

# The nature of luminous Ly $\alpha$ emitters at $z \sim 2\text{--}3$ : maximal dust-poor starbursts and highly ionizing AGN

David Sobral,<sup>1,2\*</sup> Jorryt Matthee,<sup>2</sup> Behnam Darvish,<sup>3</sup> Ian Smail,<sup>4</sup> Philip N. Best,<sup>5</sup> Lara Alegre,<sup>5,6</sup> Huub Röttgering,<sup>2</sup> Bahram Mobasher,<sup>7</sup> Ana Paulino-Afonso,<sup>1,6</sup> Andra Stroe<sup>8†</sup> and Iván Oteo<sup>5,8</sup>

<sup>1</sup>Department of Physics, Lancaster University, Lancaster, LA1 4YB, UK

<sup>2</sup>Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands

<sup>3</sup>Cahill Center for Astrophysics, California Institute of Technology, 1216 East California Boulevard, Pasadena, CA 91125, USA

<sup>4</sup>Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE UK

<sup>5</sup>Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

<sup>6</sup>Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, P-1749-016 Lisbon, Portugal

<sup>7</sup>Department of Physics and Astronomy, University of California, 900 University Ave., Riverside, CA 92521, USA

<sup>8</sup>European Southern Observatory, Karl-Schwarzschild-Str. 2, D-85748 Garching, Germany

Accepted 2018 March 21. Received 2018 March 17; in original form 2018 February 27

## ABSTRACT

Deep narrow-band surveys have revealed a large population of faint Ly  $\alpha$  emitters (LAEs) in the distant Universe, but relatively little is known about the most luminous sources ( $L_{\text{Ly}\alpha} \gtrsim 10^{42.7} \text{ erg s}^{-1}$ ;  $L_{\text{Ly}\alpha} \gtrsim L_{\text{Ly}\alpha}^*$ ). Here we present the spectroscopic follow-up of 21 luminous LAEs at  $z \sim 2\text{--}3$  found with panoramic narrow-band surveys over five independent extragalactic fields ( $\approx 4 \times 10^6 \text{ Mpc}^3$  surveyed at  $z \sim 2.2$  and  $z \sim 3.1$ ). We use WHT/ISIS, Keck/DEIMOS, and VLT/X-SHOOTER to study these sources using high ionization UV lines. Luminous LAEs at  $z \sim 2\text{--}3$  have blue UV slopes ( $\beta = -2.0_{-0.1}^{+0.3}$ ) and high Ly  $\alpha$  escape fractions ( $50_{-15}^{+20}$  per cent) and span five orders of magnitude in UV luminosity ( $M_{\text{UV}} \approx -19$  to  $-24$ ). Many (70 per cent) show at least one high ionization rest-frame UV line such as C IV, N V, C III], He II or O III], typically blue-shifted by  $\approx 100\text{--}200 \text{ km s}^{-1}$  relative to Ly  $\alpha$ . Their Ly  $\alpha$  profiles reveal a wide variety of shapes, including significant blue-shifted components and widths from 200 to  $4000 \text{ km s}^{-1}$ . Overall,  $60 \pm 11$  per cent appear to be active galactic nucleus (AGN) dominated, and at  $L_{\text{Ly}\alpha} > 10^{43.3} \text{ erg s}^{-1}$  and/or  $M_{\text{UV}} < -21.5$  virtually all LAEs are AGNs with high ionization parameters ( $\log U = 0.6 \pm 0.5$ ) and with metallicities of  $\approx 0.5 - 1 Z_{\odot}$ . Those lacking signatures of AGNs ( $40 \pm 11$  per cent) have lower ionization parameters ( $\log U = -3.0_{-0.9}^{+1.6}$  and  $\log \xi_{\text{ion}} = 25.4 \pm 0.2$ ) and are apparently metal-poor sources likely powered by young, dust-poor ‘maximal’ starbursts. Our results show that luminous LAEs at  $z \sim 2\text{--}3$  are a diverse population and that  $2 \times L_{\text{Ly}\alpha}^*$  and  $2 \times M_{\text{UV}}^*$  mark a sharp transition in the nature of LAEs, from star formation dominated to AGN dominated.

**Key words:** galaxies: active – galaxies: evolution – galaxies: high-redshift – galaxies: ISM – galaxies: starburst – cosmology: observations.

## 1 INTRODUCTION

There has been a steady increase in the number of sources identified at moderately high redshifts ( $z \gtrsim 2\text{--}3$ ) over the past 20 yr (e.g. Ellis et al. 2013; Bowler et al. 2014; Atek et al. 2015; Bouwens et al. 2015a,b; Finkelstein et al. 2015; Santos, Sobral & Matthee 2016).

\* E-mail: d.sobral@lancaster.ac.uk

† ESO Fellow.

Most have been found using the Lyman-break technique (e.g. Koo & Kron 1980; Guhathakurta, Tyson & Majewski 1990; Steidel & Hamilton 1993; Giavalisco, Steidel & Macchetto 1996), combined with deep, multiband imaging with the *Hubble Space Telescope* (*HST*). However, spectroscopy of faint continuum-selected candidates has progressed at a much slower pace as spectroscopic continuum detections are difficult for such faint sources. A different approach is to find and study sources with bright Ly  $\alpha$  in emission (e.g. Ouchi et al. 2008; Oesch et al. 2015; Sobral et al. 2015; Zitrin et al. 2015; Hu et al. 2016; Jiang et al. 2017; Matthee et al. 2017c)

across a range of redshifts (e.g. Møller & Warren 1993; Harikane et al. 2017; Sobral et al. 2018). While these can now be found with wide-field ground-based surveys (e.g. Matthee et al. 2015; Hu et al. 2016; Santos et al. 2016; Zheng et al. 2017; Shibuya et al. 2018), much is unknown about their nature, metallicity, and stellar populations (see e.g. Nilsson et al. 2007; Wold, Barger & Cowie 2014; Matthee et al. 2017d; Sobral et al. 2018).

Significant spectroscopic progress has also been made by targeting lensed galaxies (e.g. Swinbank et al. 2010; Vieira et al. 2013; Stark et al. 2014, 2015a), allowing the detection and study of intrinsically very faint lines. Combining results from intrinsically bright and faint sources reveals a picture in which galaxies appear to have ubiquitous high equivalent width (EW) nebular emission lines at high redshift (e.g. Smit et al. 2014; Sobral et al. 2014; Khostovan et al. 2016; Mármol-Queraltó et al. 2016; Smit et al. 2016). This may be due to high redshift galaxies producing more ionizing photons per UV luminosity (e.g. Topping & Shull 2015; Bouwens et al. 2016; Matthee et al. 2017a; Stanway 2017). Recent results also highlight that high redshift galaxies can have hard ionizing spectra, as indicated by UV emission lines such as C III], C IV, O III] or He II (e.g. Stark et al. 2015b; Sobral et al. 2015; Jaskot & Ravindranath 2016; Laporte et al. 2017; Schmidt et al. 2017). This is in agreement with more detailed, larger statistics results, which show an increase in the ionization parameter with redshift, traced by e.g. [O III]/[O II] (e.g. Erb et al. 2010; Nakajima & Ouchi 2014; Khostovan et al. 2016; Nakajima et al. 2016; Vanzella et al. 2016, 2017) or evidence for hard spectra from iron-poor stellar populations (Steidel et al. 2016). Understanding the origin of the strong evolution relative to local galaxies is still a major open question. Many mechanisms/physical processes have been proposed, including binary stars, bursty star formation histories, AGNs, and several effects at low metallicity (e.g. Eldridge, Izzard & Tout 2008; Eldridge & Stanway 2009, 2012; Gräfenor & Vink 2015; Ma et al. 2016; Stanway, Eldridge & Becker 2016).

Furthermore, the common approach of selecting Lyman-break galaxies (LBGs) has resulted in statistically large samples, but failed to reveal a convincing population of Lyman continuum (LyC) leakers (e.g. Siana et al. 2007; Guaita et al. 2016; Marchi et al. 2017b; Marchi, Pentericci et al. 2017a), required for re-ionization (e.g. Faisst 2016). This may be a consequence of selection, as LBGs are biased against sources leaking significant amounts of LyC photons (e.g. Cooke et al. 2014). LBGs tend to be more evolved and show little to no Ly $\alpha$  in emission (e.g. Cassata et al. 2015). In addition, recent theoretical and observational work suggests an important link between LyC and Ly $\alpha$  escape fractions (e.g. Verhamme et al. 2015, 2017; Dijkstra, Gronke & Venkatesan 2016; Izotov et al. 2018; Vanzella et al. 2018), implying that Ly $\alpha$  emitters (LAEs) contribute significantly to the LyC luminosity density. The faint end slope of the Ly $\alpha$  luminosity function (LF) is steep ( $\alpha \sim -2$ ), implying a large number of faint LAEs (e.g. Rauch et al. 2008; Dressler et al. 2015; Gronke et al. 2015; Santos et al. 2016; Drake et al. 2017b) and LAEs produce typically many more ionizing photons per UV luminosity than LBGs (e.g. Nakajima et al. 2016; Harikane et al. 2017; Matthee et al. 2017a). The combination of recent results on LAEs is particularly promising and further enhances the motivation to study those: the high EW LAEs may be exactly what is needed to re-ionize the Universe, having ‘high’ escape fractions (e.g. Verhamme et al. 2017), high production of ionizing photons (e.g. Nakajima & Ouchi 2014; Nakajima et al. 2016, 2018), and a large number of densities (e.g. Drake et al. 2017b; Sobral et al. 2018).

Despite recent progress, little is known about the nature and redshift evolution of the most luminous LAEs ( $L_{\text{Ly}\alpha} \sim$

$10^{43-44.5} \text{ erg s}^{-1}$ ). At  $z \sim 2-3$  spectroscopic studies have mostly focused on the follow-up of fainter, more numerous LAEs typically below the characteristic Ly $\alpha$  ( $L_{\text{Ly}\alpha} \lesssim 10^{43} \text{ erg s}^{-1}$ ) and UV luminosities at that redshift (e.g. Fynbo, Møller & Thomsen 2001; Fynbo et al. 2003; Ouchi et al. 2008; Grove et al. 2009; Trainor et al. 2015; Hathi et al. 2016), or by blindly finding and studying faint LAEs with deep but small volume searches with IFU instruments such as MUSE (e.g. Drake et al. 2017a; Hashimoto et al. 2017; Caruana et al. 2018). Important progress has also been made by studying extremely rare quasars found with SDSS (e.g. Richards et al. 2006),  $\approx 7-10$  orders of magnitude brighter in the UV and revealing large Ly $\alpha$  haloes with  $L_{\text{Ly}\alpha} \sim 10^{43-44} \text{ erg s}^{-1}$  (e.g. Borisova et al. 2016). Still, little is known about the general population of bright LAEs, although there is some evidence for these sources to have a significant AGN contribution (e.g. Ouchi et al. 2008; Nilsson et al. 2009; Matthee et al. 2017b; Sobral et al. 2017b).

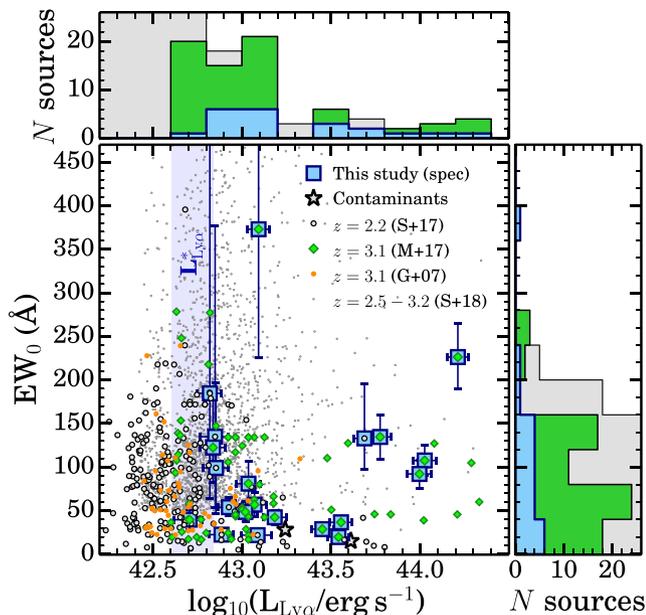
In this paper we study bright LAEs at  $z \sim 2-3$  through relatively deep spectroscopic follow-up, in order to unveil their nature and physical properties. These bright sources span the parameter space between faint LAEs and rare quasars. The paper is organized as follows. In Section 2 we present the observations of our sample of bright Ly $\alpha$  candidates with WHT/ISIS, Keck/DEIMOS, and VLT/X-SHOOTER, followed by the respective data reduction. In Section 3 we present the measurements and analysis. Results are presented in Section 4. The nature of bright LAEs is studied in Section 5, based on high ionization UV lines. We discuss our results in Section 6. Conclusions are presented in Section 7. Throughout this paper, we use AB magnitudes (Oke & Gunn 1983), a Salpeter (Salpeter 1955) IMF, and a  $\Lambda$ CDM cosmology with  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_M = 0.3$ , and  $\Omega_\Lambda = 0.7$ .

## 2 OBSERVATIONS AND DATA REDUCTION

### 2.1 Sample selection

Our luminous LAEs are selected from wide-field narrow-band Ly $\alpha$  surveys. Candidate bright LAEs at  $z \sim 2.2$  are selected from the  $1.43 \text{ deg}^2$  CALYMHA survey ( $\text{EW}_0 > 5 \text{ \AA}$ , covering the UDS and COSMOS fields; see Matthee et al. 2016; Sobral et al. 2017b, 2018). We select those with Ly $\alpha$  luminosities above  $10^{42.8} \text{ erg s}^{-1}$  (up to  $10^{43.6} \text{ erg s}^{-1}$ ); see Fig. 1. Luminous  $z \sim 3.1$  Ly $\alpha$  candidates are selected from a total of  $\sim 3.1 \text{ deg}^2$  over the GOODS-N, SA22 and Boötes fields ( $\text{EW}_0 > 25 \text{ \AA}$ ; see Matthee et al. 2017b) to have luminosities in excess of  $10^{42.7} \text{ erg s}^{-1}$  (up to  $10^{44.2} \text{ erg s}^{-1}$ ). Our targets have number densities of  $10^{-4}-10^{-6} \text{ Mpc}^{-3}$  (e.g. Matthee et al. 2017b; Sobral et al. 2018), and have been found in a total volume of  $\approx 4 \times 10^6 \text{ Mpc}^3$ . We note that due to the selection applied to obtain candidate LAEs from the full sample of emission line candidates (see Sobral et al. 2018), our samples will be biased against potential (galaxy-galaxy) lensed sources, as the continuum of the foreground lens will lead to classifying lensed LAEs as low redshift interlopers.

In total, we have targeted 23 sources as candidate luminous LAEs. We present their Ly $\alpha$  luminosities and rest-frame Equivalent Widths ( $\text{EW}_0$ ) in Fig. 1, compared to the parent sample of more typical, lower luminosity LAEs at both  $z \sim 2.2$  (Sobral et al. 2017b) and  $z \sim 3.1$  (Matthee et al. 2017b). As Fig. 1 shows, our sample is representative of the most luminous LAEs, covering 1.5 dex in Ly $\alpha$  luminosities ( $10^{42.7-44.2} \text{ erg s}^{-1}$ ) and 1.2 dex in rest-frame Ly $\alpha$   $\text{EW}_0$  (20–400  $\text{\AA}$ ).



**Figure 1.** Comparison between the distribution of Ly $\alpha$  luminosities and rest-frame EWs (both quantities derived from narrow-band imaging) of the  $z = 2.2$  (from the CALYMA survey, S+17; Matthee et al. 2017b; Sobral et al. 2017b) and  $z = 3.1$  parent samples (M+17; Matthee et al. 2017b). We also show a sample of fainter LAEs at  $z = 3.1$  probing a smaller volume (G+07; Gronwall et al. 2007) and the large sample of Ly $\alpha$  emitters at  $z \sim 2.5-3.2$  (SC4K; Sobral et al. 2018).  $L_{\text{Ly}\alpha}^*$  (the knee/typical luminosity of the Ly $\alpha$  LF; Sobral et al. 2018) is indicated at  $z \sim 2-3$  as a shaded region, representing the errors and positive evolution from  $z \sim 2$  to  $z \sim 3$ . Our spectroscopic sample is representative of  $L_{\text{Ly}\alpha} \gtrsim L_{\text{Ly}\alpha}^*$  LAEs, probing a large range in luminosities and EWs.

## 2.2 Spectroscopic observations

### 2.2.1 WHT/ISIS

We used the Intermediate dispersion Spectrograph and Imaging System (ISIS) on the William Herschel Telescope (WHT) to observe  $z \sim 2-3$  Ly $\alpha$  candidates. ISIS is a double-armed, medium-resolution spectrograph, which allows simultaneous observing with the blue and red arms.

A total of six  $z \sim 2-3$  Ly $\alpha$  candidates were selected from GOODS-N and Boötes. We observed each source for roughly 1 h during 2016 May 4–6 and June 30–July 1; (see Table A1). All observations were done under clear conditions and the seeing varied between  $\sim 0.4$  arcsec and  $\sim 1.5$  arcsec (typical seeing  $\approx 0.9''$ ). We used a 1 arcsec slit for all observations.

The blue and red arms were used with the R600 gratings, with central wavelengths 4800 Å ( $R \sim 2400$ ) and 7000 Å ( $R \sim 3900$ ) on the first three nights and with the R300B/R316R gratings with central wavelengths 5000 Å ( $R \sim 1200$ ) and 8000 Å ( $R \sim 2200$ ) for the final two nights. We obtained biases, arcs and flats for each grating, arm and central wavelength, at the start and end of each night.

Our targets are all too faint (typically  $I \sim 23-24$  mag) in the continuum to be directly acquired using WHT/ISIS, and thus we identified a relatively nearby star ( $I \sim 13-16$  mag,  $\sim 20-60$  arcsec away from our target) which we used for acquisition. We conduct our observations by first acquiring the offset star, taking a 30–60 s (depending on its magnitude) exposure on the star, then blindly offsetting the telescope to the target and taking one 900–1000 s

exposure of the target. After each science exposure, we go back to the star, offset along the slit by  $\pm 10$  arcsec, acquire the star again, take another star spectrum, and then offset to the science target to take another exposure. This procedure allows us to use the star as a flux calibrator and slit loss estimator. Our procedure also allows us to extract the trace and its curvature in an optimal way, which we use to combine and extract the spectra (see Section 2.3.1).

### 2.2.2 Keck/DEIMOS

DEIMOS was used to spectroscopically target eight luminous Ly $\alpha$  candidates at  $z \sim 3$  in five different masks for the Boötes field, and in one mask for GOODS-N over three different nights in 2016 June 2, July 6, and July 29. We have also observed other line emitters (see Table A3 and Matthee et al. 2017b). Observations were conducted under clear conditions. The seeing was in the range  $\approx 0.6-0.9$  arcsec. We used a central wavelength of 7200 Å and the 600L grating ( $R \sim 2400$ ), with a pixel scale of  $0.65 \text{ Å pix}^{-1}$ , which allowed us to probe from 4550 Å to 9850 Å. We used the 0.75 arcsec slit, in the same mode as used by Darvish et al. (2015).

We obtained biases, arcs, and flats at the start of each observing night. Typical science exposures of 900 s were obtained with a  $\pm 2.5$  arcsec dithering for two of the nights and without dithering for the final observing night on 2016 July 30. Observations are listed in Table A3.

### 2.2.3 VLT/X-SHOOTER

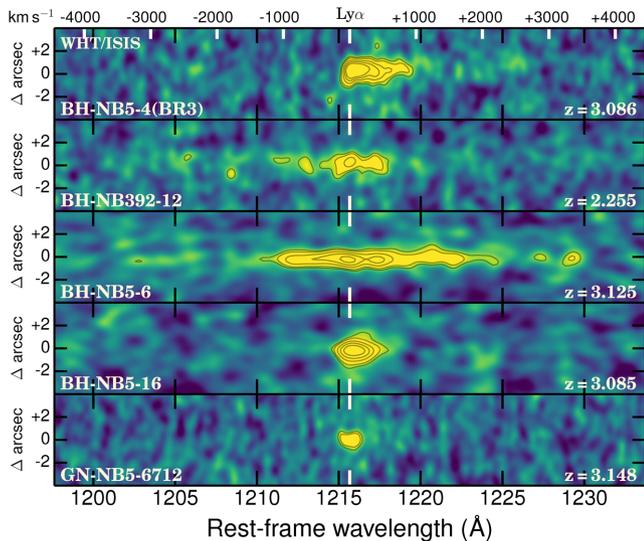
We used X-SHOOTER (Vernet et al. 2011) to observe 11 candidate luminous  $z = 2.2-3.1$  LAEs in 2016 October, 2017 January, June–July, December, and 2018 January (see Table A4). X-SHOOTER allows us to simultaneously obtain a relatively high-resolution spectrum with the UV/blue (UVB), visible (VIS), and near-infrared (NIR) arms, providing a coverage from 3000 to 24 800 Å. The seeing varied between 0.8 arcsec and 1.5 arcsec (median seeing 0.9 arcsec) and observations were done under clear conditions. We used the low read-out speed without binning. Typical exposure times of 200 s, 300 s, and  $4 \times 80$  s in the UV, optical, and NIR arms were used, respectively (see Table A4). We first acquired a star (with  $I$ -band magnitudes 16–17 AB) and applied a blind offset to the target. We nodded along the slit from an A to a B position (typically 4–6 arcsec apart), including a small jitter box in order to always expose on different pixels. We used 1 arcsec slits for both the optical and near-infrared arms (resolution of  $R \sim 2500$  and  $R \sim 4400$ , for the optical and near-infrared arms, respectively) and the 1.2 arcsec slit in the UV ( $R \sim 4400$ ). The total exposure times for each source are given in Table A4.

## 2.3 Data reduction

### 2.3.1 WHT/ISIS: the William Herschel ISIS pipeline

We developed a PYTHON pipeline to reduce ISIS spectra<sup>1</sup> and exploit our observing setup/methodology. We reduce the data on a night-by-night basis, and only combine data from different nights for extracted 1D spectra. Our pipeline follows standard steps: we bias-subtract, flat-field, and wavelength-calibrate the spectra using appropriate calibration frames. Additional wavelength calibration

<sup>1</sup> The William Herschel ISIS pipeline (see Appendix A1).



**Figure 2.** The reduced 2D Ly $\alpha$  spectra, displayed in the S/N space, smoothed with a Gaussian kernel of 2 pixels, from our WHT/ISIS observations, labelled with ID and redshift, and ordered by Ly $\alpha$  flux top to bottom (brightest on top). We show  $1.5\sigma$ ,  $2\sigma$ ,  $3\sigma$ ,  $4\sigma$ ,  $5\sigma$  contours and use contrast cutoffs at  $-1$  and  $+2\sigma$  to display the 2D spectra. The top axis shows the velocity shifts from the Ly $\alpha$  redshift. We find one broad-line emitter (BH-NB5-6), and one with a blue wing/complicated dynamics (BH-NB392-12). The brightest source observed with WHT/ISIS is BR3, which shows the most asymmetric Ly $\alpha$  line profile. The faintest source targeted with Ly $\alpha$  coverage is GN-NB5-6712, which is also one of the sources showing the narrowest Ly $\alpha$  profile.

is performed by exploiting the sky lines in deep spectroscopic exposures. We describe the processes in more detail in Appendix A1. In Fig. 2 we show our reduced 2D spectra, centred on the Ly $\alpha$  line within a window of  $\pm 4000$  km s $^{-1}$ .

We flux calibrate our spectra by using the blind offset stars and their SDSS magnitudes. The resolution (FWHM) of each arm/grating is estimated with un-blended sky lines and found to be 2 and 1.8 Å for the R600B/R600R gratings spectra and 4.0 and 3.6 Å FWHM for the R300B/R316R gratings. We bin our 1D spectra by roughly one-third of the corresponding FWHM, which, in practice, results in binning by two spectral pixels. We show the typical rms per binned resolution element in Table A2.

### 2.3.2 Keck/DEIMOS: spec2d pipeline

All Keck/DEIMOS data were reduced using the DEIMOS SPEC2D pipeline (Cooper et al. 2012), which follows very similar steps to our WHT/ISIS data reduction. Briefly, all observed spectra were flat-fielded, cosmic ray-removed, sky-subtracted, and wavelength-calibrated (see Darvish et al. 2015). We used standard Kr, Xe, Ar, and Ne arc lamps for the wavelength calibration. We also checked the wavelength solution at the end of the pipeline reduction, taking advantage of the numerous OH sky lines. For the observations obtained with a dithering pattern, we find significant offsets between the observed sky line positions and the correct OH line wavelengths of over 10–20 Å, varying in a non-linear way in the blue and in the red. Therefore, for all observations for which we dithered, we identify  $\sim 100$  unblended OH lines and obtain a final wavelength calibration, which produces an rms of  $\approx 0.5$  Å, about  $\sim 6$ –7 times better than the resolution.

For the final night of observations with Keck/DEIMOS, no dithering pattern was used for sky subtraction. Without dithering, the wavelength calibration provided by the pipeline was found to be accurate within 0.5 Å, and no further correction was necessary. We show the reduced 2D spectra in Fig. 3.

The pipeline also generates the 1D spectrum extraction from the reduced 2D spectrum, following the optimal extraction algorithm of Horne (1986). This extraction creates a one-dimensional spectrum of the target, containing the summed flux at each wavelength in an optimized window. We flux calibrate the data with bright enough sources within the masks, and also observations of the Feige 66 standard star. We bin our spectra to one-third of the observed resolution (FWHM) of 3 Å (which results in binning 2 spectral pixels,  $\approx 1$  Å). Our final resolution binned spectra have a typical rms per binned resolution element ( $1\sigma$ ) of  $3 \times 10^{-19} - 5 \times 10^{-19}$  erg s $^{-1}$  cm $^{-2}$  Å $^{-1}$  (see Table A2) and thus typically a factor of  $\sim 10$  deeper than WHT/ISIS.

### 2.3.3 VLT/X-SHOOTER: ESO pipeline

We use the ESO X-SHOOTER pipeline (Modigliani et al. 2010) to reduce the UVB, VIS, and NIR spectra separately (see also Matthee et al. 2017c). The data reduction steps follow closely those implemented for both ISIS and DEIMOS data, with the necessary differences, particularly for the X-SHOOTER NIR arm which allow us to obtain [O III] and H $\alpha$  for the  $z \sim 2.2$  LAEs (NIR spectral properties will be fully discussed in Matthee et al., in preparation).

In short, we start by grouping frames. We then identify and process the bias frames (to create master biases in the UVB/VIS arms) and dark current (NIR arm). We produce master flats per arm and flat-field the data. Finally, sky subtraction and wavelength calibration are done. In order to be able to flux calibrate the spectra, several standard stars have been observed (see Table A4). We use the X-SHOOTER pipeline to reduce the standard stars in the same way as the science targets and combine the exposures from single observing blocks.

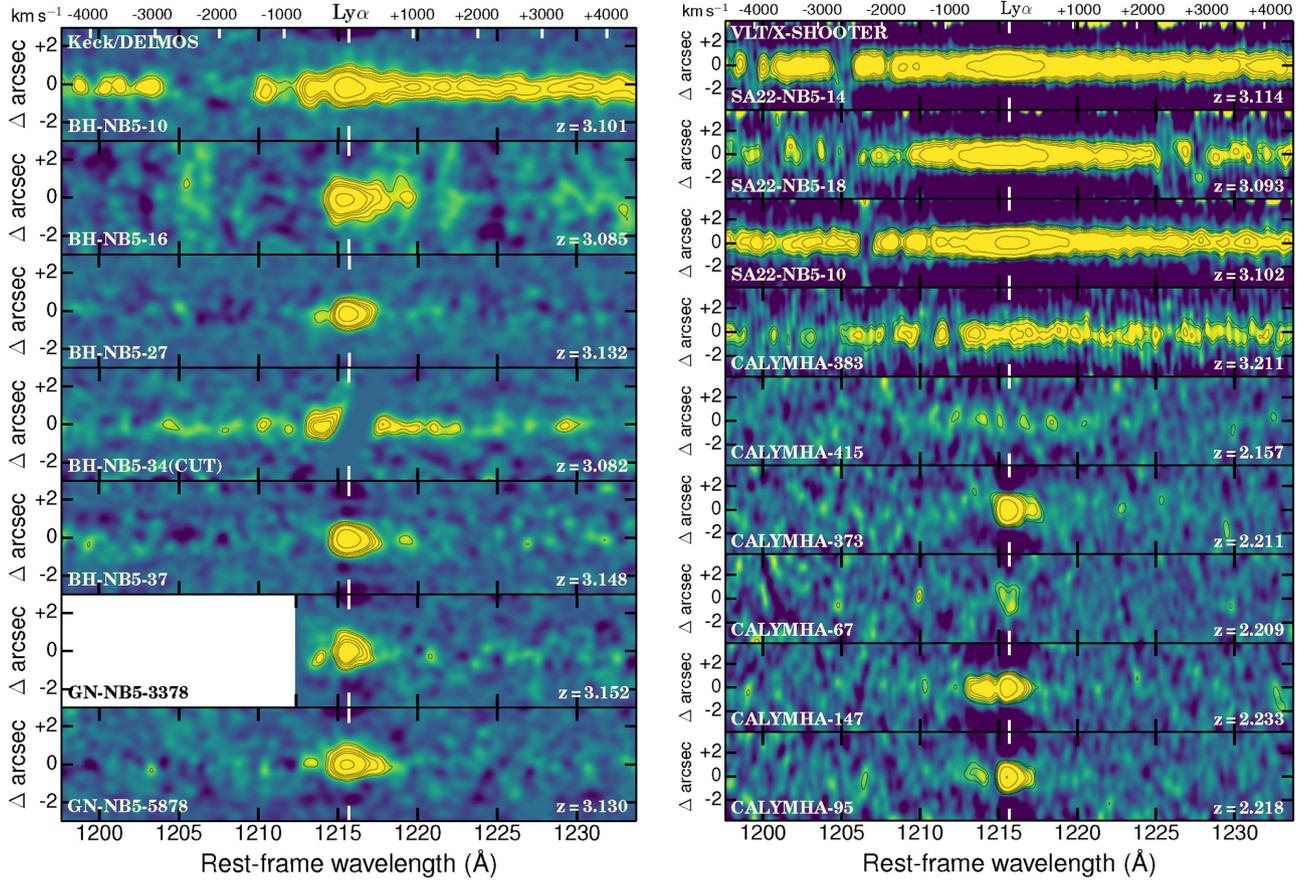
In the case that a source has been observed by multiple observing blocks, we co-add the frames by weighting each frame with the inverse of the variance (noise) in 2D. Corrections for slight positional variations based on the position of the peak of observed Ly $\alpha$  lines were applied. We show the reduced 2D spectra in Fig. 3, ordered by their narrow-band-derived Ly $\alpha$  luminosity.

We extract 1D spectra by summing the counts in an optimized spatial window, which typically corresponds to a total of  $\sim 10$  spatial pixels (1.8 arcsec). We use isolated sky lines together with arc lines to find that the resolution of our spectra (FWHM) is 1.6 Å in the UVB and VIS arms and 3.6 Å in the NIR arm. We finally bin our spectra to one-third of the FWHM which, in practice, corresponds to binning 3 and 2 spectral pixels for the UVB/VIS and NIR arms, respectively. Table A2 provides the typical rms per binned resolution element for our binned spectra.

## 3 MEASUREMENTS AND ANALYSIS

### 3.1 Spectroscopic redshifts and line identification

All our spectra are calibrated to air wavelengths. Before we determine redshifts, we first convert from air wavelengths to vacuum (see Morton 1991), and then use vacuum rest-frame wavelengths for the rest of the paper (see Table 1).



**Figure 3.** The reduced 2D spectra of luminous LAEs from our DEIMOS (*left*) and X-SHOOTER (*right*) observations, smoothed with a Gaussian kernel of 2 pixels, labelled with ID and redshift, listed from the brightest (top) to the faintest (bottom) per instrument. We show  $1.5\sigma$ ,  $2\sigma$ ,  $3\sigma$ ,  $4\sigma$ ,  $5\sigma$ ,  $10\sigma$ ,  $20\sigma$  contours. Our LAEs present a wide variety of Ly $\alpha$  line profiles with some broad lines (FWHMs in excess of  $1000 \text{ km s}^{-1}$ ) at the highest Ly $\alpha$  luminosities, and narrower Ly $\alpha$  lines at lower Ly $\alpha$  luminosities. Five sources show a blue-shifted Ly $\alpha$  component, with CALYMHA-147 being the most noticeable example and showing the highest blue-to-red ratio ( $\approx 0.6 \pm 0.1$ ). LAEs with significant UV continuum also reveal the Ly $\alpha$  forest. Note that the ‘negative flux’ regions are a consequence of jittering along the slit.

**Table 1.** A list of high ionization rest-frame UV and optical lines used in this paper (see also Veilleux 2002). We list them in vacuum wavelengths (see Morton 1991). For spectra with high enough resolution to resolve doublet lines, we also attempt to fit a double Gaussian. For simplicity, and unless the S/N is high enough, we fix the doublet separation and impose that both lines should have the same FWHM.

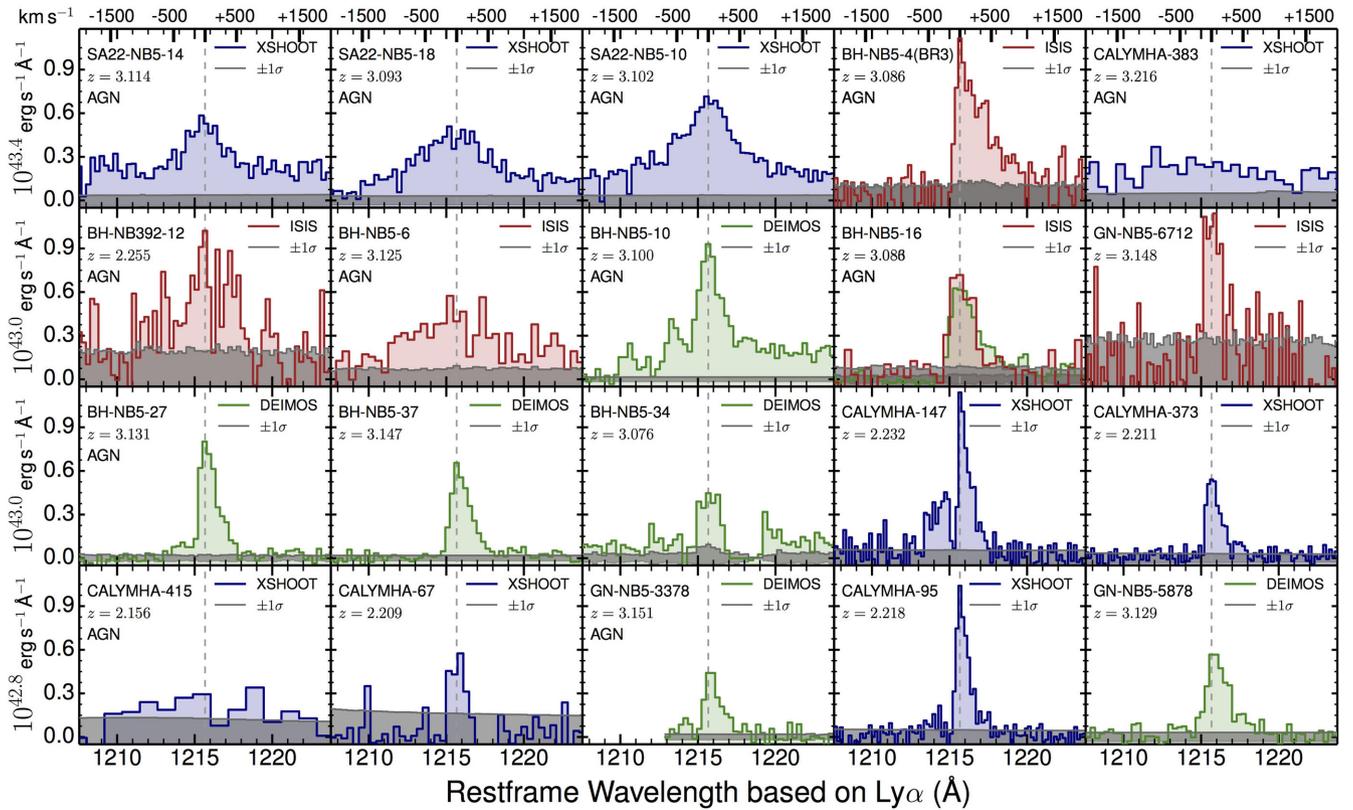
Emission line	$\lambda_{\text{vacuum}}$ (Å)	Ionization Energy (eV)	# Detections (99.4 per cent conf.)
Ly $\alpha$	1215.67	13.6	20 (100 per cent)
Nv	1238.8, 1242.8	77.4	6 (33 per cent)
O IV]	1401, 1407	54.9	1 (11 per cent)
N IV]	1483.4, 1486.6	47.4	2 (11 per cent)
C IV]	1548.2, 1550.8	47.9	8 (40 per cent)
He II	1640.47	54.4	5 (25 per cent)
O III]	1661, 1666	35.1	1 (5 per cent)
N III]	1749.7, 1752.2	29.6	2 (11 per cent)
C III]	1907, 1910	24.4	4 (25 per cent)

We start by obtaining a redshift using the Ly $\alpha$  emission only ( $z_{\text{Ly}\alpha}$ ) and identifying its peak (see Figs 2, 3, and 4). We store these as the redshift of the Ly $\alpha$  emission line, and use those as a first approximation of the redshift of the source. We then look for other lines (see Table 1) by searching and fitting Gaussians (or double

Gaussians) and deriving a redshift for them, allowing for velocity offsets up to  $\pm 1000 \text{ km s}^{-1}$ .

### 3.2 Contaminants

We confirm that the vast majority (91 per cent) of our 23 candidates are luminous LAEs, but we also find two contaminants. These include one low redshift [O II] emitter at  $z = 0.056$  (CALYMHA-438) and a star with a large number of absorption bands (CALYMHA-85), both with very low EWs (see Fig. 1). These imply an overall contamination of  $\approx 10$  per cent, likely lower at  $z = 3.1$  than at  $z = 2.23$  (a consequence of the very low EW cut used for  $z = 2.23$ ; see discussion in Sobral et al. 2017b). We further study the potential contamination and completeness of the larger sample of bright narrow-band selected sources from the parent samples (Matthee et al. 2017b; Sobral et al. 2017b) by using the follow-up of other sources with Keck/DEIMOS (see Matthee et al. 2017b). Based on one missed real LAE and 2 extra [O II] contaminants (see Matthee et al. 2017b, for further details), we estimate a conservative contamination of  $\approx 16$  per cent (four contaminants in a sample of 25 sources) at  $z \sim 2-3$  for our parent sample of bright LAEs and a completeness of  $>90$  per cent (only one real LAE missed by our selection).



**Figure 4.** Our 1D extracted, flux calibrated Ly $\alpha$  spectra for the three instruments (shown with different colours) used for our entire sample of 20 luminous LAEs with Ly $\alpha$  coverage. We order sources from the most to the least luminous in Ly $\alpha$  (based on the NB estimate prior to spectroscopy) from the top to bottom and left to right (note a decrease of 0.6 dex from the upper to lower panels), and we find significant differences in the line widths and profiles with decreasing luminosity. Overall, bright LAEs present a wide variety of Ly $\alpha$  line profiles, with FWHMs in excess of 1000 km s $^{-1}$  at the highest Ly $\alpha$  luminosities, and narrow Ly $\alpha$  lines at lower Ly $\alpha$  luminosities. Five sources (25 per cent) show a blue-shifted Ly $\alpha$  component and CALYMHA-147 presents the highest blue-to-red ratio of  $\approx 0.6 \pm 0.1$ . Note that BH-NB5-55 does not have spectral coverage at the Ly $\alpha$  wavelength and thus it is not shown. We show the spectra of BH-NB5-16 observed with both DEIMOS and ISIS, showing good agreement. Some sources with very low exposure time of just a few minutes (e.g. CALYMHA-415) yield relatively low S/N per spectral element, but their redshifts and other properties are securely determined by higher S/N detections of at least another line. Further properties of our Ly $\alpha$  sources are provided in Table 2.

### 3.3 Final Ly $\alpha$ sample

Out of the 23 targeted sources, we spectroscopically confirm 21 as LAEs,<sup>2</sup> with seven sources at  $z \approx 2.2$  and 14 sources at  $z \approx 3.1$ . Our final sample of bright LAEs ( $L_{\text{Ly}\alpha} \gtrsim 10^{42.7} \text{ erg s}^{-1}$ ;  $L_{\text{Ly}\alpha} \gtrsim L_{\text{Ly}\alpha}^*$ ) is thus composed of 21 sources. We present the sample in Table 2 and provide both the catalogue and the reduced 1D spectra on-line, with the published version of the paper.

### 3.4 Line measurements and upper limits

We explore our 1D spectra to fit single Gaussians and/or double Gaussians, depending on whether the emission is a doublet or not (Table 1). When fitting a double Gaussian, we fix the doublet separation and set the FWHM to be the same for both lines. In order to avoid problems with OH sky lines and very low signal-to-noise (S/N) regions of the spectra, we mask regions with the strongest OH lines. We perturb each spectrum in a window of  $\pm 4000 \text{ km s}^{-1}$  around each line being evaluated (see Table 1) by independently varying each spectral element/data point along their Gaussian probability (flux) and re-fit. We do this 10000 times per source and

per line and take the median, 16th, and 84th percentiles as the best values and the lower and upper errors. We also compute the 0.6th and 99.4th percentiles, roughly corresponding to  $-2.5\sigma$  to  $+2.5\sigma$ . Whenever the flux of the 0.6th percentile of a given line is consistent with zero or below, we assign that line its 99.4th flux percentile as an upper limit. Table 1 provides the list of the lines that we use. As a further step, we visually inspect all lines and the best fits.

Not surprisingly, our most significant line is Ly $\alpha$ , with a median S/N (integrated) of  $\sim 20$  (ranging from  $\sim 4$  to  $\sim 60$ ) across the entire sample and all instruments. The second most common line at high S/N is C IV (see Table 1) significantly detected in about half of the sample with an S/N from  $\sim 3$  to  $\sim 30$ . Other high ionization UV lines like N V, He II and C III] are detected in  $\sim 4$ – $6$  sources at typical S/N of  $\sim 3$ – $10$  (BR3 yields S/N ratios of up to 50 in these lines). N IV is only significantly detected with an S/N  $\sim 7$  in BH-NB5-4 (BR3), although tentatively detected at S/N  $\sim 2.5$  in another source (see Tables 1 and B1).

### 3.5 UV luminosities and UV $\beta$ slope

We compute the UV luminosity at rest-frame  $\approx 1500 \text{ \AA}$  ( $M_{\text{UV}}$ ) for all our luminous LAEs. We use the magnitude in the closest observed

<sup>2</sup> For one LAE, BH-NB392-55, we do not have coverage of the Ly $\alpha$  line itself, but rather confirm it through strong C IV emission (see Fig. A1).

**Table 2.** Our spectroscopic sample of luminous LAEs, ranked by redshift and then by Ly $\alpha$  luminosity of the parent sample (high to low luminosity top to bottom). We present the main measured photometric and Ly $\alpha$  properties and also provide the most likely nature classification of each source, based on the FWHM of the Ly $\alpha$  line (BL-AGN versus NL emitters), but also on the ratios between the different rest-frame UV emission lines for the NL emitters (AGN versus SF; see Section 4.1.2). All errors are the 16th and 84th percentiles. Some uncertainties in the Ly $\alpha$  redshift are lower than 0.001, but here we set them 0.001 if they are below that value. We add 15 per cent uncertainties in the flux calibration or zero-point calibration in quadrature for the final error on the Ly $\alpha$  luminosity. (1) CALYMHA-147 shows a bright blue-component and the FWHM given is a fit to both components; the FWHM of the red component is  $\approx 180 \text{ km s}^{-1}$ , while the blue component has an FWHM of  $\approx 400 \text{ km s}^{-1}$ . (2) The FWHM of this source should be interpreted with care due to the lack of coverage towards the redder part of Ly $\alpha$  (see e.g. Fig. 3). This table is available on-line as a FITS catalogue without digit truncation.

ID (This paper)	R.A. (J2000)	Dec. (J2000)	$z_{\text{spec}}$ (Ly $\alpha$ )	$M_{\text{UV}}$ (AB)	$\beta_{\text{UV}}$	$\log L_{\text{Ly}\alpha}$ (erg s $^{-1}$ )	EW $_0$ ( $\text{\AA}$ )	FWHM (km s $^{-1}$ )	Class.
BH-NB392-12	14 30 28.55	+33 33 29.0	2.255 $^{+0.001}_{-0.001}$	-21.9 $^{+0.2}_{-0.2}$	-1.4 $^{+0.2}_{-0.2}$	43.69 $^{+0.06}_{-0.06}$	130 $^{+50}_{-40}$	1320 $^{+240}_{-200}$	BL-AGN
CALYMHA-415	02 16 33.29	-05 17 57.8	2.153 $^{+0.004}_{-0.003}$	-21.1 $^{+0.2}_{-0.2}$	-2.0 $^{+0.1}_{-0.1}$	43.18 $^{+0.07}_{-0.07}$	22 $^{+5}_{-4}$	2010 $^{+450}_{-570}$	BL-AGN
BH-NB392-55	14 30 40.31	+34 03 20.6	2.197 $^{+0.568}_{-0.003}$	-22.4 $^{+0.2}_{-0.2}$	-2.0 $^{+0.1}_{-0.1}$	43.09 $^{+0.08}_{-0.08}$	22 $^{+7}_{-6}$	—	Unclass.
CALYMHA-373	02 17 46.13	-05 02 55.5	2.211 $^{+0.001}_{-0.001}$	-20.9 $^{+0.2}_{-0.2}$	-2.1 $^{+0.2}_{-0.2}$	42.93 $^{+0.06}_{-0.06}$	54 $^{+11}_{-10}$	260 $^{+20}_{-20}$	NL-SF
CALYMHA-147	10 00 13.91	+01 39 24.3	2.232 $^{+0.001}_{-0.001}$	-19.9 $^{+0.2}_{-0.2}$	-1.9 $^{+0.4}_{-0.4}$	42.85 $^{+0.07}_{-0.08}$	100 $^{+80}_{-40}$	620 $^{+60}_{-40}$ (1)	NL-SF
CALYMHA-67	10 01 36.21	+02 15 16.8	2.209 $^{+0.001}_{-0.001}$	-19.7 $^{+0.2}_{-0.2}$	-2.1 $^{+0.3}_{-0.3}$	42.85 $^{+0.07}_{-0.08}$	130 $^{+200}_{-70}$	280 $^{+70}_{-60}$	NL-SF
CALYMHA-95	10 01 06.55	+01 45 45.5	2.218 $^{+0.001}_{-0.001}$	-19.2 $^{+0.3}_{-0.3}$	-1.6 $^{+0.2}_{-0.2}$	42.81 $^{+0.07}_{-0.07}$	180 $^{+210}_{-90}$	230 $^{+20}_{-20}$	NL-SF
SA22-NB5-14	22 15 22.60	+01 31 06.7	3.114 $^{+0.001}_{-0.001}$	-23.7 $^{+0.2}_{-0.2}$	-2.4 $^{+0.1}_{-0.1}$	44.21 $^{+0.06}_{-0.06}$	230 $^{+40}_{-40}$	4220 $^{+300}_{-290}$	BL-AGN
SA22-NB5-18	22 22 02.72	-00 27 19.7	3.096 $^{+0.001}_{-0.001}$	-23.4 $^{+0.2}_{-0.2}$	-2.5 $^{+0.1}_{-0.1}$	44.03 $^{+0.06}_{-0.06}$	110 $^{+20}_{-20}$	1820 $^{+80}_{-80}$	BL-AGN
SA22-NB5-10	22 09 19.01	-00 06 16.3	3.102 $^{+0.001}_{-0.001}$	-23.3 $^{+0.2}_{-0.2}$	-1.1 $^{+0.1}_{-0.1}$	43.99 $^{+0.07}_{-0.07}$	90 $^{+16}_{-15}$	1790 $^{+100}_{-90}$	BL-AGN
BH-NB5-4(BR3)	14 32 30.55	+33 39 57.3	3.086 $^{+0.001}_{-0.001}$	-22.7 $^{+0.2}_{-0.2}$	-2.4 $^{+0.5}_{-0.5}$	43.78 $^{+0.06}_{-0.06}$	150 $^{+30}_{-30}$	690 $^{+80}_{-70}$	NL-AGN
CALYMHA-383	02 17 32.39	-05 12 50.8	3.222 $^{+0.004}_{-0.004}$	-24.3 $^{+0.2}_{-0.2}$	-1.7 $^{+0.1}_{-0.1}$	43.54 $^{+0.06}_{-0.06}$	26 $^{+4}_{-4}$	4300 $^{+700}_{-700}$	BL-AGN
BH-NB5-6	14 33 21.85	+33 54 20.3	3.125 $^{+0.001}_{-0.001}$	-23.2 $^{+0.2}_{-0.2}$	0.7 $^{+0.5}_{-0.5}$	43.56 $^{+0.06}_{-0.06}$	37 $^{+6}_{-6}$	2700 $^{+400}_{-400}$	BL-AGN
BH-NB5-10	14 33 23.82	+33 38 47.2	3.100 $^{+0.001}_{-0.001}$	-23.2 $^{+0.2}_{-0.2}$	-2.1 $^{+1.5}_{-1.5}$	43.45 $^{+0.06}_{-0.06}$	29 $^{+5}_{-5}$	470 $^{+20}_{-10}$	NL-AGN
BH-NB5-16	14 31 06.50	+34 04 23.8	3.085 $^{+0.002}_{-0.002}$	-21.7 $^{+0.2}_{-0.2}$	-2.3 $^{+0.9}_{-0.9}$	43.19 $^{+0.06}_{-0.06}$	42 $^{+8}_{-8}$	460 $^{+20}_{-20}$	NL-AGN
BH-NB5-27	14 30 32.58	+33 59 22.1	3.131 $^{+0.001}_{-0.001}$	-21.2 $^{+0.2}_{-0.2}$	-0.6 $^{+1.5}_{-1.5}$	43.07 $^{+0.07}_{-0.07}$	57 $^{+14}_{-13}$	340 $^{+10}_{-10}$	NL-AGN
GN-NB5-6712	12 36 07.98	+62 23 14.1	3.148 $^{+0.001}_{-0.001}$	-20.7 $^{+0.2}_{-0.2}$	-2.1 $^{+0.7}_{-0.7}$	43.10 $^{+0.06}_{-0.06}$	380 $^{+300}_{-150}$	400 $^{+110}_{-80}$	NL-SF
BH-NB5-34	14 31 03.87	+33 34 46.2	3.076 $^{+0.001}_{-0.001}$	-20.9 $^{+0.2}_{-0.2}$	-1.2 $^{+1.4}_{-1.5}$	43.02 $^{+0.07}_{-0.07}$	50 $^{+12}_{-11}$	420 $^{+30}_{-20}$ (2)	NL-SF
BH-NB5-37	14 33 24.35	+33 39 38.5	3.147 $^{+0.001}_{-0.001}$	-20.3 $^{+0.3}_{-0.3}$	-2.4 $^{+3.1}_{-3.1}$	43.03 $^{+0.07}_{-0.07}$	80 $^{+30}_{-20}$	360 $^{+10}_{-10}$	NL-SF
GN-NB5-3378	12 37 21.68	+62 13 50.2	3.151 $^{+0.001}_{-0.001}$	-20.4 $^{+0.3}_{-0.3}$	-3.2 $^{+1.1}_{-1.2}$	42.84 $^{+0.07}_{-0.07}$	120 $^{+70}_{-40}$	300 $^{+20}_{-20}$	NL-AGN
GN-NB5-5878	12 36 19.47	+62 15 01.8	3.129 $^{+0.001}_{-0.001}$	-21.3 $^{+0.2}_{-0.2}$	-1.7 $^{+0.5}_{-0.5}$	42.69 $^{+0.08}_{-0.08}$	40 $^{+10}_{-10}$	390 $^{+20}_{-20}$	NL-SF

band ( $m$ ) to rest-frame 1500  $\text{\AA}$ , which approximately corresponds to  $g$  for  $z = 2.2$  and  $R$  for  $z = 3.1$  and compute

$$M_{\text{UV}} = m - [5 \log_{10}(D_L[z]) - 5 - 2.5 \log_{10}(1+z)] + f, \quad (1)$$

where  $D_L$  is the luminosity distance in parsec and  $f$  is a correction factor from the fact that our filters may not trace 1500  $\text{\AA}$  exactly and have different widths. We use the results from Ilbert et al. (2009) and Ly $\alpha$  selected sources from Sobral et al. (2018) to compare  $M_{\text{UV}}$  values computed using  $g$  and  $R$  and  $M_{\text{UV}}$  computed from the full SED fitting. We find averages of  $f \approx 0$  for  $z \approx 2.2$  ( $g$ ) and  $f \approx -0.2$  for  $z \sim 3.1$  ( $R$ ). The dispersion on  $f$  is  $\approx 0.2$ . Errors on  $M_{\text{UV}}$  are computed by perturbing the appropriate  $m$  and  $f$  along their Gaussian distributions independently 10 000 times and calculating the 16th and 84th percentiles (see Table 2).

We estimate the rest-frame UV  $\beta$  slope ( $f_\lambda \propto \lambda^\beta$ ; e.g. Meurer et al. 1995, 1997; Burgarella, Buat & Iglesias-Páramo 2005; Bouwens et al. 2009; Ono et al. 2010b; Dunlop et al. 2012) with

$$\beta = -\frac{m_1 - m_2}{2.5 \log_{10}(\lambda_{m_1}/\lambda_{m_2})} - 2, \quad (2)$$

by using  $R$  ( $m_1$ ) and  $I$  ( $m_2$ ) for  $z \approx 2.2$  and  $I$  ( $m_1$ ) and  $z$  ( $m_2$ ) for  $z \approx 3.1$  (observed  $\lambda_R \approx 6100 \text{\AA}$ ,  $\lambda_I = 7600 \text{\AA}$ ,  $\lambda_z = 9000 \text{\AA}$ , rest-frame  $\approx 1800, 2200 \text{\AA}$  at each of the two redshift groups). We use these filters as they avoid strong contamination from C IV, which would

make strong C IV emitters artificially blue if we obtain  $\beta$  using rest-frame  $\approx 1500, 2200 \text{\AA}$ . Errors on  $\beta$  are computed with the same methodology as for  $M_{\text{UV}}$ . Due to the slight differences in the filters and rest-frame, we add 0.1 mag in quadrature to the final error (see Table 2).

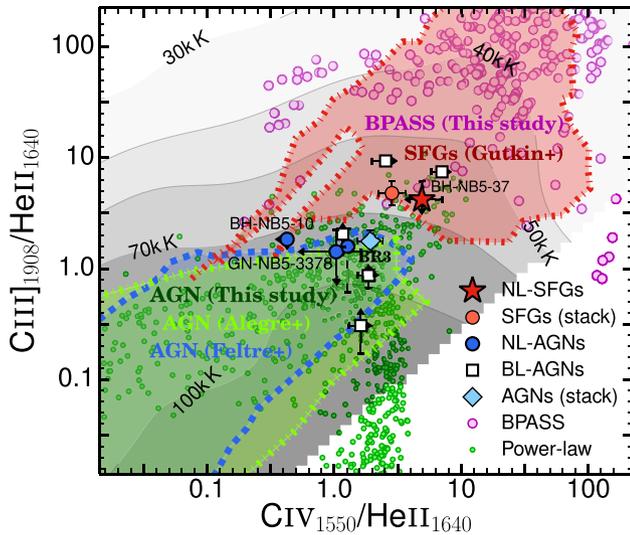
### 3.6 Estimating the Ly $\alpha$ escape fraction and $\xi_{\text{ion}}$

We obtain a rough proxy for Ly $\alpha$   $f_{\text{esc,UV}}$  based on  $M_{\text{UV}}$  (a more detailed investigation will be presented in Sobral & Matthee 2018, exploring H $\alpha$  and Balmer decrements). We obtain  $f_{\text{esc,UV}}$  by (1) converting  $M_{\text{UV}}$  to SFR (in  $M_\odot \text{ yr}^{-1}$ ) following Kennicutt (1998) with a Salpeter IMF:<sup>3</sup>

$$\text{SFR} = (1.4 \times 10^{-28})(4\pi \times 9.521 \times 10^{38})10^{-0.4(M_{\text{UV}}+48.6)} \quad (3)$$

and by correcting for dust extinction; (2) converting Ly $\alpha$  to SFR assuming a 100 per cent escape fraction, case B and Salpeter IMF (see e.g. Sobral et al. 2018), and (3) obtaining the ratio between (2) and (1). For comparison, we also compute  $f_{\text{esc}}$  using the Ly $\alpha$  rest-frame EW as in Sobral et al. (2017b) (see also Matthee et al. 2017c) and find excellent agreement.

<sup>3</sup> Note that the choice of IMF is irrelevant due to step 2.



**Figure 5.** Ratios between C III], C IV and He II for the sources where we can constrain at least one of the lines in each of the ratios. We show the range of emission line ratios typical of power-law AGN (Feltre et al. 2016, Alegre et al., in preparation) and those more typical of star-forming galaxies (SFGs; Gutkin, Charlot & Bruzual 2016), together with an indication of the effective temperatures based on our blackbody models. We also show the stack of all sources that we classify as SFGs, showing that they are within the region expected for SFGs and contrast them with the stack of all sources classified as AGN.

Here we correct for dust extinction (see discussions in Matthee et al. 2016; An et al. 2017) based on our  $\beta$  measurements. We use  $A_{UV} \approx 4.43 + 1.99\beta$  (Meurer, Heckman & Calzetti 1999), resulting in a bootstrapped  $A_{UV} \approx 0.5^{+0.6}_{-0.2}$  (corresponding to the median  $\beta = -2.0 \pm 0.2$ ,<sup>4</sup> or  $E(B - V) \approx 0.05$ ; Meurer et al. 1999). This implies a correction to UV SFRs of our LAEs of  $1.5^{+1.1}_{-0.3}$ , and thus Ly  $\alpha$  escape fractions (for our observed Ly  $\alpha$  luminosities) lower by these factors compared to the no-dust correction assumption.

Finally, we also estimate the ionization efficiency ( $\xi_{ion}$ ) of LAEs by following Matthee et al. (2017c), using Ly  $\alpha$  luminosity and the escape fraction to estimate the LyC luminosity and compare it to the dust corrected UV luminosity.

## 4 RESULTS

### 4.1 AGN among the most luminous LAEs

Some of the most luminous LAEs may be powered by AGN activity. We identify AGN within our luminous LAEs at  $z \sim 2-3$  as sources with (i) broad Ly  $\alpha$  lines (FWHM  $> 1000 \text{ km s}^{-1}$ ; e.g. Fig. 4 and Table 2), (ii) X-ray detections, or (iii) emission line ratios associated with AGN activity, by exploring high ionization lines (Nv, He II, C IV and C III]); see Fig. 5 (see also e.g. Feltre, Charlot & Gutkin 2016; Nakajima et al. 2017).

#### 4.1.1 Broad-line and X-ray AGN

We start by identifying AGN using the broadness of the Ly  $\alpha$  line (see further details in Appendix B). We find that 7 out of the

<sup>4</sup> We exclude AGN and compute the bootstrapped median of  $\beta$  as  $-2.0 \pm 0.2$ . If we use the entire sample, we obtain a bootstrapped median of  $\beta = -2.1 \pm 0.2$

20 with Ly  $\alpha$  FWHM measurements have FWHM clearly above  $1000 \text{ km s}^{-1}$  (see Fig. 4 and also Figs 2 and 3 for 2D information), and thus we class them as broad-line AGN<sup>5</sup> (BL-AGN) (see Table 2). These sources have FWHMs ranging from  $\sim 1300$  to  $\sim 4000 \text{ km s}^{-1}$  (Fig. 4). Our results imply a BL-AGN fraction of  $35 \pm 11$  per cent within our full sample.

Our SA22 sources do not have X-ray coverage, as they are significantly away ( $\sim 1$  deg in three different directions) from the SSA22 *Chandra* follow-up (Lehmer et al. 2009; Saez et al. 2015), but they all show very broad Ly  $\alpha$  emission (FWHM  $\gtrsim 1800 \text{ km s}^{-1}$ ), and thus are cleanly identified as AGN. Our GOODS-N sources have good *Chandra* coverage (Alexander et al. 2003), but none of our three sources is detected in the 2 Ms depth data. Our three sources in COSMOS are also not detected in deep X-ray observations (see Civano et al. 2016), with the same holding for our three UDS sources, although for UDS only shallower *XMM* observations ( $\approx 3$  times shallower than COSMOS) are available (Ueda et al. 2008). Finally, for our Boötes sources, we find one X-ray detection (BH-NB5-6, a broad-line AGN), while the other sources are not detected in the X-rays (for a flux limit of  $4 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$ , about 5 times shallower than COSMOS X-ray data; see Murray et al. 2005). Therefore, using the available X-ray data, we find a single X-ray source (BH-NB5-6) out of 18 which have X-ray coverage, but we note that the available data are very heterogeneous, do not cover our most luminous LAEs in SA22, and are generally not deep enough to detect lower accretion and lower black mass AGN. Furthermore, the short duty cycles of X-rays can also lead to underestimating the AGN fraction using X-rays only (Shankar, Weinberg & Miralda-Escudé 2009).

#### 4.1.2 Narrow-line AGN

Our spectroscopy allows us to explore the nature of the sources further and search for AGN activity by using other emission lines besides Ly  $\alpha$ . We start by identifying sources with significant Nv (ionization energy of 77.4 eV) which, with observed ratios of  $Nv/Ly \alpha \gtrsim 0.1$  we robustly class as AGN. This is because even the most extreme stellar populations fail to produce significant Nv emission ( $Nv/Ly \alpha < 0.001$ ) due to the sharp cut-off of ionizing photons at energies significantly below 77.4 eV (e.g. Feltre et al. 2016). This makes Nv one of the cleanest lines to identify AGN activity. We find that sources SA22-NB5-18, BH-NB5-6, -10, -16, -27, and CALYMHA-415 all show significant Nv emission (observed ratios  $Nv/Ly \alpha \approx 0.1 - 0.6$ ). Half of these sources are BL-AGN (see Section 4.1.1), while BH-NB5-6 is also an X-ray source. We therefore find an AGN fraction of  $32 \pm 11$  per cent based solely on high  $Nv/Ly \alpha$  ratios ( $Nv/Ly \alpha \gtrsim 0.1$ ).

For the sources without Nv detections and with narrow Ly  $\alpha$  lines ( $< 1000 \text{ km s}^{-1}$ ), we explore other UV rest-frame lines of lower ionization energy. In Fig. 5 we show emission line ratios of high ionization UV lines which are in principle capable of further distinguishing between sources which are dominated by either AGN or star formation. We show four of our narrow-line emitters where we can measure/constrain both line ratios plotted. For comparison and interpretation, we also show the locations of power-law ionizing

<sup>5</sup> Stellar winds may typically lead to de-projected gas velocities of  $\sim 100$  to  $\sim 1000 \text{ km s}^{-1}$ , although in principle they can reach up to  $\approx 3000 \text{ km s}^{-1}$  (Heckman 2003). However, outflows are expected to lead to highly asymmetric lines, while the Ly  $\alpha$  profiles of the broad-line LAEs are all very symmetric.

sources (AGN) from Feltre et al. (2016), along with the location for SFGs (Gutkin et al. 2016). We also provide emission line ratios computed with CLOUDY to sample both populations (see Section 5.2). We show our luminous broad-line AGNs for which we can constrain the line ratios for comparison. Apart from BH-NB5-10, which was already classed as an NL-AGN due to the high  $N\text{v}/\text{Ly}\alpha$  ratio of  $\approx 0.17 \pm 0.03$ , we identify two further narrow-line sources which are consistent with being NL-AGN: BR3, the brightest narrow-line LAE in our sample and GN-NB5-3378, with a much fainter Ly $\alpha$  luminosity. Both BH-NB5-10 and BR3 have emission-line FWHMs of  $\approx 600-700 \text{ km s}^{-1}$ , higher than the other NL-emitters ( $\approx 100-300 \text{ km s}^{-1}$ ), also being brighter in UV and Ly $\alpha$ . We find one source, BH-NB5-37, which is consistent with being dominated by star formation, but suggests a relatively high effective temperature (see Fig. 5) and could still have an AGN source. For the other narrow-line sources without  $N\text{v}$  where we do not detect high ionization UV lines, we classify them as SF as our observations should have detected high ionization lines if they had a significant AGN contribution. Our stack of all these sources (see Fig. 5) suggests that this is a reasonable assumption, as it places the overall sample of sources classified as SF in the SFG region in Fig. 5. We note, nonetheless, that even the stack of SFGs suggests high effective temperatures, and it is possible that even those LAEs have at least some AGN component.

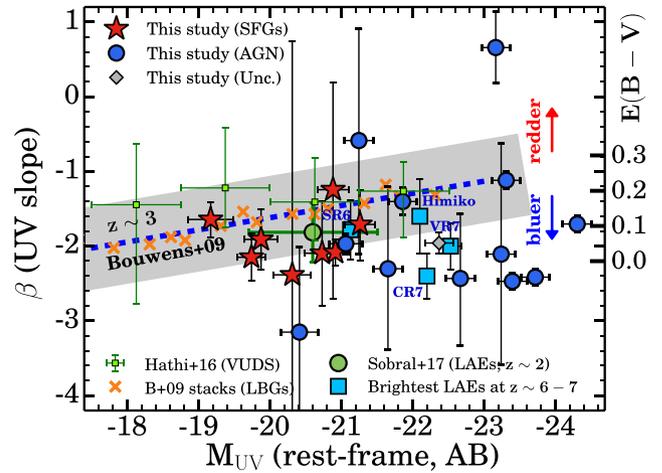
#### 4.1.3 Final AGN and SF classifications

Overall, we find seven broad-line AGN and five narrow-line AGN, with a total of 12 AGNs out of the sample of 20 classifiable sources. Our results therefore imply a total AGN fraction of  $60 \pm 9$  per cent for luminous LAEs at  $z \sim 2-3$ . Relatively high AGN fractions ( $\sim 20-50$  per cent) have also been found for bright LAEs at  $z \sim 0.3-1$  (e.g. Wold et al. 2014, 2017) and at  $z \sim 2-3$  (e.g. Ouchi et al. 2008; Konno et al. 2016; Sobral et al. 2017b; Matthee et al. 2017b), while for the much fainter, numerous LAEs the AGN fraction is typically no more than a few per cent (e.g. Wang et al. 2004; Nilsson et al. 2007; Ouchi et al. 2008; Oteo et al. 2015; Calhau et al. 2018; Sobral et al. 2018).

#### 4.2 The UV properties of luminous LAEs

On average, we find a bootstrapped median for the sample of SFGs of  $M_{\text{UV}} = -20.5^{+0.4}_{-0.3}$  and  $M_{\text{UV}} = -21.5 \pm 0.4$  for the full sample, with  $M_{\text{UV}}$  ranging from  $-19.2$  to  $-23.7$  (see Table 2 and Fig. 6). We note that while luminous LAEs have Ly $\alpha$  luminosities of  $L_{\text{Ly}\alpha}^*$  or higher, they also have UV luminosities of roughly  $M_{\text{UV}}^*$  ( $M_{\text{UV}}^* \approx -20.5$  at  $z \sim 2-3$ , e.g. Arnouts et al. 2005; Reddy & Steidel 2009; Parsa et al. 2016) and have a range of  $M_{\text{UV}}$  similar to the full sample of Sobral et al. (2017b) and the sample of  $\gtrsim L_{\text{Ly}\alpha}^*$  sources presented in Sobral et al. (2018), thus sharing properties with larger, photometric samples of LAEs.

We find a bootstrapped median<sup>6</sup> of  $\beta = -2.0^{+0.3}_{-0.1}$  (the bluest source having  $\beta \approx -3.2$  and the reddest having  $\beta \approx 0.7$ ; see Fig. 6), revealing that our luminous LAEs are very blue on average, but with one outlier. The median  $\beta$ , which we find for luminous LAEs, is in agreement with the full sample of LAEs at  $z \sim 2.2$  from Sobral et al. (2017b), and consistent with other studies focusing on lower luminosity Ly $\alpha$  selected sources (e.g. Venemans et al. 2005; Ouchi et al. 2008; Matthee et al. 2016). Interpreting  $\beta$  as a dust extinction



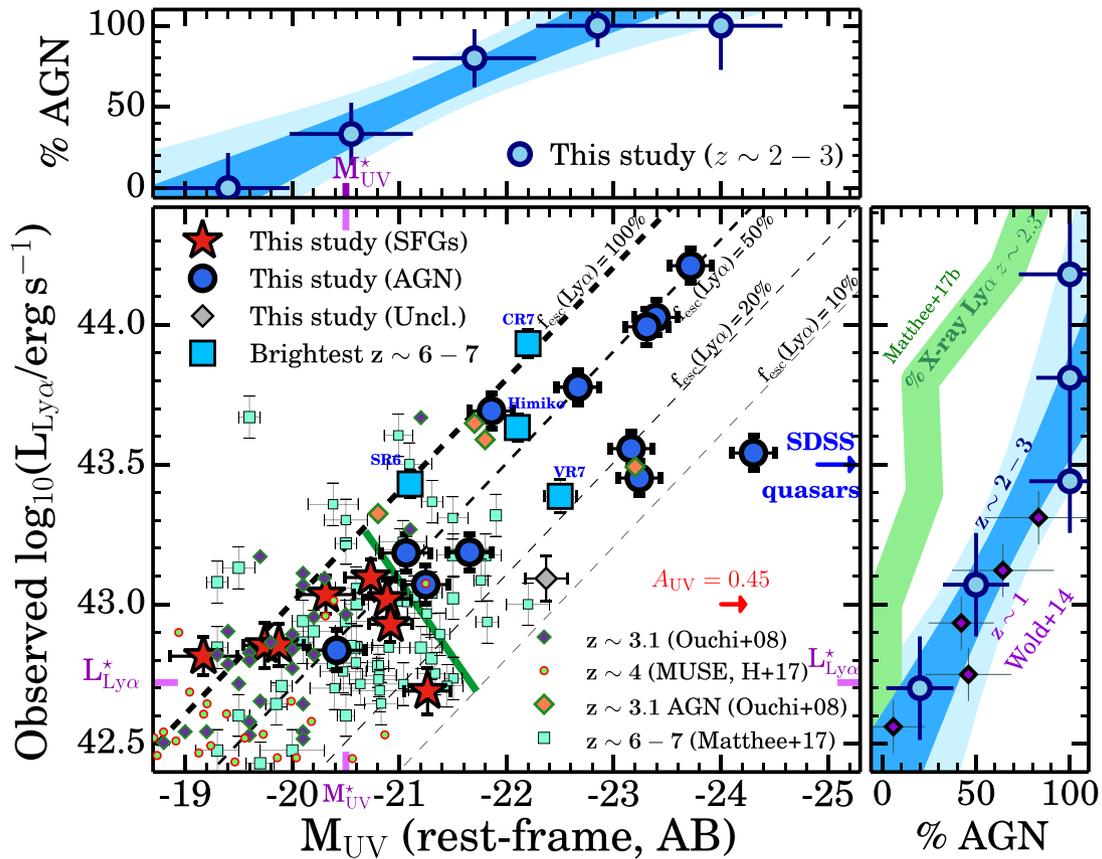
**Figure 6.** The relation between rest-frame UV luminosities ( $M_{\text{UV}}$ ) and the UV  $\beta$  slope for our luminous LAEs at  $z \sim 2-3$  and comparison with other studies: LBGs at  $z \sim 2-4$  (Bouwens et al. 2009), LAEs at  $z \sim 2.2$  (Sobral et al. 2017b), LAEs within the UV continuum-selected VUDS sample at  $z \sim 2-2.5$  (Hathi et al. 2016), and also the brightest  $z = 6-7$  LAEs (Matthee et al. 2017c). We find that our sources are in general bluer than the population of Lyman-break selected sources at a similar redshift, with the deviation becoming stronger at the highest UV luminosities in the regime dominated by AGN, suggesting younger ages, low dust extinction, and relatively unobscured AGN activity for the AGN. We find no significant relation between  $\beta$  and  $M_{\text{UV}}$  for luminous LAEs with the slope of the potential relation being consistent with 0 ( $-0.07^{+0.11}_{-0.15}$ ).

indicator for the non-AGN sources suggests  $E(B - V) = 0.05^{+0.08}_{-0.05}$  (Meurer et al. 1999) or  $A_{\text{UV}} \approx 0.5^{+0.6}_{-0.2}$  (see Section 3.6), implying that the bulk of our luminous LAEs have little to no dust, and thus are potentially similar to fainter LAEs (e.g. Nilsson et al. 2007; Ono et al. 2010a; Nakajima et al. 2012). This is also consistent with Balmer decrement measurements/constraints for some of our sources (Matthee et al., in preparation). We list the individual  $M_{\text{UV}}$  and  $\beta$  values for each source in Table 2.

In Fig. 6 we show the relation between  $M_{\text{UV}}$  and  $\beta$  for luminous LAEs, and also compare our sources with the literature (Bouwens et al. 2009; Matthee et al. 2015, 2017c; Sobral et al. 2015, 2017b; Hathi et al. 2016). We find little to no correlation between  $M_{\text{UV}}$  and  $\beta$  for our luminous LAEs as a whole, with our sources being generally very blue, irrespective of UV luminosity (Fig. 6). We further test this by fitting a linear relation to 10 000 realizations, each varying each individual pair of  $(M_{\text{UV}}, \beta)$  randomly and independently within their full probability distributions. For the full sample (SFGs only), we find a slope of  $-0.07^{+0.11}_{-0.15}$  ( $-0.1^{+0.4}_{-0.2}$ ) for the relation between  $M_{\text{UV}}$  and  $\beta$ , fully consistent with no relation for either the full sample or just the sample of star-forming galaxies. We also find no significant relation when studying the AGNs only.

Our results are in very good agreement with Hathi et al. (2016), who found a slope of  $0.00 \pm 0.04$  for their global sample, consistent with our results of  $-0.07^{+0.11}_{-0.15}$  for a much more luminous sample. Our luminous LAEs are significantly brighter in Ly $\alpha$  than essentially all LAEs in Hathi et al. (2016) which come from a sample that is UV continuum-selected. We also note that we find luminous LAEs to be even bluer than sources studied in Hathi et al. (2016). While our results for SFGs are in reasonable agreement with those of Bouwens et al. (2009) for LBGs within the error bars (see Fig. 6), LAEs tend to be below the relation found for LBGs. We also find that luminous LAEs deviate most from the Bouwens et al. (2009)

<sup>6</sup> The simple average for the full sample is  $\beta = -1.9 \pm 0.7$ .



**Figure 7.** The relation between rest-frame UV and Ly $\alpha$  luminosities. We find that Ly $\alpha$  roughly scales with UV for our sample at  $z \sim 2-3$ , consistent with relatively high Ly $\alpha$  escape fractions, which we show (as dashed lines), from 10 per cent to 100 per cent assuming  $M_{UV}$  traces star-formation and by correcting for dust with  $\beta = -2.0$  (see the arrow that shows how UV luminosity is increased by correcting for dust extinction and how the lines of constant Ly $\alpha$  escape fraction have shifted in the opposite direction). Our sources span a relatively unexplored region in the parameter space at  $z \sim 2-3$ , in between the brightest LAEs found with MUSE (Drake et al. 2017a; Hashimoto et al. 2017) or the brightest typical LAEs (e.g. Ouchi et al. 2008; Trainor et al. 2015) and almost up to UV luminosities of the faintest quasars found at  $z \sim 2-3$  with e.g. SDSS (e.g. Richards et al. 2006). Our AGNs present the highest Ly $\alpha$  and UV luminosities within our sample; a simple dividing line ( $[-20.7, 43.25], [-21.7, 42.7]$ ) is able to isolate all SFGs and all but one AGN from each other at  $z \sim 2-3$ . The AGN fraction of luminous LAEs increases with both luminosities as  $f_{AGN} = (-0.30 \pm 0.07)(M_{UV} + 20.5) + (0.35 \pm 0.11)$  and  $f_{AGN} = (0.78 \pm 0.22)(\log_{10}(L_{Ly\alpha}) - 42.72) + (0.24 \pm 0.14)$ ; we show the corresponding 1 $\sigma$  and 2 $\sigma$  contours. Qualitatively similar results have been found in recent studies (see Ouchi et al. 2008; Wold et al. 2014; Matthee et al. 2017b; Calhau et al. 2018). We also compare our results with a literature compilation obtained by Matthee et al. (2017c) at higher redshift.

relation at the brightest UV luminosities, but this may be a simple consequence of the powering nature for these sources (e.g. AGN).

#### 4.3 The relation between Ly $\alpha$ and UV at $z \sim 2-3$

In Fig. 7 we compare the Ly $\alpha$  luminosities of our sources with their rest-frame UV luminosities. Our sources span the relatively unexplored range between the more numerous LAEs, typically an order of magnitude fainter in Ly $\alpha$  and UV luminosities (e.g. Ouchi et al. 2008; Trainor et al. 2015; Drake et al. 2017a; Hashimoto et al. 2017) and luminous quasars at  $z \sim 2-3$  (Richards et al. 2006), which can have similar Ly $\alpha$  luminosities to our luminous LAEs (e.g. Borisova et al. 2016) but UV luminosities  $\sim -25$  to  $\sim -30$ , and thus significantly brighter in the continuum.

The brightest LAEs in terms of  $M_{UV}$  within our sample are also the brightest in Ly $\alpha$ , but the relation shows significant scatter, leading to a 2 dex spread in  $M_{UV}$  for an  $\approx 1.7$  dex spread in Ly $\alpha$  luminosity. Such scatter could be due to variations in dust content, the powering source of the ionizing photons and/or due to different ionization efficiencies (see discussions in e.g. Matthee et al. 2017a;

Sobral et al. 2018). One also expects that sources may have different Ly $\alpha$  escape fractions, even though it is known that LAEs (Ly $\alpha$ -selected) tend to have typically high Ly $\alpha$  escape fractions (Nilsson et al. 2009; Nilsson & Møller 2009; Wardlow et al. 2014; Trainor et al. 2015; Sobral et al. 2017b). We show indicative lines of constant Ly $\alpha$  escape fraction (correcting for dust extinction by using the UV  $\beta$  slope; see Section 3.6), which, under very simple assumptions, can easily explain the observed scatter. Many of our Ly $\alpha$  sources are consistent with having very high Ly $\alpha$  escape fractions which can be up to  $\approx 100$  per cent. The high escape fractions are likely an important factor in explaining why these sources are so Ly $\alpha$ -bright. Alternatively, they may (also) have other powering sources hidden in the UV and/or a higher ionization efficiency,  $\xi_{ion}$  (e.g. Matthee et al. 2017a). By estimating  $\xi_{ion}$  for the bright LAEs, we find a bootstrapped<sup>7</sup>  $\xi_{ion} = 10^{25.4 \pm 0.2} \text{ Hz erg}^{-1}$ , with individual ionization efficiencies for the star-forming LAEs varying from  $\xi_{ion} \approx 10^{25.0} \text{ Hz erg}^{-1}$  to  $\xi_{ion} \approx 10^{25.9} \text{ Hz erg}^{-1}$ . These compare with

<sup>7</sup> By correcting for dust extinction, we obtain  $\xi_{ion} = 10^{25.1 \pm 0.2} \text{ Hz erg}^{-1}$ .

values of  $\xi_{\text{ion}} = 10^{25.0-25.3} \text{ Hz erg}^{-1}$  for continuum selected LBGs at  $z \sim 2-5$  (Bouwens et al. 2016; Nakajima et al. 2017; Shivaei et al. 2018),  $\xi_{\text{ion}} = 10^{24.8} \text{ Hz erg}^{-1}$  for highly star-forming H $\alpha$  emitters at  $z \approx 2$  (Matthee et al. 2017a) and  $\xi_{\text{ion}} \approx 10^{25.5} \text{ Hz erg}^{-1}$  for low luminosity LAEs at  $z \sim 3-6$  (Harikane et al. 2017; Matthee et al. 2017a; Nakajima et al. 2018).

Correcting for dust (see Section 3.6) assuming  $A_{\text{UV}} = 0.45$  ( $A_{\text{UV}} = 0.0$ ), we find an average Ly $\alpha$  escape fraction of  $50_{-15}^{+20}$  per cent ( $70_{-20}^{+30}$  per cent) for our luminous SFLAEs at  $z \sim 2-3$ . As an independent estimate, we also use the relation between Ly $\alpha$  escape fraction and the Ly $\alpha$  rest-frame EW presented in Sobral et al. (2017b), which is independent of any assumptions regarding dust extinction, to find a bootstrapped Ly $\alpha$  escape fraction that is in perfect agreement ( $50_{-15}^{+20}$  per cent).<sup>8</sup> These high escape fractions are comparable with those found for fainter LAEs. Some of the latest measurements include a Ly $\alpha$  escape fraction of  $37 \pm 7$  per cent for a more general sample of fainter LAEs at  $z \sim 2$  (Sobral et al. 2017b) and  $\approx 30$  per cent at  $z \sim 2-4$  (e.g. Wardlow et al. 2014; Kusakabe et al. 2015). These results are much higher than the low escape fractions of  $\approx 2$  per cent<sup>9</sup> for a more general population of star-forming galaxies selected through H $\alpha$  at  $z \sim 2$  (Matthee et al. 2016),  $\approx 4-5$  per cent (Hayes et al. 2011; Sobral et al. 2018) from a comparison of the evolution of the Ly $\alpha$  and UV LFs at  $z \sim 2-3$ , or  $\approx 5$  per cent for UV selected sources at  $z \sim 2-3$  (e.g. VUDS; Cassata et al. 2015); see also An et al. (2017) and Sobral et al. (2017b). Our results thus show that luminous LAEs have Ly $\alpha$  escape fractions at least comparable to or higher than the numerous low luminosity LAEs, hinting that at least the SF-dominated ones are likely ‘scaled-up’ or ‘maximal’ versions of the faint LAEs.

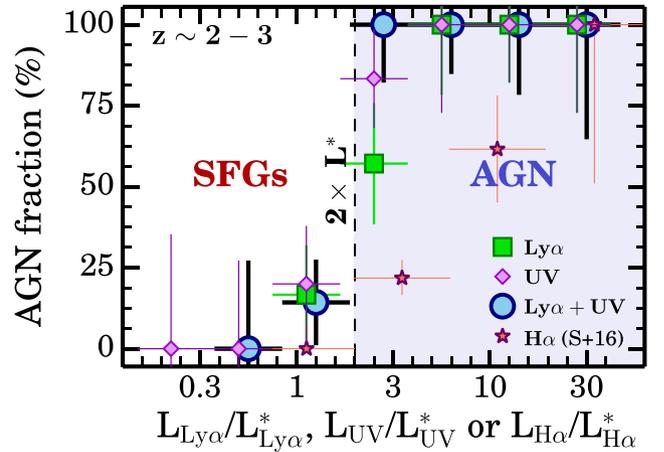
#### 4.4 Ly $\alpha$ and UV luminosities as AGN activity predictors for LAEs at $z \sim 2-3$

Overall, our sample of luminous LAEs at  $z \sim 2-3$  reveals a high AGN fraction of  $\approx 60$  per cent. We now investigate how this fraction depends on both Ly $\alpha$  and UV luminosities. As Fig. 7 shows, LAEs are essentially all AGNs above an Ly $\alpha$  luminosity of  $\approx 10^{43.3} \text{ erg s}^{-1}$  or for UV luminosities brighter than  $M_{\text{UV}} \approx -21.5$ , qualitatively similar to results at different redshifts (e.g. Ouchi et al. 2008; Wold et al. 2014, 2017; Konno et al. 2016; Matthee et al. 2017b; Sobral et al. 2017b). These include both the narrow-line emitters, which are brighter in Ly $\alpha$  than in the UV, and BL-AGN, which are typically brighter in the UV. In Fig. 7 we also show the rough location of the knee of the Ly $\alpha$  (Sobral et al. 2018) and UV LFs (e.g. Arnouts et al. 2005; Reddy & Steidel 2009; Parsa et al. 2016). We find a rise of the AGN fraction from virtually 0 to 100 per cent happening just under 1 dex in UV and Ly $\alpha$  luminosities, with the fraction below the knee of each LF being  $\approx 0$  per cent. We quantify the relation between the AGN fraction and UV or Ly $\alpha$  luminosity in a similar way to other relations in this paper. We find that the AGN fraction among bright LAEs rises with  $M_{\text{UV}}$  significantly ( $\sim 4\sigma$ ) and is well parametrized (from  $-19$  to  $-23$ ) by  $f_{\text{AGN}} = (-0.30 \pm 0.07)(M_{\text{UV}} + 20.5) + (0.35 \pm 0.11)$ . The relation between  $f_{\text{AGN}}$  and Ly $\alpha$  luminosity (significant at an  $\sim 3\sigma - 4\sigma$ ) is well parametrized (from  $\log_{10}(L_{\text{Ly}\alpha})$  42.5–44.0) by  $f_{\text{AGN}} = (0.78 \pm 0.22)(\log_{10}(L_{\text{Ly}\alpha}) - 42.72) + (0.24 \pm 0.14)$ .

We also compare our results with Wold et al. (2014, 2017) at  $z \sim 1$  and Matthee et al. (2017b) at  $z \sim 2.3$  (Fig. 7), which show

<sup>8</sup> We find a bootstrapped  $\text{EW}_0 = 90_{-23}^{+44} \text{ \AA}$ .

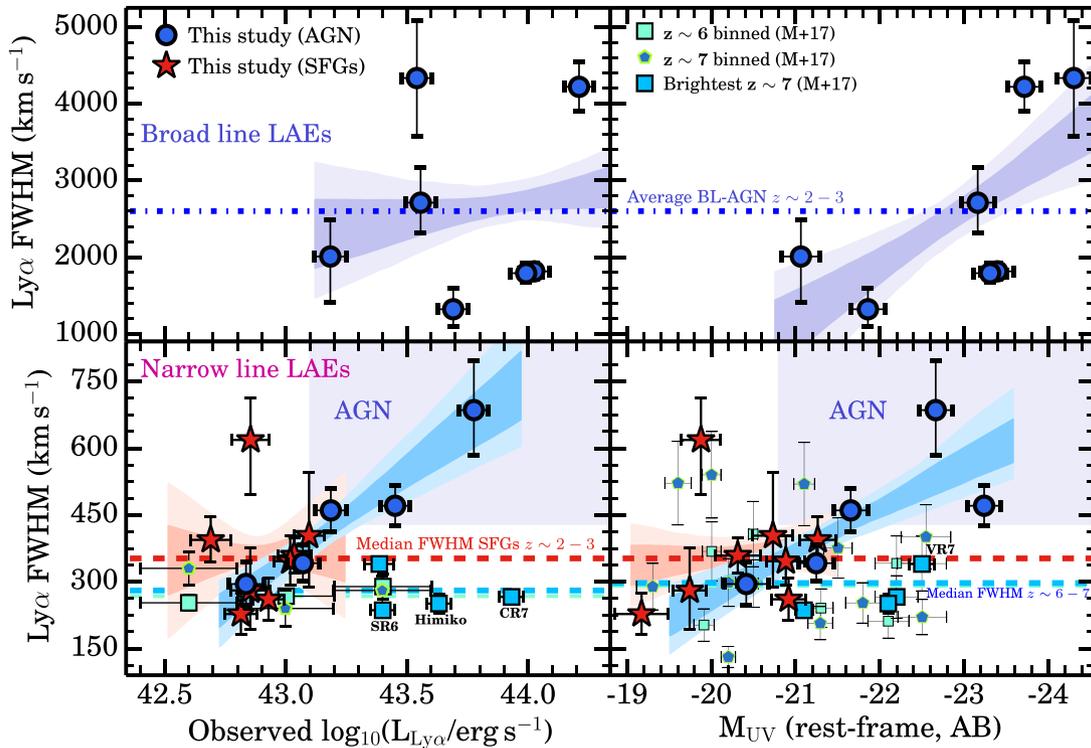
<sup>9</sup> Directly measured with H $\alpha$ .



**Figure 8.** The AGN fraction of luminous LAEs at  $z \sim 2-3$  as a function of the ratio between luminosity and the typical luminosity ( $L^*$ ) for Ly $\alpha$ , UV, or the average of the two, and comparison to Sobral et al. (2016) for H $\alpha$  emitters with H $\alpha$  luminosity. We find that the transition from star-forming to AGN-dominated happens above twice the typical luminosity, with increasing Ly $\alpha$  luminosity resulting in a smoother growth than UV. When combining Ly $\alpha$  and UV, we find a much sharper transition, well modelled by a step function. Our results provide a simple way to distinguish between AGN and SFGs and suggest an important physical transition at  $\approx 2L^*$ , likely linked with the maximal unobscured luminosity by a starburst at  $z \sim 2-3$  explaining the difference between LAEs and H $\alpha$  emitters.

very similar trends (see also Ouchi et al. 2008; Konno et al. 2016). By comparing our results with  $z \sim 1$  (see Fig. 7), we find little to no evolution in the relation between the AGN fraction and the Ly $\alpha$  luminosity. This is apparently at odds with results from X-rays and radio (Konno et al. 2016; Matthee et al. 2017b; Calhau et al. 2018), which would hint for a strong evolution between  $z \sim 1$  and  $z \sim 2-3$ . However, X-ray and radio selections are only able to recover a fraction of LAEs with AGN activity. Rest-frame UV spectra are crucial to cleanly identify sources with AGN signatures, as many do not show X-ray or radio emission down to even the deepest limits currently achieved over large fields. Our results further stress that X-ray identification of AGN (e.g. Konno et al. 2016; Sobral et al. 2017b; Matthee et al. 2017b) must be taken as a lower limit to the AGN fraction, due to both short duty cycles and the typically high accretion rates needed for detection in the X-rays at high redshift.

In order to further explore the relation between the AGN fraction and UV or Ly $\alpha$  luminosities, we present those together in Fig. 8. We convert UV absolute magnitudes to luminosity and normalize by the corresponding  $L_{\text{UV}}^*$  ( $M_{\text{UV}}^* = -20.5$ ; Parsa et al. 2016). We also normalize the Ly $\alpha$  luminosities by  $L_{\text{Ly}\alpha}^* = 10^{42.72} \text{ erg s}^{-1}$  (Sobral et al. 2018). The rise in the AGN fraction is found to be smoother with increasing Ly $\alpha$  luminosity than with UV luminosity (Fig. 8), but the transition happens at the corresponding  $2 \times L^*$  in both cases and the AGN fraction starts rising above  $L^*$ . Using the average  $L/L^*$  from combining UV and Ly $\alpha$  leads to the AGN fraction rising even sharply with increasing relative luminosity, from  $11 \pm 10$  per cent below  $2 \times L^*$  to  $100_{-7}^{+0}$  per cent above  $2 \times L^*$ , suggesting a difference above  $5\sigma$ . By perturbing each bin and re-fitting with (i) a linear relation and (ii) a step function with a transition at  $2 \times L^*$ , we find that step functions result in reduced  $\chi^2$  typically 2–3 times smaller than linear fits, with 92 per cent of all fits preferring a step function. We note that combining the information on Ly $\alpha$  FWHMs with either Ly $\alpha$  or UV luminosities also leads to a sharper transition



**Figure 9.**  $\text{Ly}\alpha$  FWHM versus  $\text{Ly}\alpha$  and UV luminosities for  $z=2-3$ . On average, we find that  $\text{Ly}\alpha$  FWHMs tend to increase with both  $\text{Ly}\alpha$  and UV luminosities. However, we find that the trend is driven by AGN and that NL-AGN have the most statistically significant relations between luminosity and  $\text{Ly}\alpha$  FWHM. We show the 2.3, 16, 84, and 97.7 percentiles from a Monte Carlo fitting routine. For BL-AGN, we find a weak to no correlation between FWHM and  $\text{Ly}\alpha$  luminosity but a significant relation between FWHM and  $M_{\text{UV}}$ . We find no significant relation between FWHM and any of the two luminosities for SFGs, although we note that a significant fraction of SFGs would be well fitted by an extrapolation of the relations found for NL-AGN. For comparison with sources selected with the same methods at higher redshift, we include a compilation presented in Matthee et al. (2017c) and also highlight the most luminous sources found at  $z\sim 6-7$ , comparable in luminosities to some of our sources. Our results show a significant difference in the FWHMs of equally selected bright LAEs at fixed  $M_{\text{UV}}$  and  $L_{\text{Ly}\alpha}$  between  $z\sim 2-3$  and  $z\sim 6-7$ .

at  $2 \times L^*$  than when using a single luminosity to predict the AGN fraction.

Our results show that there is a relatively sharp transition in the nature of luminous LAEs at  $2 \times L^*$ . Such transition suggests that either above  $L^*$  AGN activity becomes more and more prominent (see Fig. 8 and e.g. Sobral et al. 2016), and/or there is a relatively well-defined physical limit in the output/observed luminosities from star-forming galaxies at  $z\sim 2-3$  which does not apply to AGN. This limit would likely be linked with (i) the intrinsic maximal starburst before radiation pressure halts further activity (see Crocker et al. 2018) and perhaps even more importantly with (ii) the maximal ‘observable’ or unobscured starburst which requires the highest possible output of LyC and UV photons without significant dust obscuration, so it can be seen as bright as possible in  $\text{Ly}\alpha$  and UV. Thus, while there are star-forming galaxies that intrinsically produce much higher UV and  $\text{Ly}\alpha$  luminosities than the observed  $2 \times L^*$  in the UV (Casey, Narayanan & Cooray 2014a), a significant increase in SFR is linked with a similar or even higher increase of dust obscuration (e.g. Hopkins et al. 2001; Garn & Best 2010; Sobral et al. 2012), leading to an observed UV and  $\text{Ly}\alpha$  luminosity that can even decrease, thus limiting the maximum observed luminosities (see e.g. Nilsson et al. 2011). If this is the case, the only sources that can still be observable as LAEs brighter than  $(1-2) \times L^*$  require AGN activity. In these cases, due to the physics of the accretion disc or powerful outflows clearing out material, the AGN activity can lead to even larger observed UV and  $\text{Ly}\alpha$  luminosities. Results

from  $\text{H}\alpha$  (non-resonant and less affected by dust) selected sources (Sobral et al. 2016) show an increase of the AGN fraction with increasing  $L_{\text{H}\alpha}/L_{\text{H}\alpha}^*$  above  $2 \times L_{\text{H}\alpha}^*$  at  $z\sim 2$  (Fig. 8), but such an increase is significantly slower and the AGN fraction only reaches 100 per cent by  $20 \times L_{\text{H}\alpha}^*$ , a factor of 10 larger than seen for LAEs. We discuss these results further and implications for a physical limit for the UV and  $\text{Ly}\alpha$  starburst observed luminosities in Section 6.1.

#### 4.5 The relation between $\text{Ly}\alpha$ FWHMs, UV, and $\text{Ly}\alpha$ luminosities at $z\sim 2-3$

In Fig. 9 we investigate how the  $\text{Ly}\alpha$  FWHM depends on  $\text{Ly}\alpha$  and rest-frame UV luminosities for bright LAEs at  $z\sim 2-3$ . We start by evaluating a potential simple linear relation between FWHM and each luminosity. We find a slope that is significantly ( $>5\sigma$ ) away from zero (no-relation), with no realization ( $p < 10^{-4}$ ) consistent with a zero slope. However, the strong trend could be a consequence of a sharp change in the population from narrow to broad-line emitters from low to high luminosities. In order to further investigate the trends, we split the sample in (i) SFGs, (ii) NL-AGN, and (iii) BL-AGN and repeat the measurements/analysis. Our results show that there is no significant correlation between  $\text{Ly}\alpha$  FWHM and either  $\text{Ly}\alpha$  or UV luminosity for SFGs alone (see Fig. 9), with the best-fitting slope being fully consistent with zero within the 16 and 84 percentiles. For NL-AGN we find a strong correlation between FWHM and both  $\text{Ly}\alpha$  and UV luminosities (see Fig. 9),

with the median slopes of all fits being  $\approx 4\sigma-5\sigma$  away from zero and with 100 per cent of realizations resulting in a positive slope. This relation may be related with outflows becoming more prominent and ejecting material at higher velocities for higher luminosity NL-AGN. Alternatively, high velocity outflows may lead to higher Ly $\alpha$  escape fractions, thus explaining the tighter correlation between Ly $\alpha$  luminosity and FWHM than UV luminosity with FWHM. It is worth noting, nonetheless, that there is a significant fraction of SFG LAEs that are well fitted by the range of fits performed to the NL-AGN, with many of these SFGs lying well within the  $1\sigma$  range of best fits to the NL-AGN. This could imply that the physics behind the correlation is similar in both NL-SFGs and NL-AGN, and could be driven by low-velocity SF-driven outflows which may be roughly proportional to the LyC production (SFRs).

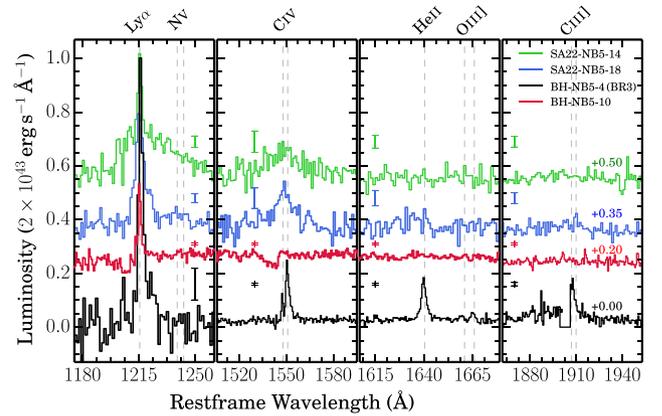
Our results for BL-AGN reveal no significant correlation between Ly $\alpha$  FWHM and Ly $\alpha$  luminosity, with  $>30$  per cent of realizations being consistent with a zero or even an anti-correlation, while we find a relatively significant correlation between FWHM and  $M_{UV}$  for BL-AGN, with the slope being  $\sim 3\sigma$  away from zero. For BL-AGN, our results may be explained by the physics of the accretion disc and the mass of the super-massive black holes: more massive black holes will produce higher FWHMs and higher UV luminosities, but the observed Ly $\alpha$  output both may be variable and will depend on orientation angle and the likely complex radiation transfer. This could explain why the FWHM correlates well with UV luminosity, but little with observed Ly $\alpha$ .

We also compare our results at  $z \sim 2-3$  with the literature at higher redshift for LAEs with comparable Ly $\alpha$  and UV luminosities (Fig. 9) and selected in the same way as our sources (Matthee et al. 2017c). We find that at  $z \approx 5.7-6.6$  there is little to no relation between FWHM and UV or Ly $\alpha$  luminosity. Furthermore, we find an offset between the median FWHM at  $z \sim 2-3$  and  $z \sim 6-7$ . Such an offset is particularly striking at the highest Ly $\alpha$  luminosities ( $>10^{43.3}$  erg s $^{-1}$ ) and the highest UV luminosities ( $<-21.5$ ), as the FWHMs at  $z \sim 6-7$  are relatively narrow ( $\sim 200-300$  km s $^{-1}$ ; Matthee et al. 2017c), while the median at  $z \sim 2-3$  is above 1000 km s $^{-1}$ . This suggests that the powering sources (and the ISM/IGM conditions) needed to result in exceptionally high Ly $\alpha$  and UV luminosities are different at  $z \sim 6-7$  and  $z \sim 2-3$ , and that high Ly $\alpha$  and UV luminosities can be achieved at high redshift without potentially very energetic/high-velocity outflows (see Section 6.2). These may be connected with, e.g. higher  $\xi_{ion}$ , higher Ly $\alpha$  escape fractions (see discussions in Sobral et al. 2018) and potentially lower metallicity environments which ultimately may result in higher luminosities for unobscured starbursts at high redshift (see Section 6.2 for further discussion).

## 5 THE NATURE OF THE BRIGHTEST LAES WITH REST-FRAME UV LINES

### 5.1 Velocity offsets of luminous LAEs at $z \sim 2-3$

We use non-resonant lines (when detected) to derive the systemic redshift of each source, and then use them to calculate the shift of the peak of Ly $\alpha$  from each line (see Appendix B). We find such shifts to be generally positive, i.e. for Ly $\alpha$  to generally be redshifted compared to the other emission lines (e.g. Trainor et al. 2016), but we find some notable exceptions, particularly for sources where C IV and/or He II are redshifted more than Ly $\alpha$  by  $\sim 100$  km s $^{-1}$  (e.g. GN-NB5-3378). In general, we find C IV to be redshifted by  $\approx 100$  km s $^{-1}$  compared to other lines which should be tracing the



**Figure 10.** Comparison between some of our brightest LAEs at  $z = 3.1$  (all classified as AGNs), showing a diversity of FWHMs, continuum and also other rest-frame UV lines. We show the  $1\sigma$  error on the 1D for each rest-frame wavelength window so it is easier to distinguish between noise and real emission and absorption features. Note that spectra have been shifted up to avoid overlap. The brightest broad LAE (SA22-NB5-14) shows significant transmission blue-ward of Ly $\alpha$ , while the second brightest (SA22-NB5-18) shows much less transmission, and the narrow line emitters show little to no transmission just blue-ward of Ly $\alpha$ .

systemic redshift better. This is not surprising since C IV is also a resonant line, and thus has a behaviour similar to Ly $\alpha$ .

For the narrow LAEs at  $z \sim 2.2$ , we find low-velocity offsets between Ly $\alpha$  and other rest-frame UV lines from  $\approx 30$  km s $^{-1}$  to  $\approx 250$  km s $^{-1}$  with an average of  $140^{+80}_{-30}$  km s $^{-1}$ ; these are sources with the most symmetric and narrowest Ly $\alpha$  FWHM ( $\approx 200-300$  km s $^{-1}$ ) and also include sources with significant blue wings (see also e.g. Verhamme et al. 2017). For the sample at  $z \sim 3$ , we find a larger variety of velocity offsets, ranging from  $\approx -250$  km s $^{-1}$  to  $\approx 800$  km s $^{-1}$  and an average of  $\approx 100^{+300}_{-240}$  km s $^{-1}$ . The higher incidence of blue-shifted Ly $\alpha$  in emission for some of the  $z \sim 3$  sources might be linked with their higher luminosities and potentially higher velocity outflows.

Overall, we find typical Ly $\alpha$  to systemic velocity offsets of  $\approx 120$  km s $^{-1}$  but with a number of offsets (for seven LAEs) consistent with zero, implying that Ly $\alpha$  is likely escaping close to systemic. The velocity offsets we obtain can be compared with typical velocity offsets of  $\approx 450 \pm 50$  km s $^{-1}$  from UV selected galaxies at similar redshifts (e.g. Erb et al. 2010; Steidel et al. 2010) and  $\sim 200$  km s $^{-1}$  for much fainter Ly $\alpha$  selected sources (see Trainor et al. 2015; Guaita et al. 2017; Verhamme et al. 2017). We conclude that luminous LAEs have Ly $\alpha$  velocity offsets in respect to systemic which are similar, or even smaller, to those of the faint Ly $\alpha$  selected sources, and much lower than that of the more general Lyman-break population at similar UV luminosities. Such low-velocity offsets may be a consequence of the relatively high Ly $\alpha$  escape fractions (see e.g. Behrens, Dijkstra & Niemeyer 2014; Verhamme et al. 2015; Rivera-Thorsen et al. 2017), which in turn provide conditions for the high observed Ly $\alpha$  luminosities.

### 5.2 The nature of the brightest LAEs with individual emission-line detections

In Fig. 10 we show our most luminous narrow-line LAE (BR3), which shows strong C IV, He II, C III] and weaker O III], revealing a highly ionizing powering source. No Nv is detected and all lines have FWHMs similar to Ly $\alpha$ , of  $\approx 600$  km s $^{-1}$ . We also show even

**Table 3.** Parameters and ranges used for our photo-ionization CLOUDY (Ferland et al. 1998, 2013) modelling. We vary density, metallicity, and the ionization parameter ( $\log U$ ) for both the star-like ionization (here modelled with BPASS, but also with black bodies of varying temperature from 20 to 160K) and for an AGN-like ionization (here modelled with a power-law slope).

Parameter	Range used for all models
Density ( $n_{\text{H}} \text{ cm}^{-3}$ )	100, 300, 1000
Metallicity ( $\log Z_{\odot}$ )	-2 to +0.5 (steps of 0.05)
Ion. param. ( $\log U$ )	-5 to +2 (steps of 0.2)
<b>Type of model</b>	<b>Range used</b>
Blackbody (Temp., K)	20k–160k (steps of 1k)
Power-law (slope)	-2.0 to -1.0 (steps of 0.05)
BPASS ( $\log \text{Age}$ , yr)	6.0–9.0 (steps of 0.1)

brighter broad-line emitters, and one of the brightest in the UV (BH-NB5-10), which shows large blue-shifted absorption features at C IV and Ly  $\alpha$ . Fig. 10 reveals the large variety in the rest-frame UV spectra of the most luminous LAEs at  $z \sim 2-3$ , which we find to be consistent with being powered by AGN.

In order to interpret the spectra, we conduct photo-ionization modelling (see also Nakajima et al. 2017) using (i) power-law spectra (to model AGN), (ii) BPASS (Eldridge & Stanway 2009, 2012; Stanway et al. 2016) stellar models, including binary evolution, and (iii) black-bodies with a range of effective temperatures. We explore a relatively wide range of physical conditions by using the CLOUDY (v 13.05) photo-ionization code (Ferland et al. 1998, 2013). Further details are given in Alegre et al. (in preparation) and in Appendix C. We summarize the physical conditions we explore in Table 3 (see also Feltre et al. 2016; Gutkin et al. 2016; Nakajima et al. 2017).

We estimate physical conditions within bright LAEs by using the line ratios presented in Table B1. In addition, we use upper limits to provide either upper or lower limits on line ratios. We then explore the multidimensional line ratio parameter space to find all CLOUDY models that produced line ratios consistent with observations within  $3\sigma$  of each line ratio, upper or lower limits of the line ratios. We then compute the best physical parameter (see Table 3) and typical uncertainties by taking the median, 16 and 84 percentiles of each physical parameter within all models that result in line ratios within  $3\sigma$  of the observations. This means that our approach is statistically more robust than simply trying to find the ‘best-fit’ of CLOUDY outputs to observations and more conservative in uncertainties than focusing on e.g. a single combination of line ratios. We also compute the fraction of emission line ratios that each type of model is able to reproduce within  $1\sigma$  of the constrained line ratios; we use these to discuss which model best described each source or stack of sources.

BR3 is best described by hot blackbody models or power-law models. BPASS photoionization models fail to reproduce 40 per cent of the emission line ratios available, re-enforcing BR3’s AGN nature. Power-law models suggest that BR3 has a power-law slope of  $\alpha = -1.2^{+0.2}_{-0.3}$ , an ionization parameter of  $\log U = -2.2^{+0.2}_{-0.1}$  and a gas metallicity of  $\log(\text{O}/\text{H}) + 12 = 8.5 \pm 0.2$ . Blackbody models suggest  $T_{\text{eff}} = 135 \pm 10 \text{ kK}$ , an ionization parameter of  $\log U = -2.0^{+1.3}_{-0.3}$  and a gas metallicity of  $\log(\text{O}/\text{H}) + 12 = 8.3^{+0.4}_{-0.1}$ . The results suggest that BR3 is a highly ionizing source with a high effective temperature, while its gas is consistent with being mildly sub-solar at roughly  $0.5 Z_{\odot}$ .

We perform the same analysis for sources in which we detect at least two other UV emission lines in addition to Ly  $\alpha$ , allowing

for line ratios between those three or more lines (and constraints on the others) to be used. It is worth noting that requiring at least three UV lines to be detected is expected to select sources with the highest ionization parameters, AGN nature, and/or high effective temperatures.

In general, we find that our AGN-classified sources have line ratios which are only reproducible with power-law models or black-bodies with high temperatures around  $\sim 120-140 \text{ kK}$ , particularly due to their N V and He II fluxes. These sources have line ratios which imply ionization parameters ( $\log U$ ) from  $\approx -2$  to  $\approx 1$  and gas-phase metallicities between  $\approx 0.4 Z_{\odot}$  and  $\approx 2 Z_{\odot}$ , mostly driven by the strong N V detections (see e.g. Matsuoka et al. 2009), but also by C IV. These results show that there is a diversity of properties within the AGN population of bright LAEs at least in terms of gas metallicities and ionization parameters. The derived properties of the AGN in our sample are also very similar to those derived for high redshift radio galaxies (see Miley & De Breuck 2008).

For the only source with at least two UV lines which is likely dominated by star formation (GN-NB5-6712), we find that BPASS models are able to reproduce the line ratios, although blackbody models with temperatures  $\approx 70-100 \text{ kK}$  also reproduce the constraints on the line ratios. The CLOUDY-BPASS models suggest an  $\sim 20$ -Myr-old burst, an ionization parameter of  $\log U = -3.0^{+1.7}_{-0.8}$  and sub-solar to solar gas-phase metallicity ( $0.6^{+0.8}_{-0.4} Z_{\odot}$ ), and thus similar to properties derived for faint LAEs (e.g. Nakajima & Ouchi 2014; Suzuki et al. 2017). We note, nonetheless, that GN-NB5-6712 is not necessarily representative of the population of bright LAEs dominated by star formation, particularly due to the detection of both C IV and He II.

### 5.3 Results from stacks

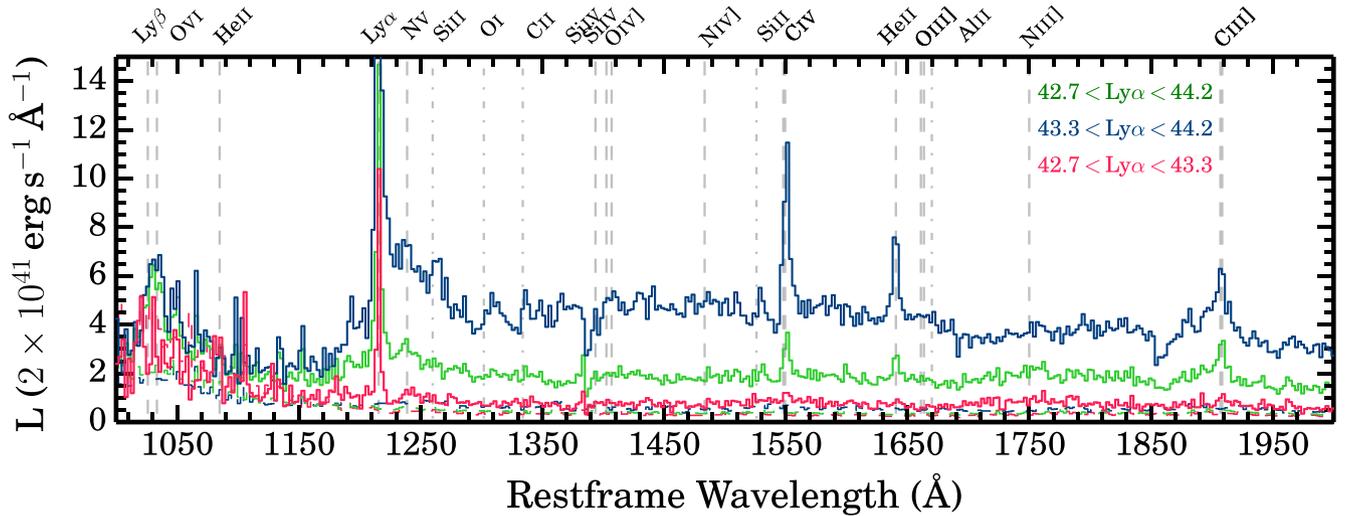
For a fraction of the LAEs the physical parameters are unconstrained due to many lines being below the detection limit. This can in principle be addressed through stacking similar sources, allowing us to better constrain the general nature of the sources within the population of luminous LAEs as a whole. We stack the sub-samples listed in Table 4, and constrain the emission lines following the same methodology as for individual measurements. We stack using the Ly  $\alpha$  redshift. As we are only interested in obtaining the stacked flux or luminosity of each line, we bin each spectrum so that we eliminate most of the scatter in velocity offsets between Ly  $\alpha$  and other lines (to avoid blurring them out) and also make doublets unresolved. In practice, we bin each spectrum to a resolution of  $750 \text{ km s}^{-1}$  at rest-frame  $1500 \text{ \AA}$ , motivated by the largest velocity offsets in our sample. This also allows us to measure emission-line fluxes with single Gaussians for all stacks.

In order to stack the spectra, we transform flux density into luminosity density taking into account the redshift of each source and average combine spectra weighting by the inverse of the variance, as well as masking strong OH sky lines. We also obtain median stacks which provide similar results but with lower S/N.

We show our results for a range of our stacking analysis as a function of Ly  $\alpha$  luminosity in Fig. 11 and the full results in Table 4. We find that the strength of rest-frame UV lines roughly scales with Ly  $\alpha$  and UV luminosities (mostly driven by AGN activity), but that the dependence with increasing Ly  $\alpha$  FWHM appears different, particularly due to the spectra of the broadest line emitters which show relatively symmetric Ly  $\alpha$  (implying an average line-of-sight volume with very low column density), symmetric strong Ly  $\beta$  and

**Table 4.** Results from the stacking of our sources in different sub-samples and using all constrained UV lines (see Table B1) in order to extract likely physical conditions. We provide both line ratios for the stacks and potential physical conditions from our CLOUDY modelling; see Section 5.2 and Table 3. Effective temperatures ( $T_{\text{eff}}$ ) are indicative, coming from the warm ionized inter-stellar medium and from our blackbody ionizing sources. Gas metallicities are given in  $\log(\text{O}/\text{H})+12$  with solar being 8.7.

Stack	Nv/Ly $\alpha$	C IV/He II	C III]/He II	$\log U$	Gas metallicity	Burst Age (Myrs)	Power-law $\alpha$	$T_{\text{eff}}$ (kK)
Full sample	$0.11 \pm 0.01$	$1.7 \pm 0.2$	$5.2 \pm 0.4$	$-0.7^{+0.5}_{-0.4}$	$8.7^{+0.3}_{-0.1}$	$6^{+25}_{-4}$	$-1.4^{+0.3}_{-0.3}$	$150^{+10}_{-50}$
All SFGs	$<0.16$	$2.9 \pm 0.8$	$4.8 \pm 1.4$	$-3.0^{+1.6}_{-0.9}$	$8.2^{+0.5}_{-0.3}$	$20^{+40}_{-15}$	—	—
All AGNs	$0.11 \pm 0.02$	$2.3 \pm 0.2$	$3.0 \pm 0.2$	$0.6^{+0.5}_{-0.5}$	$8.8^{+0.1}_{-0.1}$	—	$-1.4^{+0.4}_{-0.2}$	$70^{+70}_{-10}$
$42.6 < L_{\text{Ly}\alpha} < 43.3$	$0.05 \pm 0.02$	$2.5 \pm 0.9$	$2.5 \pm 0.8$	$-1.3^{+0.1}_{-0.4}$	$7.5^{+0.5}_{-0.1}$	$4^{+2}_{-2}$	$-1.7^{+0.3}_{-0.1}$	$130^{+20}_{-20}$
$43.3 < L_{\text{Ly}\alpha} < 44.2$	$0.22 \pm 0.02$	$2.0 \pm 0.1$	$2.6 \pm 0.2$	$-0.6^{+0.1}_{-0.1}$	$8.7^{+0.1}_{-0.1}$	$3^{+1}_{-2}$	$-1.5^{+0.3}_{-0.3}$	$155^{+5}_{-5}$
$-21.5 < M_{\text{UV}} < -19.1$	$<0.12$	$<1.9$	$2.1 \pm 0.7$	$-2.0^{+1.3}_{-0.6}$	$8.7^{+0.1}_{-0.5}$	$60^{+250}_{-55}$	$-1.5^{+0.3}_{-0.3}$	$120^{+20}_{-20}$
$-24.4 < M_{\text{UV}} < -21.5$	$<0.18$	$2.2 \pm 0.2$	$3.1 \pm 0.3$	$-1.1^{+0.4}_{-1.1}$	$8.3^{+0.2}_{-0.1}$	$13^{+1}_{-2}$	$-1.0^{+0.1}_{-0.6}$	$90^{+50}_{-20}$
$200 < \text{FWHM} < 1000$	$0.10 \pm 0.01$	$<0.70$	$7.2 \pm 1.7$	$-0.0^{+0.2}_{-0.2}$	$8.9^{+0.1}_{-0.1}$	$80^{+300}_{-70}$	$-1.5^{+0.2}_{-0.2}$	$130^{+5}_{-20}$
$1000 < \text{FWHM} < 3000$	$<0.21$	$>4.3$	$>2.1$	$-1.4^{+0.9}_{-0.2}$	$8.8^{+0.2}_{-0.1}$	—	$-1.4^{+0.2}_{-0.2}$	$120^{+30}_{-20}$



**Figure 11.** Examples of weighted (by the inverse of the noise) average stacks by splitting our sample of  $z \sim 2-3$  luminous LAEs in two based on the average Ly $\alpha$  luminosity and also by stacking the full sample. We indicate the rest-frame wavelengths of the main rest-frame UV nebular and ISM lines with dashed and dot-dashed lines, respectively. The horizontal dashed lines show the  $1\sigma$  error per resolution element of each stack. We find a strong increase in the high ionization UV emission lines with increasing Ly $\alpha$  luminosity, while blue-shifted ISM absorption lines are also clearer and more blueshifted for the stack of the most luminous LAEs.

broad high ionization metal lines like Nv, C IV and C III] but weak to undetectable e.g. He II. The most luminous LAEs are also those with the most symmetric Ly $\alpha$  lines and show the strongest C IV, He II and C III] lines, along with significantly blue-shifted ISM lines likely tracing strong outflows of  $\sim 750-1000 \text{ km s}^{-1}$  (see also Erb 2015) (see Fig. 11).

We find that the ionization parameter increases with increasing Ly $\alpha$  and UV luminosities, being the highest for the stack of AGN sources ( $\log U = 0.6 \pm 0.5$ ) and the lowest for SFGs ( $\log U = -3.0^{+1.6}_{-0.9}$ ). It is therefore likely that trends with luminosity are driven by the prevalence of AGN sources at bright luminosities and at higher FWHMs (see Figs 8 and 9) and not due to a change in properties of star-forming dominated LAEs. We also find evidence for gas-phase metallicities of LAEs to be lower ( $\approx 0.1 - 0.3 Z_{\odot}$ ) for SFGs (see Table 4) and for lower luminosity LAEs (similar to e.g. Stark et al. 2014; Nakajima et al. 2018), and closer to solar or higher for the most luminous LAEs (Table 4), but again this is likely caused by the sharp transition between star-forming and AGN-dominated (see Fig. 8).

## 6 DISCUSSION

### 6.1 Bright LAEs at $z \sim 2-3$ : the SF-AGN transition and the physical interpretation

Our results show that the AGN fraction of luminous LAEs strongly depends on both UV and Ly $\alpha$  luminosity at  $z \sim 2-3$ . The brightest LAEs in Ly $\alpha$  or in the UV are AGN, and above  $L_{\text{Ly}\alpha} > 10^{43.3} \text{ erg s}^{-1}$  or  $M_{\text{UV}} < -21.5$  virtually all luminous LAEs are AGN, causing an abrupt change in the physical properties of the ionized gas at  $2 \times L^*$ , as seen in Section 5.3. This may be explained by the fact that only AGN can reach the highest observed luminosities in either UV or Ly $\alpha$ , and suggests a limiting observed SFR<sup>10</sup> for star-forming dominated LAEs of  $\approx 20 M_{\odot} \text{ yr}^{-1}$  at  $z \sim 2-3$ . In principle, such an SFR limit could be a consequence of the exponential cut-offs

<sup>10</sup> Obtained converting the observed maximal  $M_{\text{UV}} = -21.5$  or  $L_{\text{Ly}\alpha} = 10^{43.3} \text{ erg s}^{-1}$  to an SFR using a Salpeter IMF.

in the galaxy mass and the gas mass fraction functions, and the properties of dust formation in massive starbursts, with both effects potentially evolving with redshift (see Section 6.2). These likely combine to create a strong threshold on UV luminosity of SF-dominated systems.

If dust is the main driver of a sharp UV and Ly $\alpha$  observed luminosity limit for starbursts (leading to a sharp transition between SF and AGN dominated at  $2 \times L^*$ ), one would expect the limit to be different when looking at H $\alpha$  luminosity of H $\alpha$  selected sources, which is significantly less affected by dust. Results from Sobral et al. (2016) show that there is a rise in the AGN fraction of H $\alpha$  selected sources as a function of H $\alpha$  luminosity above  $2 \times L_{\text{H}\alpha}^*$ , but such rise is much slower and not as abrupt as the one found for LAEs (see Fig. 8) with Ly $\alpha$  or UV luminosities. H $\alpha$  emitters only become 100 per cent AGN dominated at an observed H $\alpha$  luminosity of about  $10^{43.8} \text{ erg s}^{-1}$ , corresponding to an observed SFR limit of  $\approx 500 M_{\odot} \text{ yr}^{-1}$  in H $\alpha$  luminosity, 25 times larger than the limit for the observed UV and Ly $\alpha$  SFRs for LAEs. We therefore interpret the sharp shift from star-forming to AGN dominated sources at  $2 \times L^*$  for LAEs as likely a combination of two effects: a rise in the AGN fraction with intrinsic luminosity (Sobral et al. 2016), as seen for H $\alpha$  selected samples, and the dominating factor of increased obscured SFRs as a function of increasing intrinsic SFR (e.g. Garn & Best 2010; Swinbank et al. 2004; Whitaker et al. 2017), leading to a maximal unobscured starburst.

While intrinsically star-formation dominated sources can still reach even higher SFRs and produce higher UV and Ly $\alpha$  luminosities, such intense starbursts happen in dusty environments or starbursts themselves produce the dust at  $z \sim 2-3$  (showing up as dusty star-forming galaxies; see Casey et al. 2014a), thus limiting or even reducing the observed UV and Ly $\alpha$  luminosities. The high intrinsic production of UV and Ly $\alpha$  photons of dusty star-forming galaxies suggests that any escape channel created by intense feedback in those galaxies will still lead to non-negligible UV and Ly $\alpha$  luminosities which are observed in a variety of studies (see e.g. Chapman et al. 2005; Nilsson & Møller 2009; Casey et al. 2014b; Matthee et al. 2016), but that given our results will not exceed  $2 \times L^*$  at  $z \sim 2-3$ .

We conclude that the clear separation between star-forming dominated and AGN dominated LAEs at  $2 \times L^*$  is caused primarily by obscured star-formation becoming dominant above observed SFRs of  $\approx 20 M_{\odot} \text{ yr}^{-1}$ . Prominent dust obscuration will efficiently absorb UV photons and directly or indirectly (via scattering) lead to low Ly $\alpha$  escape fractions (e.g. Atek et al. 2008; Matthee et al. 2016) such that the observed luminosities saturate. Ly $\alpha$  and UV luminosities observed above  $2 \times L^*$  are thus likely to be produced by different physical processes linked with AGN activity, which in principle do not create dust. These may include strong AGN-driven outflows and radiation from different regions within the accretion disc, as seen from the evidence of the UV bright broad line emitters, while shocks may also play a role. The existence of UV blue, Ly $\alpha$  bright quasars (e.g. Borisova et al. 2016) implies that the physical processes in accretion discs of massive black holes are scalable to much higher UV luminosities without dust imposing a strong limitation in the observed luminosities. This is likely due to outflows and high ionization radiation fields which we find to be present ( $\log U$  from  $\approx -2$  to  $\approx 1$ ; see Sections 5.2 and 5.3) that may destroy dust or open channels for Ly $\alpha$  and UV photons to escape (see also Venemans et al. 2007; Miley & De Breuck 2008). Nevertheless, it should be noted that our most luminous LAEs already have Ly $\alpha$  luminosities as high (e.g. Borisova et al. 2016) as the ones found for the most UV luminous quasars (e.g. Richards et al. 2006), and

reaching down to luminosities below those of fainter AGN (e.g. Gavignaud et al. 2006). Therefore, while the UV luminosity seems to be scalable up from even our brightest UV sources by a few orders of magnitude, there may be a saturation in the observable Ly $\alpha$  luminosity at close to  $10^{44-45} \text{ erg s}^{-1}$  for AGN at  $z \sim 2-3$ .

## 6.2 Implications for higher redshift: evolution?

Our results and physical interpretation at  $z \sim 2-3$  may allow us to shed further light into the nature and evolution of higher redshift LAEs selected in the same way. Recently, Sobral et al. (2018) obtained a large sample of  $\sim 4000$  LAEs from  $z \sim 2$  to  $z \sim 6$  and explored it to infer a strong evolution (a factor of  $\approx 4-5$ ) in the typical Ly $\alpha$  escape fraction of star-forming galaxies, rising as  $(1+z)^2$ . This increase is in addition to the likely increase of  $\xi_{\text{ion}}$  (tracing high burstiness and/or an average change in stellar populations) by a factor of  $\sim 2$ , leading to a total observed rise of  $(1+z)^3$  in the Ly $\alpha$  to UV luminosity density ratio with increasing redshift (see also Hayes et al. 2011). Furthermore, Sobral et al. (2018) find a factor of  $\sim 5$  rise in  $L_{\text{Ly}\alpha}^*$  from  $z \sim 2$  to  $z \sim 6$ , which, under our interpretation of  $2 \times L_{\text{Ly}\alpha}^*$  being the limit of the observed Ly $\alpha$  luminosity of a starburst would imply that by  $z \sim 6$  the limiting observable Ly $\alpha$  luminosity for a starburst will be  $\approx 10^{44} \text{ erg s}^{-1}$ . If this is the case, by  $z \sim 6$  the limiting unobscured SFR would be  $\approx 100 M_{\odot} \text{ yr}^{-1}$ , although the UV observed SFR may be limited to a lower value due to the high ratio of LyC to UV photons ( $\xi_{\text{ion}}$ ) expected by  $z \sim 6$  (see e.g. Matthee et al. 2017a). These results would provide a natural explanation for the exceptionally high Ly $\alpha$  luminosities of equally selected LAEs at  $z \sim 6-7$  (Ouchi et al. 2009; Sobral et al. 2015; Hu et al. 2016; Matthee et al. 2017c) with  $L_{\text{Ly}\alpha} \approx 10^{43.4-43.9} \text{ erg s}^{-1}$  that so far show no convincing evidence for being dominated by AGN activity (see also Pallottini et al. 2015; Mas-Ribas, Dijkstra & Forero-Romero 2016; Bagley et al. 2017; Pacucci et al. 2017). For example, these luminous  $z \sim 6-7$  LAEs show Ly $\alpha$  profiles which are typically narrow ( $\approx 200-300 \text{ km s}^{-1}$ ; Matthee et al. 2017c), contrarily to the high FWHMs we find at  $z \sim 2-3$  for the same luminosities (see Fig. 9). According to our interpretation of an evolving  $L_{\text{Ly}\alpha}^*$  by a factor of 5, the  $z \sim 6-7$  bright LAEs would easily be consistent with being star-forming dominated, with CR7's Ly $\alpha$  luminosity ( $\approx 10^{43.9} \text{ erg s}^{-1}$ ; Sobral et al. 2015) being the closest to our inferred maximal unobscured starburst luminosity of  $\approx 10^{44} \text{ erg s}^{-1}$  at  $z \sim 6$ . Recent deep ALMA observations of CR7 (Matthee et al. 2017e) are fully consistent with this interpretation, with CR7's obscured SFR being below  $< 4 M_{\odot} \text{ yr}^{-1}$ .

Potential mechanisms which may be able to raise the maximal unobscured luminosity of early luminous LAEs are (major) mergers (seen as a single source) and/or significant gas inflows without reducing the Ly $\alpha$  escape fraction and/or without leading to significant dust obscuration. The necessary conditions likely require dust destruction and/or inefficient dust formation at low metallicities (see e.g. Behrens et al. 2018), which may be more common at higher redshift, including the possibility of witnessing a first significant burst of star-formation. Interestingly, there is evidence for the brightest LAEs at  $z \sim 7$  being multiple component/mergers (Ouchi et al. 2010; Sobral et al. 2015, 2017a; Matthee et al. 2017c), while those at  $z \sim 2-3$  are typically very compact and single component dominated (e.g. Paulino-Afonso et al. 2018). At  $z \sim 2-3$  major mergers lead to very high SFRs but that are significantly obscured (see Casey et al. 2014a).

Our scenario of an evolving maximal unobscured starburst is consistent with results from Calhau et al. (2018), who find that

the relation between the X-ray AGN fraction of LAEs and Ly $\alpha$  luminosity declines with redshift or shifts to higher luminosities at higher redshift. Our discussed scenario, along with its physical implications, can easily be tested with the *James Webb Space Telescope* (*JWST*). Using *JWST*'s IFU capabilities in the NIR and MIR regimes, one can observe CR7, Himiko, VR7, MASOSA, and other luminous LAEs (see e.g. Matthee et al. 2017c) and establish them as either star-forming or AGN dominated, along with measuring their resolved dust obscuration using the Balmer decrement. Most interestingly, IFU data will allow us to start understanding the physical reasons behind a potentially higher maximal unobscured starburst at high redshift and implications for dust formation and stellar populations in early luminous LAEs.

## 7 CONCLUSIONS

We have presented the spectroscopic follow-up of 23 luminous ( $L_{\text{Ly}\alpha} \gtrsim L_{\text{Ly}\alpha}^*$ ) Ly $\alpha$  candidates at  $z \sim 2-3$  found with large narrow-band surveys in Boötes, COSMOS, GOODS-N, SA22, and UDS. We have used WHT/ISIS, Keck/DEIMOS, and VLT/X-SHOOTER to spectroscopically confirm 21 of them as luminous LAEs. We exploit the wide wavelength coverage of our data and the high-resolution spectra to measure line profiles, investigate the nature of the sources, and explore UV emission line ratios. Our main results are as follows:

(i) Luminous LAEs at  $z \sim 2-3$  present a striking diversity in terms of their Ly $\alpha$  and UV properties. They also present a wide diversity in the shape of the Ly $\alpha$  line, ranging from roughly symmetric to highly asymmetric and including 25 percent of sources with significant blue-shifted Ly $\alpha$  components or double-peaked lines, typically linked with high Ly $\alpha$  escape fractions.

(ii) Luminous LAEs are in general very blue (UV slope of  $\beta = -2.0_{-0.1}^{+0.3}$ ) and span five orders of magnitude in rest-frame UV luminosities, from  $\sim 0.3 M_{\text{UV}}^*$  to  $\sim 30 M_{\text{UV}}^*$ . They likely have a low dust content, consistent with  $E(B - V) \approx 0.05$ , and have high Ly $\alpha$  escape fractions of  $50_{-15}^{+20}$  percent after correcting for dust extinction.

(iii) Besides Ly $\alpha$ , the most prevalent high ionization UV line in luminous LAEs is C IV, present in 40 percent of the sample, but we also find N V, He II, C III] and O III] as relatively common lines, particularly at the brightest UV and Ly $\alpha$  luminosities.

(iv) Ly $\alpha$  FWHMs of luminous LAEs vary from  $\sim 200$  to  $\sim 4000 \text{ km s}^{-1}$ , with the prevalence of broad lines rising significantly with increasing Ly $\alpha$  and UV luminosities. We find significant relations between the Ly $\alpha$  FWHM and Ly $\alpha$  and UV luminosities for NL-AGN, likely driven by outflows.

(v) The narrow LAEs have small velocity offsets between Ly $\alpha$  and other lines of only  $\sim 100-200 \text{ km s}^{-1}$ , consistent with Ly $\alpha$  photons escaping relatively close to systemic and providing further evidence for high Ly $\alpha$  escape fractions in bright LAEs.

(vi) We find a significant increase in the AGN fraction of LAEs with both Ly $\alpha$  and UV luminosities from  $\sim 0$  to  $\sim 100$  percent. Above  $L_{\text{Ly}\alpha} > 10^{43.3} \text{ erg s}^{-1}$  and/or  $M_{\text{UV}} < -21.5$  all luminous LAEs at  $z \sim 2-3$  are AGN. We find simple relations to predict the AGN fraction of LAEs (within the ranges studied):  $f_{\text{AGN}} = (-0.30 \pm 0.07)(M_{\text{UV}} + 20.5) + (0.35 \pm 0.11)$  and  $f_{\text{AGN}} = (0.78 \pm 0.22)(\log_{10}(L_{\text{Ly}\alpha}) - 42.72) + (0.24 \pm 0.14)$ .

(vii) The transition from star-forming dominated to AGN dominated happens at  $2 \times L^*$  and is likely linked with a physical limit for the maximum observed output of an unobscured starburst of  $\sim 20 M_{\odot} \text{ yr}^{-1}$  at  $z \sim 2-3$ . For higher intrinsic UV and Ly $\alpha$  lumi-

nosities dust likely prevents such sources from being observed as brighter than  $2 \times L^*$ .

(viii) AGN LAEs reveal high ionization parameters ( $\log U = 0.6 \pm 0.5$ ) and sub-solar to solar metallicities. Luminous LAEs lacking signatures of AGN ( $40 \pm 11$  percent) are less ionizing ( $\log U = -3.0_{-0.9}^{+1.6}$  and  $\log \xi_{\text{ion}} = 25.4 \pm 0.2$ ), apparently metal-poor ( $12 + \log(\text{O}/\text{H}) = 8.2_{-0.3}^{+0.5}$ ) sources likely powered by young, dust-poor ‘maximal’ starbursts.

(ix) The ionization parameter rises with Ly $\alpha$  and UV luminosities by typically 1 dex from the low to high luminosities, while gas-phase metallicities rise with Ly $\alpha$  luminosity from  $12 + \log(\text{O}/\text{H}) = 7.5_{-0.1}^{+0.5}$  to  $12 + \log(\text{O}/\text{H}) = 8.7_{-0.1}^{+0.1}$ . We interpret these results as a consequence of the transition between star-forming dominated systems to AGN dominated which we find to be particularly abrupt at  $2 \times L^*$ .

## ACKNOWLEDGEMENTS

We thank the anonymous reviewer for their timely and constructive comments that greatly helped us to improve the manuscript. DS acknowledges financial support from the Netherlands Organization for Scientific research (NWO) through a Veni fellowship and from Lancaster University through an Early Career Internal Grant A100679. JM acknowledges the support of a Huygens PhD fellowship from Leiden University. BD acknowledges financial support from NASA through the Astrophysics Data Analysis Program (ADAP), grant number NNX12AE20G, and the National Science Foundation, grant number 1716907. IRS acknowledges support from the ERC Advanced Grant DUSTYGAL (321334), STFC (ST/P000541/1), and a Royal Society/Wolfson Merit Award. PNB is grateful for support from STFC via grant ST/M001229/1. We thank Anne Verhamme, Kimihiko Nakajima, Ryan Trainor, Sangeeta Malhotra, Max Gronke, James Rhoads, Fang Xia An, Matthew Hayes, Takashi Kojima, Mark Dijkstra, and Anne Jaskot for many helpful and engaging discussions, particularly during the SnowCLAW Ly $\alpha$  workshop. We thank Bruno Ribeiro, Stephane Charlot, and Joseph Caruana for comments on the manuscript. The authors would also like to thank Ingrid Tengs, Meg Singleton, Ali Khostovan, and Sara Perez for participating in part of the observations. We also thank João Calhau, Leah Morabito, Sérgio Santos, and Aayush Saxena for their assistance with the narrow-band observations which allowed to select some of the sources. Based on observations obtained with the William Herschel Telescope, program: W16AN004; the Very Large Telescope, programs: 098.A-0819 & 099.A-0254; and the Keck II telescope, program: C267D. Based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under ESO programme IDs 294.A-5018, 294.A-5039, 092.A-0786, 093.A-0561, 097.A-0943, 098.A-0819, 099.A-0254 and 179.A-2005. The authors acknowledge the award of service time (SW2014b20) on the WHT. WHT and its service programme are operated on the island of La Palma by the Isaac Newton Group in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias. The authors would also like to thank all the extremely helpful observatory staff that have greatly contributed towards our observations, particularly Fiona Riddick, Lilián Dominguez, Florencia Jimenez, and Ian Skillen.

We have benefited greatly from the publicly available programming language PYTHON, including the NUMPY & SCIPY (Van Der Walt, Colbert & Varoquaux 2011; Jones et al. 2001), MATPLOTLIB (Hunter 2007), ASTROPY (Astropy Collaboration et al. 2013), and the TOPCAT analysis program (Taylor 2013). This research has made use of the Vizier catalogue access tool, CDS, Strasbourg, France.

The samples and parent samples used for this paper are publicly available (Matthee et al. 2017b; Sobral et al. 2017b, 2018). We also make all our 1D (binned to one-third of the resolution, flux calibrated, without any slit correction) spectra public on-line with the final refereed version of this paper, together with the catalogue of luminous LAEs and electronic versions of relevant tables (see Appendix A2).

## REFERENCES

- Alexander D. M. et al., 2003, *AJ*, 126, 539  
 An F. X., Zheng X. Z., Hao C.-N., Huang J.-S., Xia X.-Y., 2017, *ApJ*, 835, 116  
 Arnouts S. et al., 2005, *ApJL*, 619, L43  
 Astropy Collaboration, 2013, *A&A*, 558, A33  
 Atek H., Kunth D., Hayes M., Östlin G., Mas-Hesse J. M., 2008, *AAP*, 488, 491  
 Atek H. et al., 2015, *ApJ*, 814, 69  
 Bagley M. B. et al., 2017, *ApJ*, 837, 11  
 Behrens C., Dijkstra M., Niemeyer J. C., 2014, *A&A*, 563, A77  
 Behrens C., Pallottini A., Ferrara A., Gallerani S., Vallini L., 2018, *MNRAS*, 477, 552  
 Borisova E. et al., 2016, *ApJ*, 831, 39  
 Bouwens R. J. et al., 2009, *ApJ*, 705, 936  
 Bouwens R. J. et al., 2015a, *ApJ*, 803, 34  
 Bouwens R. J. et al., 2015b, *ApJ*, 803, 34  
 Bouwens R. J., Smit R., Labbé I., Franx M., Caruana J., Oesch P., Stefanon M., Rasappu N., 2016, *ApJ*, 831, 176  
 Bowler R. A. A. et al., 2014, *MNRAS*, 440, 2810  
 Burgarella D., Buat V., Iglesias-Páramo J., 2005, *MNRAS*, 360, 1413  
 Calhau J., Sobral D., Matthee J., Paulino-Afonso A., Santos S., Stroe A., 2018, *MNRAS*, submitted  
 Caruana J. et al., 2018, *MNRAS*, 473, 30  
 Casey C. M., Narayanan D., Cooray A., 2014a, *Phys. Rep.*, 541, 45  
 Casey C. M. et al., 2014b, *ApJ*, 796, 95  
 Cassata P. et al., 2015, *AAP*, 573, A24  
 Chapman S. C., Blain A. W., Smail I., Ivison R. J., 2005, *ApJ*, 622, 772  
 Civano F. et al., 2016, *ApJ*, 819, 62  
 Cooke J., Ryan-Weber E. V., Garel T., Díaz C. G., 2014, *MNRAS*, 441, 837  
 Cooper M. C., Newman J. A., Davis M., Finkbeiner D. P., Gerke B. F., 2012, *Astrophysics Source Code Library*, record ascl:1203.003  
 Crocker R. M., Krumholz M. R., Thompson T. A., Clutterbuck J., 2018, *MNRAS*, preprint, (arXiv:1802.03117)  
 Darvish B., Mobasher B., Sobral D., Hemmati S., Nayyeri H., Shivaeei I., 2015, *ApJ*, 814, 84  
 Dijkstra M., Gronke M., Venkatesan A., 2016, *ApJ*, 828, 71  
 Drake A. B. et al., 2017a, *MNRAS*, 471, 267  
 Drake A. B. et al., 2017b, *A&A*, 608, A6  
 Dressler A., Henry A., Martin C. L., Sawicki M., McCarthy P., Villaneuva E., 2015, *ApJ*, 806, 19  
 Dunlop J. S., McLure R. J., Robertson B. E., Ellis R. S., Stark D. P., Cirasuolo M., de Ravel L., 2012, *MNRAS*, 420, 901  
 Eldridge J. J., Stanway E. R., 2009, *MNRAS*, 400, 1019  
 Eldridge J. J., Stanway E. R., 2012, *MNRAS*, 419, 479  
 Eldridge J. J., Izzard R. G., Tout C. A., 2008, *MNRAS*, 384, 1109  
 Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, *PASA*, 34, e058  
 Ellis R. S. et al., 2013, *ApJL*, 763, L7  
 Erb D. K., 2015, *Nature*, 523, 169  
 Erb D. K., Pettini M., Shapley A. E., Steidel C. C., Law D. R., Reddy N. A., 2010, *ApJ*, 719, 1168  
 Faisst A. L., 2016, *ApJ*, 829, 99  
 Feltre A., Charlot S., Gutkin J., 2016, *MNRAS*, 456, 3354  
 Ferland G. J., Korista K. T., Verner D. A., Ferguson J. W., Kingdon J. B., Verner E. M., 1998, *PASP*, 110, 761  
 Ferland G. J. et al., 2013, *Rev. Mex. Astron. Astrofis.*, 49, 137  
 Finkelstein S. L. et al., 2015, *ApJ*, 810, 71  
 Fynbo J. U., Møller P., Thomsen B., 2001, *A&A*, 374, 443  
 Fynbo J. P. U., Ledoux C., Møller P., Thomsen B., Burud I., 2003, *A&A*, 407, 147  
 Garn T., Best P. N., 2010, *MNRAS*, 409, 421  
 Gavignaud I. et al., 2006, *A&A*, 457, 79  
 Giavalisco M., Steidel C. C., Macchetto F. D., 1996, *ApJ*, 470, 189  
 Gräfenor G., Vink J. S., 2015, *A&A*, 578, L2  
 Gronke M., Dijkstra M., Trenti M., Wyithe S., 2015, *MNRAS*, 449, 1284  
 Gronwall C. et al., 2007, *ApJ*, 667, 79  
 Grove L. F., Fynbo J. P. U., Ledoux C., Limousin M., Møller P., Nilsson K. K., Thomsen B., 2009, *A&A*, 497, 689  
 Guaita L. et al., 2016, *A&A*, 587, A133  
 Guaita L. et al., 2017, *A&A*, 606, A19  
 Guhathakurta P., Tyson J. A., Majewski S. R., 1990, *ApJ*, 357, L9  
 Gutkin J., Charlot S., Bruzual G., 2016, *MNRAS*, 462, 1757  
 Harikane Y. et al., 2017, *ApJ*, preprint (arXiv:1711.03735)  
 Hashimoto T. et al., 2017, *A&A*, 608, A10  
 Hathi N. P. et al., 2016, *A&A*, 588, A26  
 Hayes M., Schaerer D., Östlin G., Mas-Hesse J. M., Atek H., Kunth D., 2011, *ApJ*, 730, 8  
 Heckman T. M., 2003, *Rev. Mex. Astron. Astrofis.*, 17, 47  
 Hopkins A. M., Connolly A. J., Haarsma D. B., Cram L. E., 2001, *AJ*, 122, 288  
 Horne K., 1986, *PASP*, 98, 609  
 Hu E. M., Cowie L. L., Songaila A., Barger A. J., Rosenwasser B., Wold I. G. B., 2016, *ApJL*, 825, L7  
 Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90  
 Ilbert O. et al., 2009, *ApJ*, 690, 1236  
 Izotov Y. I., Schaerer D., Worseck G., Guseva N. G., Thuan T. X., Verhamme A., Orlitová I., Fricke K. J., 2018, *MNRAS*, 474, 4514  
 Jaskot A. E., Ravindranath S., 2016, *ApJ*, 833, 136  
 Jiang L. et al., 2017, *ApJ*, 846, 134  
 Jones E., Oliphant T., Peterson P. et al., 2001, *SciPy: Open Source Scientific Tools for Python*, <http://www.scipy.org/>  
 Kennicutt Jr R. C., 1998, *ARAA*, 36, 189  
 Khostovan A. A., Sobral D., Mobasher B., Smail I., Darvish B., Nayyeri H., Hemmati S., Stott J. P., 2016, *MNRAS*, 463, 2363  
 Konno A., Ouchi M., Nakajima K., Duval F., Kusakabe H., Ono Y., Shimasaku K., 2016, *ApJ*, 823, 20  
 Koo D. C., Kron R. T., 1980, *PASP*, 92, 537  
 Kusakabe H., Shimasaku K., Nakajima K., Ouchi M., 2015, *ApJ*, 800, L29  
 Laporte N., Nakajima K., Ellis R. S., Zitrin A., Stark D. P., Mainali R., Roberts-Borsani G. W., 2017, *ApJ*, 851, 40  
 Lehmer B. D. et al., 2009, *ApJ*, 691, 687  
 Ma X., Hopkins P. F., Kasen D., Quataert E., Faucher-Giguère C.-A., Kereš D., Murray N., Strom A., 2016, *MNRAS*, 459, 3614  
 Marchi F. et al., 2017a, *A&A*, preprint (arXiv:1710.10184)  
 Marchi F. et al., 2017b, *A&A*, 601, A73  
 Mármol-Queraltó E., McLure R. J., Cullen F., Dunlop J. S., Fontana A., McLeod D. J., 2016, *MNRAS*, 460, 3587  
 Mas-Ribas L., Dijkstra M., Forero-Romero J. E., 2016, *ApJ*, 833, 65  
 Matsuoka K., Nagao T., Maiolino R., Marconi A., Taniguchi Y., 2009, *A&A*, 503, 721  
 Matthee J., Sobral D., Santos S., Röttgering H., Darvish B., Mobasher B., 2015, *MNRAS*, 451, 400  
 Matthee J., Sobral D., Oteo I., Best P., Smail I., Röttgering H., Paulino-Afonso A., 2016, *MNRAS*, 458, 449  
 Matthee J., Sobral D., Best P., Khostovan A. A., Oteo I., Bouwens R., Röttgering H., 2017a, *MNRAS*, 465, 3637  
 Matthee J., Sobral D., Best P., Smail I., Bian F., Darvish B., Röttgering H., Fan X., 2017b, *MNRAS*, 471, 629  
 Matthee J., Sobral D., Darvish B., Santos S., Mobasher B., Paulino-Afonso A., Röttgering H., Alegre L., 2017c, *MNRAS*, 472, 772  
 Matthee J. et al., 2017d, *ApJ*, 851, 145  
 Matthee J. et al., 2017e, *ApJ*, 851, 145  
 Meurer G. R., Heckman T. M., Leitherer C., Kinney A., Robert C., Garnett D. R., 1995, *AJ*, 110, 2665

- Meurer G. R., Heckman T. M., Lehnert M. D., Leitherer C., Lowenthal J., 1997, *AJ*, 114, 54
- Meurer G. R., Heckman T. M., Calzetti D., 1999, *ApJ*, 521, 64
- Miley G., De Breuck C., 2008, *A&A Rev.*, 15, 67
- Modigliani A. et al., 2010, *Proc. SPIE*, 7737, 773728
- Morton D. C., 1991, *ApJS*, 77, 119
- Murray S. S. et al., 2005, *ApJS*, 161, 1
- Møller P., Warren S. J., 1993, *A&A*, 270, 43
- Nakajima K., Ouchi M., 2014, *MNRAS*, 442, 900
- Nakajima K. et al., 2012, *ApJ*, 745, 12
- Nakajima K., Ellis R. S., Iwata I., Inoue A. K., Kusakabe H., Ouchi M., Robertson B. E., 2016, *ApJL*, 831, L9
- Nakajima K. et al., 2017, *A&A*, preprint ([arXiv:1709.03990](https://arxiv.org/abs/1709.03990))
- Nakajima K., Fletcher T., Ellis R. S., Robertson B. E., Iwata I., 2018, *MNRAS*, 477, 2098
- Nilsson K. K., Møller P., 2009, *A&A*, 508, L21
- Nilsson K. K. et al., 2007, *AAP*, 471, 71
- Nilsson K. K., Tapken C., Møller P., Freudling W., Fynbo J. P. U., Meisenheimer K., Laursen P., Östlin G., 2009, *A&A*, 498, 13
- Nilsson K. K., Östlin G., Møller P., Möller-Nilsson O., Tapken C., Freudling W., Fynbo J. P. U., 2011, *A&A*, 529, A9
- Oesch P. A. et al., 2015, *ApJL*, 804, L30
- Oke J. B., Gunn J. E., 1983, *ApJ*, 266, 713
- Ono Y. et al., 2010a, *MNRAS*, 402, 1580
- Ono Y., Ouchi M., Shimasaku K., Dunlop J., Farrah D., McLure R., Okamura S., 2010b, *ApJ*, 724, 1524
- Oteo I., Sobral D., Ivison R. J., Smail I., Best P. N., Cepa J., Pérez-García A. M., 2015, *MNRAS*, 452, 2018
- Ouchi M. et al., 2008, *ApJs*, 176, 301
- Ouchi M. et al., 2009, *ApJ*, 696, 1164
- Ouchi M. et al., 2010, *ApJ*, 723, 869
- Pacucci F., Pallottini A., Ferrara A., Gallerani S., 2017, *MNRAS*, 468, L77
- Pallottini A. et al., 2015, *MNRAS*, 453, 2465
- Parsa S., Dunlop J. S., McLure R. J., Mortlock A., 2016, *MNRAS*, 456, 3194
- Paulino-Afonso A. et al., 2018, *MNRAS*, 476, 5479
- Rauch M. et al., 2008, *ApJ*, 681, 856
- Reddy N. A., Steidel C. C., 2009, *ApJ*, 692, 778
- Richards G. T. et al., 2006, *AJ*, 131, 2766
- Rivera-Thorsen T. E. et al., 2017, *A&A*, 608, L4
- Saez C. et al., 2015, *MNRAS*, 450, 2615
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Santos S., Sobral D., Matthee J., 2016, *MNRAS*, 463, 1678
- Schmidt K. B. et al., 2017, *ApJ*, 839, 17
- Shankar F., Weinberg D. H., Miralda-Escudé J., 2009, *ApJ*, 690, 20
- Shibuya T. et al., 2018, *PASJ*, 70, S14
- Shivvaei I. et al., 2018, *ApJ*, 855, 42
- Siana B. et al., 2007, *ApJ*, 668, 62
- Smit R. et al., 2014, *ApJ*, 784, 58
- Smit R., Bouwens R. J., Labbé I., Franx M., Wilkins S. M., Oesch P. A., 2016, *ApJ*, 833, 254
- Sobral D., Best P. N., Matsuda Y., Smail I., Geach J. E., Cirasuolo M., 2012, *MNRAS*, 420, 1926
- Sobral D., Best P. N., Smail I., Mobasher B., Stott J., Nisbet D., 2014, *MNRAS*, 437, 3516
- Sobral D., Matthee J., Darvish B., Schaerer D., Mobasher B., Röttgering H. J. A., Santos S., Hemmati S., 2015, *ApJ*, 808, 139
- Sobral D., Kohn S. A., Best P. N., Smail I., Harrison C. M., Stott J., Calhau J., Matthee J., 2016, *MNRAS*, 457, 1739
- Sobral D. et al., 2017a, *MNRAS*, preprint ([arXiv:1710.08422](https://arxiv.org/abs/1710.08422))
- Sobral D. et al., 2017b, *MNRAS*, 466, 1242
- Sobral D., Matthee J., 2018, *A&A*, preprint ([arXiv:1803.08923](https://arxiv.org/abs/1803.08923))
- Sobral D., Santos S., Matthee J., Paulino-Afonso A., Ribeiro B., Calhau J., Khostovan A. A., 2018, *MNRAS*, 476, 4725
- Stanway E. R., 2017, in Eldridge J. J., Bray J. C., McClelland L. A. S., Xiao L., eds, *Proc. IAU Symp.* 329, *The Lives and Death-Throes of Massive Stars*. Cambridge Univ. Press, Cambridge, p. 305
- Stanway E. R., Eldridge J. J., Becker G. D., 2016, *MNRAS*, 456, 485
- Stark D. P. et al., 2014, *MNRAS*, 445, 3200
- Stark D. P. et al., 2015a, *MNRAS*, 450, 1846
- Stark D. P. et al., 2015b, *MNRAS*, 454, 1393
- Steidel C. C., Hamilton D., 1993, *AJ*, 105, 2017
- Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, *ApJ*, 717, 289
- Steidel C. C., Strom A. L., Pettini M., Rudie G. C., Reddy N. A., Trainor R. F., 2016, *ApJ*, 826, 159
- Suzuki T. L. et al., 2017, *ApJ*, 849, 39
- Swinbank A. M., Smail I., Chapman S. C., Blain A. W., Ivison R. J., Keel W. C., 2004, *ApJ*, 617, 64
- Swinbank A. M. et al., 2010, *Nature*, 464, 733
- Taylor M., 2013, *Starlink User Note*, 253
- Topping M. W., Shull J. M., 2015, *ApJ*, 800, 97
- 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
- Ueda Y. et al., 2008, *ApJS*, 179, 124
- Van Der Walt S., Colbert S. C., Varoquaux G., 2011, *Comput. Sci. Eng.*, 13, 22
- Vanzella E. et al., 2016, *ApJL*, 821, L27
- Vanzella E. et al., 2017, *MNRAS*, 467, 4304
- Vanzella E. et al., 2018, *MNRAS*, 476, L15
- Veilleux S., 2002, in Green R. F., Khachikian E. Y., Sanders D. B., eds, *ASP Conf. Ser. Vol. 284, IAU Colloq. 184: AGN Surveys*. Astron. Soc. Pac., San Francisco, p. 111
- Venemans B. P. et al., 2005, *A&A*, 431, 793
- Venemans B. P. et al., 2007, *A&A*, 461, 823
- Verhamme A., Orlitová I., Schaerer D., Hayes M., 2015, *A&A*, 578, A7
- Verhamme A., Orlitová I., Schaerer D., Izotov Y., Worseck G., Thuan T. X., Guseva N., 2017, *A&A*, 597, A13
- Vernet J. et al., 2011, *A&A*, 536, A105
- Vieira J. D. et al., 2013, *Nature*, 495, 344
- Wang J. X. et al., 2004, *ApJ*, 608, L21
- Wardlow J. L. et al., 2014, *ApJ*, 787, 9
- Whitaker K. E., Pope A., Cybulski R., Casey C. M., Popping G., Yun M. S., 2017, *ApJ*, 850, 208
- Wold I. G. B., Barger A. J., Cowie L. L., 2014, *ApJ*, 783, 119
- Wold I. G. B., Finkelstein S. L., Barger A. J., Cowie L. L., Rosenwasser B., 2017, *ApJ*, 848, 108
- Zheng Z.-Y. et al., 2017, *ApJ*, 842, L22
- Zitrin A. et al., 2015, *ApJL*, 810, L12

## SUPPORTING INFORMATION

Supplementary data are available at [MNRAS](https://www.mnras.org) online.

**Table 2.** Our spectroscopic sample of luminous LAEs, ranked by redshift and then by Ly $\alpha$  luminosity of the parent sample (high to low luminosity top to bottom).

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

## APPENDIX A: OBSERVATIONS AND DATA REDUCTION

Tables [A1](#), [A3](#), and [A4](#) present the observing logs for WHT/ISIS, Keck/DEIMOS, and VLT/X-SHOOTER observations, respectively. In Table [A2](#) we present the best estimates of the spectral and noise properties of all reduced data for WHT/ISIS, Keck/DEIMOS, and VLT/X-SHOOTER.

**Table A1.** Observing log for WHT/ISIS (as part of W16AN004, PI: Sobral) of our luminous LAEs at  $z \sim 2$  (from Sobral et al. 2017b) and  $z \sim 2-3$  (from Matthee et al. 2017b) observed in the Boötes and GOODS-N fields. We order sources based on the predicted Ly  $\alpha$  luminosity from the corresponding narrow-band estimate (brightest on top, faintest at the bottom). Exposure times (ET) are the same for the blue and red arms, and thus we do not give them individually. \*BR3 and BH-NB5-16 were also observed with Keck/DEIMOS. All the sources were confirmed as real LAEs at either  $z \sim 3.1$  or  $z \sim 2.2$ , although note that BH-NB392-55 does not have Ly  $\alpha$  coverage in the observational set-up and it was confirmed by the strong presence of C IV; see e.g. Fig. A1.

Target (ID)	Grism	ET (target) (ks)	ET (star) (ks)	Calib Star (SDSS ID)	Date (2016)	Seeing (")	Sky	Moon
BH-NB5-4 (BR3)*	R600	3.0	0.1	J143232.23+333903.7	5 May	0.9	Clear	Dark
BH-NB5-6	R300/R316	5.4	0.2	J143317.66+335348.9	1 Jul	0.8	Clear	Dark
BH-NB392-12	R600	3.0	0.1	J143026.70+333455.4	4 May	1.0	Clear	Dark
BH-NB5-16*	R300/R316	6.3	0.2	J143102.13+340329.1	30 Jun	0.6	Clear	Dark
GN-NB501-6712	R600	11.5	0.4	J123617.51+622445.9	4-5 May	1.0	Clear	Dark
BH-NB392-55	R600	4.0	0.1	J143043.56+340349.4	5 May	1.0	Clear	Dark

**Table A2.** Final spectral and median noise properties of the binned (typically binned by  $\sim 2$  pixels) spectra ( $\pm 100 \text{ \AA}$  of  $\lambda$  given and excluding strong OH lines) of all reduced data for WHT/ISIS, Keck/DEIMOS, and VLT/X-SHOOTER. The  $1\sigma$  noise levels are in units of  $\text{erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ . Note that in the case of X-SHOOTER, the noise in the NIR arm can vary significantly at the position of the strongest sky lines.

Telescope/ Instrument	Arm/ Grism	$\lambda$ (nm)	FWHM ( $\text{km s}^{-1}$ )	$1\sigma$ noise ( $\times 10^{-18}$ )
WHT/ISIS	R600B	400	150	15
WHT/ISIS	R600R	500	120	8
WHT/ISIS	R600R	700	77	4
WHT/ISIS	R300B	500	240	2
WHT/ISIS	R316R	700	154	1
Keck/DEIMOS	B 600L	500	180	0.5
Keck/DEIMOS	R 600L	700	130	0.3
VLT/XSHOOT	UVB	400	120	6
VLT/XSHOOT	UVB	500	96	3
VLT/XSHOOT	VIS	700	69	5
VLT/XSHOOT	NIR	1600	67	3

## A1 A PYTHON WHT/ISIS pipeline: WHIS

Here we present and explain in detail our PYTHON WHT/ISIS data reduction pipeline. On a night-by-night basis, the pipeline starts by identifying the different calibration and science frames, and groups them appropriately. Bias frames are median combined to produce a master bias. All frames are then bias subtracted. Flats within the linear regime are combined, per arm, grating, and central wavelength, after normalizing each flat by its median, and then median combined to produce different normalized master flats. All individual raw frames are then flat fielded using the appropriate normalized master flat. Arcs are median combined per arm, grating and central wavelength to produce master arcs. We combine arcs separately for the morning and evening arcs because we notice that small shifts can happen across the night.

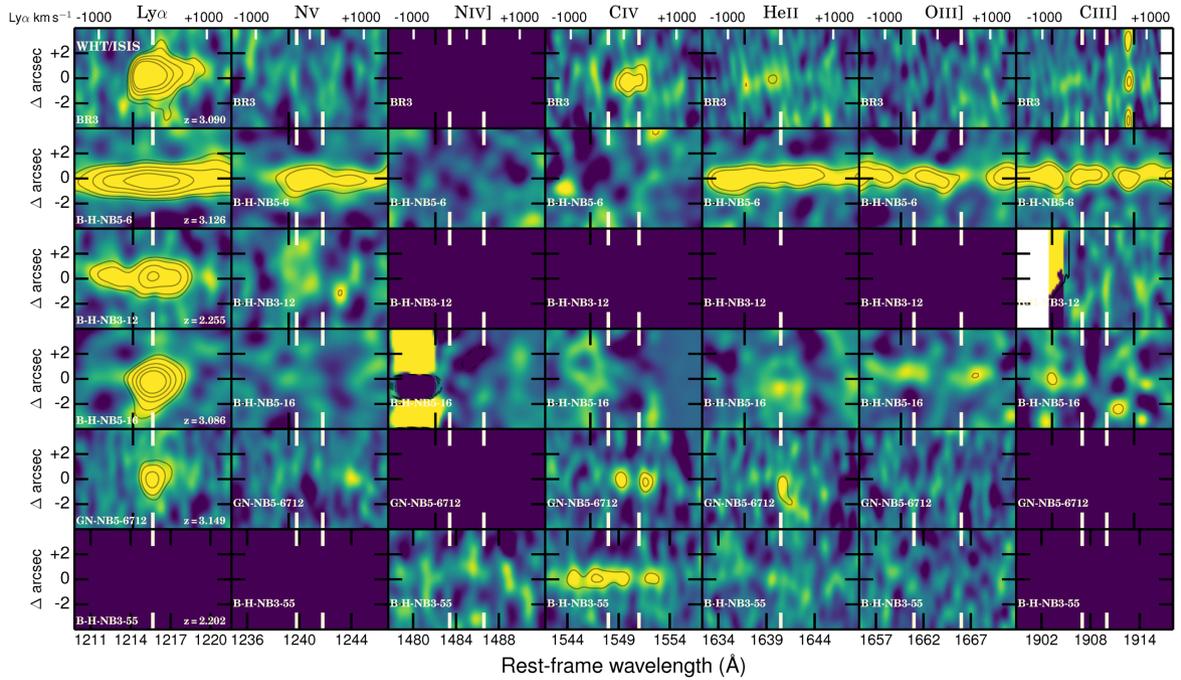
Master arcs are used to extract the 1D arc spectra per night, per grating, and central wavelength. Arcs lines are matched to air wavelength CU+CAr lines, matching  $\sim 30$  spectral lines covering the full spectral range per arm. We fit third-order polynomials and find the rms in the wavelength calibration solution to always be between 0.1 and 0.2  $\text{\AA}$  without any noticeable trend. We perform a final check

**Table A3.** Observing log for the different three observing nights with Keck DEIMOS (as part of program C267D; PI Darvish) for luminous  $z \sim 3$  LAEs in the Boötes and GOODS-N fields from Matthee et al. (2017b). Observing nights of 2016 June 3 and 2016 July 7 used a dither pattern of  $\pm 2.5$  arcsec, while for the 2016 July 30 run no dithering was applied. The three sources after the main eight targets were targeted as part of C267D out of the sample of all emission line candidates (ELCs) which were not Ly  $\alpha$  selected; one turned out to be a real LAE, while the other two are powerful quasars, selected due to significant transmission at  $\sim 5000 \text{ \AA}$ . We also provide information and the reduced spectra for these final three sources, but due to the different selection, we do not include them in the main analysis of Ly  $\alpha$  selected sources. The final two sources were targeted initially as potential bright Ly  $\alpha$  candidates, and then confirmed to be [O II] emitters at  $z \sim 0.3$ , but were not included in Matthee et al. (2017b) as LAE candidates; we list them here for completeness. Feige 66 was used as a flux calibrator along with a source with bright enough continuum, B-NB921-198 (Matthee et al. 2017b). \*\*For BH-NB5-34 the Ly  $\alpha$  coverage is partially missing along a spectral-spatial direction.

Target (ID)	Mask (name)	ET (targets) (ks)	Date (2016)	Seeing (")	Sky	Moon	Nature ( $z_{\text{spec}}$ and emission line)
BH-NB5-4 (BR3)*	btm-1	5.4	3 June	0.7	Clear	Dark	$z = 3.1$ LAE
BH-NB5-10	btm-1	5.4	3 Jun	0.7	Clear	Dark	$z = 3.1$ LAE
BH-NB5-16*	btm-5	7.2	30 Jul	0.6	Clear	Dark	$z = 3.1$ LAE
BH-NB5-27	btm-2	4.5	3 Jun	0.6	Clear	Dark	$z = 3.1$ LAE
BH-NB5-34**	btm-3	5.4	7 Jul	0.8	Clear	Dark	$z = 3.1$ LAE
BH-NB5-37	btm-1	5.4	3 Jun	0.7	Clear	Dark	$z = 3.1$ LAE
GN-NB501-3378	gdsn	7.2	3 Jun	1.0	Clear	Dark	$z = 3.2$ LAE
GN-NB501-5878	gdsn	7.2	3 Jun	1.0	Clear	Dark	$z = 3.1$ LAE
B-NB501-31679	btm-1	4.5	3 Jun	0.6	Clear	Dark	ELC $z \sim 3$ quasar (Ly $\alpha$ )
B-NB501-25550	btm-1	4.5	3 Jun	0.6	Clear	Dark	ELC $z \sim 4$ quasar (Ly $\alpha$ forest selected)
B-NB501-24390	btm-1	4.5	3 Jun	0.6	Clear	Dark	ELC $z = 3.1$ Ly $\alpha$
BH-NB5-9	btm-1	4.5	3 Jun	0.6	Clear	Dark	$z = 0.3$ [O II] emitter
BH-NB5-13	btm-1	4.5	7 Jul	0.6	Clear	Dark	$z = 0.3$ [O II] emitter

**Table A4.** Observing log for the VLT/X-SHOOTER programs (098.A-0819 and 099.A-0254, PIs: Sobral, Matthee) for luminous LAEs at  $z \sim 2.2-3.1$  from Sobral et al. (2017b) and Matthee et al. (2017b) in COSMOS, UDS, and SA22. Sources are ordered based on the predicted Ly $\alpha$  luminosity measured from the narrow-band data, if all were LAEs. Exposure times (ETs) are given separately for each X-SHOOTER arm/spectrograph as they vary slightly due to the different read-out times. The final column presents what each source was found to be and the approximate redshift using Ly $\alpha$ . Overall, out of the 11 targets, we confirm 9 as LAEs, and 2 as contaminants, with one being a strong [O II] emitter with no detectable continuum, and another one being a star with narrow and broad-band features which mimic those of some quasars.

Target (ID)	ET UVB (ks)	NIR (ks)	VIS (ks)	Calib. Star (name)	Dates	Seeing (")	Sky	Moon	Nature ( $z_{spec}$ )
SA22-NB5-14	2.4	2.6	1.6	Hip108612	28 Jun 2017	1.3	Clear	Grey	$z = 3.1$ Ly $\alpha$
SA22-NB5-18	2.4	2.6	1.6	Hip108612	28 Jun 2017	1.2	Clear	Grey	$z = 3.1$ Ly $\alpha$
SA22-NB5-10	2.4	2.6	1.6	F-110	19 Jun 2017	1.0	Clear	Grey	$z = 3.1$ Ly $\alpha$
CALYMHA-85	1.2	1.2	0.8	Hip033300	20 Jan 2017	1.3	Thin	Grey	$z = 0.0$ star
CALYMHA-383	0.6	0.6	0.4	Hip033300	25 Oct 2016	1.1	Clear	Dark	$z = 3.2$ Ly $\alpha$
CALYMHA-415	1.2	1.2	0.8	GD71	4 Oct 2016	1.3	Clear	Dark	$z = 2.2$ Ly $\alpha$
CALYMHA-438	1.2	1.2	0.8	Hip033300	25 Oct 2016	1.4	Clear	Dark	$z = 0.1$ [O II]
CALYMHA-373	12	12	8	F-110/EG274	2/22/30-31 Jul 2017	0.9	Clear	Dark	$z = 2.2$ Ly $\alpha$
CALYMHA-67	1.2	1.2	0.8	Hip033300	20 Jan 2017	0.9	Thin	Grey	$z = 2.2$ Ly $\alpha$
CALYMHA-147	12	12	8	Hip039540	20 Jan, 21-23 Dec 2017	0.7	Clear	Dark	$z = 2.2$ Ly $\alpha$
CALYMHA-95	12	12	8	Feige110	28 Dec 2017 12-13 Jan 2018	0.8	Clear	Dark	$z = 2.2$ Ly $\alpha$

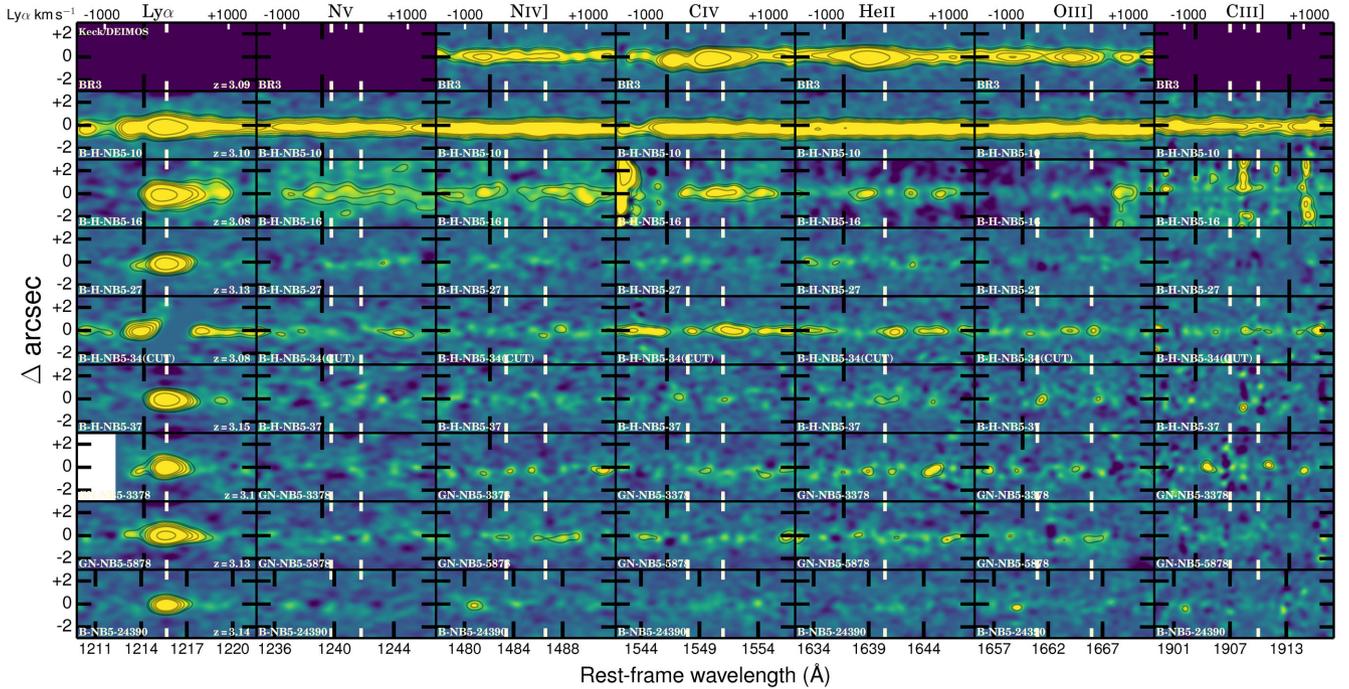


**Figure A1.** Reduced 2D spectra from our WHT/ISIS observations and zoom into the major emission lines studied here. We order sources according to their Ly $\alpha$  luminosity, from the more luminous at the top, to the faintest at the bottom. We use low and high cut-offs corresponding to  $-1\sigma$  and  $+2\sigma$ . We show contours corresponding to  $1.5\sigma$ ,  $2\sigma$ ,  $3\sigma$ ,  $4\sigma$ ,  $5\sigma$ ,  $10\sigma$  in all 2D zoom-ins. Velocity offsets are given with respect to Ly $\alpha$ . Windows of the 2D spectra in dark represent spectral regions without spectral coverage, where no information about the specific line is available for the specific instrument/source combination. Apart from significant Ly $\alpha$  detections, we find one source with strong Nv (BH-NB5-6), three sources with C IV, another three sources with He II and 1–2 sources with O III] or C III]. GN-NB5-6712 is particularly interesting, showing redshifted C IV emission (relative to Ly $\alpha$ ) and no He II offset; this could mean Ly $\alpha$  is escaping roughly at systemic, with C IV likely indicating that this is due to an outflow likely clearing the path to Ly $\alpha$  photons and potentially LyC photons (Sobral & Matthee 2018).

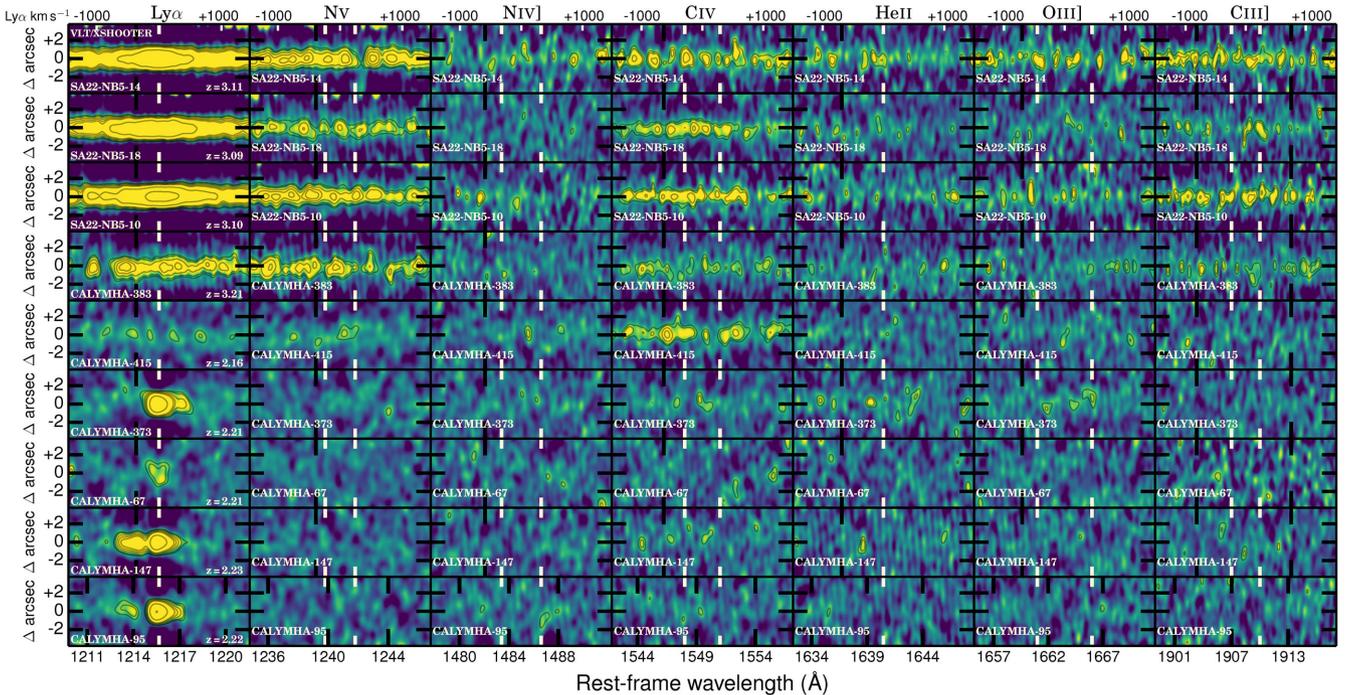
and, if needed, a small correction to the wavelength calibration on a source-by-source basis (typically  $<0.5 \text{ \AA}$ ) by exploring the available sky lines obtained from the same frames as the science targets, allowing us to trace any small changes.

For each science target, we start by using the corresponding acquisition star frames and bin the 2D spectrally and spatially in order to automatically find the trace of the star in each individual 2D spectra. We find that fitting a line to the trace is able to trace it without any offsets along the wavelength direction. We then use the

position of the trace in the spectral centre of the 2D as the centre of the trace for the target. We use those to find the exact location of the trace in the multiple offset spectra and to compute the required offsets in pixels. We use these to subtract both star and science spectra in sets of ABBA and to combine them. We then obtain an average stack (we also obtain median stacks and stacks without sky subtraction), removing individual pixels which are more than  $5\sigma$  away from the average counts, to obtain a final average stacked 2D of the science target and its appropriate star. We then automatically



**Figure A2.** Reduced 2D spectra from our DEIMOS observations and zoom in into the major emission lines studied here. Sources are ordered by decreasing Ly $\alpha$  luminosity (top to bottom). We use low and high cut-offs corresponding to  $-1\sigma$  and  $+2\sigma$ . We show contours corresponding to  $1.5\sigma$ ,  $2\sigma$ ,  $3\sigma$ ,  $4\sigma$ ,  $5\sigma$ ,  $10\sigma$ ,  $20\sigma$  in all 2D zoom-ins. Velocity offsets are given in respect to Ly $\alpha$ .



**Figure A3.** Reduced 2D spectra from our VLT/X-SHOOTER observations and zoom in into the major emission lines studied here. Sources are ordered by decreasing Ly $\alpha$  luminosity (top to bottom). We use low and high cut-offs corresponding to  $-1\sigma$  and  $+2\sigma$ . We show contours corresponding to  $1.5\sigma$ ,  $2\sigma$ ,  $3\sigma$ ,  $4\sigma$ ,  $5\sigma$ ,  $10\sigma$ ,  $20\sigma$  in all 2D zoom-ins. Velocity offsets are given in respect to Ly $\alpha$ . Note that there are significant negative regions (in both flux and signal-to-noise) directly above and below the continuum and emission line detections: this is a simple consequence of jittering along the slit, and a further check for the reliability of the detections (real detections require the negatives up and down). We find that the most luminous LAEs are dominated by broad-line AGN, with strong high ionization rest-frame UV lines, while such lines are much weaker or not present in the faintest among our sample, which show much narrower Ly $\alpha$  profiles.

find the trace in the star 2D spectrum and use it to extract both the science and the star spectra. We also extract the sky spectrum. We use the sky spectra to check the wavelength calibration and to apply any small wavelength offsets needed based on the calibrated ISIS spectra of OH lines. As a final sky subtraction step, we minimize residuals from sky lines by measuring the average counts away from the location of the source (where the average should be zero), and subtract those counts per resolution element. We show 2D spectra in Fig. A1.

We flux calibrate 1D spectra by using the calibration star and taking full advantage of our observation methodology (of always observing a calibration star prior to each science exposure). We use SDSS magnitudes to produce an artificial spectrum of each star. We then compute the ratio between total counts measured and the flux densities derived from SDSS for each calibration star. We do this after masking all positions of strong lines (e.g. Ca HK, H $\delta$ , H $\beta$ , H $\alpha$ ) and linearly interpolate between those wavelengths. We find the flux calibration to be accurate within  $\sim 10$  per cent. By definition, due to the way we extract our science targets (using the trace of the stars which we use for flux calibration) all point sources will be slit corrected. We use our narrow-band imaging to comment on the need for any extra corrections for some of our sources which are not point-like. Our spectra have a typical rms/depth of  $(1 - 2) \times 10^{-17}$  erg s $^{-1}$  for a 200 km s $^{-1}$  line (rest-frame) in both blue and red arms. We show examples of 1D extracted spectra in Fig. 4.

## A2 Reduced 2D and 1D spectra

We release the binned 1D spectra and we show in Figs A1, A3, and A2 all the 2D spectra for our LAEs, zoomed in at the location of the major emission lines studied in this paper. All velocity offsets shown on the top of each figure are in respect to Ly $\alpha$  (here we use the redshift determined by fitting a Gaussian to the Ly $\alpha$  line). Note that for some lines and for some instrument configurations there is no coverage.

## APPENDIX B: MEASURING LINES, FWHMS, AND VELOCITY OFFSETS

By fitting Gaussian line profiles, we use the width of the line ( $\sigma_{\lambda_0}$ ), fitted in  $\text{\AA}$  (rest-frame), and convert it to km s $^{-1}$  using

$$\sigma = c \frac{\sigma_{\lambda_0}}{\lambda_0} \text{ km s}^{-1} \quad (\text{B1})$$

where  $c$  is the speed of light, 299 792.458 km s $^{-1}$  and  $\lambda_0$  is the rest-frame wavelength:  $\lambda_0 = \lambda/(1+z)$ . The observed full width at half-maximum,  $\text{FW}_{\text{obs}}$ , is

$$\text{FW}_{\text{obs},0} = 2\sqrt{2 \times \ln(2)} \times \sigma \text{ km s}^{-1}. \quad (\text{B2})$$

We correct our  $\text{FW}_{\text{obs}}$  measurements for the appropriate instrument dispersion/resolution,<sup>11</sup>  $\text{FW}_{\text{inst}}$  (for Ly $\alpha$ : 100–200 km s $^{-1}$ )

using:

$$\text{FWHM} = \sqrt{\text{FW}_{\text{obs},0}^2 - \text{FW}_{\text{inst}}^2} \text{ km s}^{-1}. \quad (\text{B3})$$

For the Ly $\alpha$  line these lead to corrections of  $<100$  km s $^{-1}$  from observed to intrinsic FWHM due to the instrument dispersion. Our results are presented in Table 2 and in Section 4.5.

We use the best-fitting central wavelength of each line which is significantly detected to obtain a redshift per line ( $z_{\text{line}}$ ) and to compute a velocity offset with respect to Ly $\alpha$  by using

$$v_{\text{offset,obs}} = c \left( \frac{1 + z_{\text{Ly}\alpha}}{1 + z_{\text{line}}} - 1 \right) \text{ km s}^{-1} \quad (\text{B4})$$

We provide tables with the measurements presented and used in this work, e.g. Table B1 in a fits format with the final refereed paper, and also the binned spectra.

## APPENDIX C: CLOUDY MODELLING

We explore stellar-like ionizing spectra by using a large range of black bodies, with effective temperatures extending from 20 000 K to 160 000 K in steps of 1000 K. We explore a wide range of metallicities, from 0.01  $Z_{\odot}$  to 3  $Z_{\odot}$  in 0.05 steps in log space. We also vary the ionization parameter  $U$ , from 0.000 01 to 10 ( $-5$  to 2 in log in steps of 0.2), from lower to high ionization parameters. We follow known relations between how C/O varies with metallicity, but also run CLOUDY for different C/O ratios allowed by current data (see Nakajima et al. 2017). We use densities of 100, 300, and 1000 cm $^{-2}$ . Here we focus on the emerging emission line ratios. We note that our aim is not to claim ‘true’ physical conditions, but rather to broadly identify which physical conditions may be present. In our derived best physical conditions, we also take all the models within  $1\sigma$  of the observed line ratios, or that are allowed by the line ratio limits.

In order to compare with more realistic stellar models, we use BPASS (Eldridge & Stanway 2009, 2012; Stanway et al. 2016)<sup>12</sup> models that include stellar rotation and binary evolution. These allow us to better constrain realistic ionizing fluxes from stellar populations and to provide a link to indicative stellar ages. Furthermore, we also explore power-law ionizing sources, which we associate with AGNs, and run those models for the same physical conditions. For many emission line ratios the hottest blackbody ionizing sources become very similar to the result of a power-law ionizing source, but for the even higher ionization lines, and for appropriate choices of lines, there is still a good separation. We note nonetheless that naturally some of the line ratios used to separate stellar from an AGN nature that have used standard stellar populations do not work in the cases where stars can have much higher effective temperatures (due to binary interaction, winds exposing the deeper components of the star, or rotation).

<sup>11</sup> We estimate the instrument resolution by measuring non-blended arc or OH lines.

<sup>12</sup> See also Eldridge et al. (2017).

**Table B1.** Measurements and constraints on rest-frame UV emission line fluxes for the sources in our final spectroscopic sample of luminous LAEs. All emission line fluxes are given in  $10^{-17}$  erg s $^{-1}$  cm $^{-2}$ . Measurements and the upper and lower errors are derived from perturbing spectra 10 000 times on each individual spectral element and (re-)fitting Gaussian/double Gaussians and obtaining the median, 84th, and 16th percentiles, respectively. For sources in which the 0.6 percentile of all 10 000 realizations results in a zero or negative flux we assign the 99.4 percentile as the upper limit, roughly corresponding to a  $2.5\sigma$  limit. Fluxes do not include any slit correction. For some sources we either do not have coverage for a specific line and/or the 99.4 percentile flux limit is above  $\approx 10^{-15}$  erg s $^{-1}$  cm $^{-2}$  and thus we label those as ‘—’ due to being completely unconstrained.  $v_{\text{offset, obs}}$  values are measured using detected lines in relation to Ly  $\alpha$ , including those detected in the IR arm in XSHOOTER data for CALYMHA-147, 373, and 95 (based on [O III] and H  $\alpha$ ; Matthee et al., in preparation).

ID: Line fluxes	Ly $\alpha$ (erg s $^{-1}$ cm $^{-2}$ )	Nv	N IV	C IV	He II	O III]	N III]	C III]	$v_{\text{offset, obs}}$ (km s $^{-1}$ )
SA22-NB5-14	92.0 $^{+4.0}_{-4.1}$	—	<14	39.5 $^{+7.0}_{-6.7}$	<5.7	<13	<11	<42	640 $^{+340}_{-340}$
SA22-NB5-18	66.4 $^{+1.9}_{-1.8}$	21.3 $^{+4.5}_{-3.7}$	<8.6	44.0 $^{+9.7}_{-8.2}$	<28	<9.4	<25	8.8 $^{+3.1}_{-3.8}$	780 $^{+380}_{-320}$
SA22-NB5-10	86.3 $^{+2.7}_{-2.7}$	—	<7.6	35.9 $^{+6.8}_{-6.3}$	<20	<16	<36	17.0 $^{+4.0}_{-3.8}$	-90 $^{+230}_{-240}$
BH-NB5-4(BR3)	71.4 $^{+5.1}_{-5.0}$	<55	1.1 $^{+0.4}_{-0.2}$	14.3 $^{+0.5}_{-0.5}$	11.1 $^{+0.2}_{-0.2}$	1.5 $^{+0.6}_{-0.2}$	1.2 $^{+0.2}_{-0.2}$	24.0 $^{+3.6}_{-13.6}$	-20 $^{+120}_{-130}$
CALYMHA-383	55.2 $^{+6.3}_{-6.5}$	<80	<21	<67	<15	<14	<29	<46	—
BH-NB392-12	83.6 $^{+10.5}_{-10.0}$	<32	—	—	—	—	—	—	—
BH-NB5-6	13.2 $^{+1.2}_{-1.1}$	6.2 $^{+1.0}_{-1.1}$	—	<14	<12	<10	<9.9	25.1 $^{+6.0}_{-5.8}$	-90 $^{+400}_{-400}$
BH-NB5-10	12.7 $^{+0.2}_{-0.2}$	2.1 $^{+3.2}_{-0.3}$	<8.2	<1.2	2.6 $^{+0.3}_{-0.3}$	—	<0.7	<5.3	150 $^{+170}_{-500}$
BH-NB5-16	10.6 $^{+0.3}_{-0.3}$	6.6 $^{+0.7}_{-0.7}$	2.6 $^{+0.6}_{-1.5}$	<9.0	<0.6	<1.3	2.0 $^{+9.1}_{-1.2}$	<9.1	110 $^{+580}_{-180}$
GN-NB5-6712	26.6 $^{+4.7}_{-4.0}$	<65	—	9.2 $^{+2.1}_{-2.3}$	6.6 $^{+2.9}_{-2.3}$	<23	<17	—	20 $^{+160}_{-110}$
BH-NB392-55	—	—	<3.2	34.7 $^{+99.0}_{-8.6}$	<14	<15	—	—	—
BH-NB5-27	13.1 $^{+0.3}_{-0.3}$	1.3 $^{+0.4}_{-0.3}$	<13	<0.7	<4.1	<1.9	<2.8	<1.3	-160 $^{+230}_{-170}$
BH-NB5-37	10.7 $^{+0.2}_{-0.2}$	<1.8	<22	<3.8	0.7 $^{+0.3}_{-0.2}$	<3.7	<0.5	<3.2	-110 $^{+90}_{-100}$
BH-NB5-34	4.1 $^{+0.4}_{-0.4}$	<8.5	<0.5	8.2 $^{+0.6}_{-0.6}$	<3.3	<0.4	<1.7	—	460 $^{+170}_{-190}$
CALYMHA-147	43.0 $^{+2.9}_{-3.7}$	<18	<12	<7.8	<15	<23	<37	—	250 $^{+10}_{-10}$
CALYMHA-373	12.8 $^{+0.7}_{-0.7}$	<4.4	<2.6	<2.0	<2.2	<5.0	<55	<12	150 $^{+10}_{-10}$
CALYMHA-415	25.5 $^{+6.9}_{-7.2}$	12.7 $^{+4.5}_{-4.0}$	<17	15.8 $^{+3.6}_{-3.2}$	<6.3	<5.2	—	<59	-20 $^{+350}_{-350}$
CALYMHA-67	8.4 $^{+1.9}_{-1.7}$	<11	<15	<11	<5.7	<5.2	<39	<56	—
GN-NB5-3378	2.7 $^{+0.1}_{-0.1}$	<0.6	<1.0	<1.1	0.6 $^{+0.3}_{-0.2}$	<0.3	<1.9	<1.3	-250 $^{+760}_{-650}$
CALYMHA-95	15.6 $^{+0.8}_{-0.7}$	<11	<7.6	<7.6	<1.7	<6.0	<33	<10	30 $^{+70}_{-60}$
GN-NB5-5878	4.7 $^{+0.2}_{-0.2}$	<0.5	<0.8	<0.2	<0.3	<0.5	<1.0	<1.8	—
#Lines detected	20 (100 per cent)	6 (33 per cent)	2 (11 per cent)	8 (40 per cent)	5 (25 per cent)	1 (5 per cent)	2 (11 per cent)	4 (25 per cent)	—

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.