b: binned/averaged quantity used for fitting; B: bootstrap analysis when fitting each of the 10,000 times; G: each data bin is perturbed along its Gaussian probability distribution.]

Sample	A (Å^{-1})	B	[notes]
$z \sim 0 - 0.3$	$0.0041^{+0.0006}_{-0.0004}$	$0.00^{+0.03}_{-0.02}$	[i,B]
$z \sim 2.2$	$0.0056^{+0.0012}_{-0.0011}$	$0.00^{+0.05}_{-0.05}$	[b,G]
$z \sim 2.6$	$0.0054^{+0.0016}_{-0.0015}$	$0.01^{+0.11}_{-0.11}$	[b,G]
$z \sim 0 - 2.2$	$0.0045^{+0.0008}_{-0.0007}$	$0.00^{+0.06}_{-0.06}$	[b,G]
$z \sim 2.2 - 2.6$	$0.0056^{+0.0012}_{-0.0012}$	$0.00^{+0.07}_{-0.08}$	[b,G]
$z \sim 0 - 2.6$	$0.0048^{+0.0007}_{-0.0007}$	$0.00^{+0.05}_{-0.05}$	[b,G]

efficiency of ionising photons compared to the UV luminosity, ξ_{ion} (see e.g. Matthee et al. 2017a; Shivaei et al. 2017).

While a relation between $f_{esc,Ly\alpha}$ and EW_0 is expected, we can investigate if it simply follows what would be predicted given that both the UV and Lyα trace SFRs. In order to predict $f_{esc,Ly\alpha}$ based on Lyα EW_0 we follow Dijkstra & Westra (2010) who use the Kennicutt (1998) SFR calibrations for a Salpeter IMF. As in Dijkstra & Westra (2010), we assume two different UV slopes: $\beta = -2.0$ and $\beta = -1.0$, which encompass the majority of LAEs (note that a steeper β results in an even more significant disagreement with observations) and can predict that $f_{esc,Ly\alpha} = C \times \frac{EW_0}{E}$, with $E = 76$ Å and $C = \frac{\nu_{Ly\alpha}}{\nu_{UV}}^{-2-\beta}$. We use $C = 0.89$ and $C = 0.75$ for the different β slopes as in Dijkstra & Westra (2010). Note that this methodology implicitly results in assuming a "canonical", constant $\xi_{ion} = 1.3 \times 10^{25}$ Hz erg^{-1}
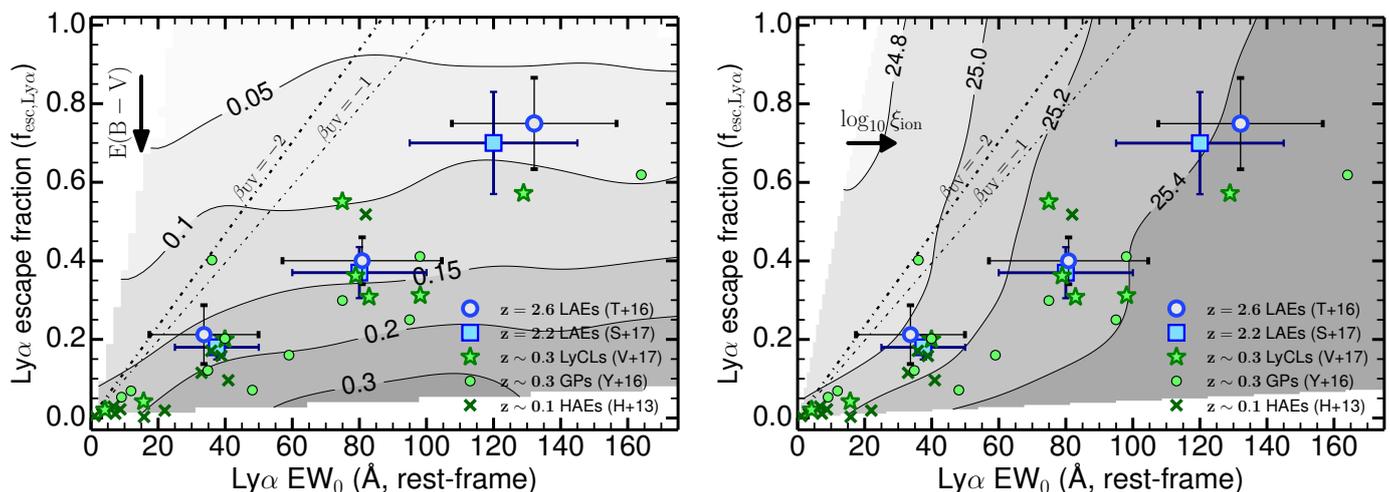
Fig. 2. *Left:* The predicted relation between $f_{esc,Ly\alpha}$ and $Ly\alpha$ EW_0 for different $E(B-V)$ (contour levels) with our toy model (see §3.3 and Appendix A). We find that dust extinction drives the simple predicted relation down, with data at $z \sim 0 - 2.6$ hinting for lower dust extinction at the highest EW_0 and higher dust extinction at the lowest EW_0, but with the range being relatively small overall and around $E(B-V) \approx 0.1 - 0.2$. *Right:* The predicted relation between $f_{esc,Ly\alpha}$ and $Ly\alpha$ EW_0 by varying ξ_{ion} (contours). We find that while increasing $E(B-V)$ mostly shifts the relation down, increasing ξ_{ion} moves the relation to the right. Observations thus hint for an increase in the typical ξ_{ion} for LAEs with increasing EW_0.

(Kennicutt 1998)[4], and a unit ratio between $Ly\alpha$ and UV SFRs (see Sobral et al. 2018b).

Predicting $f_{esc,Ly\alpha}$ based on the ratio of $Ly\alpha$ to UV using EW_0 (see Dijkstra & Westra 2010) significantly overestimates $f_{esc,Ly\alpha}$ (as indicated by the dot-dashed lines in Figure 1). Observations reveal higher $Ly\alpha$ EW_0 (by a factor of just over ~ 2) than expected for a given $f_{esc,Ly\alpha}$, with the offset between the simple prediction and observations potentially becoming larger for increasing EW_0. These results reveal processes that can boost the ratio between $Ly\alpha$ and UV (boosting EW_0), particularly by boosting $Ly\alpha$, or processes that reduce $f_{esc,Ly\alpha}$. Potential explanations include scattering, (differential) dust extinction, excitation due to shocks originating from stellar winds and/or AGN activity, and short time-scale variations in SFRs, leading to a higher ξ_{ion} (see Figure 1). High ξ_{ion} values ($\xi_{ion} \approx 3 \times 10^{25}$ Hz erg^{-1}) seem to be typical for LAEs (e.g. Matthee et al. 2017a; Nakajima et al. 2018) and may explain the observed relation, even more so if ξ_{ion} rises with increasing EW_0 (e.g. Matthee et al. 2017a), but dust extinction likely also plays a role (Figure 1).

3.3. The $f_{esc,Ly\alpha}$-EW_0 relation: physical interpretation

In order to further interpret the role of dust ($E(B-V)$) and ξ_{ion} on the observed $f_{esc,Ly\alpha}$-EW_0 and what the relation may be telling us, we produce a simple analytical model (see details in Appendix A). We independently vary SFRs, $E(B-V)$ and ξ_{ion}. The toy model follows our framework using a Calzetti et al. (2000) dust law and the Kennicutt (1998) calibrations and relations between UV and Hα. We also vary some assumptions independently, which include the intrinsic $Ly\alpha$/Hα ratio and $f_{esc,LyC}$. Furthermore, we introduce an extra parameter to further vary $f_{esc,Ly\alpha}$ and mimic processes which are hard to model, such as scattering, which can significantly reduce or even boost $f_{esc,Ly\alpha}$ (Neufeld 1991). We compute observed $Ly\alpha$ EW_0 and compare them with $f_{esc,Ly\alpha}$ for 20,000 galaxy realisations. Further details are given in Appendix A.

The key results from our toy model are shown in Figure 2. We find that both $E(B-V)$ and ξ_{ion} likely play a role in setting

the $f_{esc,Ly\alpha}$-EW_0 relation and changing it from simple predictions to the observed relation (see §3.2). As the left panel of Figure 2 shows, observed LAEs on the $f_{esc,Ly\alpha}$-EW_0 relation seem to have low $E(B-V) \approx 0.1 - 0.2$, with the lowest EW_0 sources displaying typically higher $E(B-V)$ of 0.2-0.3 and the highest EW_0 sources likely having lower $E(B-V)$ of < 0.1. Furthermore, as the right panel of Figure 2 shows, high EW_0 LAEs have higher ξ_{ion}, potentially varying from $\log_{10}(\xi_{ion}/\text{Hz erg}^{-1}) \approx 25$ to $\log_{10}(\xi_{ion}/\text{Hz erg}^{-1}) \approx 25.4 - 25.5$. Our toy model interpretation is consistent with recent results (e.g. Trainor et al. 2016; Matthee et al. 2017a; Nakajima et al. 2018) for high EW_0 LAEs. Overall, a simple way to explain the $f_{esc,Ly\alpha}$-EW_0 relation at $z \sim 0 - 2.6$ is for LAEs to have narrow ranges of low $E(B-V) \approx 0.1 - 0.2$, that decrease slightly as a function of EW_0 and a relatively narrow range of high ξ_{ion} values that increase with EW_0.

Our toy model explores the full range of physical conditions independently without making any assumptions on how parameters may correlate, in order to interpret the observations in a simple unbiased way. However, the fact that observed LAEs follow a relatively tight relation between $f_{esc,Ly\alpha}$ and EW_0 suggests that there are important correlations between e.g. dust, age and ξ_{ion}. By selecting simulated sources in our toy model that lie on the observed relation (see Appendix A.1), we recover a tight correlation between ξ_{ion} and $E(B-V)$, while the full generated population in our toy model shows no correlation at all by definition (see Figure A.1). This implies that the observed $f_{esc,Ly\alpha}$-EW_0 is likely a consequence of an evolutionary ξ_{ion}-$E(B-V)$ sequence for LAEs. For further details, see Appendix A.1.

3.4. Estimating $f_{esc,Ly\alpha}$ with a simple observable: $Ly\alpha$ EW_0

We find that LAEs follow a simple relation between $f_{esc,Ly\alpha}$ and $Ly\alpha$ EW_0 roughly independently of redshift (for $z \leq 2.6$). Motivated by this, we propose the following empirical estimator (see Table 1) for $f_{esc,Ly\alpha}$ as a function of $Ly\alpha$ EW_0 (Å):

$$f_{esc,Ly\alpha} = 0.0048^{+0.0007}_{-0.0007} EW_0 \pm 0.05 \quad [0 < EW_0 < 160]. \quad (3)$$

This relation may hold up to $EW_0 \approx 210$ Å, above which we would predict $f_{esc,Ly\alpha} \approx 1$. This relation suggests that it is possible to estimate $f_{esc,Ly\alpha}$ for LAEs within 0.2 dex even if only

---

[4] $\xi_{ion} = 1.3 \times 10^{25} \frac{\text{SFR}_{H\alpha}}{\text{SFR}_{UV}}$ (Hz erg^{-1}).
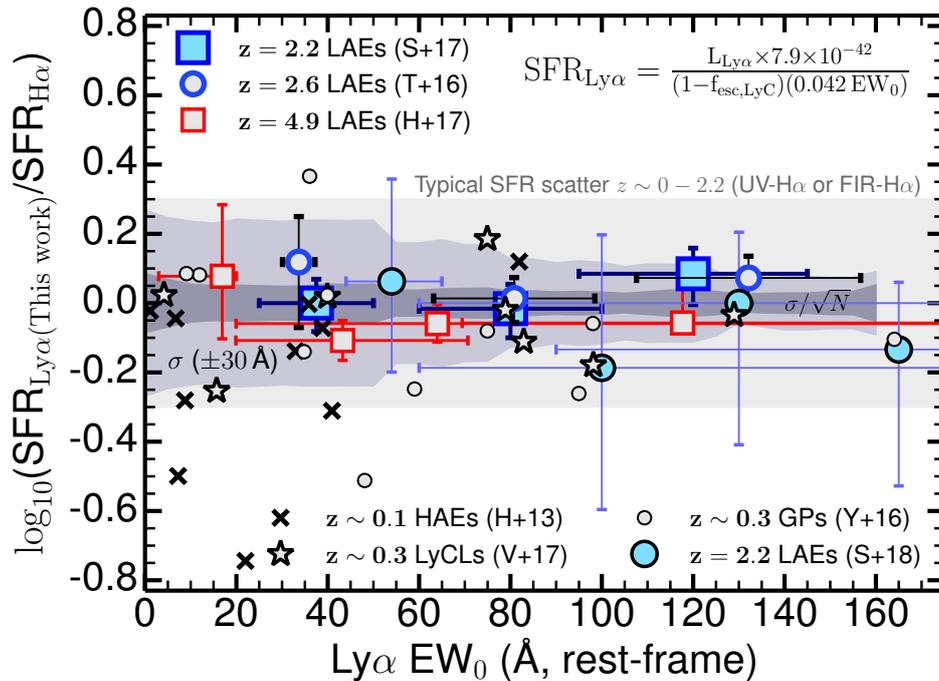
Fig. 3. The logarithmic ratio between SFRs computed with Equation 4 using Lyα luminosity and EW_0 and the "true" SFR, measured directly from dust-corrected Hα luminosity (given our definitions, $\log_{10}(SFR_{Ly\alpha(This\ work)}/SFR_{H\alpha}) = \log_{10}(f_{esc,Ly\alpha(H\alpha)}/f_{esc,Ly\alpha(This\ work)})$). We find a relatively small scatter which may decrease for higher EWs and that is at the global level of ±0.12 dex for the typical definition of LAE at higher redshift ($EW_0 > 20$ Å), but rises to ≈ 0.2 dex at the lowest EWs, likely due to a larger range of dust properties. We also provide a comparison of the typical scatter between UV and FIR SFRs in relation to Hα at $z \sim 0 - 2$ (≈ 0.3 dex; see e.g. Domínguez Sánchez et al. 2012; Oteo et al. 2015).

the Lyα EW_0 is known/constrained. It also implies that the observed Lyα luminosities are essentially equal to intrinsic Lyα luminosities for sources with EW_0 as high as ≈ 200 Å. We propose a linear relation for its simplicity and because current data do not suggest a more complex relation. Larger data-sets with Hα and Lyα measurements, particularly those covering a wider parameter space may lead to the necessity of a more complicated functional form. A departure from a linear fit may also provide further insight of different physical processes driving the relation (e.g. winds, orientation angle, burstiness or additional ionisation processes such as fluorescence).

We further test the validity of Equation 3 by measuring the ratio between the real (Hα-based) $f_{esc,Ly\alpha}$ fraction and that inferred from the simple predicting relation. We conclude that while the escape of Lyα photons can depend on a range of properties in a very complex way (see e.g. Hayes et al. 2010; Matthee et al. 2016; Yang et al. 2017), using EW_0 and Equation 3 leads to predicting $f_{esc,Ly\alpha}$ within ≈ 0.1 − 0.2 dex of real values. This compares with a larger scatter of ≈ 0.3 dex for relations with derivative or more difficult quantities to measure such as dust extinction or the red peak velocity of the Lyα line (e.g. Yang et al. 2017). Equation 3 may thus be applied to estimate $f_{esc,Ly\alpha}$ for a range of LAEs in the low and higher redshift Universe. For example, J1154+2443 (Izotov et al. 2018), has a measured $f_{esc,Ly\alpha}$ directly from dust corrected Hα luminosity of ≈ 0.7−0.8[5], while Equation 3 would imply ≈ 0.6 − 0.7 based on the $EW_0 ≈ 133$ Å for Lyα, thus implying a difference of only 0.06-0.1 dex. Furthermore, in principle, Equation 3 could also be explored to transform EW_0 distributions (e.g. Hashimoto et al. 2017, and references therein) into distributions of $f_{esc,Ly\alpha}$ for LAEs.

3.5. Lyα as a SFR indicator: empirical calibration and errors

Driven by the simple relation found up to $z \sim 2.6$, we derive an empirical calibration to obtain SFRs based on two simple, direct observables for LAEs at high redshift: 1) Lyα EW_0 and 2) observed Lyα luminosity. This calibration is based on observables, but predicts the dust-corrected SFR. Based on Equations 2 and 3, for a Salpeter (Chabrier) IMF we can derive[6]:

$$SFR_{Ly\alpha}\ [M_\odot\ yr^{-1}] = \frac{L_{Ly\alpha} \times 7.9\ (4.4) \times 10^{-42}}{(1 - f_{esc,LyC})(0.042\ EW_0)}\ (\pm 15\%) \qquad (4)$$

The current best estimate of the scatter in Equation 3 (the uncertainty in the relation to calculate $f_{esc,Ly\alpha}$ is ±0.05) implies a ±0.07 dex uncertainty in the extinction corrected SFRs from Lyα with our empirical calculation. In order to investigate other systematic errors, we conduct a Monte Carlo analysis by randomly varying $f_{esc,LyC}$ (0.0 to 0.2) and the case B coefficient (from 8.0 to 9.0), along with perturbing $f_{esc,Ly\alpha}$ from −0.05 to +0.05. We assume that all properties are independent, and thus this can be seen as a conservative approach to estimate the uncertainties. We find that the uncertainty in $f_{esc,Ly\alpha}$ is the dominant source of uncertainty (12%) with the uncertainty on $f_{esc,LyC}$ and the case B coefficient contributing an additional 3% for a total of 15%. This leads to an expected uncertainty of Equation 4 of 0.08 dex.

3.6. Lyα as a SFR indicator: performance and implications

In Figure 3 we apply Equation 4 to compare the estimated SFRs (from Lyα) with those computed with dust corrected Hα luminosities. We also include individual sources at $z \sim 2.2$ (S18; Sobral et al. 2018a) and recent results from Harikane et al. (2017) at

---

[5] This may be up to ≈ 0.98 if Hβ is used; see (Izotov et al. 2018).

[6] Note that the constant 0.042 has units of $Å^{-1}$, and results from $8.7 \times 0.0048\ Å^{-1}$.

$z = 4.8$ which were not used in the calibration, and thus provide an independent way to test our new calibration. We find a global scatter of ≈ 0.12 dex, being apparently larger for lower EW_0, but still lower than the typical scatter between SFR indicators after dust corrections (e.g. UV-Hα or FIR-Hα; see Domínguez Sánchez et al. 2012; Oteo et al. 2015), as shown in Figure 3. The small scatter and approximately null offset between our calibration's prediction and measurements presented by Harikane et al. (2017) at $z \sim 5$ suggest that Equation 4 may be applicable at higher redshift with similarly competitive uncertainties (see §3.7 and §3.8).

3.7. Application to bright and faint LAEs at high redshift

Our new empirical calibration of Lyα as a SFR indicator allows to estimate SFRs of LAEs at high redshift. The global Lyα luminosity function at $z \sim 3 - 6$ has a typical Lyα luminosity of $10^{42.9}$ erg s^{-1} (Sobral et al. 2018b), with these LAEs having $EW_0 \approx 80$ Å (suggesting $f_{esc,Ly\alpha} = 0.38 \pm 0.05$ with Equation 3), which implies SFRs of ≈ 20 M$_\odot$ yr^{-1}. If we explore the public SC4K sample of LAEs at $z \sim 2-6$ (Sobral et al. 2018b), limiting it to sources with up to $EW_0 = 210$ Å and that are consistent with being star-forming galaxies ($L_{Ly\alpha} < 10^{43.2}$ erg s^{-1}; see Sobral et al. 2018a), we find a median SFR for LAEs of 12^{+9}_{-5} M$_\odot$ yr^{-1}, ranging from ≈ 2 M$_\odot$ yr^{-1} to ≈ 90 M$_\odot$ yr^{-1} at $z \sim 2 - 6$. These reveal that "typical" to luminous LAEs are forming stars below and up to the typical SFR (SFR$^\star \approx 40 - 100$ M$_\odot$ yr^{-1}) at high redshift (see Smit et al. 2012; Sobral et al. 2014).

Deep MUSE Lyα surveys (e.g. Drake et al. 2017; Hashimoto et al. 2017) are able to sample the faintest LAEs with a median $L_{Ly\alpha} = 10^{41.9\pm0.1}$ erg s^{-1} and $EW_0 = 87 \pm 6$ (Hashimoto et al. 2017) at $z \sim 3.6$. We predict a typical $f_{esc,Ly\alpha} = 0.42 \pm 0.05$ and $SFR_{Ly\alpha} = 1.7\pm0.3$ M$_\odot$ yr^{-1} for those MUSE LAEs. Furthermore, the faintest LAEs found with MUSE have $L_{Ly\alpha} = 10^{41}$ erg s^{-1} (Hashimoto et al. 2017), implying SFRs of ≈ 0.1 M$_\odot$ yr^{-1} with our calibration. Follow-up *JWST* observations targeting the Hα line for faint MUSE LAEs are thus expected to find typical Hα luminosities of 2×10^{41} erg s^{-1} and as low as $\approx 1-2\times10^{40}$ erg s^{-1} for the faintest LAEs. Based on our predicted SFRs, we expect MUSE LAEs to have UV luminosities from $M_{UV} \approx -15.5$ for the faintest sources, to $M_{UV} \approx -19$ for more typical LAEs, thus potentially linking faint LAEs discovered from the ground with the population of SFGs that dominate the faint end of the UV luminosity function (e.g. Fynbo et al. 2003; Gronke et al. 2015; Dressler et al. 2015).

3.8. Comparison with UV and implications at higher redshift

Equations 3 and 4 can be applied to a range of spectroscopically confirmed LAEs in the literature. We also extend our predictions to sources within the epoch of re-ionisation[7]. We explore a recent extensive compilation by Matthee et al. (2017c) of both Lyα- and UV-selected LAEs with spectroscopic confirmation and Lyα measurements (e.g. Ouchi et al. 2008, 2009; Ono et al. 2012; Sobral et al. 2015; Zabl et al. 2015; Stark et al. 2015c; Ding et al. 2017; Shibuya et al. 2018). These include published $L_{Ly\alpha}$, EW_0 and M_{UV}. In order to correct UV luminosities we use the UV β slope, typically used to estimate A_{UV}[8]. For UV-selected sources we assume $\beta = -1.6 \pm 0.2$ dex (typical for their UV luminosity;

---

[7] See Laursen et al. (2011) for important caveats on how the transmission at line-centre is affected by an increasing IGM neutral fraction
[8] We use $A_{UV} = 4.43 + 1.99\beta$; see Meurer et al. (1999).
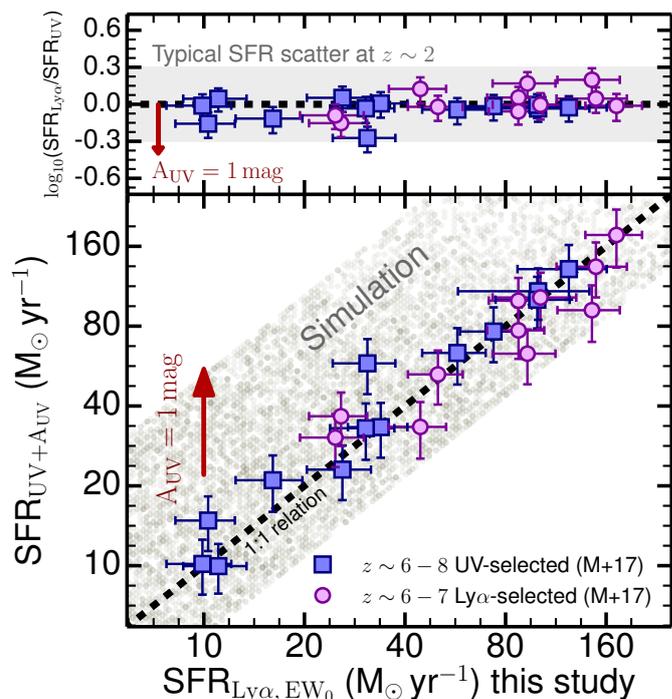
Fig. 4. Comparison between SFRs computed with our new empirical calibration for Lyα as a SFR indicator (Equation 4) and those computed based on dust corrected UV luminosity (see §3.8) for a compilation of $z \sim 5 - 8$ sources (see Matthee et al. 2017c, and references therein). Our simple empirical calibration of Lyα as a SFR is able to recover dust corrected UV SFRs with a typical scatter of ≈ 0.2 dex, being slightly higher for the more luminous LAEs than for the continuum selected LAEs which probe down to lower SFRs (scatter ≈ 0.08 dex which is very close to the systematic scatter expected; see §3.5). We also compute SFRs in the same way with observables from our toy model and show the results of our simulation. We find that the scatter in our toy model is much larger, with this being driven by $E(B - V)$ being able to vary from 0.0 to 0.5.

e.g. Bouwens et al. 2009), while for the luminous LAEs we use $\beta = -1.9 \pm 0.2$ dex. We predict their SFRs using $L_{Ly\alpha}$ and EW_0 only (Equation 4) and compare with SFRs measured from dust-corrected UV luminosities (Kennicutt 1998); see Table A.2. We make the same assumptions and follow the same methodology to transform the observables of our toy model into SFRs (see Figure 4). We note that, as our simulation shows, one expects a correlation even if our calibration of Lyα as a SFR indicator is invalid at high redshift. Therefore, we focus our discussion on the normalisation of the relation and particularly on the scatter, not on the existence of a relation. We also note that our calibration is based on dust corrected Hα luminosities at $z \sim 0 - 2.6$, and that UV luminosities are not used prior to this Section.

Our results are shown in Figure 4 (see Table A.2 for details on individual sources), which contains sources at a variety of redshifts, from $z \sim 6$ to $z \sim 8$ (e.g. Oesch et al. 2015; Stark et al. 2017). We find a remarkable agreement between our predicted Lyα SFRs based solely on Lyα luminosities and EW_0 and the dust corrected UV SFRs for a range of sources at $z \sim 6 - 8$. We find that the scatter between UV-based and Lyα based SFRs to be ≈ 0.2 dex. Interestingly, we find a larger scatter for sources selected as LAEs (0.23 dex) than those that were selected using UV continuum using e.g. *HST* (although they are also LAEs), for which we find a scatter of only 0.08 dex.

Overall, our results and application to higher redshift reveals that Equation 4 is able to retrieve SFRs with very simple observ-

ables even for LAEs within re-ionisation (e.g. Ono et al. 2012; Stark et al. 2015c, 2017; Schmidt et al. 2017). In the early Universe the fraction of sources that are LAEs is higher, thus making our calibration applicable to a larger fraction of the galaxy population, perhaps with an even smaller scatter due to the expected narrower range of physical properties. Our calibration of Lyα as a SFR indicator is simple, directly calibrated with Hα, and should not have a significant dependence on e.g. metallicity, unlike other proposed SFRs tracers at high redshift such as [CII] luminosity or other weak UV metal lines.

It is nonetheless surprising that our calibration apparently still works even at $z \sim 7-8$ for luminous LAEs. This seems to indicate that the IGM may not play a significant role for these Lyα-visible sources, potentially due to early ionised bubbles (Matthee et al. 2015) or velocity offsets of Lyα with respect to systemic (see e.g. Stark et al. 2017).

3.9. A tool for re-ionisation: predicting the LyC luminosity

Based on our results and assumptions (see §2.4), we follow Matthee et al. (2017a)[9] and derive a simple expression to predict the number of produced LyC photons per second, Q_{ion} (s^{-1}) with direct Lyα observables (L$_{Ly\alpha}$ and EW$_0$):

$$Q_{ion,Ly\alpha} \, [s^{-1}] = \frac{L_{Ly\alpha}}{c_{H\alpha} \, (1 - f_{esc,LyC}) \, (0.042 \, EW_0)}, \qquad (5)$$

where $c_{H\alpha} = 1.36 \times 10^{-12}$ erg (e.g. Kennicutt 1998; Schaerer 2003), under our case B recombination assumption (see §2.4).

Recent work by e.g. Verhamme et al. (2017) show that LyC leakers are strong LAEs, and that f$_{esc,Ly\alpha}$ is linked and/or can be used to predict f$_{esc,LyC}$ (see Chisholm et al. 2018). Equation 5 provides an extra useful tool: an empirical simple estimator of Q_{ion} for LAEs given observed Lyα luminosities and EW$_0$. Note that Equation 5 does not require measuring UV luminosities or ξ_{ion}, but instead direct, simple observables. Matthee et al. (2017c) already used a similar method to predict ξ_{ion} at high redshift. Coupled with an accurate estimate of the escape fraction of LyC photons from LAEs, which can be obtained with *HST*, a robust estimate of the full number density of LAEs from faint to the brightest sources (Sobral et al. 2018b) and their redshift evolution, Equation 5 may provide a simple tool to further understand if LAEs were able to re-ionise the Universe.

This allows the prediction of f$_{esc,Ly\alpha}$ based on a simple direct observable, and thus to compute the intrinsic Lyα luminosity of LAEs at high redshift.

- The observed f$_{esc,Ly\alpha}$-EW$_0$ implies a tight ξ_{ion}-$E(B-V)$ sequence for LAEs, with higher ξ_{ion} at lower $E(B-V)$ and vice versa. Both ξ_{ion} and $E(B-V)$ seem to depend on Lyα EW$_0$ (Figure 2). Our results imply that the higher the EW$_0$ selection, the higher the ξ_{ion} and the lower the $E(B-V)$.

- The f$_{esc,Ly\alpha}$-EW$_0$ relation reveals a scatter of only 0.1-0.2 dex for LAEs, and there is evidence for the relation to hold up to $z \sim 5$ (Figure 3). The scatter is higher towards lower EW$_0$, consistent with a larger range in dust properties for sources with the lowest EW$_0$. At the highest EW$_0$, on the contrary, the scatter may be as small as ≈ 0.1 dex, consistent with high EW$_0$ LAEs being an even more homogeneous population of dust-poor, high ionisation star-forming galaxies.

- We use our results to calibrate Lyα as a SFR indicator for LAEs (Equation 4) and find a global scatter of 0.2 dex between measurements using Lyα only and those using dust-corrected Hα luminosities. Our results also allow us to derive a simple estimator of the number of LyC photons produced per second (Equation 5) with applications to studies of the epoch of re-ionisation.

- Equation 4 implies that star-forming LAEs at $z \sim 2-6$ have SFRs typically ranging from 0.1 to 20 M$_\odot$ yr^{-1}, with MUSE LAEs expected to have typical SFRs of 1.7 ± 0.3 M$_\odot$ yr^{-1}, and more luminous LAEs having SFRs of 12^{+9}_{-5} M$_\odot$ yr^{-1}.

- SFRs based on Equation 4 are in very good agreement with dust corrected UV SFRs even within the epoch of re-ionisation and for a range of sources, hinting for it to be applicable in the very early Universe. If shown to be the case, our results have implications for the minor role of the IGM in significantly changing Lyα luminosities and EW$_0$ for luminous LAEs within the epoch of re-ionisation, and show that measuring L$_{Ly\alpha}$ and EW$_0$ provide apparently reliable SFRs.

Our results provide a simple interpretation of the tight f$_{esc,Ly\alpha}$-EW$_0$ relation. Most importantly, we provide simple and practical tools to estimate f$_{esc,Ly\alpha}$ at high redshift with two direct observables and thus to use Lyα as a SFR indicator and to measure the number of ionising photons from LAEs. The empirical calibrations presented here can be easily tested with future observations with *JWST* which can obtain Hα and Hβ measurements for high-redshift LAEs.

4. Conclusions

Lyα is intrinsically the brightest emission-line in active galaxies, and should be a good SFR indicator. However, the uncertain and difficult to measure f$_{esc,Ly\alpha}$ has limited the interpretation and use of Lyα luminosities. In order to make progress, we have explored samples of LAEs at $z = 0 - 2.6$ with direct Lyα escape fractions measured from dust corrected Hα luminosities which do not require any SED fitting, ξ_{ion} or other complex assumptions based on derivative quantities. Our main results are:

- There is a simple, linear relation between f$_{esc,Ly\alpha}$ and Lyα EW$_0$: f$_{esc,Ly\alpha}$ = 0.0048 EW$_0$[Å] $\pm$ 0.05 (Equation 3) which is shallower than simple expectations, due to both more ionising photons per UV luminosity (ξ_{ion}) and declining dust extinction ($E(B-V)$) for LAEs with increasing EW$_0$ (Figure 1).

---

[9] We assume f$_{dust}$ $\approx$ 0 (see Matthee et al. 2017a), i.e., we make the assumption that for LAEs the dust extinction to LyC photons within HII regions is $\approx$ 0.

References

An, F. X., Zheng, X. Z., Hao, C.-N., Huang, J.-S., & Xia, X.-Y. 2017, ApJ, 835, 116
Astropy Collaboration et al. 2013, A&A, 558, A33
Atek, H., Kunth, D., Hayes, M., Östlin, G., & Mas-Hesse, J. M. 2008, A&A, 488, 491
Bacon, R., Brinchmann, J., Richard, J., et al. 2015, AAP, 575, A75
Bouwens, R. J., Illingworth, G. D., Franx, M., et al. 2009, ApJ, 705, 936
Cai, Z., Fan, X., Jiang, L., et al. 2015, ApJ, 799, L19
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Cardamone, C. et al. 2009, MNRAS, 399, 1191

Cassata, P., Tasca, L. A. M., Le Fèvre, O., et al. 2015, A&A, 573, A24
Cassata, P. et al. 2011, A&A, 525, A143
Charlot, S. & Fall, S. M. 1993, ApJ, 415, 580
Chisholm, J., Gazagnes, S., Schaerer, D., et al. 2018, ArXiv e-prints [arXiv:1803.03655]
Ciardullo, R., Zeimann, G. R., Gronwall, C., et al. 2014, ApJ, 796, 64
Dijkstra, M. 2017, ArXiv e-prints [arXiv:1704.03416]
Dijkstra, M. & Westra, E. 2010, MNRAS, 401, 2343
Ding, J., Cai, Z., Fan, X., et al. 2017, ApJ, 838, L22
Domínguez Sánchez, H. et al. 2012, MNRAS, 426, 330
Drake, A. B. et al. 2017, MNRAS, 471, 267
Dressler, A., Henry, A., Martin, C. L., et al. 2015, ApJ, 806, 19
Fynbo, J. P. U., Ledoux, C., Møller, P., Thomsen, B., & Burud, I. 2003, A&A, 407, 147
Gawiser, E., Francke, H., Lai, K., et al. 2007, ApJ, 671, 278
Gronke, M., Dijkstra, M., Trenti, M., & Wyithe, S. 2015, MNRAS, 449, 1284
Hagen, A., Zeimann, G. R., et al. 2016, ApJ, 817, 79
Harikane, Y., Ouchi, M., Shibuya, T., et al. 2017, ArXiv e-prints [arXiv:1711.03735]
Hashimoto, T. et al. 2017, A&A, 608, A10
Hayes, M., Östlin, G., Schaerer, D., et al. 2013, ApJL, 765, L27
Hayes, M., Schaerer, D., Östlin, G., et al. 2011, ApJ, 730, 8
Hayes, M. et al. 2010, Nature, 464, 562
Hayes, M. et al. 2014, ApJ, 782, 6
Henry, A., Scarlata, C., Martin, C. L., & Erb, D. 2015, ApJ, 809, 19
Hu, E. M., Cowie, L. L., Barger, A. J., et al. 2010, ApJ, 725, 394
Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90
Izotov, Y. I., Orlitová, I., Schaerer, D., et al. 2016a, Nature, 529, 178
Izotov, Y. I., Schaerer, D., Thuan, T. X., et al. 2016b, MNRAS, 461, 3683
Izotov, Y. I., Schaerer, D., Worseck, G., et al. 2018, MNRAS, 474, 4514
Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open source scientific tools for Python
Kennicutt, Jr., R. C. 1998, ARAA, 36, 189
Laursen, P., Sommer-Larsen, J., & Razoumov, A. O. 2011, ApJ, 728, 52
Mainali, R., Kollmeier, J. A., Stark, D. P., et al. 2017, ApJ, 836, L14
Malhotra, S. & Rhoads, J. E. 2004, ApJL, 617, L5
Martin, C. L. & Sawicki, M. 2004, ApJ, 603, 414
Matthee, J., Sobral, D., Best, P., et al. 2017a, MNRAS, 465, 3637
Matthee, J., Sobral, D., Boone, F., et al. 2017b, ApJ, 851, 145
Matthee, J., Sobral, D., Darvish, B., et al. 2017c, MNRAS, 472, 772
Matthee, J., Sobral, D., Oteo, I., et al. 2016, MNRAS, 458, 449
Matthee, J., Sobral, D., Santos, S., et al. 2015, MNRAS, 451, 400
Meurer, G. R., Heckman, T. M., & Calzetti, D. 1999, ApJ, 521, 64
Miley, G. & De Breuck, C. 2008, AAPR, 15, 67
Momose, R., Ouchi, M., Nakajima, K., et al. 2014, MNRAS, 442, 110
Nakajima, K., Fletcher, T., Ellis, R. S., Robertson, B. E., & Iwata, I. 2018, ArXiv e-prints [arXiv:1801.03085]
Neufeld, D. A. 1991, ApJ, 370, L85
Oesch, P. A., van Dokkum, P. G., Illingworth, G. D., et al. 2015, ApJ, 804, L30
Ono, Y. et al. 2012, ApJ, 744, 83
Oteo, I., Sobral, D., Ivison, R. J., et al. 2015, MNRAS, 452, 2018
Ouchi, M. et al. 2008, ApJs, 176, 301
Ouchi, M. et al. 2009, ApJ, 696, 1164
Oyarzún, G. A., Blanc, G. A., González, V., Mateo, M., & Bailey, III, J. I. 2017, ApJ, 843, 133
Partridge, R. B. & Peebles, P. J. E. 1967, ApJ, 147, 868
Pritchet, C. J. 1994, PASP, 106, 1052
Rauch, M. et al. 2008, ApJ, 681, 856
Rhoads, J. E., Malhotra, S., Dey, A., et al. 2000, ApJL, 545, L85
Schaerer, D. 2003, AAP, 397, 527
Schmidt, K. B. et al. 2017, ApJ, 839, 17
Shibuya, T. et al. 2018, PASJ, 70, S15
Shivaei, I., Reddy, N. A., et al. 2017, ArXiv e-prints [arXiv:1711.00013]
Smidt, J., Wiggins, B. K., & Johnson, J. L. 2016, ArXiv e-prints [arXiv:1603.00888]
Smit, R., Bouwens, R. J., Franx, M., et al. 2012, ApJ, 756, 14
Sobral, D., Best, P. N., Matsuda, Y., et al. 2012, MNRAS, 420, 1926
Sobral, D., Best, P. N., Smail, I., et al. 2014, MNRAS, 437, 3516
Sobral, D., Matthee, J., Best, P., et al. 2017, MNRAS, 466, 1242
Sobral, D., Matthee, J., Darvish, B., et al. 2015, ApJ, 808, 139
Sobral, D., Matthee, J., Darvish, B., et al. 2018a, ArXiv e-prints [arXiv:1802.10102]
Sobral, D., Santos, S., Matthee, J., et al. 2018b, ArXiv e-prints [arXiv:1712.04451]
Song, M., Finkelstein, S. L., et al. 2014, ApJ, 791, 3
Stark, D. P., Ellis, R. S., Charlot, S., et al. 2017, MNRAS, 464, 469
Stark, D. P., Richard, J., Charlot, S., et al. 2015a, MNRAS, 450, 1846
Stark, D. P., Walth, G., Charlot, S., et al. 2015b, MNRAS, 454, 1393
Stark, D. P. et al. 2015c, MNRAS, 450, 1846
Steidel, C. C., Bogosavljević, M., Shapley, A. E., et al. 2011, ApJ, 736, 160
Taylor, M. 2013, Starlink User Note, 253
Tilvi, V., Pirzkal, N., Malhotra, S., et al. 2016, ApJ, 827, L14
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. 2016, ApJ, 832, 171
van Breukelen, C., Jarvis, M. J., & Venemans, B. P. 2005, MNRAS, 359, 895
Van Der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science & Engineering, 13, 22
Vanzella, E., Pentericci, L., Fontana, A., et al. 2011, ApJ, 730, L35
Verhamme, A., Orlitová, I., Schaerer, D., et al. 2017, A&A, 597, A13
Verhamme, A., Schaerer, D., Atek, H., & Tapken, C. 2008, A&A, 491, 89
Wisotzki, L., Bacon, R., et al. 2016, A&A, 587, A98
Yang, H., Malhotra, S., Gronke, M., et al. 2016, ApJ, 820, 130
Yang, H., Malhotra, S., Gronke, M., et al. 2017, ApJ, 844, 171
Zabl, J., Nørgaard-Nielsen, H. U., Fynbo, J. P. U., et al. 2015, MNRAS, 451, 2050

Table A.1. The parameters varied in our simple toy model of 20,000 sources to interpret the observational results (see Appendix A).

Property	Minimum	Maximum	Δ param.
SFR ($M_\odot$ yr^{-1})	0.1	100	0.01 dex
$\log_{10}(\xi_{ion}/\mathrm{Hz\,erg}^{-1})$	24.7	25.7	0.01 dex
$f_{esc,LyC}$	0.0	0.15	0.01
Lyα/Hα	8.0	9.0	0.01
$E(B-V)$	0.0	0.5	0.01
Extra $f_{esc,Ly\alpha}$	0.0	1.3	0.01

Appendix A: Toy-model for $f_{esc,Ly\alpha}$ dependencies

We construct a simple analytical toy-model to produce observable Hα, UV and Lyα luminosities and EW$_0$ from a range of input physical conditions (see Table A.1). We independently sample in steps of 0.01 or 0.01 dex combinations of SFR, $f_{esc,LyC}$, case B Lyα/Hα intrinsic ratio, $\log_{10}(\xi_{ion}/\mathrm{Hz\,erg}^{-1})$, $E(B-V)$ with a Calzetti et al. (2000) dust law and a parameter to control $f_{esc,Ly\alpha}$ (from e.g. scattering leading to higher dust absorption or scattering Lyα photons away from or into the observers' line of sight) which acts as a further factor affecting $f_{esc,Ly\alpha}$; see Table A.1 for the range in parameters explored independently. We follow Kennicutt (1998) and all definitions and assumptions mentioned in this paper. We publicly release our simple PYTHON script which can be used for similar studies and/or to study different ranges in the parameter space, or conduct studies in which properties are intrinsically related/linked as one expects for realistic galaxies.

Appendix A.1: The $f_{esc,Ly\alpha}$-EW$_0$ results from a tight ξ_{ion}-$E(B-V)$ sequence for LAEs

We use our simple analytical model to further interpret the observed relation between $f_{esc,Ly\alpha}$-EW$_0$ and its tightness. We take all artificially generated sources and select those that satisfy the observed relation given in Equation 3, including its scatter (see Figure A.1). We further restrict the sample to sources with Lyα EW$_0$ > 25 Å. We find that along the observed $f_{esc,Ly\alpha}$-EW$_0$ relation, LAEs become less affected by dust extinction as a function of increasing EW$_0$, while ξ_{ion} increases, as already shown in §3.3 and Figure 2.

In the right panel of Figure A.1 we show the full parameter range explored in ξ_{ion}-$E(B-V)$. By constraining the simulated sources with the observed $f_{esc,Ly\alpha}$-EW$_0$ relation, we obtain a tight (± 0.1 dex), linear relation between $\log_{10} \xi_{ion}$ and $E(B-V)$ given by $\log_{10}(\xi_{ion}/\mathrm{Hz\,erg}^{-1}) \approx -1.76 \times E(B-V) + 25.6$. This means that in order for simulated sources to reproduce observations, LAEs should follow a very well defined ξ_{ion}-$E(B-V)$ sequence with high ξ_{ion} values corresponding to very low $E(B-V)$ (mostly at high EW$_0$ and high $f_{esc,Ly\alpha}$) and higher $E(B-V)$ to lower ξ_{ion} (mostly at low EW$_0$ and high $f_{esc,Ly\alpha}$). Our results thus hint for the $f_{esc,Ly\alpha}$-EW$_0$ to be driven by the physics (and diversity) of young and metal poor stellar populations and their evolution.

Appendix B: Data used for the high-redshift comparison between UV and Lyα SFRs

Table A.2 provides the data used for Figure 4, including individual measurements per source, their name and reference. Note that the data is taken from a compilation from Matthee et al. (2017c) with minor modifications for a few LAEs, as a indicated in Table A.2.
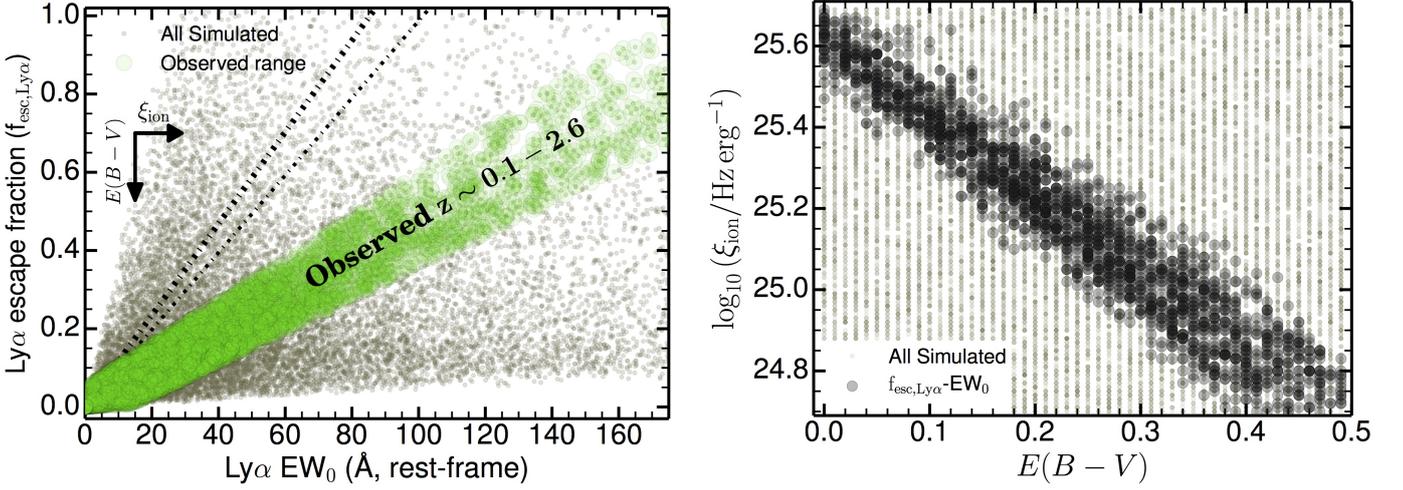
Fig. A.1. *Left:* The predicted relation between $f_{esc,Ly\alpha}$ and Lyα EW_0 for our toy model, which shows little to no correlation by sampling all physical parameters independently (see Table A.1). We also show the observed range ($\approx \pm 3\sigma$) which is well constrained at $z \sim 0 - 2.6$. We use simulated sources that are consistent with observations of LAEs to explore the potential reason behind the observed tight $f_{esc,Ly\alpha}$-EW_0 correlation for LAEs. *Right:* By restricting our toy model to the observed relation and its scatter, we find a relatively tight ξ_{ion}-$E(B-V)$ sequence for LAEs ($EW_0 > 20 - 25$ Å): $\log_{10}(\xi_{ion}/\text{Hz erg}^{-1}) \approx -1.76 \times E(B-V) + 25.6$. The highest observed EW_0 correspond to the highest ξ_{ion} and the lowest $E(B-V)$, while lower EW_0 leads to a lower ξ_{ion} and a higher $E(B-V)$. Our results thus show that the tight $f_{esc,Ly\alpha}$-EW_0 correlation for LAEs at $z \sim 0 - 2.6$ is likely driven by a ξ_{ion}-$E(B-V)$ sequence that may be related with important physics such as the age of the stellar populations, their metallicity, dust production and how those evolve together.

Table A.2. Application to high redshift UV-continuum and Lyα selected LAEs (see compilation by Matthee et al. 2017c). Errors on Lyα luminosity and EW_0 are assumed to be ≈ 0.1 dex, while errors on M_{UV} are taken as ≈ 0.2 dex. We compute the UV SFRs (SFR$_{UV}$, dust corrected) using Kennicutt (1998) and $\beta = -1.6 \pm 0.2$ for UV-selected and $\beta = -1.9 \pm 0.2$ for Lyα selected sources. Lyα SFRs (SFR$_{Ly\alpha}$; calibrated to be dust-corrected) are computed with our Equation 4. Notes: 1: EW_0 have been recomputed and rest-framed when compared to original reference. 2: M_{UV} have been recomputed when compared to original reference. 3: Values used are from Zabl et al. (2015). 4: Computed as in Matthee et al. (2017b). This table is also provided in FITS format.

Name (UV selected)	z	$\log_{10}(L_{Ly\alpha})$ [erg s^{-1}]	EW_0 [Å]	M_{UV} [mag]	SFR$_{UV}$ [M$_\odot$ yr^{-1}]	SFR$_{Ly\alpha}$ [M$_\odot$ yr^{-1}]	Reference
A383-5.2	6.03	42.8	138	−19.3	10^{+2}_{-3}	11^{+2}_{-3}	Stark et al. (2015c)
RXCJ2248.7-4431-ID3	6.11	42.5	40	−20.1	21^{+5}_{-6}	16^{+4}_{-5}	Mainali et al. (2017)
RXCJ2248.7-4431	6.11	42.9	68	−20.2	23^{+5}_{-7}	25^{+5}_{-7}	Schmidt et al. (2017)
SDF-46975	6.84	43.2	43	−21.5	76^{+18}_{-23}	76^{+17}_{-24}	Ono et al. (2012)
IOK-1	6.96	43.0	42	−21.3	63^{+14}_{-19}	57^{+13}_{-18}	Ono et al. (2012)
BDF-521	7.01	43.0	64	−20.6	34^{+8}_{-10}	34^{+7}_{-9}	Cai et al. (2015)
A1703 zd6	7.04	42.5	65	−19.3	10^{+2}_{-3}	10^{+2}_{-3}	Stark et al. (2015b)
BDF-3299	7.11	42.8	50	−20.6	33^{+8}_{-10}	30^{+6}_{-9}	Vanzella et al. (2011)
GLASS-stack	7.20	43.0	210	−19.7	15^{+3}_{-4}	10^{+2}_{-3}	Smidt et al. (2016)
EGS-zs8-2	7.48	42.7	9	−21.9	110^{+25}_{-32}	103^{+45}_{-204}	Stark et al. (2015a)
FIGS GN1 1292	7.51	42.8	49	−21.2	58^{+13}_{-17}	31^{+7}_{-9}	Tilvi et al. (2016)
GN-108036	7.21	43.2	33	−21.8	101^{+24}_{-30}	99^{+24}_{-37}	Stark et al. (2015a)
EGS-zs8-1	7.73	43.1	21	−22.1	131^{+30}_{-40}	124^{+36}_{-71}	Oesch et al. (2015)
(Lyα selected)							
SR6[1]	5.68	43.4	210	−21.1	30^{+7}_{-9}	26^{+5}_{-6}	Matthee et al. (2017c)
Ding-3	5.69	42.8	62	−20.9	25^{+6}_{-8}	25^{+5}_{-7}	Ding et al. (2017)
Ding-4	5.69	42.3	106	−20.5	18^{+4}_{-5}	4^{+1}_{-1}	Ding et al. (2017)
Ding-5	5.69	43.2	79	−21.0	28^{+6}_{-8}	44^{+9}_{-12}	Ouchi et al. (2008)
Ding-1	5.70	43.0	21	−22.2	85^{+20}_{-25}	104^{+30}_{-59}	Ding et al. (2017)
J233454[2]	5.73	43.7	210	−21.5	44^{+10}_{-13}	51^{+10}_{-12}	Shibuya et al. (2018)
J021835	5.76	43.7	107	−21.7	53^{+12}_{-16}	93^{+19}_{-24}	Shibuya et al. (2018)
VR7[1]	6.53	43.4	35	−22.5	111^{+26}_{-34}	149^{+35}_{-53}	Matthee et al. (2017c)
J162126[2]	6.54	43.9	99	−22.8	146^{+34}_{-44}	170^{+34}_{-44}	Shibuya et al. (2018)
J160234	6.58	43.5	81	−21.9	64^{+15}_{-19}	88^{+18}_{-23}	Shibuya et al. (2018)
Himiko[3]	6.59	43.6	65	−22.1	77^{+18}_{-23}	143^{+30}_{-39}	Ouchi et al. (2009)
CR7[4]	6.60	43.9	211	−22.2	84^{+19}_{-26}	87^{+17}_{-22}	Sobral et al. (2015)