



**HAL**  
open science

## The X-SHOOTER Lyman $\alpha$ survey at $z = 2$ (XLS-z2) I: what makes a galaxy a Lyman $\alpha$ emitter?

Jorryt Matthee, David Sobral, Matthew Hayes, Gabriele Pezzulli, Max Gronke, Daniel Schaerer, Rohan P. Naidu, Huub Röttgering, João Calhau, Ana Paulino-Afonso, et al.

### ► To cite this version:

Jorryt Matthee, David Sobral, Matthew Hayes, Gabriele Pezzulli, Max Gronke, et al.. The X-SHOOTER Lyman  $\alpha$  survey at  $z = 2$  (XLS-z2) I: what makes a galaxy a Lyman  $\alpha$  emitter?. Monthly Notices of the Royal Astronomical Society, 2021, 505, pp.1382-1412. 10.1093/mnras/stab1304 . insu-03672418

**HAL Id: insu-03672418**

**<https://insu.hal.science/insu-03672418>**

Submitted on 19 May 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The X-SHOOTER Lyman $\alpha$ survey at $z = 2$ (XLS- $z2$ ) I: what makes a galaxy a Lyman $\alpha$ emitter?

Jorryt Matthee<sup>1</sup>,<sup>\*</sup>† David Sobral,<sup>2</sup> Matthew Hayes<sup>3</sup>, Gabriele Pezzulli<sup>4</sup>, Max Gronke<sup>5</sup>, Daniel Schaerer,<sup>6,7</sup> Rohan P. Naidu<sup>8</sup>, Huub Röttgering,<sup>9</sup> João Calhau,<sup>10,11</sup> Ana Paulino-Afonso<sup>12</sup>, Sérgio Santos<sup>2</sup> and Ricardo Amorín<sup>13,14</sup>

<sup>1</sup>Department of Physics, ETH Zürich, Wolfgang-Pauli-Strasse 27, CH-8093 Zürich, Switzerland

<sup>2</sup>Department of Physics, Lancaster University, Lancaster LA1 4YB, UK

<sup>3</sup>Stockholm University, Department of Astronomy and Oskar Klein Centre for Cosmoparticle Physics, AlbaNova University Centre, SE-10691 Stockholm, Sweden

<sup>4</sup>Kapteyn Astronomical Institute, University of Groningen, Landleven 12, NL-9747 AD Groningen, the Netherlands

<sup>5</sup>Department of Physics & Astronomy, Johns Hopkins University, Bloomberg Center, 3400 N. Charles Street, Baltimore, MD 21218, USA

<sup>6</sup>Observatoire de Genève, Université Genève, 51 Ch. des Maillettes, CH-1290 Versoix, Switzerland

<sup>7</sup>Université de Toulouse, UPS-OMP, IRAP, IRAP/CNRS, 14, Av. E. Belin, 31400 Toulouse, France

<sup>8</sup>Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

<sup>9</sup>Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands

<sup>10</sup>Instituto de Astrofísica de Canarias, E-38200 La Laguna, Tenerife, Spain

<sup>11</sup>Departamento de Astrofísica, Universidad de La Laguna, E-38205 La Laguna, Tenerife, Spain

<sup>12</sup>Centro de Astrofísica e Gravitação – CENTRA, Departamento de Física, Instituto Superior Técnico – IST, Universidade de Lisboa – UL, Av. Rovisco Pais 1, P-1049-001 Lisboa, Portugal

<sup>13</sup>Instituto de Investigación Multidisciplinar en Ciencia y Tecnología, Universidad de La Serena, Raúl Bitrán 1305, 1700000 La Serena, Chile

<sup>14</sup>Departamento de Astronomía, Universidad de La Serena, Av. Juan Cisternas 1200 Norte, 1720236 La Serena, Chile

Accepted 2021 May 1. Received 2021 April 23; in original form 2021 February 15

## ABSTRACT

We present the first results from the X-SHOOTER Lyman  $\alpha$  survey at  $z = 2$  (XLS- $z2$ ). XLS- $z2$  is a deep spectroscopic survey of 35 Lyman  $\alpha$  emitters (LAEs) utilizing  $\approx 90$  h of exposure time with Very Large Telescope/X-SHOOTER and covers rest-frame Ly  $\alpha$  to H  $\alpha$  emission with  $R \approx 4000$ . We present the sample selection, the observations, and the data reduction. Systemic redshifts are measured from rest-frame optical lines for 33/35 sources. In the stacked spectrum, our LAEs are characterized by an interstellar medium with little dust, a low metallicity, and a high ionization state. The ionizing sources are young hot stars that power strong emission lines in the optical and high-ionization lines in the ultraviolet (UV). The LAEs exhibit clumpy UV morphologies and have outflowing kinematics with blueshifted Si II absorption, a broad [O III] component, and a red-skewed Ly  $\alpha$  line. Typically, 30 per cent of the Ly  $\alpha$  photons escape, of which one quarter on the blue side of the systemic velocity. A fraction of Ly  $\alpha$  photons escape directly at the systemic suggesting clear channels enabling an  $\approx 10$  per cent escape of ionizing photons, consistent with an inference based on Mg II. A combination of a low effective H I column density, a low dust content, and young starburst determines whether a star-forming galaxy is observed as an LAE. The first is possibly related to outflows and/or a fortunate viewing angle, while we find that the latter two in LAEs are typical for their stellar mass of  $10^9 M_{\odot}$ .

**Key words:** galaxies: formation – galaxies: ISM – galaxies: starburst – dark ages, reionization, first stars.

## 1 INTRODUCTION

In the last two decades, the Lyman  $\alpha$  emission line (Ly  $\alpha$ ;  $\lambda_0 = 1215.67$  Å) has fulfilled its long-standing promise (Partridge & Peebles 1967) of being a powerful tool to study galaxies in the early Universe (e.g. Rhoads et al. 2000; Gawiser et al. 2007; Ouchi et al. 2008; Hayes et al. 2010; Kashikawa et al. 2011; Matthee et al. 2015; Konno et al. 2016; Drake et al. 2017; Zheng et al. 2017; Taylor et al.

2020). Although perhaps not as bright as intrinsically expected (e.g. Charlot & Fall 1993; Hayes 2015), its high equivalent width (EW), its rest-frame ultraviolet (UV) wavelength, the adjacent continuum breaks in the spectrum, and the peculiar line shape have made Ly  $\alpha$  an extremely useful emission line to find and spectroscopically identify the redshifts of galaxies out to the most distant Universe (e.g. Finkelstein et al. 2013; Oesch et al. 2015; Zitrin et al. 2015).

A key uncertainty in the study of Ly  $\alpha$  emission from galaxies is the Ly  $\alpha$  escape fraction,  $f_{\text{esc,Ly}\alpha}$ , and how this depends on properties of the interstellar medium (ISM). The Ly  $\alpha$  transition has a high scattering cross-section and thus is resonant (see Dijkstra 2014, for a review). This means that only small amounts of neutral hydrogen

\* E-mail: [mattheej@phys.ethz.ch](mailto:mattheej@phys.ethz.ch)

† Zwicky Fellow.

in the ISM are needed to cause Ly  $\alpha$  photons to scatter significantly. Scattering increases the effective path-length and the likelihood of absorption by dust while also leading to a diffusion in frequency and space (Neufeld 1990; Mas-Ribas et al. 2017). How this exactly happens depends on gas turbulence, the column density distribution and clumpiness, the velocity field, and the dust content in a complex way (e.g. Verhamme, Schaerer & Maselli 2006; Gronke & Dijkstra 2016; Gronke et al. 2017).

The same characteristics that make Ly  $\alpha$  observations so attractive at high redshift also make it challenging to determine the physical properties of Ly  $\alpha$  emitters (LAEs) in great detail. First, the high EWs are often accompanied by a faint UV continuum and some LAEs are not even detected in the deepest imaging that exists (e.g. Maseda et al. 2018). Secondly, the well-understood strong rest-frame optical emission lines such as [O III]<sub>5008</sub> and H  $\alpha$  are currently difficult or even impossible to observe at  $z > 3$  due to the atmospheric emission and background in the infrared. This redshift coincides with the redshift where the majority of LAEs are found with Ly  $\alpha$  being redshifted into the optical. Therefore, many open questions remain. What is the typical  $f_{\text{esc,Ly}\alpha}$ ? Why are not all distant star-forming galaxies observed as LAEs (e.g. Hayes et al. 2010; Hagen et al. 2016; Matthee et al. 2016)? What makes a galaxy an LAE?

Pioneering studies have shown that  $f_{\text{esc,Ly}\alpha}$  is  $\approx 30$  per cent in LAEs (Blanc et al. 2011; Nakajima et al. 2012; Song et al. 2014; Trainor et al. 2016; Sobral et al. 2017), but where these photons escape in the spectral and spatial domain has been poorly explored. Samples of star-forming galaxies at  $z \approx 2$  that are selected irrespective of their Ly  $\alpha$  luminosity have much lower  $f_{\text{esc,Ly}\alpha}$  in the range of 1–5 per cent (Hayes et al. 2010; Matthee et al. 2016). Furthermore,  $f_{\text{esc,Ly}\alpha}$  has been found to correlate with nebular dust attenuation (Atek et al. 2008; Blanc et al. 2011; Yang et al. 2017), but additional independent correlations with other properties such as the H I column density, gas-phase metallicity, and possibly viewing angle have been reported as well (Shibuya et al. 2014; Henry et al. 2015; Trainor et al. 2016; Yang et al. 2017). It is likely that several processes impact  $f_{\text{esc,Ly}\alpha}$ , but it is yet to be explored in detail (e.g. Runholm et al. 2020) how these vary with mass and redshift.

Besides  $f_{\text{esc,Ly}\alpha}$ , the Ly  $\alpha$  output of a galaxy is also determined by the production rate of Ly  $\alpha$  photons, which is tightly linked to the production rate of ionizing photons. The intrinsic Ly  $\alpha$  EW is related to the production rate of ionizing photons relative to the UV continuum and therefore sensitive to the spectrum of the ionizing sources (e.g. Charlot & Fall 1993; Raiter, Schaerer & Fosbury 2010). In particular, an intrinsically high Ly  $\alpha$  EW is an indicator of galaxies with very young and extremely low-metallicity stars (e.g. Charlot & Fall 1993; Raiter et al. 2010; Sobral et al. 2015; Maseda et al. 2020). A high EW could also be indicative of additional sources of Ly  $\alpha$  emission such as fluorescence in the proximity of a luminous ionizing source (such as a quasar; e.g. Cantalupo, Lilly & Haehnelt 2012; Marino et al. 2018) or collisional excitation from gravitational collapse (e.g. Dijkstra, Haiman & Spaans 2006). In this work, we assume that the main production mechanism of Ly  $\alpha$  photons is recombination associated with star formation within galaxies.

LAEs at intermediate redshift are of interest as they may be good and practically useful analogues to the galaxies responsible for reionization. This is because the galaxies that are known to have the highest Lyman Continuum (LyC;  $\lambda_0 < 912 \text{ \AA}$ ) escape fractions are strong LAEs (e.g. Izotov et al. 2018) and because the Ly  $\alpha$  EW is observed to correlate with the LyC escape fraction (Marchi et al. 2018; Steidel et al. 2018). It is thus plausible that galaxies that are significantly leaking LyC photons can be found more easily in samples of LAEs compared to general galaxy samples.

The LyC escape fraction is a key parameter for understanding the sources of cosmic reionization (Robertson et al. 2013; Naidu et al. 2020). It is, however, very challenging to measure for individual systems at  $z > 3$  due to the stochastic opacity of the intergalactic medium (IGM; Madau 1995; Inoue et al. 2014; Vanzella et al. 2018). Such measurements are possible in low-redshift analogues of distant galaxies (e.g. Izotov et al. 2018; Jaskot et al. 2019), but this requires challenging UV spectroscopy and selection functions are typically complicated. It is furthermore unclear whether the star formation histories (SFHs) of local analogues truly resemble galaxies in the early Universe (Amorín et al. 2012a). The Ly  $\alpha$  escape fraction and line shape that emerges from the ISM are correlated with the escape fraction of LyC photons (Verhamme et al. 2015; Dijkstra, Gronke & Venkatesan 2016; Izotov et al. 2018; Gazagnes et al. 2020), but are much simpler to measure for larger samples and over a range in cosmic times. Measurements of the Ly  $\alpha$  profile over a range of galaxy properties and cosmic times are therefore a promising avenue for mapping the contribution of various galaxies to the epoch of reionization (Matthee et al. 2018) and currently  $z \approx 2$  is the highest redshift where it is possible to control for the intrinsic Ly  $\alpha$  production from the ground.

Due to its sensitivity to intervening neutral hydrogen, the evolution of the Ly  $\alpha$  luminosity function (e.g. Konno et al. 2018) and the observed distributions of Ly  $\alpha$  EWs among galaxies are used as a tracer of the evolution of the neutral fraction into the epoch of reionization (e.g. Stark et al. 2010; Jung et al. 2018; Mason et al. 2018). The impact of the (neutral) IGM, however, depends on the specific velocity at which Ly  $\alpha$  photons escape the ISM (e.g. Dijkstra et al. 2014), and the EW may also vary due to evolution in the intrinsic Ly  $\alpha$  luminosity. Therefore, to fully interpret the results from such surveys in the context of an evolving IGM, we require a complete understanding of the variation of the Ly  $\alpha$  line profile and the production and escape of Ly  $\alpha$  photons among galaxies. In particular, it is crucial to map out how these variations are dependent on galaxy properties that are not affected by the evolution of the IGM, such as rest-frame optical line strengths.

To make progress on these aspects, we have undertaken a large narrow-band survey of LAEs at  $z \approx 2$  (Sobral et al. 2017), which is currently the only redshift where it is possible to measure all the important lines in the wavelength range between Ly  $\alpha$  and H  $\alpha$  with ground-based facilities. Here, we present the first results of the X-SHOOTER Lyman  $\alpha$  Survey at redshift  $z = 2$  (XLS- $z2$ ), which constitutes the spectroscopic component of this programme. The sample is composed of 35 objects, of which 20 newly observed and 15 with archival data. This survey optimally uses the large wavelength coverage ( $\lambda = 0.3\text{--}2.5 \mu\text{m}$ ) of the X-SHOOTER instrument (Vernet et al. 2011) on the Very Large Telescope (VLT), meaning that we observe all emission lines from Ly  $\alpha$  to H  $\alpha$  simultaneously. Moreover, the spectral resolution of the Ly  $\alpha$  observations of our set-up ( $R \sim 4000$ ) is significantly higher than most Ly  $\alpha$  studies at  $z > 2$  (e.g. Kulas et al. 2012; Trainor et al. 2015; Verhamme et al. 2018; Hayes et al. 2021) and the non-resonant lines in the rest-frame optical allow stringent estimates of the systemic redshift that most high-redshift studies lack. These data allow us to connect faint spectroscopic features in the rest-frame UV to the well-known optical lines. With these data, we thus simultaneously measure the intrinsic Ly  $\alpha$  budget, the attenuation and various escape mechanisms as the lowest H I column density paths and related ionization parameter, and the presence of outflows and their velocities.

In this paper, we present the selection of the sample and discuss how representative this sample is at  $z \approx 2$  (Section 2). The observations and data reduction are detailed in Section 3. We present

our methods for extracting aperture-matched one-dimensional (1D) spectra from individual objects and stacks in Section 4.1 and how we measure systemic redshifts for the majority of the sample (Section 4.2). In order to address which properties make galaxies LAEs, we focus on a stack of LAEs that are representative for the population of LAEs at  $z = 2.2$ . In Section 5, we present measurements of various absorption and emission lines from the UV to the rest-frame optical. These are used to determine the nature of the ionizing sources, the Ly  $\alpha$  escape fraction, the star formation rate, and various properties of the ISM (Section 6). In Section 7, we discuss what these results imply for the nature of LAEs, in particular what determines whether galaxies are observed as LAEs and what these results imply for galaxies in the epoch of reionization. Section 8 summarizes our results.

We use a flat  $\Lambda$ CDM cosmology with  $\Omega_M = 0.3$  and  $H_0 = 70$  km s $^{-1}$  Mpc $^{-1}$  and a Chabrier (2003) initial mass function (IMF). For solar abundances, we use the reference values  $Z_\star = 0.0142$  and  $12 + \log_{10}(\text{O}/\text{H}) = 8.7$ . Emission-line wavelengths are presented as vacuum wavelengths. All EWs are in the rest frame. Magnitudes are in the AB system.

## 2 SAMPLE

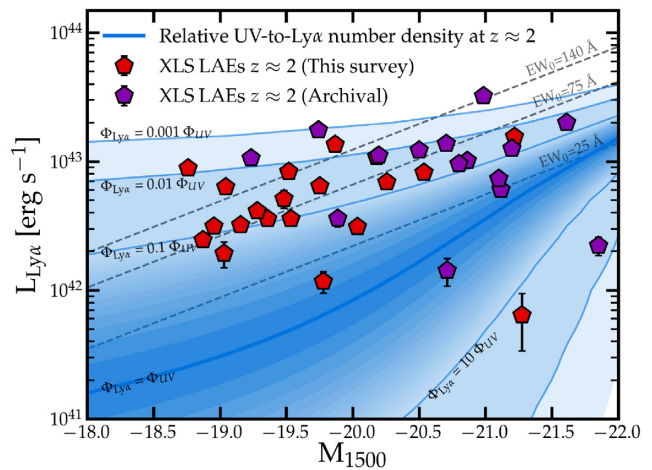
### 2.1 Selection criteria

Our full sample consists of 35 Ly  $\alpha$  flux-limited selected galaxies at redshifts  $z = 2.00$ – $2.47$  (with a median redshift  $z = 2.22$ ) that have mostly been pre-selected with narrow-band imaging and that have been observed by the X-SHOOTER spectrograph.

The main selection criterion for our sample is that targets are known Ly  $\alpha$  emitters at  $z \approx 2$  with Ly  $\alpha$  EW $_0 > 25$  Å. This redshift is high enough for observed Ly  $\alpha$  photons to be shifted beyond the atmospheric transmission cut-off in the UV, but low enough for H  $\alpha$  to avoid the high thermal background. At this particular redshift, the strong rest-frame optical H  $\alpha$ , [O III], H  $\beta$ , and [O II] lines all lie in regions with high atmospheric transmission (e.g. Nakajima et al. 2012; Sobral et al. 2013; Khostovan et al. 2015).

The majority (25/35) of our targets are directly selected based on strong Ly  $\alpha$  emission identified in narrow-band imaging spanning a combined volume of  $\approx 5 \times 10^6$  cMpc $^3$  in well-studied extragalactic fields as COSMOS, UDS, and CFHTLS-W4. This selection technique effectively implies a Ly  $\alpha$  flux and EW limit. The combination of typical narrow-band filter widths and a  $3\sigma$  excess significance typically means imposing an EW limit of  $> 25$  Å (e.g. Gronwall et al. 2007; Ouchi et al. 2008; Matthee et al. 2017b; Sobral et al. 2017). However, by prioritizing the spectroscopic follow-up to objects with a somewhat higher Ly  $\alpha$  flux ( $\gtrsim 3 \times 10^{-17}$  erg s $^{-1}$  cm $^{-2}$ ; on average  $2 \times 10^{-16}$  erg s $^{-1}$  cm $^{-2}$ ) we are effectively slightly skewed towards higher EWs at fainter UV luminosities (Fig. 1). A few targets have Ly  $\alpha$  EWs  $\lesssim 10$  Å as those were identified with a narrower filter (Matthee et al. 2016) or their initial EW measurement was overestimated. Where possible, we initially removed objects from the parent sample that are identified as being powered by active galactic nucleus (AGN) through their radio or X-ray emission (Calhau et al. 2020), or broad [full width at half-maximum (FWHM)  $> 1000$  km s $^{-1}$ ] emission lines (Sobral et al. 2018b). We note that we only performed shallower spectroscopy for a small subsample prior to this programme, meaning that the broad-line rejection could not be performed homogeneously and that due to the limiting sensitivity of the X-ray data faint AGNs could have been missed.

The sample comprises 20 targets from our own programme (ESO programme ID 102.A-0652; PI: Matthee) and 15 targets for which the



**Figure 1.** The UV and Ly  $\alpha$  luminosities of the XLS galaxies. Targets that have been observed as part of our own survey are shown in red and archival targets are shown in purple. Grey dashed lines illustrate lines at constant Ly  $\alpha$  EW assuming a fixed UV slope  $\beta = -2$ . The blue line illustrates the  $L_{\text{Ly}\alpha}$ – $M_{1500}$  relation that is the intercept along which galaxies have the same number densities per log luminosity interval according to their UV (Parsa et al. 2016) and Ly  $\alpha$  luminosity (Sobral et al. 2017) at  $z \approx 2$ . We also show the relations required for various relative number densities. While XLS LAEs are representative for the Ly  $\alpha$ -selected galaxy population, they are a rare (typically 1:15) subsample of the general UV-selected galaxy population at  $z \approx 2$ .

archival X-SHOOTER data were publicly available. The coordinates, programme IDs, and observation details of the targets are listed in Table 1. Targets were assigned an XLS-ID as follows: XLS-1 to XLS-20 originate from our main survey, while XLS-21 to 35 are archival objects. Within these two groups, the IDs are ranked by right ascension. A comparison of the XLS-ID to other IDs of these galaxies is listed in Table A1.

The main 20 targets are selected from a wide-field narrow-band survey (Sobral et al. 2017). We preferentially selected galaxies in regions with best ancillary data [e.g. *Hubble Space Telescope* (HST) coverage from the CANDELS programme]. The archival objects are selected in various ways, listed in Table 1 and described in more detail here:

(i) XLS-21, 22, and 23 were initially selected to have detections in the rest-frame UV emission lines C III] and O III], in addition to strong Ly  $\alpha$  (Amorín et al. 2017). It is unclear whether the additional selection criterion of a C III] and O III] detection necessarily implies that these objects are not representative of LAEs with similar Ly  $\alpha$  luminosity and EW. We note that these objects are all also detected by an independent Ly  $\alpha$  narrow-band survey (Sobral et al. 2018a), where they were found to be part of a larger sample with similar luminosity and EW.

(ii) XLS-24, 25, 26, 27, and 28 are identified as Ly  $\alpha$  emitters in narrow-band surveys by Nilsson et al. (2009), Nakajima et al. (2012), and Hayes et al. (2010), and were chosen for spectroscopic follow-up observations based on their high luminosity compared to the other LAEs in these respective surveys. The luminosity and EW ranges of these objects are comparable to the main 20 targets, indicating these are representative LAEs. XLS-24 is also part of our own parent catalogue and has been studied in Rhoads et al. (2014). We note that XLS-26 and 27 have additionally been observed by a programme that selected them as a candidate LyC leaker (Naidu et al. 2017; Oesch et al. 2018).

**Table 1.** Observation log. Exposure times are in ks. The spectral resolution around the Ly  $\alpha$  wavelength is based on the nominal resolution of  $R = 6700, 5400,$  and  $4100$  for slit widths  $0.8, 1.0,$  and  $1.3$  arcsec, respectively. The listed seeing is the median seeing in arcsec. Selection methods: (1) Ly  $\alpha$  flux, (2) strong rest-frame UV emission-line galaxy, (3) LyC leaker candidate, (4) BX-galaxy with H  $\alpha$  EW  $> 300 \text{ \AA}$ , and (5) high [O III]/H  $\beta$  and symmetric Ly  $\alpha$  emission at low resolution.

ID	RA (J2000)	Dec. (J2000)	$R_{\lambda=\text{Ly}\alpha}$	$t_{\text{exp,UVB}}$	$t_{\text{exp,VIS}}$	$t_{\text{exp,NIR}}$	Seeing	Programme ID	Selection
XLS-1	09:57:59.73	+02:18:04.86	4100	6.1	5.4	6.4	0.8	102.A-0652	1
XLS-2	10:00:13.91	+01:39:24.30	4100	16.8	12.8	15.0	0.7	098.A-0819, 099.A-0254, 102.A-0652	1
XLS-3	10:00:24.61	+02:27:01.07	4100	10.7	9.3	11.2	0.5	102.A-0652	1
XLS-4	10:00:26.65	+02:17:14.42	4100	10.7	9.3	11.2	0.6	102.A-0652	1
XLS-5	10:00:33.97	+02:13:15.92	4100	6.1	5.4	6.4	0.5	102.A-0652	1
XLS-6	10:00:35.73	+02:15:06.66	4100	10.7	9.3	11.2	0.7	102.A-0652	1
XLS-7	10:00:38.66	+02:09:20.72	4100	13.4	11.6	14.0	0.5	102.A-0652	1
XLS-8	10:00:42.21	+02:08:09.62	4100	8.0	7.0	8.4	0.5	102.A-0652	1
XLS-9	10:00:50.66	+02:07:42.06	4100	10.7	9.3	11.2	0.6	102.A-0652	1
XLS-10	10:00:50.87	+02:06:31.24	4100	10.7	9.3	11.2	0.7	102.A-0652	1
XLS-11	10:01:06.55	+01:45:45.47	4100	14.5	11.1	15.4	0.7	099.A-0254, 102.A-0652	1
XLS-12	10:01:36.21	+02:15:16.80	4100	11.9	10.1	12.5	0.8	098.A-0819, 102.A-0652	1
XLS-13	10:02:16.16	+02:32:18.87	4100	6.1	5.4	6.4	0.6	102.A-0652	1
XLS-14	10:02:35.32	+02:12:13.42	4100	13.6	13.2	15.1	0.7	0100.A-0213, 102.A-0652	1
XLS-15	02:17:15.52	-05:07:14.97	4100	10.7	9.3	11.2	0.7	102.A-0652	1
XLS-16	02:17:26.42	-05:13:40.95	4100	10.7	9.3	11.2	0.8	102.A-0652	1
XLS-17	02:17:41.39	-05:06:49.61	4100	10.7	9.3	11.2	0.5	102.A-0652	1
XLS-18	02:17:46.13	-05:02:55.51	4100	21.0	15.2	25.0	0.8	098.A-0819, 099.A-0254, 102.A-0652	1
XLS-19	02:17:55.77	-05:12:41.00	4100	10.7	9.3	11.2	0.8	102.A-0652	1
XLS-20	22:15:48.23	+00:23:57.46	4100	9.4	8.1	9.8	0.9	102.A-0652	1
XLS-21	10:00:10.95	+01:51:46.66	5400	10.6	11.1	10.8	0.6	0101.B-0779	2
XLS-22	10:00:39.56	+02:15:38.44	5400	10.6	11.1	10.8	0.6	0101.B-0779	2
XLS-23	10:01:20.81	+02:36:19.27	5400	7.3	7.2	7.6	0.6	0101.B-0779	2
XLS-24	10:00:49.22	+02:01:21.30	4100	3.6	3.6	3.6	0.9	084.A-0303	1
XLS-25	03:32:32.31	-28:00:52.20	5400	2.4	4.4	4.8	0.9	088.A-0672	1
XLS-26	03:32:35.48	-27:46:16.91	5400	17.3	13.3	14.4	0.7	092.A-0774, 099.A-0758	1, 3
XLS-27	03:32:46.46	-27:50:36.63	5400	10.6	10.0	10.8	0.8	099.A-0758	3
XLS-28	03:32:49.34	-27:59:52.35	5400	4.8	8.8	9.6	1.3	088.A-0672	1
XLS-29	23:46:09.06	12:47:56.00	6700	4.4	4.0	4.6	0.7	091.A-0413	4
XLS-30	23:46:18.57	12:47:47.38	6700	4.4	4.0	4.6	0.8	091.A-0413	4
XLS-31	23:46:29.43	12:49:45.54	6700	8.8	8.0	9.3	0.6	091.A-0413	4
XLS-32	02:09:49.21	-00:05:31.67	5400	2.7	2.8	3.6	0.7	097.A-0153	5
XLS-33	02:09:44.23	-00:04:13.51	5400	7.2	7.4	7.2	0.8	097.A-0153	5
XLS-34	02:09:43.15	-00:05:50.21	5400	6.3	6.5	6.3	0.9	097.A-0153	5
XLS-35	01:45:16.87	-09:46:03.47	5400	7.2	7.4	7.2	0.7	097.A-0153	5

(iii) XLS-29, 30, and 31 are selected through their high H  $\alpha$  EW ( $> 300 \text{ \AA}$ ; Terlevich et al. 2015) with respect to other objects in a catalogue of UV-selected galaxies at  $z \approx 2$  (Erb et al. 2006b). While the selection of these objects is to first order independent of the Ly  $\alpha$  emission properties, the fact that these objects have detectable Ly  $\alpha$  emission is perhaps not surprising (see also Erb et al. 2010). The H  $\alpha$  EW criterion implies that these galaxies produce significant amounts of Ly  $\alpha$  photons and the fact that the parent sample is UV selected implies that there is little dust attenuation, which is associated with a higher escape fraction of Ly  $\alpha$  photons (e.g. Atek et al. 2008).

(iv) XLS-32, 33, 34, and 35 are selected based on having strong and symmetric Ly  $\alpha$  lines in low-resolution rest-frame UV spectra (Erb et al. 2016) and initially originate from a sample of UV-selected galaxies.

Fig. 1 shows that the archival objects are typically more luminous. Several archival objects that were not directly selected as LAEs show Ly  $\alpha$  emission lines with relatively low EW. We include these in order to expand the dynamic range of EWs and Ly  $\alpha$  escape fractions.

## 2.2 How representative is the sample of the star-forming population at $z \approx 2$ ?

When defining a survey, it is crucial to first investigate how representative the galaxy sample is compared to the full galaxy sample at a similar cosmic time. Here, we address this by comparing the number densities of LAEs to the number densities of all UV-selected star-forming galaxies at  $z \approx 2$ .

As shown in Fig. 1, the Ly  $\alpha$  luminosities span the range  $(0.7\text{--}40) \times 10^{42}$  erg s $^{-1}$  [ $\approx(0.2\text{--}10) \times L^*$ ; Sobral et al. 2018a] and the UV magnitudes range from  $-18.5$  to  $-22$  [ $\approx(0.2\text{--}6) \times L^*$ ; Parsa et al. 2016]. Diagonal lines of fixed Ly  $\alpha$  EW (assuming a constant UV slope  $\beta = -2$ ) illustrate that the typical EW of our sample is  $\approx 75$  Å. When we estimate the number density associated with the Ly  $\alpha$  luminosity of each target, we find that the median number density of our targets is  $\approx 1 \times 10^{-4}$  comoving Mpc $^{-3}$  per log luminosity interval. This is a factor of 5 lower than the median number density of UV-selected galaxies associated with their UV luminosities. Fig. 1 also directly illustrates how representative our targets are in terms of their relative UV and Ly  $\alpha$  luminosities compared to the UV-selected galaxy population at  $z \approx 2$  based on their relative abundances. We derive the relation between Ly  $\alpha$  and UV luminosity that is required to match the number densities from the UV luminosity function from Parsa et al. (2016) to the Ly  $\alpha$  luminosity function from Sobral et al. (2017) at  $z \approx 2.2$ . This method assumes a linear relation between the UV and Ly  $\alpha$  luminosity. The correction factor for a sublinear relation would be in the range 1–2 for a (conservative) slope range 0.5–1.0. Even a factor of 2 would be a small correction in this context and leave our conclusions basically unchanged. The number densities are similar for objects with EWs  $\approx 10\text{--}15$  Å. For the EWs of the objects in our sample, the number density is on average 15 times lower than that of the UV-selected population. This implies that while our sample is by selection broadly representative of the population of LAEs at  $z \approx 2$ , this sample is a rare subsample of the UV-selected galaxy population at  $z \approx 2$ . This is in agreement with typical population-averaged (Hayes et al. 2011; Sobral et al. 2018a) and mass-averaged (Matthee et al. 2016) Ly  $\alpha$  escape fractions of  $\approx 2$  per cent at  $z \approx 2$ , while Ly  $\alpha$  escape fractions measured in LAEs are typically  $\approx 30$  per cent (see e.g. Hayes et al. 2010; Song et al. 2014; Trainor et al. 2016; Sobral et al. 2017 and Section 5.3).

### 2.3 Fields, photometry, and ancillary data

In this subsection, we summarize the available photometry and ancillary data of the targets. The photometry is used to derive UV luminosities, colours, and stellar masses as described in Section 2.4.

(i) XLS-1 to 14 and XLS-21 to 24 are located in the COSMOS field and are therefore covered by  $>30$  bands of near-UV to mid-IR photometry (e.g. Ilbert et al. 2009; Laigle et al. 2016). We use the aperture-corrected photometry that is described in Santos et al. (2020), which includes the latest data release (DR4) of the UltraVISTA survey in the near-infrared (McCracken et al. 2012). High-resolution *HST*/ACS imaging is available for all targets in the F814W filter (Koekemoer et al. 2007). We compile size measurements of the COSMOS targets based on these data from Paulino-Afonso et al. (2018).

(ii) XLS-15 to 19 are in the UDS/SXDS field, which is covered by very deep ground-based imaging in the optical (Furusawa et al. 2008), near-infrared (Lawrence et al. 2007; Jarvis et al. 2013), and mid-infrared (Mehta et al. 2018). We use the multiwavelength aperture-corrected photometry from Mehta et al. (2018). XLS-16 and XLS-19 are covered by the *HST* CANDELS survey, from which we compile their size measurements (van der Wel et al. 2012).

(iii) XLS-20 is located in the SA22/CFHTLS-W4 field and is covered by the deep part of the Subaru HyperSuprimeCam survey (Aihara et al. 2019). These data are about 2 mag shallower than the data in the COSMOS field. We use our own aperture-corrected photometry in the  $g$ ,  $r$ ,  $i$ ,  $z$ , and  $y$  filters with the same technique as described in detail in Santos et al. (2020).

(iv) XLS-25 to 28 targets are in the Extended Chandra Deep Field-South field. XLS-25 and 28 have high-resolution *HST*/ACS imaging from the GEMS survey (Rix et al. 2004). We compile size measurements from Häussler et al. (2007) and we use the multiwavelength ground-based photometry in  $>30$  filters from the MUSYC survey (Cardamone et al. 2010). XLS-26 and 27 are located in the *HST* extreme deep field with multiwavelength photometry from the CANDELS survey (Grogin et al. 2011; Guo et al. 2013) and we use size measurements from van der Wel et al. (2012).

(v) XLS-29 to 35 are located in extragalactic fields that were selected to have a bright background QSO (Steidel et al. 2004). For XLS-29 to 31, we collected Palomar photometry in the  $U$ ,  $G$ ,  $R$ ,  $J$ , and  $K_s$  filters from Erb et al. (2006a). We have no photometry for XLS-32 to 35, but we instead use the X-SHOOTER spectra directly to measure the UV continuum luminosity. We collected and reduced archival high-resolution *HST* imaging data for XLS-29 to 32 and XLS-35 from programmes with IDs 9133 (PI Falco), 9367 (PI Hazard), 11694 (PI Law), and 12471 (PI Erb) using the *HST* Legacy Archive.

### 2.4 Spectral energy distribution (SED) modelling

In order to obtain the rest-frame UV luminosity ( $M_{1500}$ ) and the stellar mass ( $M_{\text{star}}$ ), we model the SEDs of the galaxies using the MAGPHYS code (da Cunha, Charlot & Elbaz 2008). We use aperture-corrected photometry obtained with the same methodology as described in detail in Santos et al. (2020). For most sources, the photometric information is quite homogeneous in terms of wavelength coverage and depth (0.3–5.0  $\mu\text{m}$ ,  $\approx 28\text{--}25$  AB magnitude, with higher sensitivity in bluer bands). The COSMOS and ECDFS objects additionally benefit from several filters with medium bandwidth. For XLS-20 and XLS-29 to 31, the coverage is shallower and limited to the optical (0.3–1.0  $\mu\text{m}$ ,  $\approx 26$  AB).

In short, MAGPHYS uses dust attenuation models from Charlot & Fall (2000) and stellar populations using Bruzual & Charlot (2003) models where a Chabrier (2003) IMF with a mass range of 0.1–100  $M_{\odot}$  is assumed. The SFHs are a combination of a continuous exponentially decaying history that follows an initial rise and an additional instantaneous burst (with a duration between 30 and 300 Myr and a mass fraction of 0.1–100 of the integrated mass from the continuous SFH). As MAGPHYS does not model nebular emission, we exclude the medium and broad-band filters that are contaminated by strong Ly  $\alpha$ , H  $\beta$  + [O III], and H  $\alpha$  emission from the fitting procedure as they may lead to overestimated stellar masses (e.g. Schaerer & de Barros 2009). Except for XLS-20, 29, 30, and 31, all objects are covered by the deep *Spitzer*/IRAC data, which are particularly useful for constraining the stellar masses. For XLS-32 to 35, we use the stellar masses derived by Erb et al. (2016) and we measure the UV continuum luminosity directly from the X-SHOOTER spectrum. Independently from the SED fitting, the UV slope  $\beta$  is measured by fitting a power law of the form  $f_{\lambda} \propto \lambda^{\beta}$  to all photometric bands that cover rest-frame wavelengths from 1300 to 2100 Å. The measurements are listed in Table 2.

### 2.5 Rest-frame UV morphology

High-resolution *HST* data are available for 31 objects and we show cut-out images centred on the objects in Figs A2 and A3. For the majority of targets (XLS-1 to 14 and 21 to 24), the ACS/F814W data consist of a single orbit, but XLS-16, 19, 26, 27, 29, and 30 have deeper data (sources of these data are listed above). XLS-25 and 28 have data in a similar filter (F850LP) at moderate depth. For XLS-31

**Table 2.** General properties of the galaxies in the XLS sample. The half-light radius is determined from rest-frame UV imaging with *HST*/ACS when available.

ID	$z_{\text{sys}}$	$M_{1500}$	$\beta$	$\log_{10}(M_{\text{star}}/M_{\odot})$	$\text{SFR}_{\text{H}\alpha}/M_{\odot} \text{ yr}^{-1}$	$r_{1/2}/\text{kpc}$	$L_{\text{Ly}\alpha}/10^{42} \text{ erg s}^{-1}$	$\text{EW}_{\text{Ly}\alpha}/\text{\AA}$	$\text{EW}_{\text{H}\beta + [\text{O III}]}/\text{\AA}$
XLS-1 <sup>a</sup>	2.1961	-20.3	0.3	$10.7 \pm 0.1$	$9.3^{+85.6}_{-7.5}$	0.7	$6.9 \pm 0.7$	$98^{+10}_{-10}$	$155^{+81}_{-76}$
XLS-2	2.2296	-19.9	-2.1	$9.3 \pm 0.2$	$7.0^{+13.0}_{-2.2}$	0.8	$13.5 \pm 0.7$	$182^{+13}_{-13}$	$967^{+104}_{-103}$
XLS-3	2.2225	-19.0	-2.1	$9.6 \pm 0.3$	$1.0^{+1.5}_{-0.3}$	0.5	$3.1 \pm 0.2$	$96^{+10}_{-9}$	$184^{+13}_{-13}$
XLS-4	2.2279	-19.4	-1.9	$8.7 \pm 0.1$	$2.0^{+7.4}_{-1.3}$	1.2	$3.6 \pm 0.4$	$76^{+12}_{-11}$	$688^{+531}_{-480}$
XLS-5	2.2293	-20.5	-2.0	$9.3 \pm 0.2$	$10.3^{+30.2}_{-5.8}$	1.2	$8.2 \pm 0.5$	$63^{+4}_{-4}$	$422^{+33}_{-33}$
XLS-6	2.2218	-19.0	-2.5	$8.6 \pm 0.3$	$2.4^{+1.3}_{-0.8}$	0.6	$6.3 \pm 0.5$	$179^{+16}_{-15}$	$1391^{+164}_{-167}$
XLS-7	2.2229	-19.8	-0.9	$10.1 \pm 0.1$	$6.0^{+15.8}_{-4.1}$	1.4	$1.2 \pm 0.2$	$16^{+4}_{-3}$	$43^{+10}_{-9}$
XLS-8	2.0670	-21.3	-1.3	$9.8 \pm 0.1$	$4.8^{+5.3}_{-1.6}$	1.3	$0.1 \pm 0.3$	$1^{+1}_{-1}$	–
XLS-9	2.212 <sup>b</sup>	-19.5	-1.8	$8.9 \pm 0.3$	–	0.8	$5.1 \pm 0.8$	$95^{+15}_{-16}$	–
XLS-10	2.2158	-19.3	-1.9	$9.0 \pm 0.3$	$29.6^{+379.6}_{-27.9}$	1.0	$4.1 \pm 0.5$	$96^{+12}_{-13}$	$555^{+88}_{-91}$
XLS-11	2.2172	-19.5	-2.2	$8.8 \pm 0.3$	$3.9^{+7.7}_{-1.5}$	1.0	$8.3 \pm 0.3$	$152^{+11}_{-9}$	$1554^{+144}_{-143}$
XLS-12	2.2064	-19.8	-1.6	$9.9 \pm 0.1$	$26.4^{+18.7}_{-10.7}$	0.9	$6.4 \pm 0.5$	$99^{+11}_{-9}$	$733^{+38}_{-37}$
XLS-13	2.234 <sup>b</sup>	-19.5	-2.3	$8.8 \pm 0.1$	–	0.6	$3.6 \pm 0.3$	$64^{+7}_{-6}$	–
XLS-14	2.1418	-19.0	-0.2	$9.4 \pm 0.3$	$1.2^{+1.8}_{-0.5}$	–	$1.9 \pm 0.4$	$64^{+15}_{-15}$	$180^{+129}_{-125}$
XLS-15	2.2302	-19.2	-2.8	$8.6 \pm 0.1$	$0.9^{+0.5}_{-0.4}$	–	$3.2 \pm 0.3$	$82^{+8}_{-9}$	$1100^{+66}_{-64}$
XLS-16	2.2098	-20.0	-1.9	$8.8 \pm 0.1$	$3.0^{+1.2}_{-0.5}$	1.5	$3.1 \pm 0.4$	$39^{+5}_{-5}$	$461^{+35}_{-33}$
XLS-17	2.2015	-20.2	-2.3	$8.7 \pm 0.2$	$2.6^{+0.3}_{-0.3}$	–	$10.9 \pm 0.4$	$100^{+3}_{-3}$	$466^{+59}_{-62}$
XLS-18	2.2095	-21.2	-2.2	$9.4 \pm 0.1$	$25.2^{+18.7}_{-9.5}$	–	$15.6 \pm 0.5$	$62^{+2}_{-2}$	$862^{+28}_{-28}$
XLS-19	2.2186	-18.9	-2.3	$8.3 \pm 0.1$	$2.0^{+4.4}_{-1.0}$	1.1	$2.5 \pm 0.3$	$89^{+11}_{-11}$	$665^{+454}_{-440}$
XLS-20	2.2210	-18.8	-1.2	$7.9 \pm 1.2$	$4.6^{+42.4}_{-2.7}$	–	$8.9 \pm 0.5$	$294^{+104}_{-60}$	$3212^{+3469}_{-1426}$
XLS-21	2.4197	-20.7	-1.9	$9.2 \pm 0.5$	$8.6^{+5.7}_{-1.9}$	1.0	$13.8 \pm 0.4$	$88^{+6}_{-6}$	$2455^{+262}_{-221}$
XLS-22	2.4518	-20.2	-2.0	$9.4 \pm 0.1$	$1.8^{+0.8}_{-0.4}$	0.6	$11.2 \pm 0.3$	$111^{+4}_{-4}$	$1594^{+61}_{-62}$
XLS-23	2.4706	-21.0	-2.1	$9.7 \pm 0.1$	$16.2^{+13.6}_{-5.1}$	1.0	$32.3 \pm 0.4$	$156^{+3}_{-3}$	$1440^{+70}_{-75}$
XLS-24	2.2463	-19.7	-2.6	$8.8 \pm 0.1$	$5.7^{+24.8}_{-2.2}$	0.9	$17.6 \pm 0.3$	$261^{+7}_{-7}$	$994^{+193}_{-202}$
XLS-25	2.1721	-21.6	-2.1	$9.7 \pm 0.1$	$10.3^{+4.3}_{-3.5}$	0.8	$20.0 \pm 0.5$	$57^{+1}_{-1}$	$745^{+15}_{-17}$
XLS-26	2.1723	-19.2	-2.6	$8.6 \pm 0.1$	$11.5^{+12.3}_{-4.2}$	0.8	$10.6 \pm 0.3$	$209^{+7}_{-7}$	$3420^{+232}_{-204}$
XLS-27	1.9981	-19.9	-1.6	$8.8 \pm 0.1$	$9.7^{+1.1}_{-1.2}$	1.4	$3.6 \pm 0.5$	$42^{+6c}_{-5}$	$2108^{+170}_{-169}$
XLS-28	2.2051	-20.9	-1.7	$9.5 \pm 0.1$	$7.4^{+4.7}_{-1.1}$	4.1	$10.2 \pm 0.7$	$51^{+4}_{-4}$	$462^{+151}_{-124}$
XLS-29	2.3282	-21.9	-1.6	$9.9 \pm 0.3$	–	–	$1.7 \pm 0.4$	$4^{+1}_{-1}$	$49^{+3}_{-3}$
XLS-30	2.3051	-21.1	-2.2	$9.8 \pm 0.2$	–	–	$6.1 \pm 0.3$	$26^{+1d}_{-1}$	$392^{+35}_{-23}$
XLS-31	2.1737	-20.7	-2.3	$9.2 \pm 0.2$	–	–	$1.4 \pm 0.3$	$9^{+2d}_{-2}$	$392^{+98}_{-88}$
XLS-32	2.1682	-21.1	–	9.4	$36.6^{+52.0}_{-19.8}$	–	$7.3 \pm 0.8$	$32^{+4}_{-4}$	–
XLS-33	2.1922	-20.8	–	10.3	$14.9^{+6.4}_{-3.5}$	–	$9.6 \pm 0.4$	$60^{+5}_{-4}$	–
XLS-34	2.1885	-20.5	–	9.7	$3.6^{+4.4}_{-1.5}$	–	$12.2 \pm 0.6$	$127^{+20}_{-15}$	–
XLS-35	2.3567	-21.2	–	9.3	$20.4^{+10.3}_{-3.5}$	–	$12.6 \pm 0.5$	$40^{+2}_{-2}$	–

<sup>a</sup>XLS-1 is identified as an AGN.

<sup>b</sup>For XLS-9 and XLS-13, we list the redshifts of the red peak of the Ly  $\alpha$  line as we do not detect a non-resonant emission line in their spectra.

<sup>c</sup>We caution the interpretation of the Ly  $\alpha$  EW of XLS-27 as its Ly  $\alpha$  line is spatially offset from the continuum that has been used to estimate the EW by  $\approx 9$  kpc.

<sup>d</sup>A comparison to the measurements from Erb et al. (2016) suggests that it is plausible that a significant fraction of the Ly  $\alpha$  flux for XLS-30 and 31 is missing (i.e. a factor of 2–4, respectively) due to the use of a very narrow slit and a possible misalignment.

and 32, the only available *HST* data have been taken with WFC3 in the F140W and F160W NIR filters. The *HST* data of XLS-35 consist of F814W imaging with WFPC2.

The rest-frame UV morphologies of several LAEs show multiple clumps on  $\approx 1$  kpc scales. From simple visual inspection (see also Paulino-Afonso et al. 2018), we find that 12 out of the 31 objects with high-resolution *HST* data appear as multiple clumps; see Table 3.

### 3 VLT/X-SHOOTER SPECTROSCOPY

In this section, we describe the observations and data reduction of the X-SHOOTER data.

X-SHOOTER (Vernet et al. 2011) is a wide-band (0.3–2.5  $\mu\text{m}$ ) echelle spectrograph on the VLT. Two dichroics split the light into three arms, each with independent shutter and slit mask and with simultaneous exposures. These UVB, VIS, and NIR arms are optimized for their respective wavelength coverages of 300–560 nm (UVB), 560–1024 nm (VIS), and 1025–2480 nm (NIR). The exposures in the UVB and VIS arms are read out sequentially, meaning that in practice exposure times are longer in one of these arms compared to the other, depending on the observing strategy.

In the following sections, we describe the observing strategy, characteristics, and the data reduction. The data consist of typically 3 h per source, for a total of 89.5 h of on-source integration time (of

**Table 3.** The line-profile measurements of the fits to the [O III]<sub>4960, 5008</sub> doublet. Lines are fixed to the intrinsic 1:2.98 flux ratio. The FWHMs are listed in km s<sup>-1</sup> and not corrected for instrumental dispersion of 50 km s<sup>-1</sup>. In case two narrow components are identified,  $z_{\text{sys},1}$  is the redshift of the [O III] line that is spatially most closely associated with the peak of the Ly  $\alpha$  line. In the last column, we list the results of visually inspecting the *HST* images for the presence of a clumpy structure.

ID	$z_{\text{sys},1}$	$z_{\text{sys},2}$	$z_{\text{broad}}$	FWHM <sub>sys,1</sub>	FWHM <sub>sys,2</sub>	FWHM <sub>broad</sub>	$f_{\text{narrow},2}/f_{\text{tot}}$	$f_{\text{broad}}/f_{\text{tot}}$	Multiple clump <i>HST</i>
XLS-1	2.1961	–	–	596	–	–	–	–	N
XLS-2	2.2296	–	2.2295	69	–	227	–	0.31	N
XLS-3	2.2225	–	–	192	–	–	–	–	N
XLS-4	2.2279	–	–	207	–	–	–	–	N
XLS-5	2.2293	2.2283	–	94	75	–	0.17	–	Y
XLS-6	2.2218	–	–	98	–	–	–	–	N
XLS-7	2.2229	–	–	105	–	–	–	–	Y
XLS-8	2.0670	–	–	161	–	–	–	–	Y
XLS-9	–	–	–	–	–	–	–	–	N
XLS-10	2.2158	–	–	124	–	–	–	–	N
XLS-11	2.2172	2.2169	–	58	154	–	0.66	–	Y
XLS-12	2.2064	2.2078	–	137	106	–	0.54	–	Y
XLS-13	–	–	–	–	–	–	–	–	N
XLS-14	2.1418	–	–	70	–	–	–	–	N
XLS-15	2.2302	–	–	80	–	–	–	–	N
XLS-16	2.2098	2.2117	–	101	94	–	0.57	–	N
XLS-17	2.2015	–	–	169	–	–	–	–	–
XLS-18	2.2095	–	2.2095	113	–	374	–	0.40	–
XLS-19	2.2186	–	–	99	–	–	–	–	N
XLS-20	2.2210	–	2.2210	94	–	196	–	0.43	–
XLS-21	2.4197	–	2.4199	161	–	377	–	0.61	N
XLS-22	2.4518	–	2.4516	97	–	246	–	0.80	N
XLS-23	2.4706	–	2.4704	109	–	246	–	0.22	N
XLS-24	2.2463	–	2.2465	84	–	422	–	0.20	N
XLS-25	2.1721	2.1729	2.1725	113	66	397	0.15	0.60	Y
XLS-26	2.1723	2.1717	–	103	125	–	0.34	–	Y
XLS-27	1.9981	–	–	140	–	–	–	–	Y
XLS-28	2.2051	–	–	128	–	–	–	–	N
XLS-29	2.3282	–	2.3282	158	–	717	–	0.51	Y
XLS-30	2.3051	–	2.3052	102	–	309	–	0.47	Y
XLS-31	2.1737	–	2.1737	95	–	246	–	0.50	Y
XLS-32	2.1682	–	2.1681	106	–	293	–	0.28	N
XLS-33	2.1922	–	2.1923	105	–	235	–	0.48	–
XLS-34	2.1885	–	2.1886	175	–	254	–	0.84	–
XLS-35	2.3567	2.3578	2.3568	91	190	422	0.59	0.18	Y

which 62 h are from our own programme) in the UVB arm and similar times in the other arms. The typical spectral resolution around Ly  $\alpha$  is  $R = 4100\text{--}5400$ . The majority of the data reduction was performed with the standard ESO pipeline complemented with the MOLECFIT tool (Smette et al. 2015) to account for atmospheric transmission. We used our own PYTHON-based algorithms for optimally combining two-dimensional (2D) spectra observed over multiple dates on to a common barycentric velocity grid centred on the spatial peak of the Ly  $\alpha$  line.

### 3.1 Observations

Observations of our own programme were performed in service mode between 2018 October 2 and 2020 February 28. Archival data were taken between 2010 March and 2019 March. In general observations were performed in dark conditions and with *V*-band seeing of  $\approx 0.8$  arcsec. The nominal spectral resolution based on the slit width at the redshifted Ly  $\alpha$  wavelength, the total integration times in the various arms, and the typical seeing of all observations are listed in Table 1. All observations use a blind offset from an acquisition star and are nodding between two positions along the slit in an ABBA sequence.

For our own programme, we identified acquisition stars with  $R < 17$  by matching the parent catalogues to the *Gaia* DR2 catalogue (Gaia Collaboration 2018). We then selected the star within 120 arcsec from the target with the lowest proper motion and not blended with another object in projection. Offsets were calculated based on the relative position of the star and our targets in the narrow-band data, taking the small proper motion between the time of the narrow-band observation and the ESO semester into account. A distance of 3 arcsec between the two nodding positions was used. The slits are placed at the parallactic angle at the start of the first exposure.

We used 1.3, 1.2, and 0.9 arcsec slits in the UVB, VIS, and NIR arms corresponding to resolutions  $R = 4100, 6500,$  and  $5600$ , respectively. Individual exposure times were 670 s (UVB), 580 s (VIS), and  $4 \times 175$  s (NIR; using four integrations at each nodding position), which were repeated in cycles of 4 per observing block (OB) that lasted roughly 1 h. Typically targets were observed in three independent 1 h OBs (exact number of OBs ranging from 2 to 5). Most archival programmes used a similar observing strategy with slight variations in exposure times and slit widths. XLS-29, 30, and 31 were observed with a *K*-band blocking filter such that there is no coverage of H  $\alpha$  emission.



When the programme started, the majority of our narrow-band-selected targets were not yet spectroscopically confirmed, leading to the non-zero risk that they were interlopers or spurious sources. Therefore, while our observations were performed remotely in service mode, we specifically designed the execution strategy such that each target would only be observed for a maximum of 1 h during the first attempt. Remaining OBs were scheduled with the time constraint that they would follow at least 3 d later. This allowed us to reduce and analyse the data and communicate any target change in case that would be necessary. In practice, we changed target only twice. XLS-20 is a replacement (and therefore has less total exposure time) for a target that turned out to be a star with colours similar to a blue galaxy at  $z \sim 2$  and where variability boosted the narrow band mimicking an emission line. We also decided to move one OB from XLS-8, which turned out to have a very low Ly  $\alpha$  EW, to XLS-7.

For both our own programme and the archival data, we visually inspected each raw exposure in the UVB arm for any issues with the data. We removed a handful of exposures with poor acquisition, bad seeing, contamination by spurious light from within the telescope, or the laser from the AO system on UT4. The exposure times listed in Table 1 only include the data that have eventually been used.

### 3.2 Data reduction

The X-SHOOTER data are reduced as follows. For each OB of  $\sim 1$  h, we use the X-SHOOTER pipeline version 3.2.0 (Modigliani et al. 2010) implemented in ESOREX to apply the standard reduction steps: bias (UVB and VIS) or dark (NIR) subtraction, flat-fielding, flexure correction and 2D mapping, wavelength calibration, and flux calibration with standard stars. The same reduction steps are applied to telluric stars that are observed in the same nights. The MOLECFIT tool (Kausch et al. 2015; Smette et al. 2015) is used to apply telluric corrections to the science observations.

Individual OBs are co-added as follows. We first resample the 2D spectra to a new grid where we converted the wavelength calibration of the 2D spectra to vacuum wavelengths using the IAU standard and shifted each spectrum to the barycentric reference frame. To improve the accuracy of the re-sampling, the new 2D spectrum is oversampled by a factor of 2 using a nearest interpolation. We then shift the 2D spectra in all arms such that the spatial centre coincides with the spatial peak position of the Ly  $\alpha$  line. The peak position is identified by fitting a 2D Gaussian model to the 2D Ly  $\alpha$  spectrum after this has been convolved with a 2D Gaussian ( $\sigma_{\text{spatial}} = 0.32$  arcsec,  $\sigma_{\lambda} = 2$  Å) in order to improve the S/N and wash out the detailed spectral structure of the Ly  $\alpha$  profile. In most cases, the spatial peak of the Ly  $\alpha$  and UV continuum emission are found to be slightly off-centre due to slight inaccuracies in the acquisition and pointing of the telescope of the order of 0.2 arcsec (mean absolute deviation). In the case of XLS IDs 16, 25, and 27, we find that the Ly  $\alpha$  line is further offset by  $\approx 0.2$ –0.3 arcsec from the UV continuum and nebular lines. For XLS-27, we find an offset of 1.1 arcsec between Ly  $\alpha$  and the UV. Finally, we combine the OBs with an inverse-variance weighted average where the variance is determined over the 400–500 nm (UVB), 600–800 nm (VIS), and 1500–1600 nm (NIR) wavelength ranges.

## 4 METHODS

### 4.1 Extraction of 1D spectra

Here, we describe how 1D spectra were extracted from the 2D data. We first motivate the choice for the centre and width of the extraction

and then describe the way we optimize the spectrophotometric calibration and how we measure the noise level of the data.

#### 4.1.1 Centroid and aperture

The centre of the extraction is based on the peak position of the Ly  $\alpha$  line that we identified in the co-addition step in the data reduction (Section 3.2). We extract the 1D spectra using an optimal extraction (Horne 1986) assuming a Gaussian profile with a width that is optimized for each object individually. We collapse each 2D spectrum over a velocity range of  $-500$  to  $+500$  km s $^{-1}$  from the peak of the Ly  $\alpha$  line and we measure the FWHM of the Ly  $\alpha$ -light distribution. We repeat this process for a collapse of rest-frame wavelengths  $\lambda_0 = 1260$ –1500 Å to identify the FWHM of the UV continuum-light distribution. In this paper, we choose to use the continuum-based FWHM to extract the 1D profiles. These FWHM are much larger than the typical offsets between the Ly  $\alpha$  and the UV. For three objects (XLS-9, 14, and 22), we use a Ly  $\alpha$ -based aperture as the continuum is not detected with sufficient S/N. Typical FWHM of the continuum is 0.6 arcsec, ranging from 0.4 to 0.9 arcsec. The typical FWHM of the Ly  $\alpha$  line is a factor of 1.1 larger than that of the UV continuum. As described in detail below, the extraction aperture varies with wavelength in order to fix the encapsulated fraction of the flux.

For the majority of objects, there are no large shifts between the spatial peak of the UV continuum and Ly  $\alpha$ . Most of the objects with offsets appear as multiple component systems in either the UV continuum imaging or through multiple narrow components in the [O III] emission line, e.g. XLS-16, 25, and 35. The offsets between Ly  $\alpha$  and the UV are sufficiently small that the extraction windows centred on the Ly  $\alpha$  peak capture the large majority ( $> 80$  per cent) of the flux. XLS-27 is a special case where Ly  $\alpha$  is offset by  $\approx 1.1$  arcsec ( $\approx 9$  kpc) from the UV continuum (and the rest-frame optical lines). For this object, we therefore extract the Ly  $\alpha$  spectrum on the Ly  $\alpha$  position and the UV continuum and rest-frame optical spectrum on the position of the UV continuum. We note that we identify a spatial drift of the UV continuum across the slit in the UVB and VIS arms in the observations of XLS-29 to XLS-35. This is accounted for by increasing the extraction aperture by a factor of 3 at the expense of adding some noise.

As the seeing is wavelength dependent, using the same extraction size over the full UVB to NIR wavelength range would result in a higher enclosed flux in redder wavelengths compared to bluer wavelengths. We use the standard stars that have been observed with a very wide 5 arcsec slit and seeing conditions in the range of the observations to empirically obtain spectroscopy that encapsulated the same fraction of total flux over the full wavelength range. We measure the FWHM of the light distribution in the 2D spectra of the standard stars and store these in various wavelengths. Then, for each science object, we match the FWHM in the UVB arm to the closest standard star in terms of FWHM. We then match the extraction FWHM in the redder part of the spectrum to encapsulate the same fraction of the total flux and use this wavelength-dependent FWHM for our optimal Gaussian extraction.

#### 4.1.2 Spectrophotometric calibration

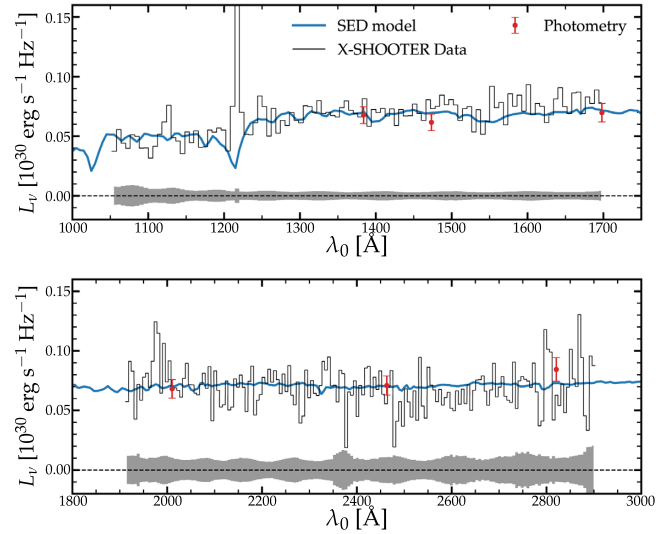
After the extraction, we optimize the spectrophotometric calibration by applying an achromatic normalization correction. The average flux in the wavelength range  $\lambda_0 = 1280$ –1500 Å that is measured in the extracted 1D spectra is thus matched to the average flux over

the same wavelength region in the SED model that is best fitted to the aperture-corrected multiwavelength photometry (Section 2.4). This final calibration step simultaneously accounts for slit losses and uncertainties in the flux calibration. The correction derived in the UV continuum is applied to the full wavelength range from UVB to NIR. The SED model fits the various photometric bands in the rest-frame UV wavelength range (observed  $B$  to  $z$  band) very well. Propagated uncertainties in the flux calibration of our spectrum that would originate from the validity of the SED model would be more important in the rest-frame optical (observed NIR). In the observed NIR, the majority of the photometric bands are contaminated by emission lines, such that the model is more dependent on choices regarding, for example, the SFH. However, as we do not use the NIR data to optimize the flux calibration, these concerns are not relevant here as long as the uncertainties in the flux calibration are achromatic.

On average, we find that the flux normalization of the spectrum is a factor of  $1.2 \pm 0.4$  higher than the photometry (where the error is the standard deviation and the extremes are 0.5 and 2.5). This suggests that uncertainties in the flux calibration dominate over slit losses. Indeed, because the sources are very compact and the seeing is typically good, we estimate slit losses  $<10$  per cent by simulating fake sources with the FWHM of the UV continuum and by measuring the fraction of the flux that is retrieved in the slit. The wavelength-collapsed UV continuum is detected with an  $S/N > 5$  in the spectra of all objects except XLS-14 and 22. We do not have aperture-corrected photometry for XLS-32 to 35. For these objects, we do not apply a correction to the flux calibration of the spectrum.

By comparing how the average fluxes vary between single OBs, we can empirically estimate the uncertainties associated with the acquisition and flux calibration. For a few sources, we are able to measure the continuum flux levels in the UVB (here we collapse  $\lambda = 400\text{--}530$  nm) and in the VIS arms (collapsing  $\lambda = 600\text{--}830$  nm) with an  $S/N > 10$  in single OBs. For these sources, we find a standard deviation of 9 per cent in the flux levels in the UVB arm and 13 per cent in the continuum level in the VIS arm. For sources where we detect the continua in individual OBs with an  $S/N$  ranging from 5 to 10, we find a typical standard deviation of 25 per cent, but it is plausible that this additional variation can be explained by measurement errors. This suggests that the uncertainties on the fluxes are  $\approx 10$  per cent. We also find that the deviations from the mean typically occur coherently for the UVB and VIS arms, which suggests that the uncertainties are achromatic.

In Fig. 2, we show a comparison of the stacked spectrum of XLS-1 to 28 (except 14, 20, 22, and 27 due to their non-detection of the UV continuum, inhomogeneous photometry, or large  $>0.3$  arcsec offset between Ly  $\alpha$  and the continuum) to the stacked SED of the same objects. The spectrum is achieved in the same way as described in Section 4.4 and is significantly binned in the wavelength direction to highlight the continuum level. While the normalizations of the stacked spectrum and the SED are matched at  $\lambda_0 = 1280\text{--}1500$  Å, we show that the same corrections also lead to consistent continuum levels at  $\lambda = 2000\text{--}2800$  Å (i.e. the VIS arm of X-SHOOTER). We also show the stacked photometric data points that were used to derive the SED fits, demonstrating that the average fit is a good fit. We cannot test how well the continuum is matched in redder wavelengths as we do not detect continuum in the NIR arm. For the 18 objects with a continuum detection in the VIS arm (collapsing  $\lambda_0 = 2000\text{--}2800$  Å) with an  $S/N > 5$ , we retrieve fully consistent corrections on a source-by-source level. Both these results validate our wavelength-dependent extraction window described above.

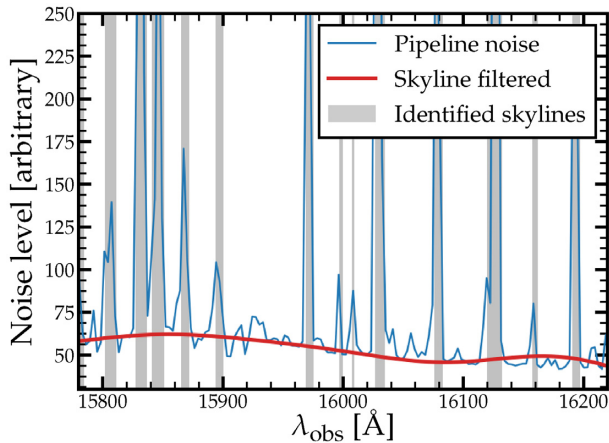


**Figure 2.** Verification of the flux calibration of the spectra. The black line shows the median stacked spectrum of XLS-1 to 28 (except 14, 20, 22, and 27). The red points show the stacked photometry in the  $B$ ,  $g$ ,  $V$ ,  $R$ ,  $I$ , and  $z$  bands. The blue line shows the median stacked SED model from the same objects. The SED model does not include nebular lines. The UVB spectrum is binned by a factor of 12 and the VIS spectrum by a factor of 25 ( $\Delta\lambda = 2.5$  and  $5$  Å, respectively). The top panel shows the spectrum in the UVB arm and the bottom panel shows the spectrum in the VIS arm. We only show the stack at rest-frame wavelengths where all objects are covered in the same arm of the X-SHOOTER spectrograph. The axis ranges are chosen to highlight the continuum emission and we note that the Ly  $\alpha$  line is a cut-off. No significant continuum is detected in the NIR arm. The grey shaded region shows the  $1\sigma$  rms of different bootstrap realizations of the stacked spectrum (Section 4.4). This is binned by the same factor as the spectrum. The normalization of the spectra is matched to the average UV continuum of the SED model over  $\lambda_0 = 1280\text{--}1500$  Å.

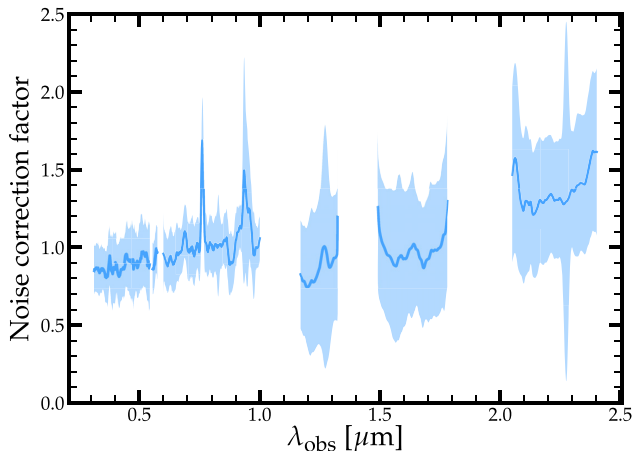
#### 4.1.3 Noise level

We estimate the noise level of the spectra by rescaling the wavelength-dependent noise level that is propagated from the pipeline with the actual noise level measured in the 2D spectra. As we only use wavelength ranges that are free from skyline emission, it is important to first identify skylines automatically, which we do in a two-step process. First, the strongest skylines are identified as inflection points in the propagated pipeline-noise model. Then, after masking these strong skylines, we use a Fourier filtering technique to reconstruct the part of the wavelength dependence of the noise that is related to instrumental throughput and thermal background. The remaining fainter skylines are identified as modes with small-scale power and can thus be removed. We illustrate this in Fig. 3, where we show an example wavelength range around the redshifted H  $\beta$  and [O III] of our target sample.

The challenge in measuring the noise level on the data itself is that there are only limited number of empty sky-pixels in the 2D spectra available. It is possible to extract the 1D spectrum of the empty sky with the same optimal extraction aperture in 6–8 apertures that are independent (depending on the width of the extraction profile) and away from the source itself or the negatives due to the nodding strategy. Then, for each wavelength interval, we could estimate the noise level from the standard deviation of these various 1D noise spectra. However, due to the low number of independent apertures this measurement is noisy. Away from skylines, where the wavelength dependence of the noise is weak and relatively smooth,



**Figure 3.** Illustration of our skyline identification method. The blue line shows the propagated noise level from the X-SHOOTER pipeline. The red line shows the reconstructed underlying wavelength dependence of the noise level after identifying and removing skylines. The identified skylines are marked as grey shaded regions.



**Figure 4.** Final correction factor applied to the pipeline-propagated noise level in order to match the noise level measured directly on the empty sky in the 2D spectra. The blue line shows the median correction factor of all objects and the shaded region shows the  $1\sigma$  range.

we can circumvent this issue by calculating the standard deviation in a running top-hat kernel of width  $20 \text{ \AA}$  using the PANDAS package. After measuring the noise in the sky regions this way, we calculate the noise-correction factor as a function of wavelength and convolve this correction factor with a Gaussian with  $\sigma = 30 \text{ \AA}$ . In Fig. 4, we show that the noise-correction factors range within  $\approx 0.5\text{--}2.0$  and are mostly important in the  $K_s$  band.

#### 4.2 Systemic redshift

It is well known that, due to resonant scattering, the peak redshift of the Ly  $\alpha$  emission does not coincide with the systemic redshift in the majority of galaxies (e.g. Steidel et al. 2010; Hashimoto et al. 2015; Verhamme et al. 2018; Muzahid et al. 2020).

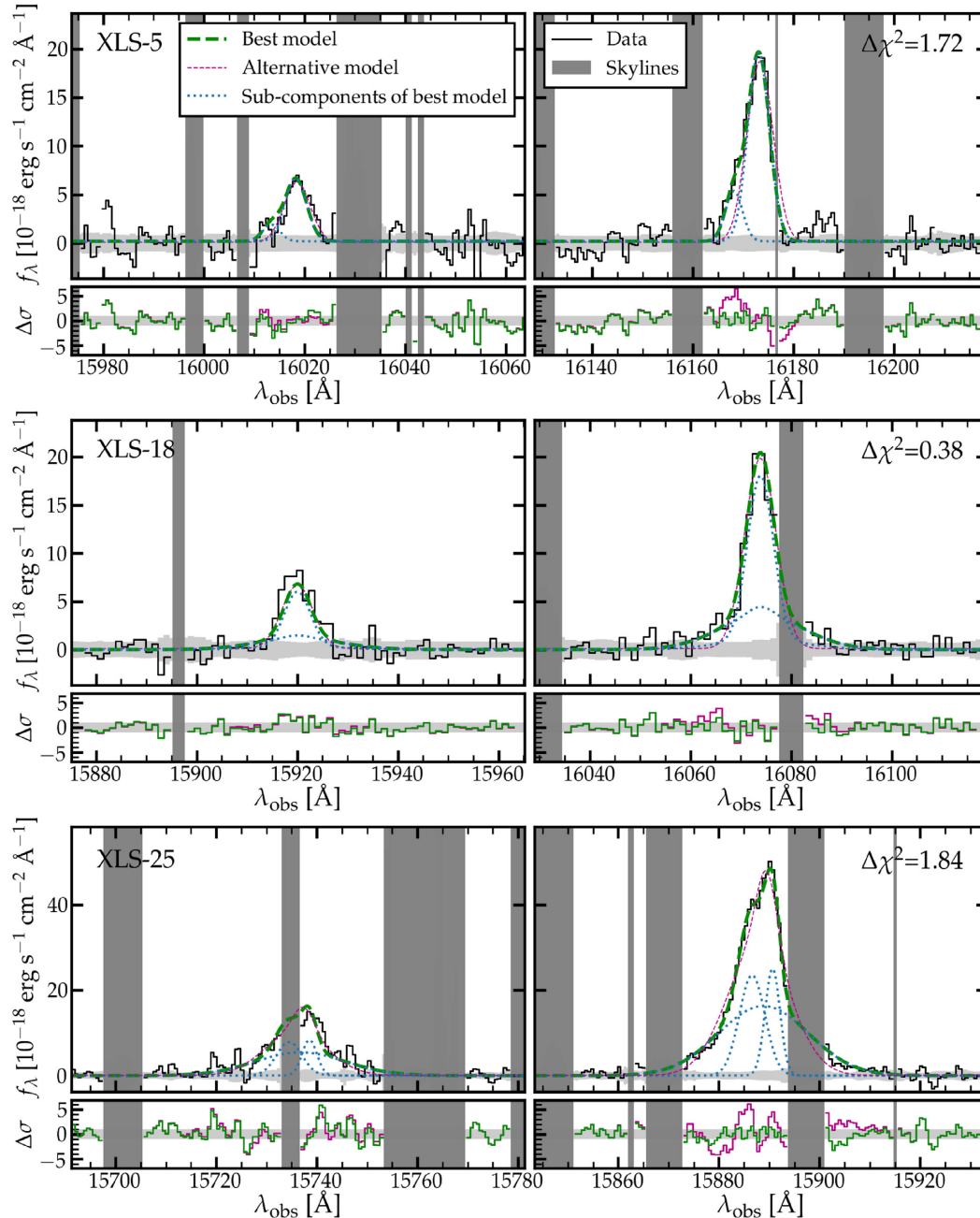
In our data, the systemic redshift is best measured with the [O III]<sub>4960,5008</sub> doublet. Unlike Ly  $\alpha$  and e.g. C IV, the [O III] doublet is not a resonant transition and it is relatively unaffected by attenuation. There are also two practical reasons why [O III] is

particularly helpful. First, after the Ly  $\alpha$  line, it is the emission line that is typically detected with highest signal-to-noise ratio. Besides, H  $\alpha$  is redshifted into the  $K_s$  band with a higher sky background compared to the observed wavelength of [O III]. Secondly, as the [O III]<sub>4960,5008</sub> doublet has a fixed flux ratio of 1:2.98, it is very useful to jointly fit both lines in the presence of skylines. For a single emission line, it often occurs that part of the line is affected by skyline residuals, challenging the measurements of the width and the peak flux in particular if the line profile is not described by a single Gaussian profile. In most cases, this limitation can be overcome by jointly fitting the [O III] doublet, because the chance that both lines are affected by skyline residuals at the same rest-frame velocity is low.

For each galaxy, we first manually obtain a first-guess redshift using one of the [O III] lines or H  $\alpha$  (in the rare case that the peaks of both [O III] lines are affected by skyline residuals). Visual inspection of the 2D spectra reveals that a significant fraction of objects consist of two (narrow) line-emitting components, in line with the high fraction of multiple component systems in the UV imaging (Section 2.5). We also notice that several bright objects show an [O III] profile with relatively broad wings, which has also been observed in low-redshift analogues of LAEs (e.g. Amorín et al. 2012b; Henry et al. 2015; Hogarth et al. 2020). These broad wings are discussed further in Section 6.2.3. Two galaxies (XLS-25 and 35) visually show two narrow components in addition to a broad component; see Fig. 5. We identify [O III] and/or H  $\alpha$  emission for 33/35 objects. These are shown in Fig. A1.

For XLS-9 and XLS-13, we were not able to detect any emission line within the full spatial area covered by the slit besides Ly  $\alpha$  in rest-frame wavelengths  $\lambda = 1000\text{--}7000 \text{ \AA}$ . We note that a faint blue Ly  $\alpha$  peak may be seen in XLS-13 at  $\approx -300 \text{ km s}^{-1}$  from the red Ly  $\alpha$  line, suggesting that the systemic redshift is at  $\approx -150 \text{ km s}^{-1}$  from the peak of the red Ly  $\alpha$  line. In order to understand why no rest-frame optical lines are detected, we determine upper limits of the strongest lines H  $\alpha$  and [O III] assuming a single Gaussian with a line-width FWHM of  $150 \text{ km s}^{-1}$  and velocity shifts of  $(-300)\text{--}(-100) \text{ km s}^{-1}$  with respect to Ly  $\alpha$ . The limiting flux of [O III] strongly depends on this shift while the expected wavelength for H  $\alpha$  is less affected by skylines. The limiting H  $\alpha$  fluxes can be translated into a lower limit of the Ly  $\alpha$  escape fraction (assuming zero attenuation). For XLS-9 and XLS-13, we find  $2\sigma$  limiting escape fractions of  $>59$  per cent and  $>131$  per cent, respectively. This implies that the non-detection of H  $\alpha$  is not particularly unexpected for XLS-9, but we would expect an H  $\alpha$  detection for XLS-13 with an  $S/N > 2.7$ . This slight discrepancy can be a statistical noise effect or indicate that the H  $\alpha$  line width of XLS-13 is broader (resulting in a higher flux limit). Regarding the (typically) brighter [O III] lines, we note that for the typical velocity shift of  $-150 \text{ km s}^{-1}$  with respect to Ly  $\alpha$ , the locations of the redshifted [O III] lines of XLS-9 and XLS-13 are both heavily affected by skylines, plausibly explaining their non-detection.

For each object with a detected [O III] line, we use the LMFIT module for PYTHON to fit the [O III] doublet both using a single Gaussian component and as a combination of two Gaussian components. We assess which fit is preferred based on the reduced  $\chi^2$ . The spectral resolution in the NIR data is  $\approx 50 \text{ km s}^{-1}$  and we do not include the instrumental dispersion in the fitting procedure. We fit three components for XLS-25 and 35 (two narrow components and a broad one). For a single Gaussian fit, we allowed the initial redshift estimate to vary by  $\pm 500 \text{ km s}^{-1}$ . Both the [O III] lines have the same line width (in  $\text{km s}^{-1}$ ) and a fixed relative flux of 1/2.98. We allow the line-width FWHM to vary from 50 to  $1000 \text{ km s}^{-1}$  with initial guess at  $150 \text{ km s}^{-1}$ . For a two-component Gaussian fit, we set the

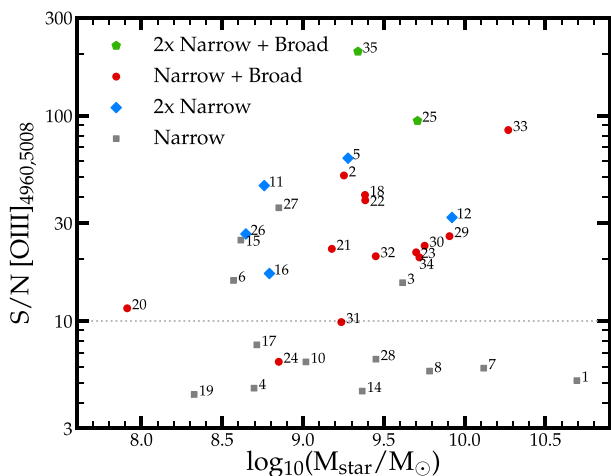


**Figure 5.** Example line-profile fits of the [O III]<sub>4960,5008</sub> doublet. In each row, the top panel shows the data and the bottom panel shows the residuals of the fits. The 1D spectrum is shown in black. The horizontal grey shaded region shows the  $1\sigma$  errors. The vertical shaded regions show the location skyline residuals that are masked in the fitting. The dashed green line shows the best-fitting (multicomponent) model, while the purple line shows the alternative model with less components. Dotted blue lines show the subcomponents.  $\Delta\chi^2$  is the difference between the reduced  $\chi^2$  of the two models that are compared. Top row: XLS-5 is best fitted by a combination of two narrow components. Middle row: XLS-18 is best fitted by a narrow component and a broad component. Bottom row: XLS-25 is best fitted by a complicated [O III] profile that consists of two narrow components and one broad component.

initial redshifts of the two components to  $\Delta z = \pm 0.001$ , respectively, and the line widths to 100 and 400  $\text{km s}^{-1}$ . The redshifts of these components are allowed to vary by 50  $\text{km s}^{-1}$  and the widths can vary freely between 50 and 1000  $\text{km s}^{-1}$  as long as the broad component is broader than the narrow component. For XLS-5, 10, 11, 12, 16, 26, and 27, where the shape of the [O III] line suggests two narrow components or where two clumps are seen in the imaging data, we fit two narrow components both with an initial width of 100  $\text{km s}^{-1}$  and a maximum allowed separation of 200  $\text{km s}^{-1}$ . In these objects, the

S/N is not sufficient to allow the detection of any additional broad component. The fits are highly sensitive to the presence of skyline residuals, which we therefore mask.

Three example [O III] fits are shown in Fig. 5. In these three example cases, a two-component fit is preferred over a single component, as can clearly be seen from the residuals in the bottom panels. XLS-25 is a good example illustrating the use of simultaneously fitting both [O III] lines due to skyline contamination. Out of the 33 objects with [O III] detections, 11 objects are fitted with a single component,



**Figure 6.** Integrated S/N ratio of the [O III]<sub>4960, 5008</sub> flux versus the stellar mass of the XLS targets. The IDs of the targets are labelled on each data point. The objects that are best fitted by a single Gaussian are shown as grey squares. Red points show objects for which an additional broad component is required. Blue diamonds show objects that are best fitted by two narrow profiles and the green pentagons mark the objects for which two narrow components and one broad component are required to accurately fit the [O III] profile. It is clear that an S/N of  $\gtrsim 10$  is typically required in order to identify complex features in the [O III] profile.

5 with two narrow components, 13 with a narrow component and a broad component, and two objects with two narrow components and one broad component. The narrow components have line-width FWHM ranging from 60 to 160 km s<sup>-1</sup>, typically 110 km s<sup>-1</sup>. This means they typically are marginally resolved. The broad components have an FWHM ranging from 200 to 700 km s<sup>-1</sup>, typically 280 km s<sup>-1</sup>. We define the redshift of the narrow component to be the systemic redshift. In case we fit two narrow components (see Table 3), we define the systemic redshift to be at the redshift of the narrow component that is closest to Ly  $\alpha$  along the spatial direction as this is likely the component or H II region that is the origin of the Ly  $\alpha$  emission. For half of these systems, the systemic redshift corresponds to a fainter component of the galaxy. However, the 2D spectrum in these cases clearly shows that the brighter components are spatially offset. If we would use the luminosity-weighted average redshift of the two components, the systemic redshift would on average change by +30 km s<sup>-1</sup> (ranging from -19 to +100 km s<sup>-1</sup>).

As shown in Fig. 6, the detection of more complex features in the [O III] spectrum depends on the integrated S/N. Below an S/N of 10 a single component is typically preferred. Above an S/N of 10, the galaxies with lower masses tend to be described by two narrow components.<sup>1</sup> For all galaxies with [O III] detection, we measure the H  $\beta$  and H  $\alpha$  fluxes assuming the same line profile as our best-fitting [O III] profiles, but we note that these have relatively low S/N ( $\approx 3$ ) in most individual objects. For comparison to other studies at  $z \approx 2$ , we list the H  $\alpha$ -based SFR (see Section 5.3.1, removing the contribution from broad emission to the H  $\alpha$  flux in case a broad component is detected assuming it is not produced by recombination radiation

<sup>1</sup>We note that the [O III] line from XLS-1 is fitted by a single component with an FWHM of 600 km s<sup>-1</sup>. XLS-1 is likely an AGN as indicated from the very broad nebular lines, the detection of broad C IV and Mg II emission, and the red SED.

associated with young stars) in Table 2.<sup>2</sup> For comparison to studies at  $z > 6$ , we also list the combined EW of H  $\beta$  and [O III] that has been measured using the continuum estimate from the SED model.

We have verified the systemic redshift in the majority of sources using other emission lines, particularly the H  $\alpha$  and H  $\beta$  lines. The typical S/N in the H  $\alpha$  line is 3 times lower than that in [O III], while the S/N in the H  $\beta$  line is 5 times lower than that in [O III]. For a few objects, we also verified the systemic redshift with detections of faint He II, O III]<sub>1661, 1666</sub>, and/or C III] line emission in the rest-frame UV, which is a useful consistency check as this ensures a stable wavelength calibration over the UVB to NIR arms.

### 4.3 Lyman $\alpha$ flux measurements

The Ly  $\alpha$  emission lines of the XLS LAEs are shown in Fig. 7. The line profiles are typically double peaked<sup>3</sup> with the redder line being the strongest and being significantly skewed. We measure the Ly  $\alpha$  flux non-parametrically by integrating the flux between  $\pm 1000$  km s<sup>-1</sup> from the systemic redshift. This velocity window captures the total Ly  $\alpha$  flux for all LAEs. For the majority of LAEs with narrower lines, the wide window will lead to very conservative uncertainties. The errors are obtained by re-measuring the flux on 1000 perturbations of the spectrum. The subtracted continuum level is estimated as the average continuum measured over the 1270–1300 Å interval as motivated below. The rest-frame Ly  $\alpha$  EW is computed as the ratio of the Ly  $\alpha$  luminosity and the average continuum luminosity density over this interval.

The continuum level around the Ly  $\alpha$  line can be fairly complicated to estimate accurately because of H I absorption in the ISM, CGM, or IGM (e.g. Laursen, Sommer-Larsen & Razoumov 2011; McKinney et al. 2019), possible absorption in A or B stars (Peña-Guerrero & Leitherer 2013), the nearby NV P Cygni profile at  $\lambda_0 \approx 1245$  Å (e.g. Chisholm et al. 2019), and the strong Si II interstellar absorption line at  $\lambda_0 = 1260$  Å (Reddy et al. 2016b). We therefore measure the continuum level throughout over the  $\lambda_0 = 1270$ –1300 Å interval where the continuum is relatively featureless. For objects with low S/N in the continuum, we estimate the continuum using SED fitting. We have verified that this results in similar Ly  $\alpha$  EWs for the objects for which we could measure the continuum as well in the spectrum.

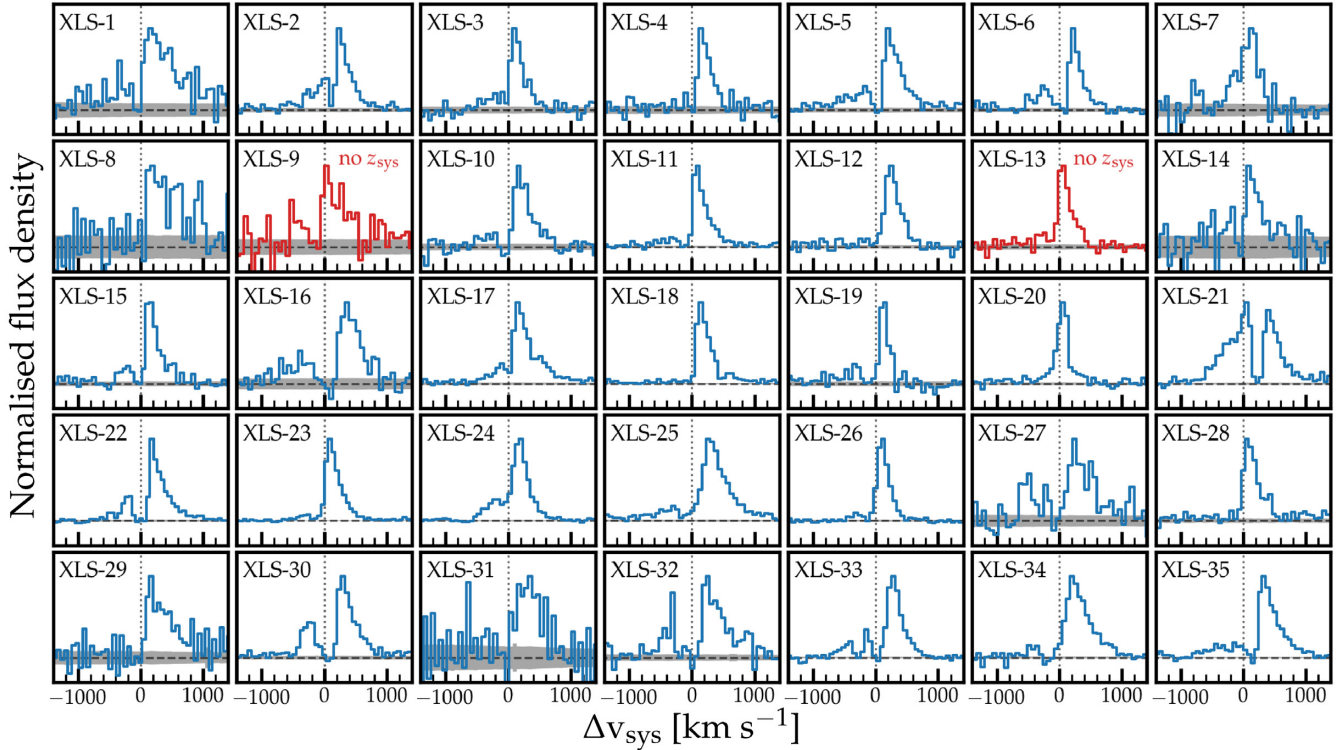
For individual objects, we do not apply an average CGM or IGM correction as it may be possible that the observed LAEs are on biased sightlines that favour higher Ly  $\alpha$  transmission. The average correction at  $z = 2.2$  only affects the blue side of the Ly  $\alpha$  line and is rather moderate compared to high redshifts (Inoue et al. 2014; Byrohl & Gronke 2020). The Ly  $\alpha$  luminosities and EWs are listed in Table 2.

### 4.4 Stacking

We use stacking to obtain the averaged spectrum of the LAEs. This stack is useful for identifying fainter features at the expense of losing information on the dispersion within the subset and complicating the analysis of line profiles.

<sup>2</sup>It is possible to revert this correction based on the relative fraction of the line flux that is in the broad component listed in Table 3.

<sup>3</sup>We will present a detailed investigation of the double peak fraction in a follow-up work, but note that 25 out of the 33 LAEs (i.e. 75 per cent) with a systemic redshift show flux on the blue side of the systemic with a signal-to-noise above 5.



**Figure 7.** Lyman  $\alpha$  profiles of the XLS sample. The spectra are binned in the velocity direction by a factor of 2 for visualization purposes. Spectra are normalized to the peak Ly  $\alpha$  flux density. The velocities are centred on the systemic redshift determined by [O III], except for XLS-9 and XLS-13 that are centred on the peak of the Ly  $\alpha$  line as no non-resonant emission line is detected. The grey shaded region shows the noise level. We note that we will explore these individual line profiles in detail in an upcoming paper.

We stack spectra in 2D because this allows us to investigate differences in the spatial extent of various wavelength regions. By stacking in 2D, we are less sensitive to positional offsets between Ly  $\alpha$  and the continuum (e.g. Hoag et al. 2019; Ribeiro et al. 2020) and uncertainties in the extraction apertures used for the 1D extraction in individual objects.

First, individual 2D spectra are mapped to two common grids in rest-frame wavelength using a linear interpolation: One grid covers the UVB and VIS arms over  $\lambda_0 = 980\text{--}3600 \text{ \AA}$  with  $\Delta\lambda_0 = 0.06 \text{ \AA}$  and the other one covers the NIR arm over  $\lambda_0 = 3650\text{--}7250 \text{ \AA}$  with  $\Delta\lambda_0 = 0.18 \text{ \AA}$ . We scale the flux density of each object to the rest-frame luminosity density and only include objects with a measured systemic redshift. We also apply the same normalization correction as described in Section 4.1.2. Observed wavelengths between 545–560 and 1000–1110 nm are masked because of bad sensitivity in the highest orders of the VIS and NIR spectrographs.

Secondly, we median combine the registered 2D spectra to obtain a typical spectrum of the specific subset. Errors are obtained through bootstrapping. We randomly resample the stacked subset 1000 times and repeat the stacking procedure for each resample to obtain the errors on the stacked spectrum.

Finally, we perform an optimal aperture-matched 1D Gaussian extraction in a slightly modified way compared to individual spectra. We measure the FWHM along the spatial direction at wavelength intervals  $[1290 \pm 50, 1390 \pm 50, 1490 \pm 50, 2200 \pm 125, 2500 \pm 125, 5008 \pm 2, 6564 \pm 2] \text{ \AA}$ . This means that our extraction is based on the size of the continuum, where we assume that the spatial extent of the nebular lines is similar to the rest-frame UV emission as we do not detect continuum emission in the NIR

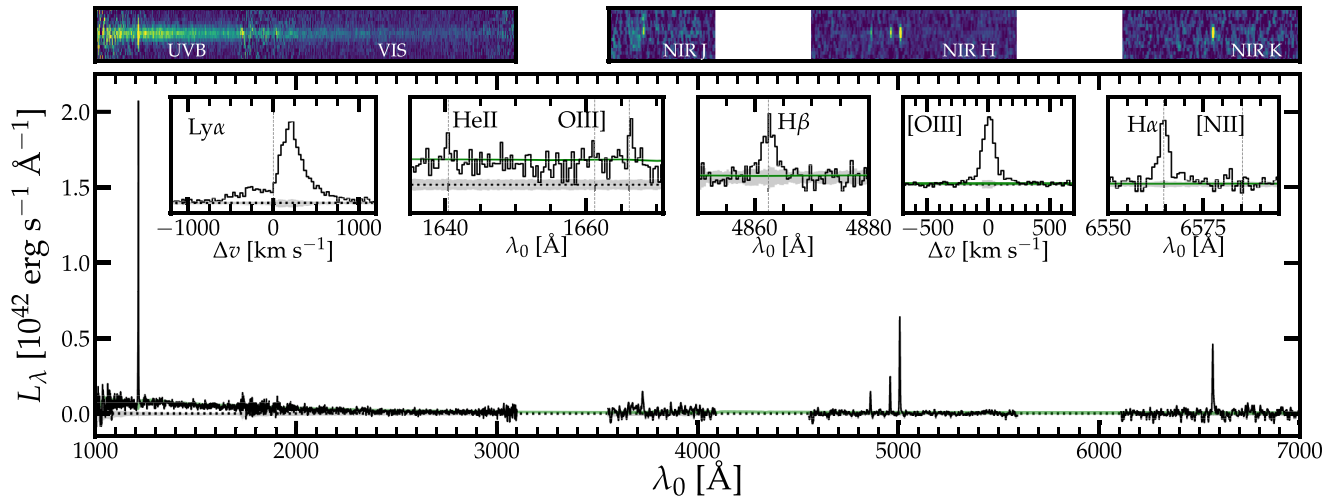
directly.<sup>4</sup> The FWHM decreases slightly with wavelength from 0.96 to 0.84 arcsec. Assuming that the wavelength dependence of the FWHM is smooth, we then fit a second-order polynomial and use that to derive the aperture-matched extraction size as a function of wavelength. The best-fitting polynomial is slightly different for stacks of different subsets as each stack consists of a different combination of atmospheric conditions. For the *representative* stack (described in Section 5), we find  $\text{FWHM} = (1.2\text{--}1.15) \times 10^{-4} \lambda + 0.89 \times 10^{-8} \lambda^2$ , where FWHM is in arcsec and  $\lambda$  is in  $\text{\AA}$ .

We have verified that the 1D extractions from 2D stacks described here are consistent with the stacks from the 1D extracted spectra of individual sources. Similarly, we also derive the median fitted SED from the photometry and its uncertainty and find good agreement with the continuum levels between 1220 and 3000  $\text{\AA}$ .

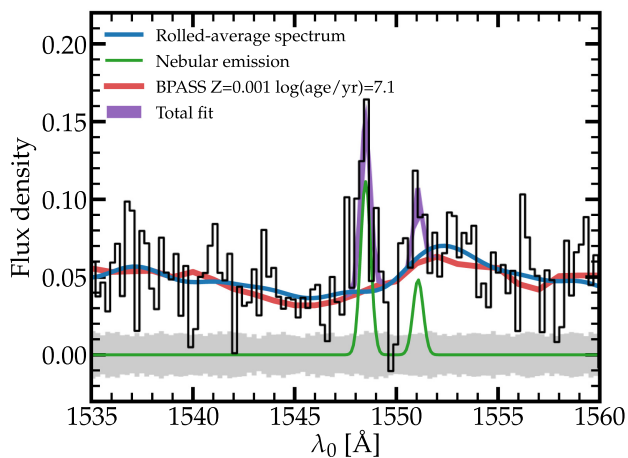
## 5 STACK OF REPRESENTATIVE LAES

Here, we present a stack of LAEs that are *representative* for LAEs at redshift  $z \approx 2$ . With this we mean specifically that we remove objects with Ly  $\alpha$  EW<sub>0</sub> < 10  $\text{\AA}$  (XLS-8) and objects that have been observed spectroscopically because of additional selection criteria (XLS-21 to 23, XLS-27, and XLS-29 to 31; see Section 2.1). We also remove XLS-1 as it is an AGN. We further remove objects for which the

<sup>4</sup>From an inspection of various 2D stacks, we note that the Ly  $\alpha$  line is slightly more extended than the UV continuum, with an FWHM that we measure to be typically  $\approx 10$  per cent higher. We find no difference in the spatial extent of the blue part of the Ly  $\alpha$  line compared to the red part of the Ly  $\alpha$  line.



**Figure 8.** Median stacked spectrum of our sample of LAEs at  $z = 2.2$ . In the top part of the figure, we show the 2D stack. In the main panel, we show the 1D extraction of this stack. The black line shows the spectrum binned by a factor of 10 (in the main panel). The grey region illustrates the uncertainty level. We show the spectrum zoomed-in on specific features: the Ly  $\alpha$  (unbinned) line, the regions around the He II and O III] lines (binned by factor of 2), the H  $\beta$  line, the [O III]<sub>5008</sub> line, and the H  $\alpha$  line (the Balmer lines binned by factor of 4, [O III] by a factor of 2). The green line shows the stacked SED model derived from the photometry.



**Figure 9.** Zoom-in on the wavelength region around the C IV feature in the stack of representative LAEs at  $z \approx 2$ . The blue line shows rolled average spectrum obtained by masking  $\pm 130 \text{ km s}^{-1}$  around the peaks of the nebular lines. The red line shows the best-fitting stellar continuum model that agrees well with the rolled average. The green lines show the best-fitting emission lines and the purple regions show the  $1\sigma$  uncertainties of the fit of the emission lines on top of the continuum.

data are not uniform: XLS-9 and XLS-13 as we did not measure a systemic redshift and XLS-29 to 31 because their H  $\alpha$  line is not covered. The *representative* subset therefore includes 20 LAEs. We show the stacked spectrum of this sample in Fig. 8. The properties of the stack are presented in Section 6.

### 5.1 Line luminosity measurements

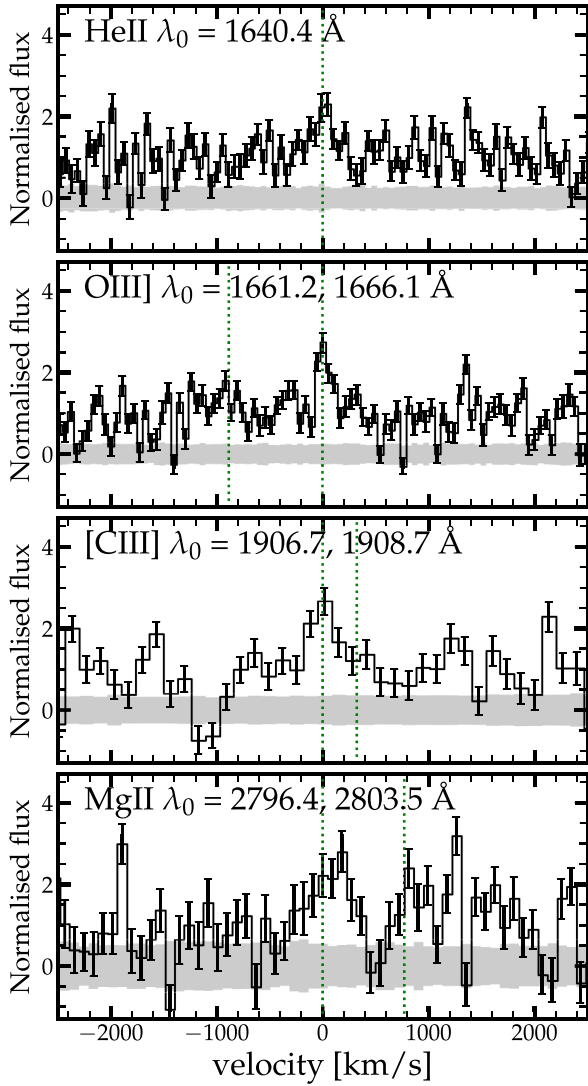
We visually inspect the stacked spectrum and find emission-line detections of Ly  $\alpha$ , C IV<sub>1548, 1551</sub>, He II<sub>1640</sub>, O III]<sub>1661, 1666</sub>, [C III]<sub>1907</sub>, C III]<sub>1909</sub>, and Mg II<sub>2796, 2803</sub> in the rest-frame UV (Figs 8, 9, and 10) and [O II]<sub>3727, 3729</sub>, [Ne III]<sub>3870</sub>, H  $\beta$ , [O III]<sub>4960, 5008</sub>, and H  $\alpha$  in the rest-frame optical (Fig. 8).

#### 5.1.1 Rest-frame optical lines

We notice that not all emission-line profiles are well described by a Gaussian profile; see the inset panels in Fig. 8. While this is not unexpected for Ly  $\alpha$ , we also notice more complex line profiles in the case of H  $\beta$ , [O III], and H  $\alpha$  that cannot be well described by the combination of a narrow component and a broad Gaussian component. This complexity is likely explained by the fact that six representative LAEs show two narrow, closely separated [O III] lines and because five LAEs have strong optical lines that also include a broad component. Indeed, we have verified that removing the identified mergers leads to slightly more symmetric lines, but we note these still do not appear Gaussian. Therefore, we measure the luminosity in these lines non-parametrically by simply integrating the luminosity density within  $\pm 1000 \text{ km s}^{-1}$  for Ly  $\alpha$ ,  $\pm 280 \text{ km s}^{-1}$  for [O III], and  $\pm 180 \text{ km s}^{-1}$  for H  $\alpha$  and H  $\beta$ . These boundaries were determined iteratively using a curve-of-growth approach. A larger window for the Balmer lines does not change the observed [O III]/H  $\beta$  ratio. We note that the wings of the [O III] line have an FWHM of  $\approx 280 \text{ km s}^{-1}$ . These wings could be present in the Balmer lines as well, but we do not detect them with the current sensitivity.

As we do not detect continuum in the NIR, we subtract the continuum measured in the median stack of the best-fitting SED models (this is shown in Fig. 8 as a green curve). This has a minimal impact on the luminosities of the emission lines in the rest-frame optical. These continuum measurements are also used when we derive the EWs of the rest-frame optical lines. The 16–84th confidence percentiles of the line luminosities and EWs are estimated by perturbing the spectrum and continuum levels with the propagated noise 1000 times. The measured luminosities and EWs are listed in Table 4.

The fainter emission lines that we detect can be well fitted by a single Gaussian, but this is probably a consequence of their lower S/N. The [O II] doublet is fitted simultaneously, fixing the two lines to have the same line width and fixing the continuum level to the level of the median SED. Similar to before, confidence intervals are estimated by perturbing the spectrum and the continuum level with their respective uncertainties. We measure similar luminosity for the



**Figure 10.** UV emission lines in the stacked XLS spectrum. Data were binned for visualization. Grey shaded regions show the noise level. Green dotted lines mark the expected positions of emission lines.

two [O II] lines and a line-width FWHM of  $120 \text{ km s}^{-1}$ . This line width is used to derive upper limits for the [N II]<sub>6585</sub> and [S II] lines and when fitting the [Ne III]<sub>3870</sub> line.

### 5.1.2 Rest-frame UV emission lines

In the rest-frame UV, we find that the non-resonant emission lines have FWHM around  $120 \text{ km s}^{-1}$ . The resonant Ly  $\alpha$  and Mg II lines are broader.

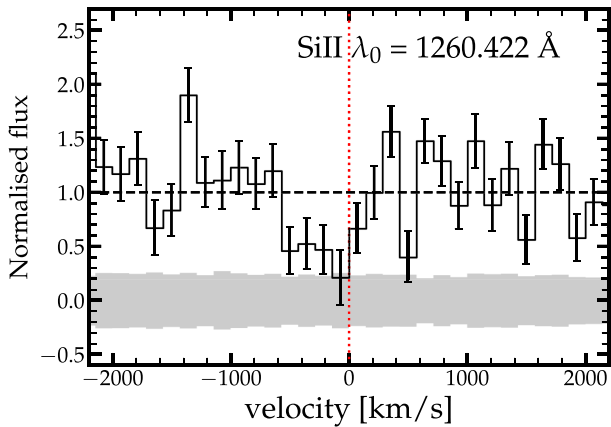
For Ly  $\alpha$ , we estimate the continuum level between 1268 and 1300 Å as described in Section 4.3. The continuum around He II, O III] and C III] is well behaved and is fitted simultaneously with the emission lines. The uncertainty of the continuum level is propagated while measuring uncertainties on the line luminosity and EW. The continuum around the C IV doublet is relatively complex due to the P-Cygni feature arising in the spectra of hot stars and possible interstellar absorption (e.g. Vidal-García et al. 2017; Chisholm et al. 2019). We therefore model the continuum by fitting a single-burst BPASS model over the wavelength ranges  $\lambda_0 = [1530\text{--}1545, 1552\text{--}$

**Table 4.** Measurements of the stacked spectrum of representative LAEs at  $z \approx 2$ . Upper limits are at the  $3\sigma$  level. EWs are all in the rest frame. Line ratios are not corrected for attenuation, unless noted specifically.

Property	Measurement stack
$M_{1500}$	$-19.9 \pm 0.2$
$\beta$	$-2.1 \pm 0.1$
$\log_{10}(M_{\text{star}}/M_{\odot})$	$9.0 \pm 0.1$
SFR <sub>SED</sub>	$4^{+5}_{-3} M_{\odot} \text{ yr}^{-1}$
Age <sub>SED</sub>	$250 \pm 200 \text{ Myr}$
$L_{\text{Ly } \alpha}$	$(6.23^{+0.08}_{-0.08}) \times 10^{42} \text{ erg s}^{-1}$
$L_{[\text{O II}]3727, 3729}$	$(0.45^{+0.05}_{-0.06}) \times 10^{42} \text{ erg s}^{-1}$
$L_{[\text{Ne III}]3869}$	$(0.18^{+0.03}_{-0.03}) \times 10^{42} \text{ erg s}^{-1}$
$L_{\text{H } \beta}$	$(0.48^{+0.03}_{-0.02}) \times 10^{42} \text{ erg s}^{-1}$
$L_{[\text{O III}]4960, 5008}$	$(2.97^{+0.07}_{-0.08}) \times 10^{42} \text{ erg s}^{-1}$
$L_{\text{H } \alpha}$	$(1.67^{+0.05}_{-0.05}) \times 10^{42} \text{ erg s}^{-1}$
$L_{[\text{N II}]6585}$	$< 0.09 \times 10^{42} \text{ erg s}^{-1}$
$L_{[\text{S II}]6718, 6733}$	$< 0.15 \times 10^{42} \text{ erg s}^{-1}$
H $\alpha$ /H $\beta$	$3.5^{+0.2}_{-0.2}$
[O II] <sub>3729</sub> /[O II] <sub>3727</sub>	$1.0^{+0.2}_{-0.2}$
O32 = $\frac{[\text{O III}]_{5008}}{[\text{O II}]_{3727, 3729}}$	$5.0^{+0.6}_{-0.5}$
O3Hb = $\frac{[\text{O III}]_{5008}}{\text{H } \beta}$	$4.7^{+0.3}_{-0.3}$
R23 = $\frac{[\text{O III}]_{4960, 5008} + [\text{O II}]_{3727, 3729}}{\text{H } \beta}$	$7.1^{+0.4}_{-0.4}$
Ne3O2 = $\frac{[\text{Ne III}]_{3870}}{[\text{O II}]_{3727, 3729}}$	$0.4^{+0.1}_{-0.1}$
$\log_{10} \text{N2Ha}$	$< -1.2$
$\log_{10} \text{S2Ha}$	$< -1.0$
$E(B - V)_{\text{gas}}$	$0.22 \pm 0.06$
SFR <sub>H <math>\alpha</math></sub>	$6^{+1}_{-1} M_{\odot} \text{ yr}^{-1}$
$\log_{10}(\xi_{\text{ion, UVdust}}/\text{Hz erg}^{-1})$	$25.3^{+0.1}_{-0.1}$
O32 <sub>int</sub>	$3.8^{+0.5}_{-0.5}$
O3Hb <sub>int</sub>	$4.5^{+0.3}_{-0.2}$
R23 <sub>int</sub>	$7.1^{+0.5}_{-0.4}$
Ne3O2 <sub>int</sub>	$0.4^{+0.1}_{-0.1}$
$\log_{10}(U)$	$-2.44 \pm 0.03 \text{ (stat.)} \pm 0.10 \text{ (sys.)}$
$n_{\text{e, [O II]}}$	$460^{+490}_{-280} \text{ cm}^{-3}$
$T_{\text{e, O}^{2+}}$	$14800^{+700}_{-800} \text{ K}$
$12 + \log(\text{O}/\text{H})_{\text{T}_e}$	$7.83^{+0.06}_{-0.05}$
$12 + \log(\text{O}/\text{H})_{\text{R23}}$	$8.39^{+0.03}_{-0.04}$
$12 + \log(\text{O}/\text{H})_{\text{O32}}$	$8.12^{+0.04}_{-0.04}$
$12 + \log(\text{O}/\text{H})_{\text{O3Hb}}$	$8.29^{+0.04}_{-0.04}$
$12 + \log(\text{O}/\text{H})_{\text{Ne3O2}}$	$8.05^{+0.08}_{-0.06}$
$12 + \log(\text{O}/\text{H})_{\text{strong-line}}$	$8.21^{+0.03}_{-0.03} \text{ (stat.)} \pm 0.14 \text{ (sys.)}$
$\log_{10}(\text{C}/\text{O})$	$-0.8^{+0.2}_{-0.2}$
$\log_{10}(\text{N}/\text{O})$	$< -0.8$
EW <sub>Ly <math>\alpha</math></sub>	$73 \pm 4 \text{ \AA}$
EW <sub>H <math>\alpha</math></sub>	$531^{+131}_{-84} \text{ \AA}$
EW <sub>H <math>\beta</math></sub>	$70^{+20}_{-12} \text{ \AA}$
EW <sub>[O III]<sub>4960, 5008</sub></sub>	$503^{+135}_{-87} \text{ \AA}$

1570] Å that are selected to mask the nebular line emission and interstellar absorption. The best-fitting model has an age of  $10^{7.1}$  yr and a metallicity  $Z = 0.001$ ; see Fig. 9. This model does not reproduce the full SED, but it serves its purpose for modelling the continuum around C IV. We find that both the 1548.19 and 1550.77 Å lines are redshifted by  $60 \pm 20 \text{ km s}^{-1}$  (indicating radiative transfer





**Figure 11.** Detection of Si II in the stacked spectrum of representative LAEs. The velocity axis is with respect to the Si II line at the systemic redshift. Data were binned.

effects; e.g. Berg et al. 2019b), and have an FWHM of  $110 \pm 25 \text{ km s}^{-1}$ . The combined EW of the lines is  $2.2 \pm 0.4 \text{ \AA}$ , where the 1548 line is  $2.5 \pm 1$  times brighter than the 1551 line.

As the Mg II doublet is redshifted into a wavelength region with several skylines, the S/N ratio of the lines and continuum is very low (the S/N values of the lines are 3.3 and 2.1, respectively). We assume a flat continuum around Mg II estimated by averaging over a  $100 \text{ \AA}$ -wide window, masking the Mg II lines. Mg II lines are fitted with single Gaussians. We find that the peaks are redshifted by  $50 \text{ km s}^{-1}$  with respect to the systemic. For the brighter Mg II<sub>2796</sub> line, we measure  $\text{FWHM} = 380 \pm 50 \text{ km s}^{-1}$  and we force the width of Mg II<sub>2803</sub> to be the same. The rest-frame EWs are  $\text{EW}_{\text{Mg II}2796} = 6.1^{+3.1}_{-2.0} \text{ \AA}$  and  $\text{EW}_{\text{Mg II}2803} = 2.5^{+1.7}_{-1.3} \text{ \AA}$ , respectively. These are relatively typical EWs given the UV luminosity of the stack (Feltre et al. 2018).

For the other UV lines, which all show no significant velocity offset compared to the systemic redshift, we subtract the continuum in a model-independent way by masking  $\pm 500 \text{ km s}^{-1}$  around the line centres and linearly interpolating the continuum level on both sides of this mask. For O III]<sub>1661, 1666</sub>, which has the highest S/N (O III]<sub>1666</sub> detection S/N = 12.8), we measure a line width  $120 \pm 30 \text{ km s}^{-1}$  and EWs  $0.6 \pm 0.2$  and  $1.3 \pm 0.3 \text{ \AA}$ , respectively. The [C III]<sub>1907</sub> and C III]<sub>1909</sub> lines are in a noisy region of the stacked spectrum as they lie in the bluest part of the VIS arm of X-SHOOTER that has a lower sensitivity than the redder parts of the UVB arm. We therefore constrain the widths of these lines to the width of the O III]<sub>1666</sub> lines and measure EWs  $2.8^{+0.7}_{-1.7}$  and  $1.5^{+0.5}_{-0.4} \text{ \AA}$ , respectively. We find an indication that the He II line is somewhat broader ( $\text{FWHM} = 210 \pm 90 \text{ km s}^{-1}$ ) than the other lines, indicating possible contribution from broad stellar He II emission (Brinchmann, Pettini & Charlot 2008). As we are interested in the nebular He II component but do not have the sufficient S/N to perform a two-component fit (He II detection S/N = 8.6), we force the width to the range of widths of the O III]<sub>1666</sub> line and find an EW of  $1.2^{+0.4}_{-0.4} \text{ \AA}$ . Allowing the line width to be larger, we would measure a line flux that is a factor of 1.3 higher.

## 5.2 Si II absorption

We detect significant absorption from the low-ionization Si II<sub>1260</sub> line in our stacked spectrum; see Fig. 11. In this figure, the continuum level is estimated from the stacked SED model, which agrees well

with the stacked spectrum. This absorption line has also been seen in several individual cases and stacks of LAEs (e.g. Shibuya et al. 2014; Rivera-Thorsen et al. 2015; Trainor et al. 2015). We do not detect other absorption lines at  $>3\sigma$  significance. The absorption EW is measured by integrating between the maximum and minimum velocities at which absorption is detected. We perturb each spectrum 1000 times to estimate the uncertainties on the EW measured this way. For Si II, the EW is  $-1.7^{+0.8}_{-0.3} \text{ \AA}$  and the absorption-weighted average velocity is  $-280^{+130}_{-70} \text{ km s}^{-1}$ .

## 5.3 Derived physical properties

We use the emission line measurements to derive the nebular attenuation,  $\text{SFR}_{\text{H}\alpha}$ , the star formation rate surface density, Ly  $\alpha$  escape fraction, the production efficiency of ionizing photons, electron density, electron temperature, and gas-phase oxygen abundance.

### 5.3.1 Dust attenuation, SFR, ionizing production efficiency, and escape of Ly $\alpha$ photons

The observed Balmer line luminosities can be used to infer  $Q_{\text{ion}}$ , the number of emitted ionizing photons per second, following  $Q_{\text{ion}} = c_{\text{H}\alpha} L_{\text{H}\alpha}$ , where  $c_{\text{H}\alpha} = 1.36 \times 10^{-12} \text{ erg}^{-1}$  (Osterbrock 1989). The conversion is sensitive to the escape fraction of ionizing photons and the dust content within H II regions, but these are assumed to be negligible here (e.g. Inoue 2001; Dopita et al. 2006). As the dominant source of ionizing photons in the average LAE is star formation (see Section 6.1), we can derive  $\text{SFR}_{\text{H}\alpha}$ .

A main uncertainty in deriving  $Q_{\text{ion}}$  is the dust attenuation affecting the observed H  $\alpha$  luminosity. We can estimate the nebular attenuation by using the observed Balmer decrement  $\tau = \log(\frac{H\alpha}{H\beta}/2.79)$ , where the assumed intrinsic line ratio of 2.79 depends slightly on electron density and temperature (we assume  $100 \text{ cm}^{-3}$  and  $15000 \text{ K}$ ; see Section 5.3.2), in combination with an assumed extinction law and dust geometry. We assume that dust is distributed as a uniform screen (cf. Scarlata et al. 2009). Following Reddy et al. (2015, 2020), we assume that the nebular extinction curve  $k(\lambda)$  follows the one from Cardelli, Clayton & Mathis (1989), such that  $E(B - V) = 0.95\tau$  and  $k_{\text{H}\alpha} = 2.52$ . For our stack, we measure  $E(B - V)_{\text{neb.}} = 0.22 \pm 0.06$  that is then used to calculate the intrinsic H  $\alpha$  luminosity following  $L_{\text{H}\alpha, \text{intr}} = L_{\text{H}\alpha, \text{obs}} \times 10^{0.4E(B-V)_{\text{neb.}}k_{\text{H}\alpha}}$ .

The conversion of the intrinsic H  $\alpha$  luminosity to SFR follows from the relations between  $Q_{\text{ion}}$  and H  $\alpha$  luminosity, and  $Q_{\text{ion}}$  and SFR. The latter conversion depends on the SFH, the IMF, the properties of massive stars (e.g. binary fraction), and the stellar metallicity. We use the conversion  $\text{SFR}_{\text{H}\alpha}/M_{\odot} \text{ yr}^{-1} = 2.29 \times 10^{-42} \times L_{\text{H}\alpha, \text{intrinsic}}/\text{erg s}^{-1}$  derived by Theios et al. (2019) for BPASS v2.2 models with a Chabrier (2003) IMF with a maximum stellar mass of  $100 M_{\odot}$ , a constant SFH with an age of  $10^8 \text{ yr}$ , and a metallicity of  $0.1 Z_{\odot}$ . Note that the conversion between SFR and intrinsic H  $\alpha$  luminosity is a factor of  $\approx 2$  smaller than the ‘standard’ conversion with the same IMF but 10 times higher metallicity (Murphy et al. 2011; Kennicutt & Evans 2012). This is due to the harder ionizing spectra of low-metallicity stars and the longer contribution to the ionizing photon flux of binary stellar populations compared to populations of single stars (e.g. Göteborg et al. 2019).

The production efficiency of ionizing photons ( $\xi_{\text{ion}}$ ) is defined as  $\xi_{\text{ion}} = \frac{Q_{\text{ion}}}{L_{\text{UV, intrinsic}}}$  (Bouwens et al. 2016). We note that  $\xi_{\text{ion}}$  is also related to the relative production of Ly  $\alpha$  photons to the UV continuum and hence to the intrinsic Ly  $\alpha$  EW (e.g. Sobral & Matthee 2019). We estimate the intrinsic UV luminosity using

$L_{UV, intrinsic} = L_{UV, observed} \times 10^{0.4E(B-V)_*k_{1500}}$ . A crucial assumption is the relation between the stellar and nebular attenuation, the latter being estimated from the Balmer decrement. Several studies that investigated the relation between the stellar and nebular attenuation at  $z \sim 2-3$  have yielded conflicting results. Some indicate that the nebular extinction and stellar extinction are similar, while others report a higher nebular attenuation (e.g. Kashino et al. 2013; Reddy et al. 2016a; Steidel et al. 2016; Faisst et al. 2019; Theios et al. 2019). Relevant for our sample is that Shivaiei et al. (2020) report a higher nebular attenuation in systems with a gas-phase metallicity below  $12 + \log(O/H) < 8.5$ . Therefore, we assume the classical  $E(B - V)_* = 0.44E(B - V)_{gas}$  (Calzetti et al. 2000) ratio here. For the UV attenuation, we use the Reddy et al. (2016a) attenuation curve, which results in  $k_{1500} = 8.68$ . The resulting  $\xi_{ion} = 10^{25.3 \pm 0.1} \text{ Hz erg}^{-1}$ . This value is consistent with the value from the BPASS model that we used to convert H  $\alpha$  luminosity to SFR. If we assume  $E(B - V)_* = E(B - V)_{gas}$ , we find a lower  $\xi_{ion} = 10^{24.9 \pm 0.1} \text{ Hz erg}^{-1}$ . The values for  $E(B - V)_*$  from the SED fitting range from 0.0 to 0.17, with a mean of 0.03. This points towards an even lower stellar attenuation compared to the nebular attenuation.

Finally, having estimated the intrinsic H  $\alpha$  luminosity, we can calculate the Ly  $\alpha$  escape fraction using  $f_{esc, Ly \alpha} = L_{Ly \alpha} / (8.7L_{H \alpha, intrinsic})$ . The intrinsic ratio between Ly  $\alpha$  and H  $\alpha$  depends slightly on gas temperature (e.g. Henry et al. 2015) but we use 8.7 for consistency with the literature. We measure  $f_{esc, Ly \alpha} = 0.27 \pm 0.04$ . This is consistent with the Ly  $\alpha$  escape fraction measured using the stacking of H  $\alpha$  narrow-band imaging data on the parent sample of LAEs (Sobral et al. 2017), which suggests that relative slit losses are unimportant.

### 5.3.2 Electron density, temperature, ionization state, and gas-phase abundances

The line ratios of the [O II]<sub>3727, 3729</sub> doublet and the [C III]<sub>1907</sub>/[C III]<sub>1909</sub> lines are sensitive to the electron density (e.g. Keenan, Feibelman & Berrington 1992; Patrício et al. 2016). Using equation (7) from Sanders et al. (2016), we measure an electron density of  $n_e = 460^{+490}_{-280} \text{ cm}^{-3}$  based on the [O II] doublet consistent with Shirazi, Brinchmann & Rahmati (2014) and Steidel et al. (2014). The uncertainties on the [C III]<sub>1907</sub>/[C III]<sub>1909</sub> ratio are too large to use it to obtain meaningful constraints on electron density.

The electron temperature is a key property of the ISM. We do not detect the temperature-sensitive [O III]<sub>4363</sub> line as this line is redshifted into a wavelength range with low atmospheric transmission. However, following the methodology from Pérez-Montero & Amorín (2017), we can estimate the temperature from the dust-corrected [O III]<sub>1661 + 1666</sub>/[O III]<sub>5008</sub> ratio. The intrinsic H  $\alpha$ /H  $\beta$  ratio that is assumed to derive the attenuation depends on the electron temperature. Therefore, we iteratively derive the attenuation and the electron temperature until convergence, which results in  $T_{e, O^{2+}} = 14\,800^{+700}_{-800} \text{ K}$ .

The relative strengths of the [O III], [O II], and H  $\beta$  lines are sensitive to the electron temperature, gas-phase metallicity, and ionization state (e.g. Nakajima & Ouchi 2014; Trainor et al. 2016). In Table 4, we list the intrinsic ratios O32, O3Hb, and R23 after correcting the emission-line luminosities for attenuation based on the Cardelli et al. (1989) curve. We find a high O32, O3Hb, and R23 of  $3.8^{+0.5}_{-0.3}$ ,  $4.5^{+0.3}_{-0.2}$ , and  $7.1^{+0.5}_{-0.4}$ , respectively. These values point towards a low metallicity parameter and a high ionization parameter. Following the iterative methodology described in Nakajima & Ouchi (2014) based on photoionization models by Kewley & Dopita (2002),

we measure  $\log_{10}(U) = -2.38^{+0.04}_{-0.05}$ . The [Ne III]-to-[O II] ratio (Ne3O2) has also been proposed to trace the ionization parameter (Levesque & Richardson 2014), similar to O32. While the lines originate from different species, the main benefit is that they are closer in wavelength and thus less susceptible to uncertainties in the attenuation correction. A caveat is that the ionization energy from [Ne III] is 41 eV, which is somewhat higher than the ionization energy of [O III] (35.1 eV). Therefore, changing the hardness of the ionization field can lead to variations in the relations between  $U$  and Ne3O2, and  $U$  and O32. Applying the calibration based on photoionization modelling of LBGs at  $z \sim 2$  by Strom et al. (2018), the observed Ne3O2 ratio implies  $\log_{10}(U) = -2.47 \pm 0.08$ , which is consistent with the value derived using their calibration for O32 [ $\log_{10}(U) = -2.49 \pm 0.03$ ]. Given the variation in  $U$  for different methods, we retrieve the average  $U$  and conservatively add 0.1 dex systematic uncertainty, i.e.  $\log_{10}(U) = -2.44 \pm 0.03 \text{ (stat.)} \pm 0.10 \text{ (sys.)}$ .

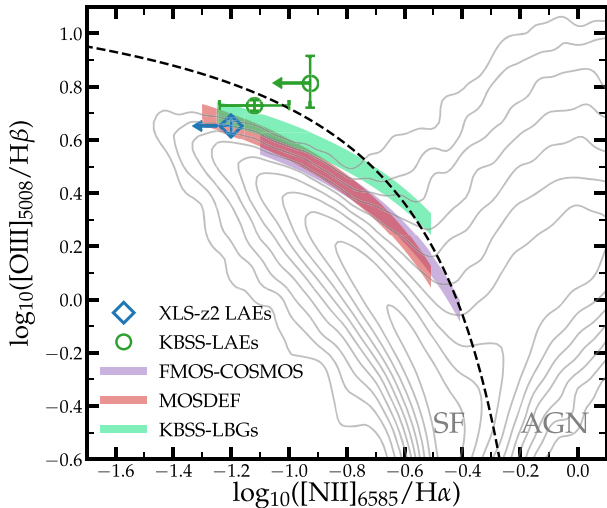
We can estimate the gas-phase metallicity directly using the electron temperature and the strength of the oxygen lines compared to H  $\beta$ . Following Pérez-Montero (2014), we measure  $12 + \log(O/H) = 7.83^{+0.06}_{-0.05}$ . We use  $n_e = 500 \text{ cm}^{-3}$ , but the metallicity would only change by 0.01 dex within the constrained range of electron densities. Alternatively, we can estimate the gas-phase metallicity using empirical strong-line calibrations derived in low-redshift analogues of high-redshift galaxies by Bian, Kewley & Dopita (2018). Besides R23, O32, and O3Hb, we also use the metallicity calibration based on the Ne3O2 ratio. Combining all our estimates (see Table 4) yields an average gas-phase metallicity of  $12 + \log(O/H)_{strong-line} = 8.21^{+0.03}_{-0.03} \text{ (stat.)} \pm 0.13 \text{ (sys.)}$ , where the systematic error comes from the variation among the different line indices. This measurement is consistent with our upper limit on the [NII]/H  $\alpha$  ratio that corresponds to  $12 + \log(O/H) > 8.25$ . The metallicity estimated from the strong lines is on average 0.4 dex higher than our estimate inferred through the  $T_e$  method, potentially indicating issues with using the UV lines with much higher critical density compared to the optical [O III] line (see also Rigby et al. 2021). Finally, following Pérez-Montero & Amorín (2017), we use the dust-corrected C IV, O III], and C III] lines to estimate a C/O abundance of  $\log_{10}(C/O) = -0.8 \pm 0.2$  and we use the upper limit on [N II] to derive  $\log_{10}(N/O) < -0.8 \text{ (} 3\sigma \text{)}$ .

## 6 RESULTS

Here, we address the average nature of LAEs at  $z \approx 2$  by predominantly synthesizing the various measurements of our stack, but we also include results from the high-resolution rest-frame morphology of individual sources. The variation of the properties among the individual sources will be explored in an upcoming paper.

### 6.1 LAEs are powered by young metal-poor stars

Several emission-line ratios that we observe in the stacked spectrum are sensitive to the source of ionization. These include not only the well-known relative strengths of [O III]/H  $\beta$  compared to [N II]/H  $\alpha$  (i.e. the BPT diagram; Baldwin, Phillips & Terlevich 1981), but also the ratio of high-ionization metal lines (such as C IV, O III], and C III]) to He II in the rest-frame UV (e.g. Feltre, Charlot & Gutkin 2016; Nakajima et al. 2018b).



**Figure 12.** Location of the stack of representative LAEs at  $z \approx 2.2$  on the ‘BPT’ diagram (blue diamond). The grey contours mark the distribution of galaxies at  $z = 0.03\text{--}0.06$  in SDSS (Brinchmann et al. 2004). The dashed line is the demarcation line between ionization due to young stars and AGN (Kauffmann et al. 2003). Green symbols show the average line ratios of LAEs and LBGs with strong Ly  $\alpha$  emission at  $z \sim 2$  measured by Trainor et al. (2016), where LAEs have the highest O3H $\beta$ . These samples are intrinsically a factor of 2 less and more luminous, respectively. The coloured shaded regions show the average locations of galaxies from various large surveys at  $z \sim 2$ . Green shows UV-selected galaxies from KBSS (Strom et al. 2017), red shows  $H$ -band-selected galaxies from MOSDEF (Shapley et al. 2015; Sanders et al. 2020b), and purple shows the slightly more massive galaxies from FMOS-COSMOS (Kashino et al. 2017). The masses of our LAEs are generally lower than those in such continuum-selected galaxy samples. Upper limits are at the  $3\sigma$  level.

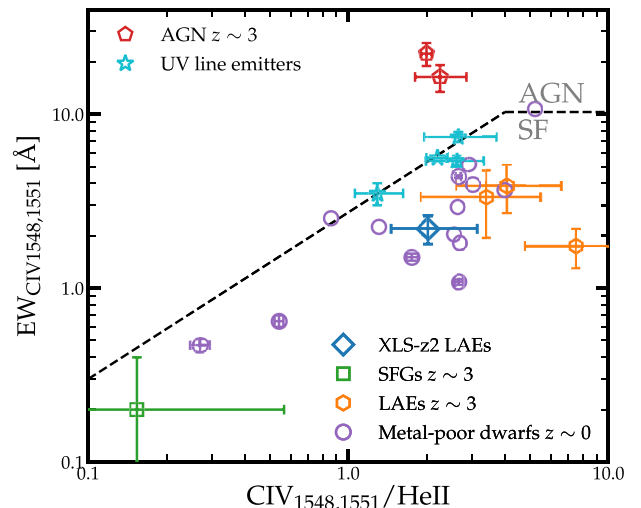
### 6.1.1 BPT diagram

In Fig. 12, we show the location of our stacked spectrum in the BPT diagram and compare this to galaxies at  $z = 0.03\text{--}0.06$  in the Sloan Digital Sky Survey (SDSS; Brinchmann et al. 2004; Alam et al. 2015). Similar to other samples of galaxies at  $z \sim 2$  (e.g. Shapley et al. 2015; Kashino et al. 2017; Strom et al. 2017), we find elevated [O III]/H  $\beta$  ratios at fixed [N II]/H  $\alpha$  that are rare in the local Universe. The line ratios are, however, still well within the part of the diagram that implies photoionization by young stars (Kauffmann et al. 2003). The observed ratios indicate a high ionization state possibly due to a hard ionizing spectrum from metal-poor stars (Steidel et al. 2016; Topping et al. 2020).

### 6.1.2 Rest-frame UV ionization diagram

In addition to using the strong rest-frame optical lines, various line ratios in the rest-frame UV have recently been proposed to be capable of identifying AGN activity (e.g. Feltre et al. 2016; Nakajima et al. 2018b).

Here, we focus on the C IV/He II ratio. While we note that measurements of nebular C IV emission may be complicated by stellar and interstellar absorption and emission as described in Section 5.1.2, the attractive feature of C IV is that its production by collisional excitation requires an energy of 47.9 eV. The main production mechanism for nebular He II is recombination, which requires an ionization energy of 54.4 eV. We therefore expect that objects with a harder ionizing spectrum have a relatively stronger He II line at



**Figure 13.** The relation versus the combined C IV<sub>1548+1551</sub> EW and the C IV/He II line ratio. The dashed line shows the demarcator between ionization by star formation or AGN (Nakajima et al. 2018b). The blue diamond shows the measurement of the typical LAEs at  $z \approx 2$ , which is consistent with photoionization by stellar atmospheres. The orange hexagons show stacks of LAEs at  $z \approx 3$  (Nakajima et al. 2018a; Feltre et al. 2020), the green square shows typical star-forming galaxies at  $z \sim 3$  from the VUDS survey (Nakajima et al. 2018b), and the purple circles show metal-poor dwarfs at  $z \sim 0$  (Senchyna et al. 2017, 2019; Berg et al. 2019a). The cyan stars are galaxies at  $z \sim 2$  that have been selected on strong rest-frame UV line emission (Amorín et al. 2017; Nakajima et al. 2018b). With red pentagons, we show the average EW and line ratio measured in stacks of UV-selected (Hainline et al. 2011) and Ly  $\alpha$ -selected AGN (Sobral et al. 2018b) at  $z \sim 2\text{--}3$ .

fixed carbon abundance. In addition to the line ratio, we also focus on the C IV EW. There naturally is a maximum EW associated with a stellar population and given the blackbody-like shape of stellar spectra that exponentially drop around 50 eV for  $T \approx 50\,000$  K, it is very challenging to reach EWs in excess of  $\sim 10$  Å, unless the stellar atmospheres are extremely hot (Nakajima et al. 2018b). Very high C IV EWs thus indicate power-law-type ionizing sources such as AGN. We show the C IV EW and the C IV/He II ratio of our stacked sample of LAEs in Fig. 13. The observed EW and line ratio can be explained by photoionization by stellar atmospheres and does not indicate significant AGN contribution, in agreement with the results from the BPT diagram. The C IV EW of typical LAEs is significantly lower than the EWs measured in stacks of Ly  $\alpha$  (Sobral et al. 2018b) and UV-selected AGN (Hainline et al. 2011) at these redshifts, validating the use of this diagram to identify AGN activity.

In addition to C IV, we use the O III]/He II line ratio to identify AGN activity (e.g. Amorín et al. 2017). Similarly, with the O III]<sub>1661+1666</sub> flux being higher than the He II flux this line ratio prefers photoionization by young stars as well (Feltre et al. 2016).

### 6.1.3 Young and metal-poor stellar populations

The EWs of the Balmer lines are sensitive to the age of stellar populations as they trace the amount of ionizing photons originating from O stars relative to the stellar continuum from lower mass stars (e.g. Leitherer et al. 1999). Likewise, the EW of the He II<sub>1640</sub> line is very sensitive to the hardness of the spectrum of a composite stellar population (e.g. Kehrig et al. 2015; Berg et al. 2018; Göteborg et al. 2019; Nanayakkara et al. 2019), which traces the stellar iron

abundance, the high-mass end of the IMF, the presence of high-mass binaries, and other detailed properties of very massive stars (e.g. Schaerer 2003; Gräfenor & Vink 2015; Szécsi et al. 2015; Stanway et al. 2020).

We explore the implications of the observed EWs of H  $\alpha$  for the age of the stellar populations using the results from Xiao, Stanway & Eldridge (2018), who performed CLOUDY modelling on single-burst BPASS models (Stanway & Eldridge 2018) with a range of metallicities ( $Z$ ) and ionization parameters ( $U$ ). As in these models the nebular metallicity is fixed to the stellar metallicity and the relative metal abundances do not vary with metallicity (cf. Steidel et al. 2016), we do not use them to interpret line ratios involving metal lines. The H  $\alpha$  EW of the stack is  $531^{+131}_{-84}$  Å. This implies a burst age of  $6^{+4}_{-2}$  Myr, which is only mildly sensitive to  $Z$ . This is in agreement with results from Leitherer et al. (1999), where such H  $\alpha$  EW corresponds to ages of 5 Myr. For a constant SFH, we find that the observed H  $\alpha$  EW corresponds to an age of  $\approx 40^{+20}_{-15}$  Myr. Independently of the H  $\alpha$  EW, the light-weighted ages inferred from the SED fitting to the broad-band continuum data are on average  $250 \pm 200$  Myr. Indeed, if these typical LAEs experienced a constant star formation rate of  $\approx 5 M_{\odot} \text{ yr}^{-1}$  they would have formed their total stellar mass in  $\approx 200$  Myr. We further note that the low C/O abundance suggests that there has not yet been time for carbon enrichment from evolved stars (e.g. Berg et al. 2019a), consistent with young ages.

We note that the He II EW of  $1.2 \pm 0.4$  Å that we observe in typical LAEs at  $z \approx 2$  suggests very hot stars with a very low stellar metallicity. However, it is challenging to quantify this metallicity as current models are known to miss sources of He II-ionizing photons (e.g. Berg et al. 2018; Saxena et al. 2020; Wofford et al. 2020), such as X-ray binaries (e.g. Schaerer, Fragos & Izotov 2019) or radiative shocks (e.g. Jaskot & Oey 2013; Stasińska et al. 2015; Plat et al. 2019). An indirect estimate of the stellar metallicity of the young stars can be obtained by assuming pure enrichment from core-collapse supernovae and that the gas-phase metallicity is the same as the stellar metallicity. Pure core-collapse enrichment yields  $O/Fe \approx 5 O/Fe_{\odot}$  (Nomoto et al. 2006) and this is warranted by observations at  $z \approx 2$  (e.g. Steidel et al. 2016; Topping et al. 2020) and hydrodynamical simulations (Matthee & Schaye 2018). Under these assumptions, the gas-phase metallicity  $12 + \log(O/H) \approx 8.0$  corresponds to a stellar metallicity ( $Z$ ) of  $\approx 10^{-3}$ , which is in agreement with the metallicity inferred through photoionization modelling of galaxies at  $z \approx 2$  with similar locations in the BPT diagram (Topping et al. 2020).

## 6.2 LAEs are clumpy and have turbulent and outflowing kinematics

### 6.2.1 Clumpy structures

As discussed in Section 2.5 and listed in Table 3, the *HST* imaging reveals multiple clumps in 12/31 objects. We further identify seven galaxies with multiple narrow emission-line components that have a velocity shift with respect to each other (Section 4.2). Except for XLS-16, these all appear clumpy in the imaging data as well. The remaining LAEs typically appear very compact, but we also note that we find a bias towards identifying multiple clumps in the more luminous systems. These results strongly suggest that the UV light and the emission lines in these LAEs are dominated by one or a few massive complexes of star formation (see also Cornachione et al. 2018), possibly associated with ongoing mergers.

### 6.2.2 Outflowing ionized gas from stellar feedback

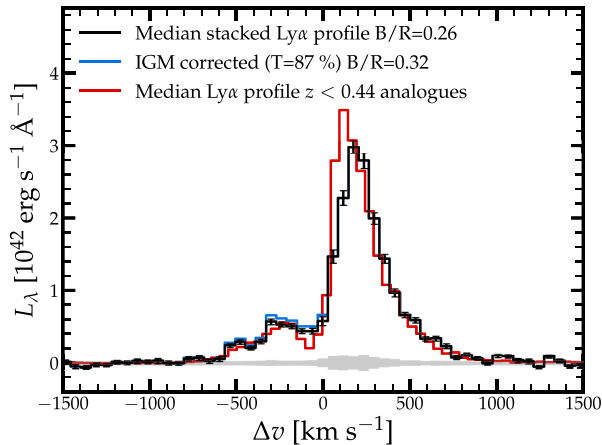
From the stack, the clearest evidence of outflowing gas is blueshifted interstellar absorption that we detect in the low-ionization Si II 1260 Å line (Fig. 11). The absorption-weighted velocity is shifted by  $-260^{+160}_{-40}$  km s $^{-1}$  with respect to the systemic velocity. This is a slightly higher velocity than measured in continuum-selected galaxies at  $z \sim 3$  (Shapley et al. 2003; Erb 2015). Such outflows could plausibly be driven by the feedback processes associated with the young starbursts in these LAEs (Ma et al. 2016). For the 17 representative LAEs that have rest-frame UV size measurements from the *HST* data (Table 2), we measure a median (mean)  $\Sigma_{\text{SFR}_{\text{H}\alpha}} = 1.2(1.8) M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$  assuming  $\Sigma_{\text{SFR}_{\text{H}\alpha}} = \text{SFR}_{\text{H}\alpha} / 2\pi r_{\text{eff,UV}}^2$  (Shibuya et al. 2019), which is well above the typical threshold of  $\Sigma_{\text{SFR}} > 0.1 M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$  for driving large-scale galactic outflows (e.g. Heckman et al. 2001).

### 6.2.3 Complex [O III] profiles: outflows, turbulence, and unresolved structures?

The [O III] line profiles in a significant fraction of the LAEs show evidence of a broader (FWHM  $\approx 280$  km s $^{-1}$ ) component. Fig. 6 shows that these broad components can only be identified in the LAEs with higher S/N detections of [O III], and that the broader components are more often seen at  $M_{\text{star}} \approx 3 \times 10^9 M_{\odot}$  than at lower masses. A broad [O III]-emitting component with an FWHM of  $\approx 280$  km s $^{-1}$  is also seen in the stack of representative LAEs, which contains the majority of LAEs for which the individual S/N of [O III] was not sufficient to detect a broad component. Broad emission lines have also been observed in local analogues of LAEs (e.g. Amorín et al. 2012b; Bosch et al. 2019; Hogarth et al. 2020) and they have been proposed to indicate (shock)-ionized outflowing material (e.g. Heckman, Armus & Miley 1990; Veilleux, Cecil & Bland-Hawthorn 2005; Freeman et al. 2019). However, these lines are typically broader than  $> 500$  km s $^{-1}$ , unlike we observe. Indeed, using stacks of galaxies with  $M_{\text{star}} \gtrsim 10^{10} M_{\odot}$  at  $z \sim 2$ , Davies et al. (2019) find that the width of the broad component increases with the star formation surface density. Following this observed correlation, the  $\Sigma_{\text{SFR}}$  of typical LAEs would imply a broader FWHM ( $\approx 600$  km s $^{-1}$ ). As broader lines are more difficult to detect, deeper data are required to resolve this tension. Alternatively, the observed broad components could be due the unresolved motions of several (potentially somewhat more evolved) H II regions within the galaxies (e.g. Östlin et al. 2015) or line emission originating from gas in turbulent mixing layers such that the width of the broad component reflects the turbulent velocity (e.g. Westmoquette et al. 2007). In order to differentiate these scenarios, we would need to measure line ratios resolved for the narrow and broad components (e.g. Hogarth et al. 2020), which the current S/N does not allow for all lines.

### 6.2.4 Outflows indicated from the Ly $\alpha$ profile

The Ly  $\alpha$  profiles of LAEs are another observable that is sensitive to the kinematics of the gas (e.g. Neufeld 1990; Verhamme et al. 2006; Kakiichi & Gronke 2019). While a detailed study of the Ly  $\alpha$  line profiles of individual LAEs will be presented in Matthee et al. (in preparation), we focus here on the stacked Ly  $\alpha$  profile that is shown in Fig. 14. The velocity axis is centred on the systemic redshift. It is clear that the median Ly  $\alpha$  profile shows a prominent skewed red line that peaks at  $\Delta v = +205 \pm 5$  km s $^{-1}$  and a flatter blue



**Figure 14.** Stacked Ly  $\alpha$  profile of the representative sample of LAEs at  $z \approx 2$  (black). 26 per cent of the flux emerges bluewards of the systemic redshift. The blue line shows the Ly  $\alpha$  profile with an inverse IGM correction from Madau (1995). This slightly increases the flux on the blue-side of the systemic velocity. The red line shows the median Ly  $\alpha$  spectra in low-redshift analogues of LAEs based on *HST/COS* observations (Hayes et al. 2021). We renormalized this spectrum by a factor of 1.5 to have the same integrated Ly  $\alpha$  luminosity. The spectra are remarkably similar, particularly as there is significant variation among the spectra of individual objects (Fig. 7). The relative strength of the red peaks with respect to the blue peaks indicates that Ly  $\alpha$  photons scatter through a mainly outflowing medium.

peak that peaks at  $\Delta v = -294 \pm 5 \text{ km s}^{-1}$ . The observed blue-to-red flux ratio is  $0.26 \pm 0.01$ . As our stacked Ly  $\alpha$  spectrum averages over many independent sightlines, we can retrieve the intrinsic flux on the blue side of the line by correcting for the average IGM attenuation at  $z = 2.2$  following Madau (1995). This correction is small and yields an intrinsic blue-to-red flux ratio of  $0.32 \pm 0.01$ . This ratio is very comparable to the average blue-to-red flux ratio observed by Trainor et al. (2015) in slightly lower resolution spectra of LAEs in quasar fields at  $z \sim 2$ . Fig. 14 also shows the average Ly  $\alpha$  profile of low-redshift analogues of LAEs at  $z < 0.44$  observed with *HST/COS* (Hayes et al. 2021). It is remarkable that the average Ly  $\alpha$  profiles at  $z \approx 2$  and the analogues are so similar, in particular as the individual Ly  $\alpha$  profiles show a lot of scatter on an individual basis (see Fig. 7 and e.g. Yang et al. 2017). The red peak appears to be less sharp and the valley appears to be slightly shallower at  $z \approx 2$  compared to the average spectrum in the low-redshift analogues. The dominant red peak is clear evidence that Ly  $\alpha$  photons scatter through an outflowing medium (e.g. Barnes et al. 2011; Gronke & Dijkstra 2016; Gurung-López, Orsi & Bonoli 2019).

### 6.3 Typical LAEs are probably moderate LyC leakers

Due to the resonant nature of the Ly  $\alpha$  transition, a connection between the escape of LyC and Ly  $\alpha$  photons is theoretically expected. These escape fractions may correlate with each other, as Ly  $\alpha$  photons scatter through the low-column density channel paths of least resistance (Dijkstra et al. 2016) that allow for LyC escape. The lowest column density channels leave an imprint on the Ly  $\alpha$  line profile (e.g. Verhamme et al. 2015; Kakiichi & Gronke 2019). As LAEs are the class of galaxies with the highest Ly  $\alpha$  escape fractions, it is plausible that they are also the types of galaxies with the highest LyC escape fraction (indeed, observations of UV-selected galaxies

indicate that galaxies with higher Ly  $\alpha$  EW have a higher escape fraction; Marchi et al. 2018; Steidel et al. 2018).<sup>5</sup>

As discussed in Section 5.3.2, the high value of O32 in the stack of typical LAEs indicates an ionization parameter  $\log_{10}(U) = -2.4$ . Combined with the gas-phase metallicity, this value indicates that the ISM, on average, is still mostly ionization bounded (Nakajima & Ouchi 2014), and it is not as high as the most extreme LyC leakers that have O32 > 10 (Izotov et al. 2018). This suggests that typical LAEs are not extreme LyC leakers.

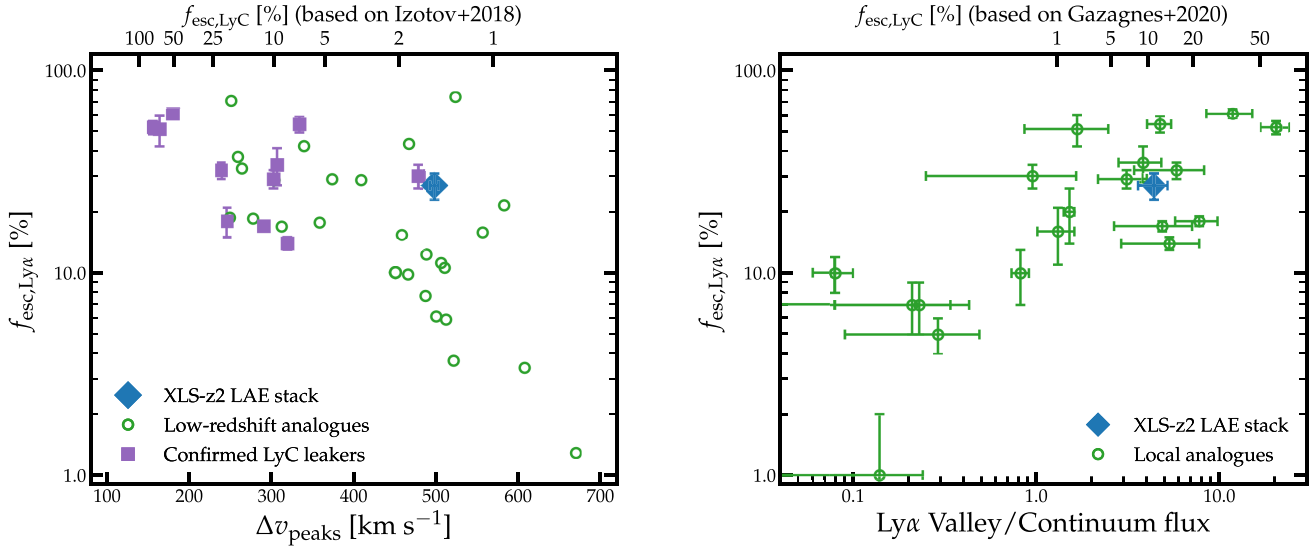
The Ly  $\alpha$  profile is a direct tracer of the H I column density. It is, however, tricky to interpret the stacked Ly  $\alpha$  line profile shown in Fig. 14. The peak separation of the Ly  $\alpha$  lines is quite large with  $498 \pm 5 \text{ km s}^{-1}$ , which suggests an escape fraction of  $f_{\text{esc,LyC}} \approx 2$  per cent based on comparison to the empirical trend identified in Izotov et al. (2018); see the left-hand panel of Fig. 15. However, the broadness of the blue part of the line compared to the broadness on the red part of the line may indicate that there is more variation in the velocity at which blue Ly  $\alpha$  photons escape compared to red Ly  $\alpha$  photons (see also Henry et al. 2015). Due to stacking, the observed ‘peak’ of the blue line is thus the result of a complex interplay of various blue peak positions and strength, and therefore less directly connected to the effective H I column density. Indeed, the left-hand panel of Fig. 15 suggests that the observed peak separation is relatively high given the  $f_{\text{esc,Ly}\alpha}$ .

On the other hand, the stacked Ly  $\alpha$  spectrum also shows clear non-zero flux at the ‘valley’ near the systemic velocity. The level of the flux in the valley traces the number of Ly  $\alpha$  photons that have not experienced significant frequency redistribution and is therefore also correlated strongly with  $f_{\text{esc,LyC}}$  (Vanzella et al. 2016; Rivera-Thorsen et al. 2017; Gazagnes et al. 2020). With a valley flux that is  $4.4 \pm 0.8$  times higher than the continuum, we infer  $f_{\text{esc,LyC}} \approx 11 \pm 6$  per cent (based on the empirical correlation in Gazagnes et al. 2020); see Fig. 15. This measurement may also be influenced by the stacking and the spectral resolution. As the majority of the flux on the red and blue parts of the line is at velocities that are  $>200 \text{ km s}^{-1}$  from the valley ( $>2.5 \times$  the resolution FWHM), the resolution effect is likely small. Therefore, while challenging to interpret, the stacked Ly  $\alpha$  profile suggests a moderate average LyC escape fraction of  $\approx 10$  per cent. This average may be driven by a few galaxies with very low H I column density ( $N_{\text{H I}} < 10^{17} \text{ cm}^{-2}$ ) channels through which a fraction of the LyC photons escape and Ly  $\alpha$  photons experience negligible velocity shift.<sup>6</sup> This average escape fraction is consistent with direct constraints from stacking LyC data of LAEs at  $z \approx 3$  (e.g. Micheva et al. 2017; Iwata et al. 2019).

The line ratio of the resonant Mg II<sub>2796, 2803</sub> doublet has been proposed as tracer of LyC escape (e.g. Chisholm et al. 2020), which is particularly relevant for the epoch of reionization where Ly  $\alpha$  suffers from IGM attenuation. The line ratio of the Mg II lines that we measure in the stack is  $1.95^{+1.82}_{-0.89}$ . While this ratio is uncertain due to the low S/N of the Mg II lines, it can still constrain the escape fraction if we cap this ratio to the intrinsic ratio of 2. Following the methodology from Chisholm et al. (2020), we infer a LyC escape

<sup>5</sup>It is possible that the LyC escape fraction is so high that the production of Ly  $\alpha$  photons is reduced. In practice, however, the Ly  $\alpha$  luminosity of such systems may still be high as the mechanisms that will lead to LyC escape may also allow Ly  $\alpha$  escape. Only extreme LyC escape fractions therefore realistically will lead to decreased Ly  $\alpha$  output.

<sup>6</sup>See, for example, XLS-2, XLS-20, and XLS-21 in Fig. 7. We will investigate these Ly  $\alpha$  spectra in detail in an upcoming paper.



**Figure 15.** The dependence of the Ly  $\alpha$  escape fraction on the line profile. In the left-hand panel, we show the separation of the red and blue peaks. For comparison, we add data points from low-redshift analogues (Yang et al. 2017) and a subset of those that have been confirmed as LyC leakers (Izotov et al. 2018). The Ly  $\alpha$  escape fraction is anticorrelated with the peak separation, although there is significant scatter. Following the empirical relation between the peak separation and the LyC escape fraction, we show the expected escape fraction on the top  $x$ -axis. The XLS- $z2$  stack of LAEs has a relatively high peak separation given the Ly  $\alpha$  escape fraction, which is a stacking effect. In the right-hand panel, we show the flux that escapes at the valley (i.e. the local minimum between the two Ly  $\alpha$  peaks close to the systemic redshift) relative to the continuum. Here, we compare to low-redshift analogues as compiled by Gazagnes et al. (2020). We also show the corresponding LyC escape fraction following the empirical relation from Gazagnes et al. (2020). While the peak separation implies a fairly low LyC escape fraction, the relative valley-to-continuum flux implies  $f_{\text{esc,LyC}} \approx 10$  per cent in typical LAEs at  $z \approx 2$ .

fraction of  $f_{\text{esc,LyC}} = 6_{-3}^{+7}$  per cent for our stack of representative LAEs at  $z \approx 2$ , consistent with our estimate based on Ly  $\alpha$ .

## 7 DISCUSSION

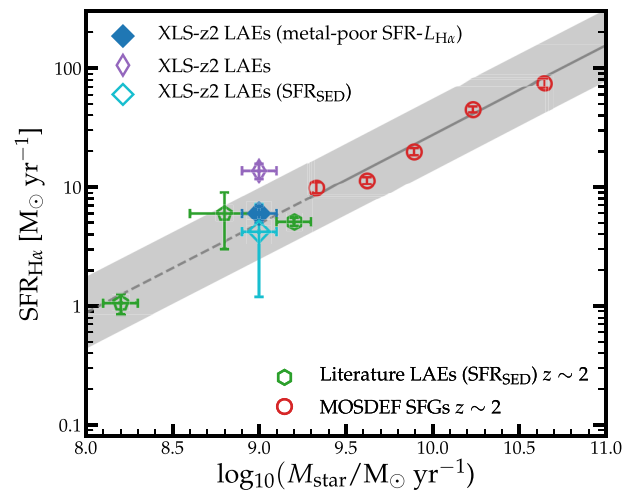
We discuss how LAEs compare to the general galaxy population at  $z \approx 2$  and what this tells us about similar galaxies in the epoch of reionization.

### 7.1 The locations of LAEs on scaling relations

#### 7.1.1 The SFR, mass, and gas-phase metallicity of LAEs

Here, we compare the location of typical LAEs to other galaxies through well-known scaling relations between SFR, mass, and metallicity. In Fig. 16, we show the location of the XLS stack on the SFR– $M_{\text{star}}$  relation. We also show the general relation found at  $z \approx 2.3$  for somewhat more massive systems (Sanders et al. 2020a) and other stacks of LAEs at  $z \sim 2$  (Guaita et al. 2011; Hao et al. 2018; Kusakabe et al. 2018). The relation from Sanders et al. (2020a) uses H  $\alpha$ -based SFRs, while the literature values on stacks of LAEs are based on SED fitting. We show the location of the XLS stack based on the SED fitting, the standard conversion between SFR and H  $\alpha$  luminosity (as also used by Sanders et al. 2020a), and the conversion assuming a lower stellar metallicity (see Section 5.3.1). The H  $\alpha$ -based SFR from the XLS stack is above the extrapolated relation between SFR and  $M_{\text{star}}$  using the standard conversion, but would be perfectly on the extrapolated relation using the metal-poor conversion that may be more realistic. The SED-based SFR places the XLS stack on a similar location as literature studies of LAEs.

A more empirical comparison of the specific SFR of LAEs with respect to the general galaxy population can be made through the rest-frame optical emission-line EWs, which scale with sSFR (in



**Figure 16.** The relation between the SFR and stellar mass at  $z \approx 2.3$ . We show the relation (grey line) for general samples of star-forming galaxies measured from the H  $\alpha$ -based red data points from Sanders et al. (2020a). We assume the typical 0.3 dex scatter on this relation (Matthee & Schaye 2019) shown as a grey band. The blue diamond shows the XLS stack assuming the metal-poor conversion factor between SFR and H  $\alpha$  luminosity, while the open purple diamond shows the standard conversion (that is also used by Sanders et al. 2020a) and the open cyan diamond shows the SED-based SFR. Green pentagons show literature measurements of SFRs of stacks of LAEs using SED fitting (Guaita et al. 2011; Hao et al. 2018; Kusakabe et al. 2018).

particular the Balmer lines). Reddy et al. (2018) provide scaling relations between these EWs and properties such as stellar mass at  $z \sim 2$ . For typical galaxies with masses of  $10^9 M_{\odot}$ , i.e. similar to our stack of LAEs, they find H  $\alpha$  and H  $\beta$  EWs of 270 and 40 Å,

respectively, with a scatter of 0.2 dex. The typical [O III] EW for this mass is 340 Å. These values are a factor of  $\approx 1.5$ –2 lower than the measured EWs in the XLS stack (Table 4), which suggests that LAEs have a slightly enhanced strength of emission lines compared to the continuum compared to galaxies with similar stellar mass, possibly due to an ongoing starburst.

Galaxies at  $z \approx 0$ –3 are observed to follow a tight relation between mass, SFR, and gas-phase metallicity (e.g. Sanders et al. 2018; Maiolino & Mannucci 2019), such that deviations from the relation between mass and SFR correlate with deviations in the mass–metallicity relation. Extrapolating the mass–metallicity relation at  $z \approx 2.3$  from Sanders et al. (2020a), we find a gas-phase metallicity that is 0.4 dex higher than the one measured for the XLS stack (using the  $T_e$  method). This implies that LAEs have relatively low gas-phase metallicities given their stellar mass. Therefore, in case LAEs follow the fundamental metallicity relation, LAEs should have elevated SFRs at a fixed mass.

### 7.1.2 Ionization conditions

As illustrated in Fig. 12, LAEs are among the galaxies with the highest ionization state known at  $z \sim 2$ . This result is in agreement with earlier results (e.g. Finkelstein et al. 2011; Nakajima et al. 2013; Song et al. 2014; Trainor et al. 2016) and also along the lines of those from Erb et al. (2016), who found that the fraction of galaxies with strong Ly  $\alpha$  emission is very high among UV-selected galaxies at  $z \approx 2$  with extreme [O III]/H  $\beta$  ratios, compared to galaxies in their sample with lower values of [O III]/H  $\beta$ . It is interesting that Sanders et al. (2020a) show that similar [O III]/H  $\beta$  ratios as those observed in LAEs are typical for low-mass galaxies at this redshift, while our measured [Ne III]/[O II] is a factor of 2 higher.

Such a high ionization state can plausibly be driven by a harder ionizing spectrum. Topping et al. (2020) show that galaxies in the high [O III]/H  $\beta$  and low [N II]/H  $\alpha$  part of the BPT diagram have young and more metal-poor stellar populations compared to galaxies that share the locus with the SDSS sample. A low stellar metallicity ( $Z_* \approx 10^{-3}$ ) is indeed not unexpected for LAEs. For example, Cullen et al. (2020) report an anticorrelation between Ly  $\alpha$  EW and stellar metallicity. The strength of high ionization UV emission lines also indicates a relatively hard ionizing spectrum in LAEs. Fig. 13 shows that the C IV EW of typical LAEs at  $z \sim 2$  is slightly lower than those seen in the most extreme systems (e.g. Amorín et al. 2017; Nakajima et al. 2018b) and fainter LAEs at  $z \sim 3$  (Nakajima et al. 2018a; Feltre et al. 2020), and comparable with local dwarf galaxies with low metallicities [ $12 + \log(\text{O}/\text{H}) \approx 7.5$ –8.0; Senchyna et al. 2017, 2019; Berg et al. 2019b]. The measured C IV EW from typical galaxies in the VUDS survey at  $z \sim 3$  (Nakajima et al. 2018b) is much lower than that for typical LAEs. Stacks of brighter continuum-selected galaxies at  $z \sim 2$ –3 (Shapley et al. 2003; Rigby et al. 2018) report no C IV in emission. The C IV EWs of typical LAEs are much lower than the EWs of  $\approx 20$ –40 Å reported in extreme LAEs at  $z \sim 7$  (Stark et al. 2015; Mainali et al. 2017).

## 7.2 A minority of low-mass galaxies at $z \sim 2$ are LAEs

We have seen that LAEs at  $z \approx 2$  have relatively low metallicities, but typical sSFRs and rest-frame optical emission-line EWs given their stellar mass. They show particularly high [O III]/H  $\beta$  values, but this result is an indirect consequence of the relatively low mass of LAEs as Sanders et al. (2020b) show that [O III]/H  $\beta$  increases monotonically with decreasing stellar mass. Ly  $\alpha$  surveys typically

pick up galaxies with a low mass (Fig. 16). This partly reflects that the Ly  $\alpha$  escape fraction and typical Ly  $\alpha$  EW anticorrelates with mass (see also Oyarzún et al. 2017; Marchi et al. 2019; Santos et al. 2020), as more massive galaxies with a higher SFR (and thus higher intrinsic Ly  $\alpha$  luminosity) would otherwise be picked up by Ly  $\alpha$  surveys as well (Matthee et al. 2016). In order to understand the relation between LAEs and the galaxy population, the question is thus whether all star-forming galaxies at  $z \approx 2$  with masses around  $10^9 M_\odot$  are LAEs.

As we showed in Section 2.2, the number density of our sample of LAEs is about 15 times lower than general UV-selected systems with similar UV luminosities. In addition, we find that the number density of H  $\beta$  + [O III] line emitters with similar line luminosities as those measured in the XLS stack is  $\approx 10^{-3} \text{ cMpc}^{-3}$  (Khostovan et al. 2015; Matthee et al. 2017b). This is about three times higher than the number density corresponding to the Ly  $\alpha$  luminosity of this stack of  $\approx 3 \times 10^{-4} \text{ cMpc}^{-3}$  (Konno et al. 2016; Sobral et al. 2017). This suggests only about one in every 3 to 10 star-forming galaxies selected either through [O III] line emission or UV emission with comparable luminosity at  $z \approx 2$  is an LAE with an EW of  $> 25$  Å (in agreement with spectroscopic follow-up of UV-selected galaxies at  $z \approx 2$ –3; Cassata et al. 2015). This result is also in agreement with the estimate of Kusakabe et al. (2018), who find that 10 per cent of galaxies with a mass of  $\sim 10^9 M_\odot$  are LAEs based on a clustering analysis at  $z = 2.2$  (see also Ouchi et al. 2010).

## 7.3 What causes the differences in observed Ly $\alpha$ output among samples?

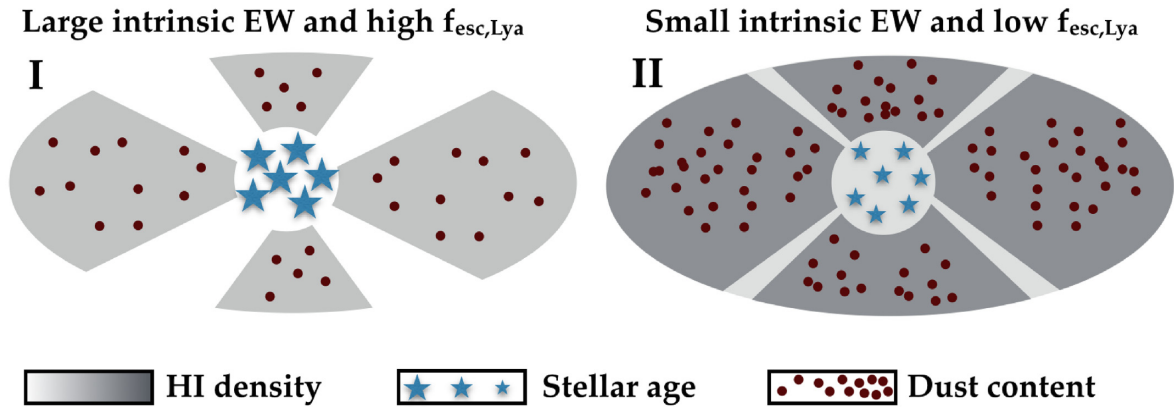
As discussed in the previous section, only about one in every 3 to 10 low-mass galaxies at  $z \approx 2$  is observed as LAE. In this section, we will discuss three possibly complementary hypotheses to explain this (note that these discussions go as far back to at least Charlot & Fall 1993). Fig. 17 sketches the various processes that are invoked in these hypotheses. Compared to other low-mass star-forming galaxies,

- A) do LAEs have a higher intrinsic EW due to a young and/or low metallicity starburst?
- B) do LAEs have a higher escape fraction due to *systematic* variations (for example, a lower dust content or the presence of strong outflows that increase the *angle-averaged* escape fraction)?
- C) do LAEs have a higher escape fraction due to a fortunate viewing angle?

We focus specifically on low-mass galaxies with  $M_{\text{star}} \sim 10^9 M_\odot$ , as it is very plausible that the low escape fraction of Ly  $\alpha$  photons in higher mass galaxies is driving their low Ly  $\alpha$  EWs and explains why massive galaxies are typically not present among samples of LAEs. Systematic variations in their dust content may partly drive this trend (Atek et al. 2008; Garn & Best 2010), but other effects such as a higher H I column density (visible through higher Ly  $\alpha$  velocity offsets in more massive systems; e.g. Steidel et al. 2010), amplifying the effect of variations in the dust content, may also be important.

### 7.3.1 Intrinsic EW?

Compared to other galaxy samples at fixed mass, one possible scenario is that LAEs have a higher intrinsic Ly  $\alpha$  luminosity compared to the UV continuum (i.e. a higher intrinsic EW; see e.g. Trainor et al. 2019). In Fig. 17, this refers to a higher density of very hot stars (i.e. a younger or more metal-poor stellar population).



**Figure 17.** Sketches of two extreme star-forming galaxies with high (I) and low (II) *likelihood* of being observed as an LAE. The orientations are random with respect to an observer in order to illustrate the relative importance of stochasticity of the viewing angle. Various physical processes that influence the observed Ly  $\alpha$  luminosity are illustrated and we stress that the relative importance of these processes remains to be fully understood and that reality is likely between these two examples. Illustrated processes are: (1) the amount of very hot stars (i.e. mostly the age and metallicity of stellar populations) indicated by the size of the blue stars that influence the intrinsic Ly  $\alpha$  EW and luminosity; (2) the amount of dust in the ISM indicated by brown dots that influences the amount of Ly  $\alpha$  photons that are destroyed; (3) the average H I column density and the effective opening angle of low-column density channels that determine the optical depth to scattering and hence impact the likelihood of dust destruction; and (4) depending on the viewing angle, the column density and dust content can vary depending on the specific sightline. Not shown in these sketches are velocity fields of the H I that also impact the Ly  $\alpha$  line profile or the cause of different column density distributions, such as possible turbulence or outflows. Note that (relatively) low-H I column density channels should not necessarily have less dust, but we do not show this for illustrative purposes.

The intrinsic Ly  $\alpha$  EW is related not only to  $\xi_{\text{ion}}$  (e.g. Sobral & Matthee 2019) but also to the EW of rest-frame optical lines. Indeed, in Section 7.1.1 we found that the rest-frame optical EWs in the XLS stack are a factor of  $\approx 1.5$ –2 higher than typical for star-forming galaxies with this mass.

Of particular use is, however, the comparison to the recent studies by Du et al. (2020) and Tang et al. (2020). These studies selected reionization-era analogues at  $z \approx 2$  based on rest-frame optical emission-line properties and measured their Ly  $\alpha$  EWs. Du et al. (2020) selected galaxies with H  $\alpha$  EW  $> 300$  Å and [O III] EW  $> 300$  Å, completed by a sample of blue, UV-selected systems. Tang et al. (2020) selected extreme [O III] emitters with EW  $> 300$  Å. The samples in these studies have a similar median stellar mass as our stack of LAEs. For comparable H  $\beta$  + [O III] EWs as our stack of representative LAEs, these studies report typical Ly  $\alpha$  EWs of  $\approx 10$  Å. These studies also find galaxies with Ly  $\alpha$  EWs as high as 70 Å (i.e. comparable to our stack). However, these Ly  $\alpha$  EWs are mostly seen for galaxies with much stronger rest-frame optical EWs ( $\gtrsim 1000$  Å). This suggests that there is significant scatter on the relation between observed Ly  $\alpha$  and H  $\beta$  + [O III] EW and that LAEs are the relatively rare outliers with high Ly  $\alpha$  flux given an H  $\beta$  + [O III] flux. Moreover, UV-selected galaxies with similar luminosity at  $z \sim 2$  have very comparable  $\xi_{\text{ion}}$  as the XLS stack (Emami et al. 2020). These comparisons suggest that the variations in observed Ly  $\alpha$  EWs among LAEs and other low-mass galaxy samples are not primarily due to significant variations in intrinsic Ly  $\alpha$  EWs and thus related to differences in the Ly  $\alpha$  escape fraction (e.g. Trainor et al. 2019).

### 7.3.2 Systematic variations in $f_{\text{esc,Ly}\alpha}$

The dust attenuation is known to be correlated to the Ly  $\alpha$  escape fraction (e.g. Atek et al. 2008; Blanc et al. 2011; Matthee et al. 2016; Runnholm et al. 2020) and variations in dust attenuation could therefore potentially explain why only a fraction of low-mass star-forming galaxies are observed as an LAE. Variations in the dust

content do not appear to dominate variations in  $f_{\text{esc,Ly}\alpha}$  at a fixed mass. Du et al. (2020) and Tang et al. (2020) estimate  $E(B - V) \approx 0.1$  based on SED fitting, while we measure  $E(B - V) = 0.2$  for LAEs, based on the Balmer decrement. This implies that dust attenuation is not the only physical driver of  $f_{\text{esc,Ly}\alpha}$ . It is possible that the (stellar) attenuation estimated from their SED fitting underestimates the nebular attenuation, but it takes a significant underestimation in order to impact this result.

Additionally, studies of low-redshift analogues show that while  $f_{\text{esc,Ly}\alpha}$  is related to the dust attenuation, the Ly  $\alpha$  peak separation is (anti)correlating equally strongly with the escape fraction (Henry et al. 2015; Yang et al. 2017). The attenuation and the peak separation are not correlated themselves in these systems. The peak separation traces the effective H I column density of the path that the Ly  $\alpha$  photons encountered while escaping the galaxy (e.g. Neufeld 1990; Hashimoto et al. 2015). The amount of low-column density channels (and their specific column density) is therefore an additional important parameter in controlling  $f_{\text{esc,Ly}\alpha}$ . This is illustrated as the variations in the greyscale in Fig. 17.

We speculate that the presence of low-column density channels could be related to the presence of turbulent kinematics or outflows (e.g. Jaskot & Oey 2014; Herenz et al. 2016). Several simulations show that galaxies can be observed as LAE preferentially slightly after starburst events, such that the stellar birth clouds could have cleared by galactic winds (e.g. Kimm et al. 2019; Smith et al. 2019). Even in the very dusty environments of extreme starbursts in the local Universe, strong galactic winds facilitate the escape of Ly  $\alpha$  photons (Martin et al. 2015). Moreover, as discussed in Section 6.2, several observations of LAEs show that turbulent gas and outflows are present in these LAEs. As shown in the simulations by Kimm et al. (2019) the blue-to-red Ly  $\alpha$  flux ratio of typical LAEs is indeed in agreement with outflowing motions around young stars that allow such low-column density pathways to be cleared. Future observations of the presence of such turbulent motions (e.g. Herenz et al. 2016; Puschnig et al. 2020) and outflows in low-mass galaxies that are not strong LAEs could test this scenario.



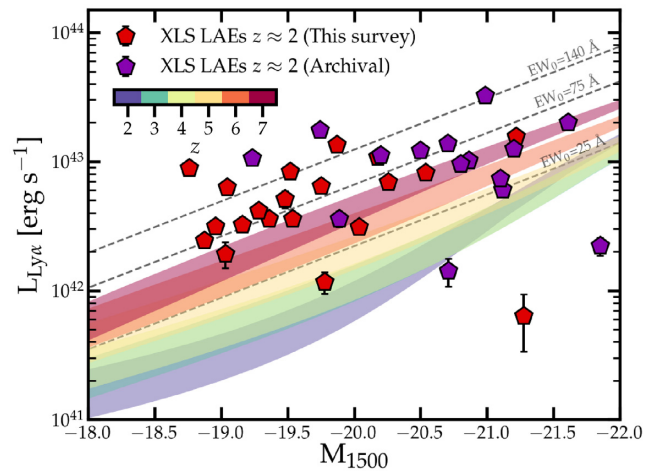
### 7.3.3 Stochastic variations in $f_{\text{esc,Ly}\alpha}$

Besides systematic variations in the dust attenuation or a physical driver of the presence of low-column density channels, the fact that we observe a galaxy as an LAE may also have a stochastic component. This could be the case when the escape fraction depends partly on the viewing angle as found in several simulations (e.g. Behrens & Braun 2014; Zheng & Wallace 2014; Smith et al. 2019). Galaxies may appear as LAEs when we are observing them along a fortunate sightline. Due to resonant scattering, the effective low-column density channels that Ly  $\alpha$  photons traversed are not necessarily single-direction ‘chimneys’ along the line of sight (e.g. Gronke & Dijkstra 2014; Eide et al. 2018; Kimm et al. 2019). Furthermore, the presence of such chimneys can boost the escape fraction significantly more than their fractional opening angle since Ly  $\alpha$  photons preferentially escape through them. Regardless, it could be that the *effective* angular covering factor of such low-column density channels introduces stochastic scatter in the observed values of  $f_{\text{esc,Ly}\alpha}$  and significantly contributes to the observed Ly  $\alpha$  luminosity of a galaxy. This idea is supported somewhat by observations from Shibuya et al. (2014), Paulino-Afonso et al. (2018), and Tang et al. (2020), who all report higher Ly  $\alpha$  EWs in galaxies at  $z \sim 2$  with smaller ellipticity (i.e. face-on systems). The fraction of 10–30 per cent of low-mass star-forming galaxies that are observed as LAE (see Section 2.2) in this picture is effectively an upper limit of the typical angular covering factor of low-column density channels.

Summarizing these various hypotheses with the sketches in Fig. 17, we find that it is plausible that several complementary processes determine why only a fraction of low-mass galaxies are observed as LAE at  $z \approx 2$ . While the dust attenuation needs to be low and there need to be sufficient amounts of young stars to yield a high intrinsic EW, these requirements are likely present in the majority of low-mass galaxies. Additionally, pathways of low column density through which Ly  $\alpha$  photons can escape need to be present. Their presence and properties may be correlated with turbulent motions and outflows powered by feedback, but as such channels do not span the full solid angle, stochasticity of our viewing angle may be induced (e.g. Hagen et al. 2016). Further studies are required to determine the relative importance of each of these effects (e.g. Nakajima et al. 2012; Matthee et al. 2016).

### 7.4 LAEs are rare galaxies at $z \approx 2$ , but representative galaxies at $z \approx 6$

Finally, we discuss how our results are relevant for galaxies in the very early Universe, in particular the epoch of reionization. Similar to the analysis in Section 2.2, we explore which relative Ly  $\alpha$  and UV luminosities are required to match the abundances on the respective luminosity functions. We use UV LFs measured by Bouwens et al. (2015) at  $z \approx 3, 4, 5, 6,$  and  $7$  and the global Ly  $\alpha$  LF at  $z = 3-6$  from Sobral et al. (2018b). The Ly  $\alpha$  LF is relatively constant over this redshift (e.g. Ouchi et al. 2008; Hayes et al. 2011; Sobral et al. 2018b; Herenz et al. 2019), while the normalization of the UV luminosity function decreases and the faint-end slope increases at higher lookback times. As a consequence, the population-averaged Ly  $\alpha$  output and the fraction of strong Ly  $\alpha$  lines among UV-selected galaxies (e.g. Stark, Ellis & Ouchi 2011; Cassata et al. 2015) increase. This may be a result of a higher escape fraction, e.g. due to a lower dust content (e.g. Hayes et al. 2011; Konno et al. 2016) and/or due to a higher intrinsic Ly  $\alpha$  EW (Sobral et al. 2018b) and/or a higher covering factor of low-column density channels.



**Figure 18.** The relation between UV and Ly  $\alpha$  luminosity for the XLS sample of LAEs at  $z \approx 2$ . The coloured shaded regions are derived as in Fig. 1. These show the relation between the UV and Ly  $\alpha$  luminosity at which the number densities of the respective LFs at  $z = 2-7$  are within a factor of 2 of each other. Most of the evolution is seen for fainter galaxies. The steepening faint end of the UV LF with increasing redshift yields a higher typical Ly  $\alpha$  EW for faint galaxies. Dashed lines indicate lines of fixed EW for a given UV slope  $\beta = -2.0$ .

The resulting UV-to-Ly  $\alpha$  luminosity relation that is required to match the abundances is shown as coloured shaded regions in Fig. 18. The evolution of the LFs implies that particularly UV-faint galaxies on average have a higher Ly  $\alpha$  luminosity with increasing redshift. This means that LAEs become more representative of the general galaxy population with increasing redshift (see also Santos et al. 2020). If we assume the constant Ly  $\alpha$  LF could be extrapolated to  $z > 6$  (ignoring any impact of the rapidly increasing opacity from the IGM; e.g. Laursen et al. 2011), we would find that a fully representative sample of the UV galaxy population at  $z \approx 7.5$  has similar Ly  $\alpha$  EWs as the XLS sample. In addition, the measured H  $\beta$  + [O III] EWs, O32 and [O III]/H  $\beta$  ratios of LAEs at  $z \approx 2$  are comparable to values expected in typical galaxies at  $z \approx 7$  (Faisst 2016; De Barros et al. 2019; Endsley et al. 2021). Moreover, the rest-frame UV sizes of LAEs are relatively constant in the population of LAEs over  $z = 2-6$  and these sizes are similar to the sizes of the general population of star-forming galaxies at  $z \approx 7$  (Paulino-Afonso et al. 2018). These comparisons suggest that the properties of the general population of star-forming galaxies in the epoch of reionization resemble the properties of typical LAEs at  $z \approx 2$ . Indeed, the SED modelling indicates light-weighted ages of  $250 \pm 200$  Myr, consistent with plausible ages for galaxies in the very early Universe.

Our discussion in Section 7.3 can also be applied to explain the relative redshift evolution of the Ly  $\alpha$  and UV LFs. It is plausible that the increasing strength of Ly  $\alpha$  compared to the UV continuum can be explained following the build-up of galaxies that constitute the LF over time. Towards higher redshift, the population-averaged dust attenuation will decrease as the fractional contribution of low-mass galaxies is higher, leading to a higher volumetric Ly  $\alpha$  escape fraction (e.g. Hayes et al. 2011). Moreover, galaxies will on average be younger, leading to a higher ionizing photon efficiency and higher intrinsic Ly  $\alpha$  EW (Matthee et al. 2017a; Sobral et al. 2018b). It is, however, also possible that there is, in addition, evolution in the H I column density (i.e. a larger covering factor of low-density channels) or presence of outflows, which has currently not been investigated in a systematic way observationally (cf. Cassata et al. 2020). Such

evolution could have implications not only on the escape fraction of ionizing photons but also on inferences of reionization that rely on assuming the Ly  $\alpha$  profile escaping the ISM of galaxies (e.g. Mason et al. 2018). Future extensions of XLS- $z2$  to  $z = 3-7$ , e.g. by combining blind spectroscopy with VLT/MUSE and *JWST*, that can measure  $f_{\text{esc,Ly } \alpha}$ , the Ly  $\alpha$  profile, and ISM conditions are required to test the relative importance of these various mechanisms.

## 8 SUMMARY

In this paper, we presented the first results of the X-SHOOTER Lyman  $\alpha$  Survey at  $z = 2$  (XLS- $z2$ ) that is a deep spectroscopic survey of 35 LAEs at  $z \approx 2$  covering the rest-frame UV to rest-frame optical ( $\lambda_0 = 0.1-0.7 \mu\text{m}$ ).

The selected LAEs have stellar masses ranging from  $10^{8-10} M_{\odot}$ , Ly  $\alpha$  luminosities  $(0.2-10) \times L_{\text{Ly}\alpha}^*$ , and UV luminosities  $(0.2-6) \times L_{\text{UV}}^*$  and are typically small ( $r_{1/2,\text{UV}} \approx 1$  kpc) and blue ( $\beta \approx -2.0$ ). The rest-frame Ly  $\alpha$  EWs range from  $\approx 1$  to 300 Å, with the majority  $>25$  and 70 Å on average. These galaxies constitute a rare,  $\approx 1/15$ , fraction of the UV-selected galaxy population at  $z \approx 2$  with UV luminosity  $M_{1500} \approx -20 \pm 1$ . Targets were observed with VLT/X-SHOOTER for typically  $\approx 3$  h. We presented the observing strategy, the data reduction and the methodology of extracting 1D spectra.

We measured systemic redshifts in 33 of the 35 targets. Most of these come from the [O III] $_{4960,5008}$  doublet with an FWHM of  $\approx 130$  km s $^{-1}$  and consistent redshifts for H  $\alpha$  are found. In the objects with best S/N, we identify complex features in the [O III] profile, such as multiple narrow or a broad (FWHM  $\approx 280$  km s $^{-1}$ ) component. Systemic redshifts are on average  $205 \pm 5$  km s $^{-1}$  bluewards of the red peak of the Ly  $\alpha$  line.

In this paper, we focused on exploring the stack of 20 *representative* LAEs to establish the various typical properties of LAEs at  $z \approx 2$  and address the question: what makes galaxies LAEs? The stack is shown in Fig. 8 and the measurements and their derived properties are listed in Tables 4 and 5. The results can be summarized as follows:

(i) The stacked spectrum of LAEs is emission line dominated with high Ly  $\alpha$ , H  $\beta$ , [O III], and H  $\alpha$  EWs. The UV continuum is blue ( $\beta = -2.1$ ) and shows several high-ionization emission lines (C IV, He II, and O III) with EWs of  $\approx 1$  Å and C III] and Mg II with EWs of  $\approx 5$  Å.

(ii) The rest-frame optical line ratios show that LAEs are powered by star formation, a picture corroborated by rest-frame UV line ratios. Based on the EWs of H  $\alpha$ , we show that the typical starburst is young (6 Myr for a single burst and 40 Myr for a constant SFH) and the emission-line ratios indicate that the stellar metallicity plausibly is as low as  $Z_{\star} \approx 10^{-3}$  ( $0.07 Z_{\odot}$ ), while the direct-method gas-phase metallicity is  $0.13 Z_{\odot}$ .

(iii) The ISM is characterized by an electron temperature of  $14\,800^{+700}_{-800}$  K, a moderately low attenuation [ $E(B - V)_{\text{gas}} = 0.22 \pm 0.06$ ], and a high ionization state. The gas-phase metallicity is low [ $12 + \log(\text{O}/\text{H})_{\text{TP}} = 7.83^{+0.06}_{-0.05}$ ] and the C/O and N/O abundances [ $\log_{10}(\text{C}/\text{O}) = -0.8^{+0.2}_{-0.2}$ ,  $\log_{10}(\text{N}/\text{O}) < -0.8$ ] are consistent with little chemical enrichment from evolved stars.

(iv) LAEs have a clumpy morphological structure in the rest-frame UV and show turbulent and outflowing kinematics. The broad [O III] components suggest outflows, but we also detect them more unambiguously through blueshifted Si II absorption. The average Ly  $\alpha$  spectrum has a clear double-peaked profile with a blue peak that is 0.3 times the flux of the dominant red peak. This shape suggests that Ly  $\alpha$  photons generally scatter through an outflowing medium while escaping galactic environments.

**Table 5.** Measurements of UV emission and absorption lines in the stack. EWs are all in the rest frame.

Property	Measurement 2D stack
$L_{\text{C IV}1548, 1551}$	$(9.9^{+1.8}_{-1.8}) \times 10^{40}$ erg s $^{-1}$
$L_{\text{He II}1640}$	$(4.9^{+1.4}_{-1.6}) \times 10^{40}$ erg s $^{-1}$
$L_{\text{O III}1661}$	$(2.6^{+1.0}_{-1.1}) \times 10^{40}$ erg s $^{-1}$
$L_{\text{O III}1666}$	$(6.1^{+1.5}_{-1.4}) \times 10^{40}$ erg s $^{-1}$
$L_{[\text{C III}]1907}$	$(10^{+5}_{-5}) \times 10^{40}$ erg s $^{-1}$
$L_{\text{C III}1909}$	$(6^{+3}_{-2}) \times 10^{40}$ erg s $^{-1}$
$L_{\text{Mg II}2796}$	$(6^{+2.2}_{-1.8}) \times 10^{40}$ erg s $^{-1}$
$L_{\text{Mg II}2803}$	$(3^{+1.3}_{-1.4}) \times 10^{40}$ erg s $^{-1}$
$\text{EW}_{\text{C IV}1548, 1551}$	$2.2^{+0.4}_{-0.4}$ Å
$\text{EW}_{\text{He II}1640}$	$1.2^{+0.4}_{-0.4}$ Å
$\text{EW}_{\text{O III}1661}$	$0.6^{+0.2}_{-0.2}$ Å
$\text{EW}_{\text{O III}1666}$	$1.3^{+0.3}_{-0.3}$ Å
$\text{EW}_{[\text{C III}]1907}$	$2.8^{+0.7}_{-1.7}$ Å
$\text{EW}_{\text{C III}1909}$	$1.5^{+0.5}_{-0.4}$ Å
$\text{EW}_{\text{Mg II}2796}$	$6.1^{+3.1}_{-2.0}$ Å
$\text{EW}_{\text{Mg II}2803}$	$2.5^{+1.7}_{-1.3}$ Å
Absorption lines	
$\text{EW}_{\text{Si II}1260}$	$-1.7^{+0.8}_{-0.3}$ Å
$\bar{v}_{\text{Si II}1260}$	$-260^{+160}_{-40}$ km s $^{-1}$
Ly $\alpha$ profile	
$f_{\text{esc,Ly } \alpha}$	$0.27 \pm 0.04$
Blue/Red	$0.26 \pm 0.03$
Valley/Cont	$4.4 \pm 0.8$
$\Delta v_{\text{red}}$	$+205 \pm 5$ km s $^{-1}$
$\Delta v_{\text{blue}}$	$-294 \pm 5$ km s $^{-1}$

(v) We measure a Ly  $\alpha$  escape fraction of  $27 \pm 4$  per cent based on dust-corrected H  $\alpha$  luminosity, consistent with earlier measurements in LAEs. A non-negligible amount of Ly  $\alpha$  photons escape directly at the systemic redshift, which suggests the presence of low-column density channels that also allow the escape of LyC photons. Based on empirical comparisons of the Ly  $\alpha$  and Mg II line profiles with those in low-redshift analogues, we estimate  $f_{\text{esc,LyC}} \approx 10$  per cent.

We use these results to discuss how LAEs compare to the general galaxy population at  $z \approx 2$ . As illustrated in Fig. 17, we identify several mechanisms that are likely important in determining whether galaxies can be observed as LAEs. These include a low dust attenuation, a high intrinsic Ly  $\alpha$  EW due to the presence of young hot stars, but also a low H I column density that may be related to gas turbulence and outflows and possibly the viewing angle.

We argue that the low mass of typical LAEs suggests that dust attenuation prevents the majority of massive galaxies to be bright LAEs. We further show, based on number density analyses, that only a fraction of  $\approx 10$  per cent of low-mass galaxies is observed as LAE at  $z \approx 2$ , consistent with duty cycles inferred from clustering analyses. LAEs appear to have a similar SFR and attenuation as typical galaxies with their mass, suggesting that effective H I column density may mostly determine whether we observe low-mass galaxies as LAEs. It is plausible that low H I column densities may be related to systematic changes in gas turbulence or outflows, but stochastic differences such as the specific viewing angle may be equally or more important. In future work, we will investigate the relative importance of the various processes by exploring the variation within the XLS sample.

Finally, we discuss that while LAEs constitute a relatively rare sample of low-mass galaxies at  $z \approx 2$ , galaxies that resemble LAEs become increasingly more common at higher redshifts. Indeed, galaxies with these Ly  $\alpha$  to UV luminosity ratios will be the norm in the epoch of reionization even if the neutral IGM may prevent us realizing that.

It is of extreme importance in understanding the early phases of galaxy formation and the late phases of the epoch of reionization to verify whether this evolution is a mere effect of the evolving stellar mass function (and the related evolution of the typical dust attenuation) or whether an increasing Ly  $\alpha$  output or higher occurrence of low-column density H I gas contributes to the observed global evolution. In upcoming XLS- $z2$  papers, we will focus on the variation in the Ly  $\alpha$  profiles and how these are related to the production and escape of Ly  $\alpha$  and LyC photons.

## ACKNOWLEDGEMENTS

We thank the referee for constructive comments and suggestions. We thank Dawn Erb, Ruari Mackenzie, Ivan Oteo, Ryan Sanders, and Johannes Zabl for useful discussions and suggestions. It is a pleasure to thank the ESO User Support, in particular Giacomo Beccari, Carlo Manara, John Pritchard, Marina Rejkuba, and Lowell Tacconi-Garman for assistance in the preparation and execution of the observations. Based on observations obtained with the VLT, programs 084.A-0303, 088.A-0672, 091.A-0413, 091.A-0546, 092.A-0774, 097.A-0153, 098.A-0819, 099.A-0758, 099.A-0254, 101.B-0779, and 102.A-0652. Based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under ESO programme ID 179.A-2005 and on data products produced by CALET and the Cambridge Astronomy Survey Unit on behalf of the UltraVISTA consortium. Based on observations made with the NASA/ESA *HST* through programs 9133, 9367, 11694, and 12471, and obtained from the Hubble Legacy Archive, which is a collaboration between the Space Telescope Science Institute (STScI/NASA), the Space Telescope European Coordinating Facility (ST-ECF/ESA), and the Canadian Astronomy Data Centre (CADM/NRC/CSA). This work is based on observations taken by the CANDELS Multi-Cycle Treasury Program with the NASA/ESA *HST*, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. MG was supported by NASA through the NASA Hubble Fellowship grant *HST*-HF2-51409 and acknowledges support from *HST* grants *HST*-GO-15643.017-A, *HST*-AR-15039.003-A, and XSEDE grant TG-AST180036. GP acknowledges support from the Netherlands Research School for Astronomy (NOVA). RA acknowledges the support of ANID FONDECYT Regular Grant 1202007. We gratefully acknowledge the PYTHON programming language, its NUMPY, MATPLOTLIB, SCIPY, LMFIT (Jones et al. 2001; Hunter 2007; van der Walt, Colbert & Varoquaux 2011), PANDAS (McKinney 2010), and ASTROPY (Astropy Collaboration 2013) packages, and the TOPCAT analysis tool (Taylor 2013). Dedicated to the memory of A. C. J. Matthee (1953–2020).

## DATA AVAILABILITY

The data underlying this article were accessed from the ESO archive. The raw ESO data can be accessed through <http://archive.eso.org/cms.html>. The derived data generated in this research will be shared on reasonable request to the corresponding author.

## REFERENCES

- Aihara H. et al., 2019, *PASJ*, 71, 114  
 Alam S. et al., 2015, *ApJS*, 219, 12  
 Amorín R., Pérez-Montero E., Vílchez J. M., Papaderos P., 2012a, *ApJ*, 749, 185  
 Amorín R., Vílchez J. M., Hägele G. F., Firpo V., Pérez-Montero E., Papaderos P., 2012b, *ApJ*, 754, L22  
 Amorín R. et al., 2017, *Nat. Astron.*, 1, 0052  
 Astropy Collaboration, 2013, *A&A*, 558, A33  
 Atek H., Kunth D., Hayes M., Östlin G., Mas-Hesse J. M., 2008, *A&A*, 488, 491  
 Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, 93, 5  
 Barnes L. A., Haehnelt M. G., Tesfari E., Viel M., 2011, *MNRAS*, 416, 1723  
 Behrens C., Braun H., 2014, *A&A*, 572, A74  
 Berg D. A., Erb D. K., Auger M. W., Pettini M., Brammer G. B., 2018, *ApJ*, 859, 164  
 Berg D. A., Erb D. K., Henry R. B. C., Skillman E. D., McQuinn K. B. W., 2019a, *ApJ*, 874, 93  
 Berg D. A., Chisholm J., Erb D. K., Pogge R., Henry A., Olivier G. M., 2019b, *ApJ*, 878, L3  
 Bian F., Kewley L. J., Dopita M. A., 2018, *ApJ*, 859, 175  
 Blanc G. A. et al., 2011, *ApJ*, 736, 31  
 Bosch G. et al., 2019, *MNRAS*, 489, 1787  
 Bouwens R. J. et al., 2015, *ApJ*, 803, 34  
 Bouwens R. J., Smit R., Labbé I., Franx M., Caruana J., Oesch P., Stefanon M., Rasappu N., 2016, *ApJ*, 831, 176  
 Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, *MNRAS*, 351, 1151  
 Brinchmann J., Pettini M., Charlot S., 2008, *MNRAS*, 385, 769  
 Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000  
 Byrohl C., Gronke M., 2020, *A&A*, 642, L16  
 Calhau J. et al., 2020, *MNRAS*, 493, 3341  
 Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, 533, 682  
 Cantalupo S., Lilly S. J., Haehnelt M. G., 2012, *MNRAS*, 425, 1992  
 Cardamone C. N. et al., 2010, *ApJS*, 189, 270  
 Cardelli J. A., Clayton G. C., Mathis J. S., 1989, *ApJ*, 345, 245  
 Cassata P. et al., 2015, *A&A*, 573, A24  
 Cassata P. et al., 2020, *A&A*, 643, A6  
 Chabrier G., 2003, *PASP*, 115, 763  
 Charlot S., Fall S. M., 1993, *ApJ*, 415, 580  
 Charlot S., Fall S. M., 2000, *ApJ*, 539, 718  
 Chisholm J., Rigby J. R., Bayliss M., Berg D. A., Dahle H., Gladders M., Sharon K., 2019, *ApJ*, 882, 182  
 Chisholm J., Prochaska J. X., Schaerer D., Gazagnes S., Henry A., 2020, *MNRAS*, 498, 2554  
 Cornachione M. A. et al., 2018, *ApJ*, 853, 148  
 Cullen F. et al., 2020, *MNRAS*, 495, 1501  
 da Cunha E., Charlot S., Elbaz D., 2008, *MNRAS*, 388, 1595  
 Davies R. L. et al., 2019, *ApJ*, 873, 122  
 De Barros S., Oesch P. A., Labbé I., Stefanon M., González V., Smit R., Bouwens R. J., Illingworth G. D., 2019, *MNRAS*, 489, 2355  
 Dijkstra M., 2014, *Publ. Astron. Soc. Aust.*, 31, e040  
 Dijkstra M., Haiman Z., Spaans M., 2006, *ApJ*, 649, 14  
 Dijkstra M., Wyithe S., Haiman Z., Mesinger A., Pentericci L., 2014, *MNRAS*, 440, 3309  
 Dijkstra M., Gronke M., Venkatesan A., 2016, *ApJ*, 828, 71  
 Dopita M. A. et al., 2006, *ApJ*, 647, 244  
 Drake A. B. et al., 2017, *A&A*, 608, A6  
 Du X., Shapley A. E., Tang M., Stark D. P., Martin C. L., Mobasher B., Topping M. W., Chevallard J., 2020, *ApJ*, 890, 65  
 Eide M. B., Gronke M., Dijkstra M., Hayes M., 2018, *ApJ*, 856, 156  
 Emami N., Siana B., Alavi A., Gburek T., Freeman W. R., Richard J., Weisz D. R., Stark D. P., 2020, *ApJ*, 895, 116  
 Endsley R., Stark D. P., Chevallard J., Charlot S., 2021, *MNRAS*, 500, 5229  
 Erb D. K., 2015, *Nature*, 523, 169

- Erb D. K., Steidel C. C., Shapley A. E., Pettini M., Reddy N. A., Adelberger K. L., 2006a, *ApJ*, 646, 107
- Erb D. K., Steidel C. C., Shapley A. E., Pettini M., Reddy N. A., Adelberger K. L., 2006b, *ApJ*, 647, 128
- Erb D. K., Pettini M., Shapley A. E., Steidel C. C., Law D. R., Reddy N. A., 2010, *ApJ*, 719, 1168
- Erb D. K., Pettini M., Steidel C. C., Strom A. L., Rudie G. C., Trainor R. F., Shapley A. E., Reddy N. A., 2016, *ApJ*, 830, 52
- Faisst A. L., 2016, *ApJ*, 829, 99
- Faisst A. L., Capak P. L., Emami N., Tacchella S., Larson K. L., 2019, *ApJ*, 884, 133
- Feltre A., Charlot S., Gutkin J., 2016, *MNRAS*, 456, 3354
- Feltre A. et al., 2018, *A&A*, 617, A62
- Feltre A. et al., 2020, *A&A*, 641, A118
- Finkelstein S. L. et al., 2011, *ApJ*, 729, 140
- Finkelstein S. L. et al., 2013, *Nature*, 502, 524
- Freeman W. R. et al., 2019, *ApJ*, 873, 102
- Furusawa H. et al., 2008, *ApJS*, 176, 1
- Gaia Collaboration, 2018, *A&A*, 616, A1
- Garn T., Best P. N., 2010, *MNRAS*, 409, 421
- Gawiser E. et al., 2007, *ApJ*, 671, 278
- Gazagnes S., Chisholm J., Schaerer D., Verhamme A., Izotov Y., 2020, *A&A*, 639, A85
- Götberg Y., de Mink S. E., Groh J. H., Leitherer C., Norman C., 2019, *A&A*, 629, A134
- Gräfenor G., Vink J. S., 2015, *A&A*, 578, L2
- Grogin N. A. et al., 2011, *ApJS*, 197, 35
- Gronke M., Dijkstra M., 2014, *MNRAS*, 444, 1095
- Gronke M., Dijkstra M., 2016, *ApJ*, 826, 14
- Gronke M., Dijkstra M., McCourt M., Oh S. P., 2017, *A&A*, 607, A71
- Gronwall C. et al., 2007, *ApJ*, 667, 79
- Guaita L. et al., 2011, *ApJ*, 733, 114
- Guo Y. et al., 2013, *ApJS*, 207, 24
- Gurung-López S., Orsi Á. A., Bonoli S., 2019, *MNRAS*, 490, 733
- Hagen A. et al., 2016, *ApJ*, 817, 79
- Hainline K. N., Shapley A. E., Greene J. E., Steidel C. C., 2011, *ApJ*, 733, 31
- Hao C.-N., Huang J.-S., Xia X., Zheng X., Jiang C., Li C., 2018, *ApJ*, 864, 145
- Hashimoto T. et al., 2015, *ApJ*, 812, 157
- Häussler B. et al., 2007, *ApJS*, 172, 615
- Hayes M., 2015, *PASA*, 32, e027
- Hayes M. et al., 2010, *Nature*, 464, 562
- Hayes M., Schaerer D., Östlin G., Mas-Hesse J. M., Atek H., Kunth D., 2011, *ApJ*, 730, 8
- Hayes M. J., Runnholm A., Gronke M., Scarlata C., 2021, *ApJ*, 908, 36
- Heckman T. M., Armus L., Miley G. K., 1990, *ApJS*, 74, 833
- Heckman T. M., Sembach K. R., Meurer G. R., Leitherer C., Calzetti D., Martin C. L., 2001, *ApJ*, 558, 56
- Henry A., Scarlata C., Martin C. L., Erb D., 2015, *ApJ*, 809, 19
- Herenz E. C. et al., 2016, *A&A*, 587, A78
- Herenz E. C. et al., 2019, *A&A*, 621, A107
- Hoag A. et al., 2019, *MNRAS*, 488, 706
- Hogarth L. et al., 2020, *MNRAS*, 494, 3541
- Horne K., 1986, *PASP*, 98, 609
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Ilbert O. et al., 2009, *ApJ*, 690, 1236
- Inoue A. K., 2001, *AJ*, 122, 1788
- Inoue A. K., Shimizu I., Iwata I., Tanaka M., 2014, *MNRAS*, 442, 1805
- Iwata I., Inoue A. K., Micheva G., Matsuda Y., Yamada T., 2019, *MNRAS*, 488, 5671
- Izotov Y. I., Worseck G., Schaerer D., Guseva N. G., Thuan T. X., Fricke Verhamme A., Orlitová I., 2018, *MNRAS*, 478, 4851
- Jarvis M. J. et al., 2013, *MNRAS*, 428, 1281
- Jaskot A. E., Oey M. S., 2013, *ApJ*, 766, 91
- Jaskot A. E., Oey M. S., 2014, *ApJ*, 791, L19
- Jaskot A. E., Dowd T., Oey M. S., Scarlata C., McKinney J., 2019, *ApJ*, 885, 96
- Jones E. et al., 2001, SciPy: Open Source Scientific Tools for Python. Available at: <http://www.scipy.org/>
- Jung I. et al., 2018, *ApJ*, 864, 103
- Kakiichi K., Gronke M., 2019, *AJ*, 908, 22
- Kashikawa N. et al., 2011, *ApJ*, 734, 119
- Kashino D. et al., 2013, *ApJ*, 777, L8
- Kashino D. et al., 2017, *ApJ*, 835, 88
- Kauffmann G. et al., 2003, *MNRAS*, 346, 1055
- Kausch W. et al., 2015, *A&A*, 576, A78
- Keenan F. P., Feibelman W. A., Berrington K. A., 1992, *ApJ*, 389, 443
- Kehrig C., Vílchez J. M., Pérez-Montero E., Iglesias-Páramo J., Brinchmann J., Kunth D., Durret F., Bayo F. M., 2015, *ApJ*, 801, L28
- Kennicutt R. C., Evans N. J., 2012, *ARA&A*, 50, 531
- Kewley L. J., Dopita M. A., 2002, *ApJS*, 142, 35
- Khostovan A. A., Sobral D., Mobasher B., Best P. N., Smail I., Stott J. P., Hemmati S., Nayyeri H., 2015, *MNRAS*, 452, 3948
- Kimm T., Blaizot J., Garel T., Michel-Dansac L., Katz H., Rosdahl J., Verhamme A., Haehnelt M., 2019, *MNRAS*, 486, 2215
- Koekemoer A. M. et al., 2007, *ApJS*, 172, 196
- Konno A., Ouchi M., Nakajima K., Duval F., Kusakabe H., Ono Y., Shimasaku K., 2016, *ApJ*, 823, 20
- Konno A. et al., 2018, *PASJ*, 70, S16
- Kulas K. R., Shapley A. E., Kollmeier J. A., Zheng Z., Steidel C. C., Hainline K. N., 2012, *ApJ*, 745, 33
- Kusakabe H. et al., 2018, *PASJ*, 70, 4
- Laigle C. et al., 2016, *ApJS*, 224, 24
- Laursen P., Sommer-Larsen J., Razoumov A. O., 2011, *ApJ*, 728, 52
- Lawrence A. et al., 2007, *MNRAS*, 379, 1599
- Leitherer C. et al., 1999, *ApJS*, 123, 3
- Levesque E. M., Richardson M. L. A., 2014, *ApJ*, 780, 100
- Ma X., Hopkins P. F., Kasen D., Quataert E., Faucher-Giguère C.-A., Kereš D., Murray N., Strom A., 2016, *MNRAS*, 459, 3614
- Madau P., 1995, *ApJ*, 441, 18
- Mainali R., Kollmeier J. A., Stark D. P., Simcoe R. A., Walth G., Newman A. B., Miller D. R., 2017, *ApJ*, 836, L14
- Maiolino R., Mannucci F., 2019, *A&AR*, 27, 3
- Marchi F. et al., 2018, *A&A*, 614, A11
- Marchi F. et al., 2019, *A&A*, 631, A19
- Marino R. A. et al., 2018, *ApJ*, 859, 53
- Martin C. L., Dijkstra M., Henry A., Soto K. T., Danforth C. W., Wong J., 2015, *ApJ*, 803, 6
- Mas-Ribas L., Hennawi J. F., Dijkstra M., Davies F. B., Stern J., Rix H.-W., 2017, *ApJ*, 846, 11
- Maseda M. V. et al., 2018, *ApJ*, 865, L1
- Maseda M. V. et al., 2020, *MNRAS*, 493, 5120
- Mason C. A., Treu T., Dijkstra M., Mesinger A., Trenti M., Pentericci L., de Barros S., Vanzella E., 2018, *ApJ*, 856, 2
- Matthee J., Schaye J., 2018, *MNRAS*, 479, L34
- Matthee J., Schaye J., 2019, *MNRAS*, 484, 915
- 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., Gronke M., Paulino-Afonso A., Stefanon M., Röttgering H., 2018, *A&A*, 619, A136
- McCracken H. J. et al., 2012, *A&A*, 544, A156
- McKinney W., 2010, in van der Walt S., Millman J., eds, Proc. 9th Python Sci. Conf. p. 56
- McKinney J. H., Jaskot A. E., Oey M. S., Yun M. S., Dowd T., Lowenthal J. D., 2019, *ApJ*, 874, 52
- Mehta V. et al., 2018, *ApJS*, 235, 36
- Micheva G., Iwata I., Inoue A. K., Matsuda Y., Yamada T., Hayashino T., 2017, *MNRAS*, 465, 316

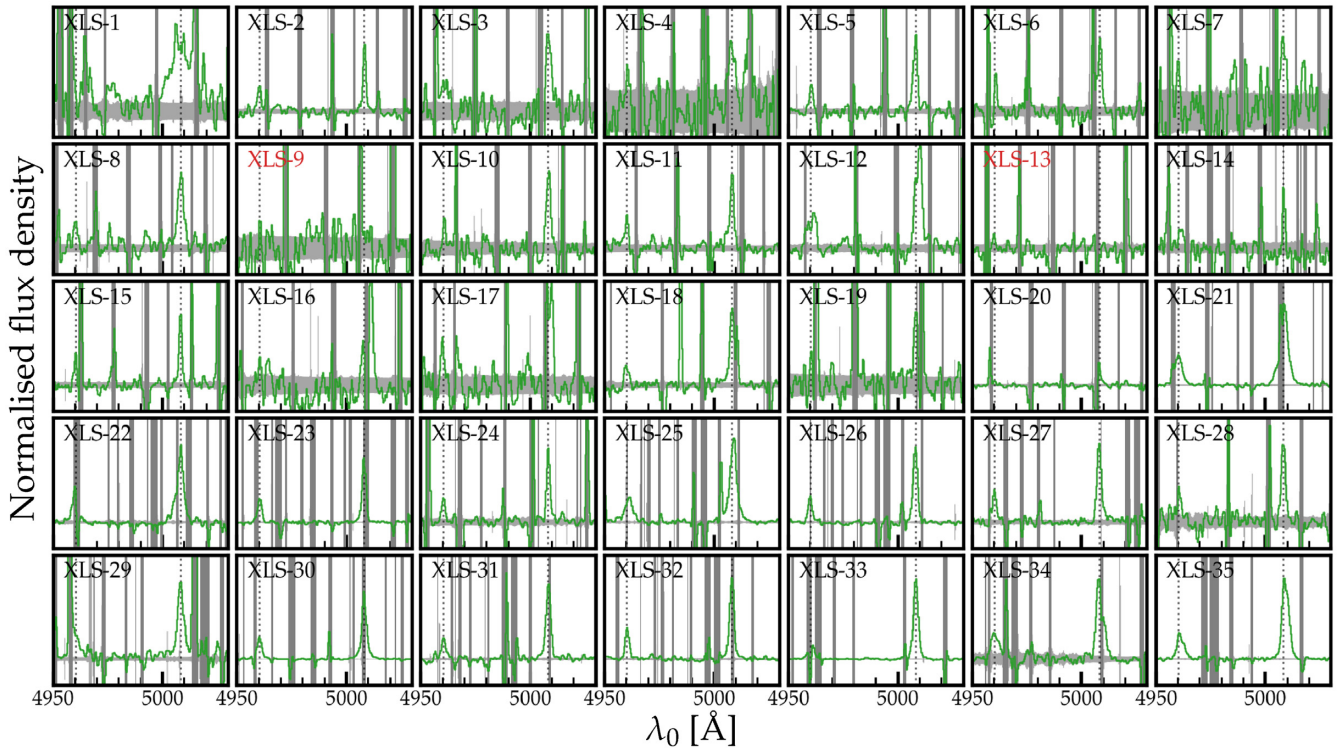
- Modigliani A. et al., 2010, in *Observatory Operations: Strategies, Processes, and Systems III*, Vol. 7737, Proceedings of the SPIE. p. 773728
- Murphy E. J. et al., 2011, *ApJ*, 737, 67
- Muzahid S. et al., 2020, *MNRAS*, 496, 1013
- Naidu R. P. et al., 2017, *ApJ*, 847, 12
- Naidu R. P., Tacchella S., Mason C. A., Bose S., Oesch P. A., Conroy C., 2020, *ApJ*, 892, 109
- Nakajima K., Ouchi M., 2014, *MNRAS*, 442, 900
- Nakajima K. et al., 2012, *ApJ*, 745, 12
- Nakajima K., Ouchi M., Shimasaku K., Hashimoto T., Ono Y., Lee J. C., 2013, *ApJ*, 769, 3
- Nakajima K., Fletcher T., Ellis R. S., Robertson B. E., Iwata I., 2018a, *MNRAS*, 477, 2098
- Nakajima K. et al., 2018b, *A&A*, 612, A94
- Nanayakkara T. et al., 2019, *A&A*, 624, A89
- Neufeld D. A., 1990, *ApJ*, 350, 216
- Nilsson K. K., Tapken C., Møller P., Freudling W., Fynbo J. P. U., Meisenheimer K., Laursen P., Östlin G., 2009, *A&A*, 498, 13
- Nomoto K., Tominaga N., Umeda H., Kobayashi C., Maeda K., 2006, *Nucl. Phys. A*, 777, 424
- Oesch P. A. et al., 2015, *ApJ*, 804, L30
- Oesch P. A. et al., 2018, *ApJS*, 237, 12
- Osterbrock D. E., 1989, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*, University Science Books, ISBN 0-935702-22-9, p. 408
- Östlin G., Marquart T., Cumming R. J., Fathi K., Bergvall N., Adamo A., Amram P., Hayes M., 2015, *A&A*, 583, A55
- Ouchi M. et al., 2008, *ApJS*, 176, 301
- Ouchi M. et al., 2010, *ApJ*, 723, 869
- Oyarzún G. A., Blanc G. A., González V., Mateo M., Bailey John I. I., 2017, *ApJ*, 843, 133
- Parsa S., Dunlop J. S., McLure R. J., Mortlock A., 2016, *MNRAS*, 456, 3194
- Partridge R. B., Peebles P. J. E., 1967, *ApJ*, 147, 868
- Patrício V. et al., 2016, *MNRAS*, 456, 4191
- Paulino-Afonso A. et al., 2018, *MNRAS*, 476, 5479
- Peña-Guerrero M. A., Leitherer C., 2013, *AJ*, 146, 158
- Pérez-Montero E., 2014, *MNRAS*, 441, 2663
- Pérez-Montero E., Amorín R., 2017, *MNRAS*, 467, 1287
- Plat A., Charlot S., Bruzual G., Feltre A., Vidal-García A., Morisset C., Chevallard J., Todt H., 2019, *MNRAS*, 490, 978
- Puschnig J. et al., 2020, *A&A*, 644, A10
- Raiter A., Schaerer D., Fosbury R. A. E., 2010, *A&A*, 523, A64
- Reddy N. A. et al., 2015, *ApJ*, 806, 259
- Reddy N. A., Steidel C. C., Pettini M., Bogosavljević M., 2016a, *ApJ*, 828, 107
- Reddy N. A., Steidel C. C., Pettini M., Bogosavljević M., Shapley A. E., 2016b, *ApJ*, 828, 108
- Reddy N. A. et al., 2018, *ApJ*, 869, 92
- Reddy N. A. et al., 2020, *ApJ*, 902, 123
- Rhoads J. E., Malhotra S., Dey A., Stern D., Spinrad H., Jannuzi B. T., 2000, *ApJ*, 545, L85
- Rhoads J. E., Malhotra S., Richardson M. L. A., Finkelstein S. L., Fynbo J. P. U., McLinden E. M., Tilvi V. S., 2014, *ApJ*, 780, 20
- Ribeiro B. et al., 2020, preprint ([arXiv:2007.01322](https://arxiv.org/abs/2007.01322))
- Rigby J. R. et al., 2018, *ApJ*, 853, 87
- Rigby J. R. et al., 2021, *ApJ*, 908, 154
- Rivera-Thorsen T. E. et al., 2015, *ApJ*, 805, 14
- Rivera-Thorsen T. E. et al., 2017, *A&A*, 608, L4
- Rix H.-W. et al., 2004, *ApJS*, 152, 163
- Robertson B. E. et al., 2013, *ApJ*, 768, 71
- Runnholm A., Hayes M., Melinder J., Rivera-Thorsen E., Östlin G., Cannon J., Kunth D., 2020, *ApJ*, 892, 48
- Sanders R. L. et al., 2016, *ApJ*, 816, 23
- Sanders R. L. et al., 2018, *ApJ*, 858, 99
- Sanders R. L. et al., 2020a, preprint ([arXiv:2009.07292](https://arxiv.org/abs/2009.07292))
- Sanders R. L. et al., 2020b, *MNRAS*, 491, 1427
- Santos S. et al., 2020, *MNRAS*, 493, 141
- Saxena A. et al., 2020, *A&A*, 636, A47
- Scarlata C. et al., 2009, *ApJ*, 704, L98
- Schaerer D., 2003, *A&A*, 397, 527
- Schaerer D., de Barros S., 2009, *A&A*, 502, 423
- Schaerer D., Fragos T., Izotov Y. I., 2019, *A&A*, 622, L10
- Senchyna P. et al., 2017, *MNRAS*, 472, 2608
- Senchyna P., Stark D. P., Chevallard J., Charlot S., Jones T., Vidal-García A., 2019, *MNRAS*, 488, 3492
- Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, *ApJ*, 588, 65
- Shapley A. E. et al., 2015, *ApJ*, 801, 88
- Shibuya T. et al., 2014, *ApJ*, 788, 74
- Shibuya T., Ouchi M., Harikane Y., Nakajima K., 2019, *ApJ*, 871, 164
- Shirazi M., Brinchmann J., Rahmati A., 2014, *ApJ*, 787, 120
- Shivaei I. et al., 2020, *ApJ*, 899, 117
- Smette A. et al., 2015, *A&A*, 576, A77
- Smith A., Ma X., Bromm V., Finkelstein S. L., Hopkins P. F., Faucher-Giguère C.-A., Kereš D., 2019, *MNRAS*, 484, 39
- Sobral D., Matthee J., 2019, *A&A*, 623, A157
- Sobral D., Smail I., Best P. N., Geach J. E., Matsuda Y., Stott J. P., Cirasuolo M., Kurk J., 2013, *MNRAS*, 428, 1128
- Sobral D., Matthee J., Darvish B., Schaerer D., Mobasher B., Röttgering H. J. A., Santos S., Hemmati S., 2015, *ApJ*, 808, 139
- Sobral D. et al., 2017, *MNRAS*, 466, 1242
- Sobral D., Santos S., Matthee J., Paulino-Afonso A., Ribeiro B., Calhau J., Khostovan A. A., 2018a, *MNRAS*, 476, 4725
- Sobral D. et al., 2018b, *MNRAS*, 477, 2817
- Song M. et al., 2014, *ApJ*, 791, 3
- Stanway E. R., Eldridge J. J., 2018, *MNRAS*, 479, 75
- Stanway E. R., Chrimes A. A., Eldridge J. J., Stevance H. F., 2020, *MNRAS*, 495, 4605
- Stark D. P., Ellis R. S., Chiu K., Ouchi M., Bunker A., 2010, *MNRAS*, 408, 1628
- Stark D. P., Ellis R. S., Ouchi M., 2011, *ApJ*, 728, L2
- Stark D. P. et al., 2015, *MNRAS*, 454, 1393
- Stasińska G., Izotov Y., Morisset C., Guseva N., 2015, *A&A*, 576, A83
- Steidel C. C., Shapley A. E., Pettini M., Adelberger K. L., Erb D. K., Reddy N. A., Hunt M. P., 2004, *ApJ*, 604, 534
- 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. et al., 2014, *ApJ*, 795, 165
- Steidel C. C., Strom A. L., Pettini M., Rudie G. C., Reddy N. A., Trainor R. F., 2016, *ApJ*, 826, 159
- Steidel C. C., Bogosavljević M., Shapley A. E., Reddy N. A., Rudie G. C., Pettini M., Trainor R. F., Strom A. L., 2018, *ApJ*, 869, 123
- Strom A. L., Steidel C. C., Rudie G. C., Trainor R. F., Pettini M., Reddy N. A., 2017, *ApJ*, 836, 164
- Strom A. L., Steidel C. C., Rudie G. C., Trainor R. F., Pettini M., 2018, *ApJ*, 868, 117
- Szécsi D., Langer N., Yoon S.-C., Sanyal D., de Mink S., Evans C. J., Dermine T., 2015, *A&A*, 581, A15
- Tang M., Stark D., Chevallard J., Charlot S., Endsley R., Congiu E., 2020, *MNRAS*, 503, 4105
- Taylor M., 2013, Starlink User Note, 253, Available at: <https://ui.adsabs.harvard.edu/abs/2013StaUN.253....T/abstract>
- Taylor A. J., Barger A. J., Cowie L. L., Hu E. M., Songaila A., 2020, *ApJ*, 895, 132
- Terlevich R., Terlevich E., Melnick J., Chávez R., Plionis M., Bresolin F., Basilakos S., 2015, *MNRAS*, 451, 3001
- Theios R. L., Steidel C. C., Strom A. L., Rudie G. C., Trainor R. F., Reddy N. A., 2019, *ApJ*, 871, 128
- Topping M. W., Shapley A. E., Reddy N. A., Sanders R. L., Coil A. L., Kriek M., Mobasher B., Siana B., 2020, *MNRAS*, 495, 4430
- 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
- Trainor R. F., Strom A. L., Steidel C. C., Rudie G. C., Chen Y., Theios R. L., 2019, *ApJ*, 887, 85
- van der Walt S., Colbert S. C., Varoquaux G., 2011, *Comput. Sci. Eng.*, 13, 22
- van der Wel A. et al., 2012, *ApJS*, 203, 24

Vanzella E. et al., 2016, *ApJ*, 821, L27  
 Vanzella E. et al., 2018, *MNRAS*, 476, L15  
 Veilleux S., Cecil G., Bland-Hawthorn J., 2005, *ARA&A*, 43, 769  
 Verhamme A. et al., 2018, *MNRAS*, 478, L60  
 Verhamme A., Schaerer D., Maselli A., 2006, *A&A*, 460, 397  
 Verhamme A., Orlitová I., Schaerer D., Hayes M., 2015, *A&A*, 578, A7  
 Vernet J. et al., 2011, *A&A*, 536, A105  
 Vidal-García A., Charlot S., Bruzual G., Hubeny I., 2017, *MNRAS*, 470, 3532  
 Westmoquette M. S., Exter K. M., Smith L. J., Gallagher J. S., 2007, *MNRAS*, 381, 894  
 Wofford A., Vidal-García A., Feltre A., Chevillard J., Charlot S., Stark D. P., Herenz E. C., Hayes M., 2020, *MNRAS*, 500, 2908  
 Xiao L., Stanway E. R., Eldridge J. J., 2018, *MNRAS*, 477, 904  
 Yang H. et al., 2017, *ApJ*, 844, 171  
 Zheng Z., Wallace J., 2014, *ApJ*, 794, 116  
 Zheng Z.-Y. et al., 2017, *ApJ*, 842, L22  
 Zitrin A. et al., 2015, *ApJ*, 810, L12

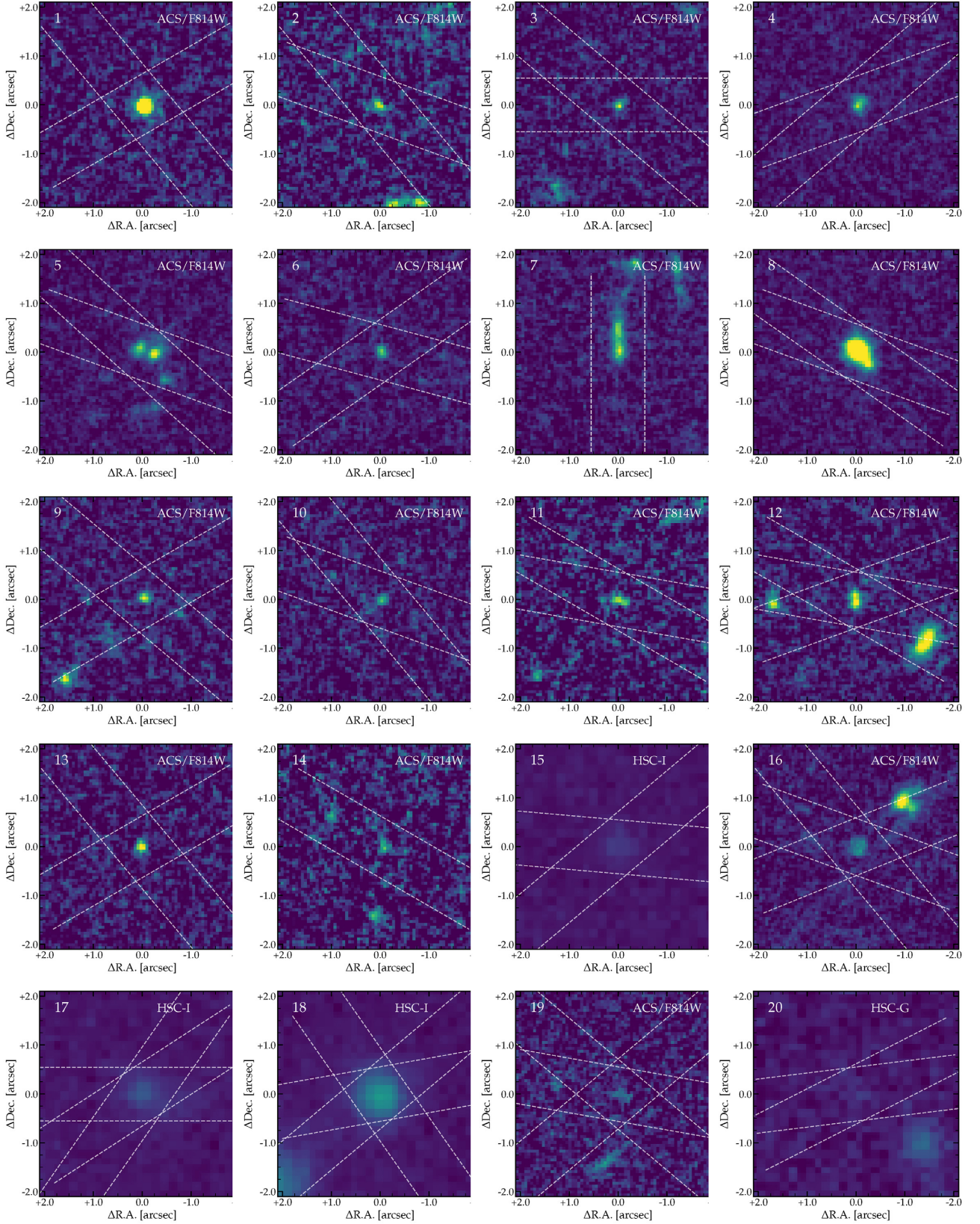
## APPENDIX

**Table A1.** Galaxy IDs that have previously been used for the galaxies in the XLS sample.

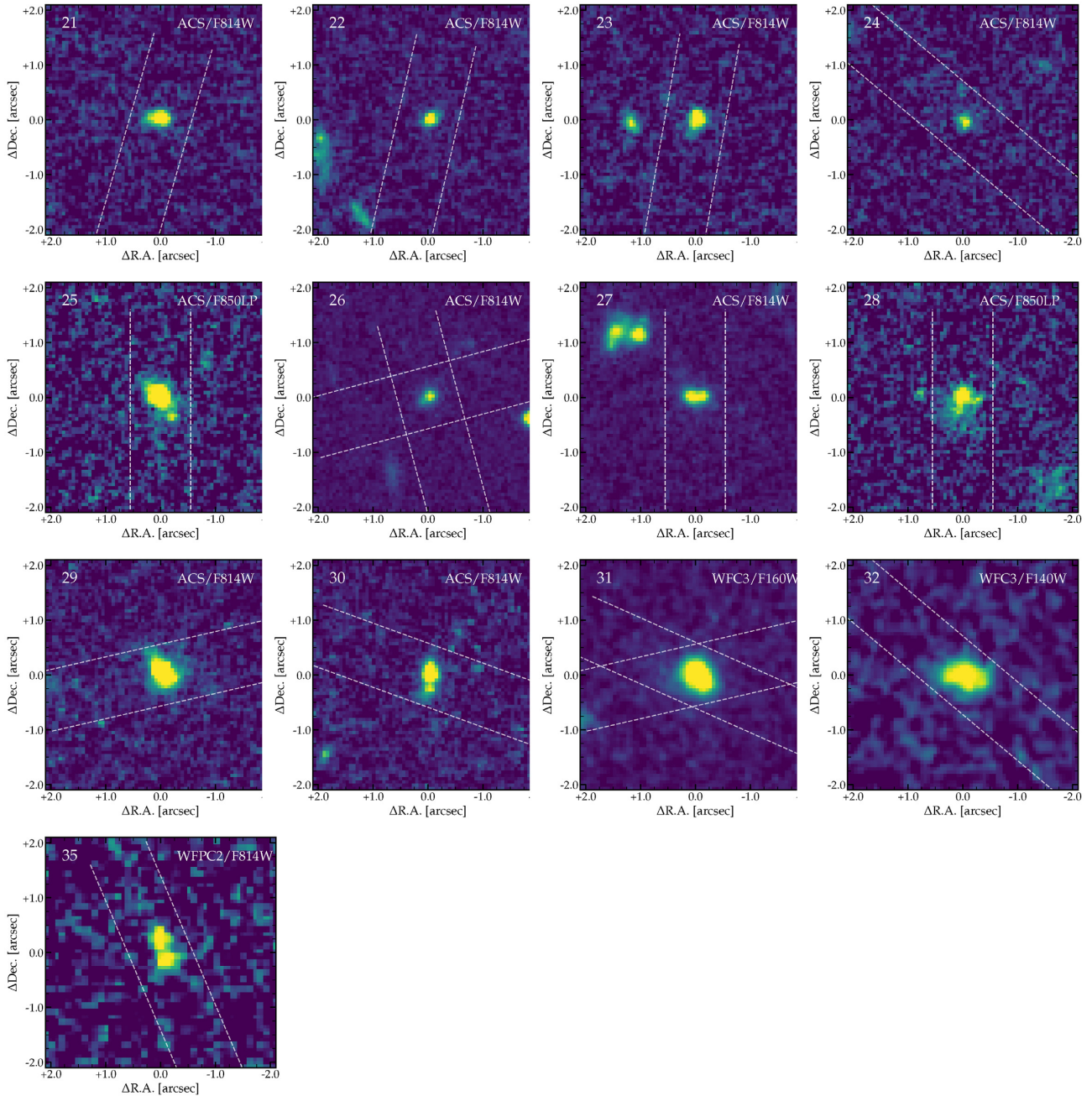
XLS-ID	Previous ID	XLS-ID	Previous ID
XLS-1	CALYMHA-S16-278	XLS-21	VUDS510583858
XLS-2	CALYMHA-S16-147	XLS-22	VUDS5100998761
XLS-3	CALYMHA-S16-138	XLS-23	VUDS5101421970
XLS-4	CALYMHA-S16-136	XLS-24	LAE47
XLS-5	CALYMHA-S16-134	XLS-25	CDFS_03865
XLS-6	CALYMHA-S16-129	XLS-26	DBL1349, GS30668
XLS-7	CALYMHA-S16-126	XLS-27	GS14633
XLS-8	CALYMHA-S16-122	XLS-28	CDFS_06482
XLS-9	CALYMHA-S16-115	XLS-29	Q2343-BX436
XLS-10	CALYMHA-S16-113	XLS-30	Q2343-BX418
XLS-11	CALYMHA-S16-95	XLS-31	Q2343-BX660
XLS-12	CALYMHA-S16-67	XLS-32	Q0207-BX144
XLS-13	CALYMHA-S16-28	XLS-33	Q0207-BX87
XLS-14	COSMOS_LYA	XLS-34	Q0207-BX74
XLS-15	CALYMHA-S16-389	XLS-35	Q0142-BX165
XLS-16	CALYMHA-S16-386	-	-
XLS-17	CALYMHA-S16-378	-	-
XLS-18	CALYMHA-S16-373	-	-
XLS-19	CALYMHA-S16-369	-	-
XLS-20	LYRS-S7	-	-



**Figure A1.** [O III]<sub>4960, 5008</sub> spectra of the XLS sample, binned by a factor of 2 and smoothed for visualization purposes. Spectra are normalized to the peak [O III]<sub>5008</sub> flux density. The locations of the [O III] lines are highlighted by dotted vertical lines. No [O III] emission is detected in XLS-9 and XLS-13 and we normalize them arbitrarily. The grey shaded region along the horizontal axis shows the noise level, while the vertical stripes mark the locations of skylines.



**Figure A2.** Cut-out images of the *HST* data in the ACS/F814W filter that traces the rest-frame UV centred on the XLS objects. White dashed lines show the orientations of the slits in the various OBs. For XLS-15, 17, 18, and 20, there are no *HST* data available such that we show ground-based data in the *g* band from the HSC survey. These data are shallower and have a lower resolution compared to the *HST* data.



**Figure A3.** As Fig. A2. For XLS-31 and 32, we can only show rest-frame optical data taken with *HST*/WFC3 instead of rest-frame UV data. The images for XLS-25 and 28 have a slightly different filter (F850LP) and the image for XLS-35 has been taken with WFPC2.

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