Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **493,** 5336–5356 (2020)
Advance Access publication 2020 March 3

doi:10.1093/mnras/staa546

Into the Ly α jungle: exploring the circumgalactic medium of galaxies at $z \sim 4-5$ with MUSE

Richard M. Bielby [1,2★] Michele Fumagalli [1,3,4] Matteo Fossati [1,3] Marc Rafelski,[5,6] Benjamin Oppenheimer,[7] Sebastiano Cantalupo [8] Lise Christensen,[9] J. P. U. Fynbo,[10,11] Sebastian Lopez [12] Simon L. Morris,[1] Valentina D'Odorico[13,14] and Celine Peroux[15,16]

Affiliations are listed at the end of the paper

Accepted 2020 February 15. Received 2020 January 24; in original form 2019 October 15

ABSTRACT

We present a study of the galaxy environment of nine strong H I + C IV absorption line systems ($16.2 < \log(N(\mathrm{HI})) < 21.2$) spanning a wide range in metallicity at $z \sim 4-5$, using MUSE integral field and X-Shooter spectroscopic data collected in a $z \approx 5.26$ quasar field. We identify galaxies within a 250 kpc and $\pm 1000\,\mathrm{km\,s^{-1}}$ window for six out of the nine absorption systems, with two of the absorption line systems showing multiple associated galaxies within the MUSE field of view. The space density of Ly α emitting galaxies (LAEs) around the H I and C IV systems is $\approx 10-20$ times the average sky density of LAEs given the flux limit of our survey, showing a clear correlation between the absorption and galaxy populations. Further, we find that the strongest C IV systems in our sample are those that are most closely aligned with galaxies in velocity space, i.e. within velocities of $\pm 500\,\mathrm{km\,s^{-1}}$. The two most metal-poor systems lie in the most dense galaxy environments, implying we are potentially tracing gas that is infalling for the first time into star-forming groups at high redshift. Finally, we detect an extended Ly α nebula around the $z \approx 5.26$ quasar, which extends up to ≈ 50 kpc at the surface brightness limit of $3.8 \times 10^{-18}\,\mathrm{erg\,s^{-1}\,cm^{-2}\,arcsec^{-2}}$. After scaling for surface brightness dimming, we find that this nebula is centrally brighter, having a steeper radial profile than the average for nebulae studied at $z \sim 3$ and is consistent with the mild redshift evolution seen from $z \approx 2$.

Key words: galaxies: evolution – galaxies: groups: general – galaxies: high-redshift – intergalactic medium.

1 INTRODUCTION

The flow of baryons on to, out of, and around galaxies is crucial to our understanding of galaxy evolution as a whole, dictating how galaxies form their stellar content. Star-forming galaxies predominantly occupy a relatively narrow (≈ 0.3 dex) main sequence, relating star-formation to stellar mass (e.g. Brinchmann et al. 2004; Daddi et al. 2007; Elbaz et al. 2007; Noeske et al. 2007; Salim et al. 2007). The gas depletion time-scales on this main-sequence are relatively short ($t_{\mathrm{depl}} \approx 10^9$ yr, e.g. Bigiel et al. 2008, 2011; Leroy et al. 2008; Tacconi et al. 2013), and to sustain the observed star-formation levels, a steady flow of cold gas from a main-sequence galaxy's surroundings is necessary. Such inflows are required to co-exist with the large-scale outflows commonly detected in star-forming galaxies across a wide range of cosmic time (e.g. Heckman, Armus & Miley 1990; Pettini et al. 2001; Shapley et al. 2003; Martin 2005; Lilly et al. 2013). Effectively, the existence of the main sequence (alongside the observed galaxy stellar mass function) necessitates that the bulk of the star-forming galaxy population exists in a quasi-steady state of gas inflow, outflow, and consumption (e.g. Bouché et al. 2010; Davé, Finlator & Oppenheimer 2012; Dekel & Mandelker 2014; Tacchella et al. 2016).

The large-scale transport of baryons is surmised to be intrinsically connected to the presence of strong H I absorption systems, and in particular Lyman limit systems (LLS) and damped Ly α systems (DLA) identified in quasar sightlines (e.g. Prochaska & Wolfe 2009; Prochaska, O'Meara & Worseck 2010; Fumagalli, O'Meara & Prochaska 2016a). Cold accretion, the dominant form of gas accretion at $z \gtrsim 1$ in simulations (e.g. Kereš et al. 2005; Dekel, Sari & Ceverino 2009), is predicted to follow collimated filamentary gas structures which would be detected at high H I column densities

★ E-mail: richard.bielby@durham.ac.uk

given a background source (e.g. Faucher-Giguère & Kereš 2011; Fumagalli et al. 2011b). Considering outflows, simulations predict that galactic winds from relatively low-mass galaxies should contain entrained cold gas clumps that would similarly be detected in absorption as high H I column density systems, likely with higher than average metallicities (e.g. Faucher-Giguère et al. 2016). Further, the observed kinematics of DLAs are seen to be reproduced in simulations only when outflows are implemented (Barnes et al. 2011; Barnes & Haehnelt 2014; Bird et al. 2015).

Detecting the host galaxies of strong H I absorbers has long been a challenge in observational astronomy. Early studies provided few, if any, detections of galaxies in close proximity of DLAs (e.g. Møller & Warren 1993; Steidel et al. 1994; Lowenthal et al. 1995; Djorgovski et al. 1996; Bunker et al. 1999; Fynbo, Burud & Møller 2000; Bouché et al. 2001; Grove et al. 2009). Significant progress on connecting strong Ly α absorbers to the galaxy population was made with the introduction of integral field spectrograph (IFS) instruments. In particular, SINFONI on the Very Large Telescope (VLT) provided some of the first significant data sets based on blind searches for H α emission in the observed near-infrared (i.e. $0.7 \lesssim z \lesssim 2$) from faint galaxies within a relatively small field of view around background quasars (e.g. Bouché et al. 2007; Péroux et al. 2011). These surveys probe the absorber environment up to ≈50 kpc, focusing on small-scale associations and so effectively probing the haloes of the closest galaxies in absorption. Such surveys produced detection rates for galaxies in the proximity of N(HI) $\gtrsim 10^{19}$ cm^{-2} absorption line systems of $\approx$ 30 per cent at $z \sim 1$, falling to just $\lesssim$ 10 per cent at $z \sim 2$ (Péroux et al. 2016).

Parallel with these studies, the introduction of the X-Shooter spectrograph on the VLT showed that the detection of associated galaxies was tied to the metallicity of the strong absorption system, with higher metallicity systems ([Si/H] > -1) showing significantly enhanced detection rates (≈ 60 per cent) of associated galaxies (Fynbo et al. 2013; Krogager et al. 2017).

With the MUSE IFS (Bacon et al. 2010) on the VLT, systematic blind surveys are now possible over larger scales (≈1 arcmin), detecting simultaneously direct associations at small impact parameters as well as the environment at larger impact parameters. Whilst past targeted surveys have revealed a handful of galaxy groups in the proximity of strong absorption line systems (e.g. Bergeron & Boisse 1986; Møller & Warren 1998; Kacprzak, Murphy & Churchill 2010), dedicated MUSE surveys are beginning to uncover greater numbers of associations with galaxy groups at low redshifts (e.g. Bielby et al. 2017a, 2019; Péroux et al. 2017; Fossati et al. 2019; Hamanowicz et al. 2020) as well as similar complex environments traced by Ly α emitters (LAEs) at $3 \lesssim z \lesssim 4$ (e.g. Fumagalli et al. 2016b, 2017b; Lofthouse et al. 2020; Mackenzie et al. 2019). Such studies are now also being complemented by ALMA observations mapping the galaxy population in [C II] and CO emission (e.g. Neeleman et al. 2017, 2018, 2019; Klitsch et al. 2018, 2019; Péroux et al. 2019).

With large surveys currently ongoing at $z \lesssim 4$, extending blind searches to higher redshifts, approaching the epoch of re-ionization, is now critical for a complete view of the gas–galaxy connection across cosmic time. To this end, redshifts of $z \approx 5-5.5$ become particularly relevant as late re-ionization models predict that the end of re-ionization may extend to $z \approx <5.5$ (Becker et al. 2015; Hartoog et al. 2015; Kulkarni et al. 2019; Nasir & D'Aloisio 2019). Moreover, differently from $z \gtrsim 6$ (e.g. Becker et al. 2012), it is still possible to identify individual absorption line systems within the thick Ly α forest with high-resolution spectroscopy, which is crucial for a detailed characterization of the hydrogen content and hence metallicity of the systems (e.g. Rafelski et al. 2014).

At present, only a handful of studies have focused on this redshift range. Cai et al. (2017) presented a narrow band search for Ly α emitters at $z \approx 5$ around C IV absorbers, finding one candidate pair within $\lesssim$40 kpc and two further candidate pairs at 160 and 200 kpc. D'Odorico et al. (2018) reported on a serendipitous CO detection of a $z \approx 5.9$ galaxy associated with a metal-poor DLA at a distance of ≈40 kpc, surmizing the DLA to be tracing either a satellite galaxy or filamentary gas structure. Likewise, Díaz et al. (2014), Díaz et al. (2015) conducted narrow band plus spectroscopic follow-up analysis of LAE galaxies around two $z > 5$ C IV absorption line systems, finding both systems resided in large-scale regions hosting galaxy overdensities.

Using more statistical techniques, Meyer et al. (2019a) argued for an excess of galaxies in the proximity of C IV absorbers leading to an associated excess of Ly α transmission in quasar sightlines through such regions. Further to these, Keating et al. (2020) discussed how LAEs should be hard to detect close to deep neutral H I absorption at these redshifts, as those may be the regions not yet ionized, and that it may be instead easier to detect galaxies in the proximity of C IV absorbers, where the local medium is perhaps more likely to be ionized by galaxies.

Leveraging the discovery of several very bright quasars at $z \gtrsim 5$ (Wang et al. 2016) for which we have collected high-quality optical and NIR spectroscopy, we can now extend these previous studies by conducting a systematic search of galaxies in high-redshift quasar fields. In this paper, we present a detailed MUSE+X-Shooter analysis along the sightline of the quasar SDSS J074749.17 + 115352.4 at $z \approx 5.26$, within which we detect multiple strong H I absorption line systems.

In Section 2, we present the details of the observations and data reduction. Section 3 presents the search for galaxies within the MUSE cube and Section 4 presents the analysis of the absorption line systems, including associated metal transitions. Section 5 introduces the simulation data used in this study. In Section 6, we present the results of combining the two data sets, whilst Section 7 provides our summary and conclusions. Throughout this paper, we assume a cosmology defined by the parameters in Planck Collaboration XIII (2016, i.e. $H_0 = 67.7$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.307$, and $\Omega_\Lambda = 0.693$) and express magnitudes in the AB system. Unless stated otherwise, distances are given in the proper coordinate frame.

2 OBSERVATIONS AND DATA REDUCTION

2.1 MUSE data

MUSE observations in the quasar field SDSS J074749.17 + 115352.4 have been collected as part of ESO programme 0102.A−0261 (PI: Bielby) between 2018 December and 2019 March, using the Wide Field Mode combined with the GALACSI (Ground Atmospheric Layer Adaptive Corrector for Spectroscopic Imaging) adaptive optics system (Ströbele et al. 2012). Conditions were generally excellent, with clear sky and sub-arcsecond image quality in dark time. We acquired a total of 24 exposures, each of 900 s, for a total on-source exposure time of 6 h. In between exposures, we applied small offsets (≈3−4 arcsec) and 15 deg rotations to reduce systematic errors.

Observations have been reduced following the methodology described in previous work (Fumagalli et al. 2016b, 2017b,a; Lofthouse et al. 2020). Briefly, we first use the standard ESO

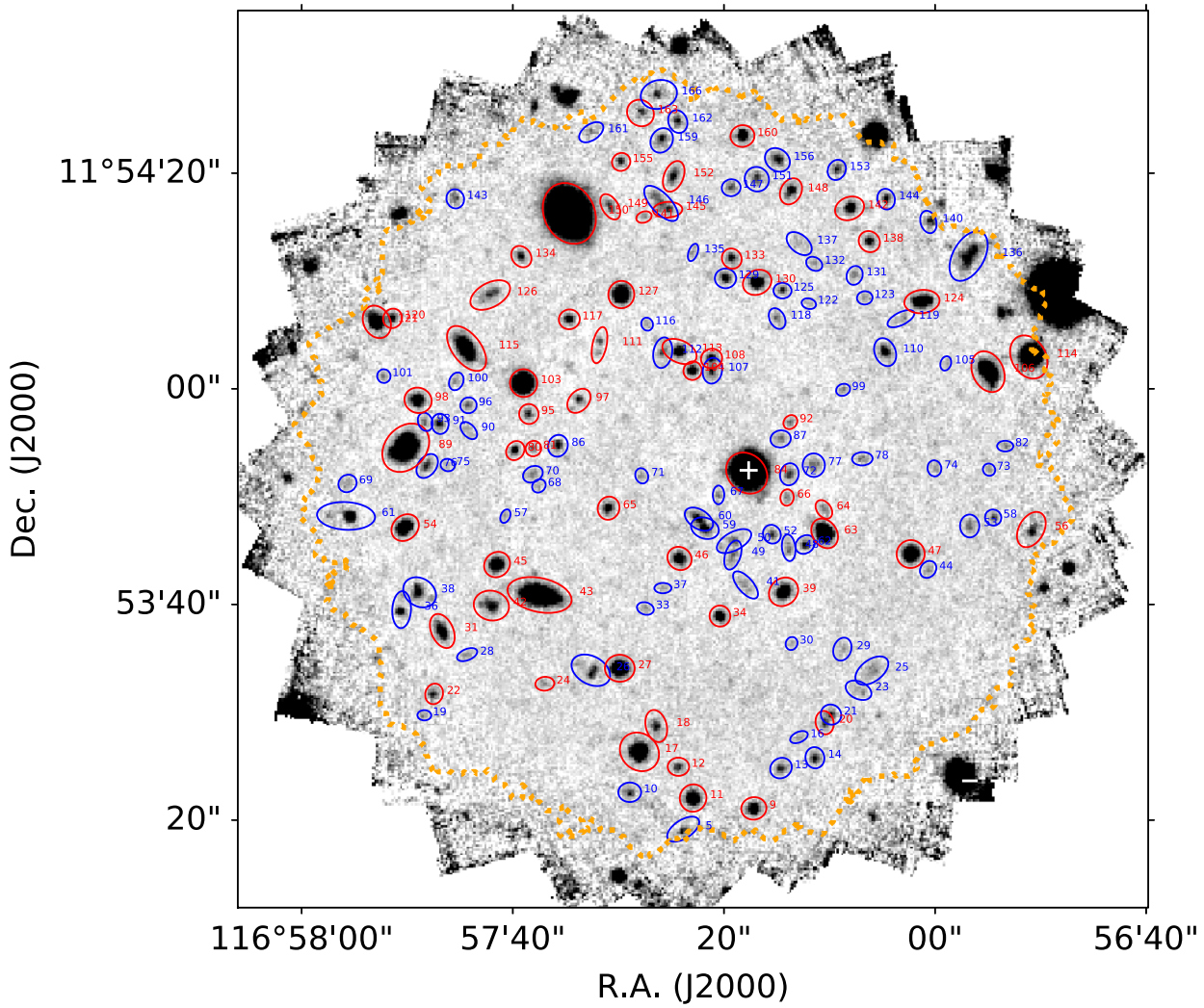

Figure 1. Reconstructed white-light image from the final data cube, showing the quasar SDSS J074749.17 + 115352.4, marked with a white cross near the centre of the field. Continuum-detected sources identified within the field are marked with apertures (red for sources with spectroscopic redshifts). The dotted orange contours enclose the region covered by at least six exposures.

pipeline (v2.4.1; Weilbacher et al. 2014) to perform basic reduction by applying calibrations. We then reconstruct datacubes for each individual exposure, which we then process with the CUBEXTRACTOR pipeline (v1.8) to improve the quality of the flat fielding and sky subtraction (see Cantalupo et al. 2019). Finally, all exposures are combined in a single cube with pixel size of 0.2 arcsec (spatial direction) and 1.25 Å (spectral direction), using both mean and median statistics. Two independent cubes, each with half the number of exposures are also produced. The white-light image reconstructed from the mean cube is shown in Fig. 1. The final image quality on this white-light image is found to be ≈ 0.62 arcsec full-width at half-maximum (FWHM) as measured by fitting Moffat profiles on point sources.

Following the procedure described in Lofthouse et al. (2020), we further re-scale the final noise cubes using bootstrap techniques of individual pixels across the 24 exposures to correctly reproduce the pixel standard deviation. At this stage, we also derive a model for the correlated noise arising from the resampling of the pixel table on to a final grid, as described in Lofthouse et al. (2020). This model is defined as a correction that needs to be applied to the propagated error for a source in an aperture of N pixels on a side, $\sigma_{\rm N}$, to recover the effective noise, $\sigma_{\rm eff}$. A second-order polynomial fit describes this correction in the form $\sigma_{\rm eff}/\sigma_{\rm N} = 1.357 + 0.128N + 0.008N^2$, as computed in a spectral window of 4 pixels between 6600 and 7900 Å (i.e. the range where we search for Ly α emission).

2.2 X-Shooter data

The background quasar spectrum (Fig. 2) comes from a recent data set obtained at the VLT using X-shooter spectrograph (Vernet et al. 2011) to observe 41 $z > 5$ bright quasar spectra (PID 98.A−0111 and 100.A−0243; PI: Rafelski, Becker et al. 2019). These quasars were identified via WISE IR colour selection and confirmed with low-resolution spectroscopy (Wang et al. 2016; Yang et al. 2016, 2017). X-shooter provides moderate resolution spectroscopy across three wavelength ranges: UVB (300–550 nm), VIS (550–1020 nm), and NIR (1020–2480 nm), although the UVB arm contains no to little flux for quasars at this redshift. The data typically have a signal-to-noise $S/N > 10$ per spectral bin and we use slits of 0.9 arcsec in the optical and 0.6 arcsec in the NIR to achieve a resolution of $R \sim 8000-9000$, sufficient to measure Ly α and accurate metallicities (Rafelski et al. 2012, 2014).

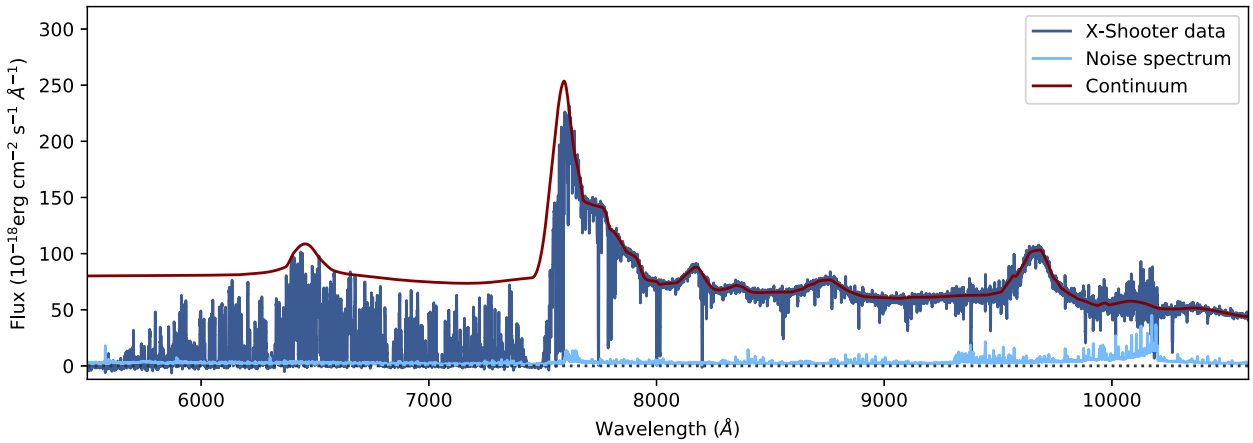

Figure 2. Reduced spectrum of SDSS J074749.17+115352.4 from the VLT X-Shooter spectrograph (blue curve), with the associated 1σ noise spectrum (pale blue) and the estimated continuum profile (solid red curve).

The X-Shooter data were reduced using standard techniques with a dedicated pipeline, as detailed in Becker et al. (2012) and López et al. (2016). A full description of the reduction of the X-Shooter spectrum, including sky subtraction, 1D extraction, and corrections for telluric absorptions is provided in Becker et al. (2019). Based on the X-shooter spectrum, we measure a redshift of $z_{\rm qso} = 5.265 \pm 0.015$ for the quasar, by fitting the Si IV and C IV emission lines. We note that these lines show velocity offsets from intrinsic redshifts of $1 \lesssim z \lesssim 6$ QSOs across a range of $\approx \pm 500\,{\rm km\,s^{-1}}$ (Meyer, Bosman & Ellis 2019b).

3 SEARCH FOR ASSOCIATED GALAXIES

As in previous analyses of MUSE data (e.g. Fumagalli et al. 2017b; Fossati et al. 2019; Lofthouse et al. 2020; Mackenzie et al. 2019), we conduct a redshift survey of galaxies detected in the continuum together with a search for line emitters in the cube.

3.1 Continuum-detected galaxies

For continuum-detected galaxies, we run SEXTRACTOR (Bertin & Arnouts 1996) on the deep white-light image, folding in the propagated pixel variance and masking regions where the number of exposures falls below 6 (i.e. the region outside the dashed yellow bounding line in Fig. 1). Only sources above five times the propagated error and with minimum area of 10 pixels are marked as detected. For each detected source, we reconstruct a 1D spectrum using all pixels within the segmentation mask, also transforming the wavelength to vacuum.

We then measure redshifts using the MARZ redshifting software (Hinton et al. 2016), which we customize[1] with high-resolution synthetic templates for passive and star-forming galaxies at $z < 2$. Following automatic template fitting, individual sources are inspected and classified by two authors (MFu and MFo) in four classes (4, secure redshift with multiple features; 3, good redshift with single but unambiguous feature; 2, possible redshift, based on a single feature; 1, unknown redshift). At $z > 4$, spectroscopic identification is based on Ly α emission alone, which often shows an asymmetric profile. Typical redshift uncertainties are $\delta z \approx 0.0003$.

[1]This version is available at https://matteofox.github.io/Marz, described in Fossati et al. (2019)

Table 1. Spectroscopic redshifts (including quality flags) and coordinates for continuum-detected sources. Only the sources at $z > 4$ are listed, with the full table available as online only-material.

ID	R.A./Dec.	Redshift	Q.F.
22	J074751.16+115331.7	4.6250	3
56	J074747.39+115346.9	4.0830	3
84	J074749.19+115352.2	5.2548	4
141	J074749.84+115415.9	5.0703	2

The continuum detected $z > 4$ objects are listed in Table 1, including the quasar itself. Fig. 3 shows the spectra of the three sources (excluding the quasar), focusing on the spectral region where Ly α is detected.

To assess the completeness of our source catalogue, we perform 10 000 repetitions of the analysis described above on mock images constructed by injecting 80 mock sources at each iteration (to avoid blending issues) in blank sky regions. We repeat this experiment twice, the first time for point sources matched to the image quality of the MUSE data (0.6 arcsec) and the second time considering exponential discs (neglecting inclination) with scale-length of 0.26 arcsec, convolved with the instrument point spread function (PSF). In Fig. 4, we show the fraction of objects recovered compared to the number of injected sources as a function of magnitude, finding that our search is 50 per cent (90 per cent) complete at 26.7 mag (26.4 mag) on the white-light image for point sources and 50 per cent (90 per cent) complete at 25.7 mag (25.3 mag) for extended sources.

3.2 Emission line galaxies

For emission line galaxies we follow the approach described in Lofthouse et al. (2020), which we only briefly summarize here. After reducing the cube to the wavelength range of interest ($\lambda = 6100-7900$ Å, covering the redshift interval $z = 4-5.5$ for Ly α), we subtract the quasar PSF and the continuum of sources using the tools distributed as part of CUBEX (for details, see e.g. Cantalupo et al. 2019).

We then run CUBEX to identify groups of at least 27 connected voxels, covering more than 9 pixel^2 once projected along the wavelength axis, and three channels in wavelength. Objects marked

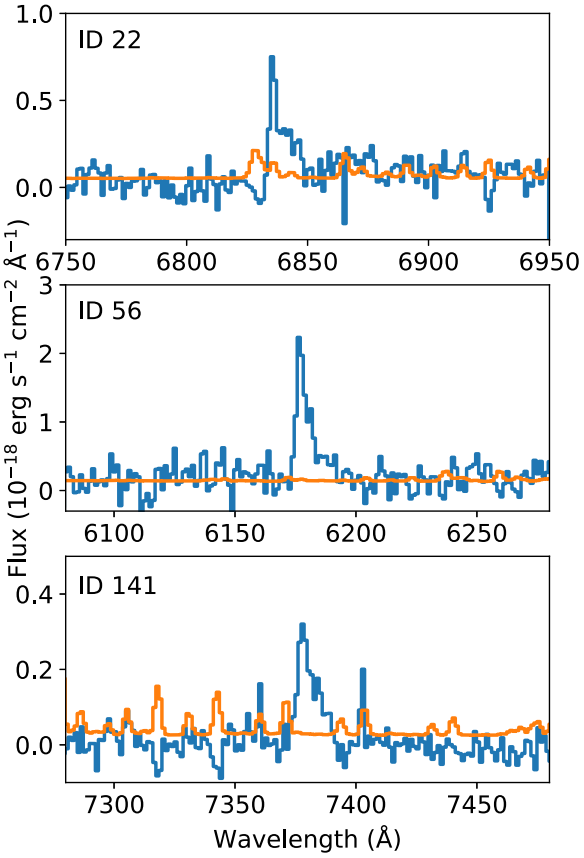

Figure 3. Spectra of all three continuum detected galaxies at $z > 4$ (excluding the quasar, see also Table 1). At these redshifts, spectroscopic identification is based on Ly α emission which often shows an asymmetric profile.

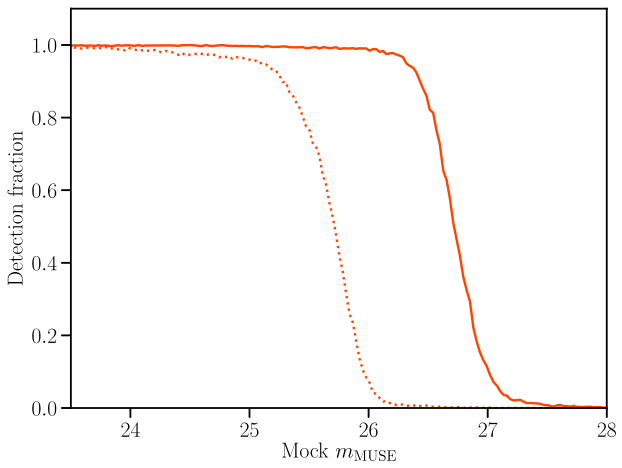

Figure 4. Completeness function for continuum sources for point sources matched to the MUSE PSF (solid line) and for extended sources with scale-length of 0.26 arcsec (dotted line). These observations are 50 per cent (90 per cent) complete at 26.7 mag (26.4 mag) for point sources, and 50 per cent (90 per cent) complete at 25.7 mag (25.3 mag) for extended sources.

for extraction are then retained if the integrated *S/N* corrected for correlated noise as discussed above is *S/N* > 6.5. In contrast to previous work conducted mostly at $\lambda < 6000$ Å, in this analysis we adopt a more conservative cut in *S/N* to ensure that we minimize the contamination arising from residuals of sky lines.

Following extraction, objects are classified in two confidence classes. Objects in class 1 have integrated $S/N \geq 7$ in the mean coadd, are detected with $S/N > 3$ within independent coadds of only half of the exposures, and the *S/N* of these two subsets agree within 50 per cent of their value. This class contains sources with the highest purity at the expense of completeness. Objects in class 2 obey a similar classification, but include the remaining objects with $6 \leq S/N < 7$, which raises the completeness at the possible expense of the purity. For this class, we also monitor the fraction of voxels within the segmentation map that is contained in a 5^3 voxel volume, which we find to be a good metric to reject spurious identifications such as very extended structures at the edge of the field where the quality of the data is significantly worse.

Next, we generate optimally extracted maps (see e.g. Borisova et al. 2016) of the mean, median cubes, and two independent-half cubes, and we extract a 1D spectrum by projecting the segmentation map in 2D and summing flux over the spatial direction as a function of wavelength. Using a dedicated GUI, we inspect these products, including the 3D segmentation map, for all the objects to remove the remaining false-positive (typically objects that present very elongated chains of voxels in their segmentation map that are not well-connected in wavelength). The remaining objects in class 1 are deemed to be real sources, with the objects in class 2 considered only candidate sources as they approach the detection limit.

At this stage, we also assign a redshift to the sources, according to the following criteria. Sources that present a clear doublet emission can be classified as either [O II] λλ3726, 3729, C III] λλ1907, 1909, or Mg II λλ2796, 2803 emitters. The classification of single line emitters is more ambiguous as multiple rest-frame lines could be in principle detected at any given redshift. For class 1 sources, we can rule out H α trivially for $\lambda < 6563$ Å and can classify H α for the remaining sources that show associated [O III] λ5007 or the [O II] doublet. C IV can be recognized by the associated strong Ly α emission (unless Ly α is absorbed or resonantly scattered) or C III] doublet emission. Finally, the identification of Ly α is also strengthened by the characteristic shape of the profile (where evident by eye in the spectra). These criteria apply also to class 2 sources, although the varying sensitivity limit across the wavelength range and the different line strengths make the classification more uncertain. Following these criteria, we identify 14 additional $z > 4$ galaxies, 3 of which are in class 2, and 6 [O II] emitters, one of which is in class 2 and a second one for which the classification is ambiguous (i.e. good signal to noise emission line, but of uncertain redshift). The properties of the sources are detailed in Table 2, a map of the Ly α emitters is shown in Fig. 5 and a gallery of the Ly α profiles is shown in Fig. 6.

The three continuum selected galaxies given in Table 1 are all also detected with the line detection method. One of these (ID 141 in continuum and ID 14 in emission) shows a small offset between the centroid of the continuum emission and the line emission, with the line emission being significantly more extended than the continuum, whilst the remaining two are more consistent in extent and centroid with the continuum detections. As such we include the line detection for this source in Table 2 for reference, but not the other two continuum sources. All three are included and treated as single objects in the analysis that follows, giving a total of 17 $z > 4$ galaxies exhibiting Ly α emission.

Similarly to the analysis of continuum sources, we quantify the completeness of our search by analysing 5000 mock cubes constructed by injecting 500 mock line emitters in empty regions of each mock cube (to avoid blending). Two types of source are considered: compact emitters with size matched to the MUSE

Table 2. Properties of the emission line galaxies detected at $S/N \geq 6$ in the MUSE cube between 6100 and 7900 Å. The area is calculated for each source out to a surface brightness level of $10^{-18.25}$ erg s^{-1} cm^{-2} arcsec^{-2}.

ID	R.A. (J2000)	Dec. (J2000)	i (mag)	M_{UV} (mag)	F_{line} (10^{-18} erg s^{-1} cm^{-2})	L_{line} (10^{40} erg s^{-1})	Redshift	Area (kpc^2)	Class	Type
1	07:47:47.399	11:53:55.96	27.0 ± 0.5	-22.0 ± 0.5	2.87 ± 0.30	72.36 ± 7.66	4.8384	137	1	LAE
2	07:47:47.705	11:53:36.66	>26.4	>-22.1	4.39 ± 0.39	113.27 ± 9.98	4.8820	225	1	LAE
3	07:47:47.988	11:53:51.94	>26.6	>-21.7	1.01 ± 0.16	22.26 ± 3.50	4.5594	155	2	LAE[a]
4	07:47:47.991	11:53:49.79	>27.4	>-21.4	2.53 ± 0.27	63.23 ± 6.85	4.8191	254	1	LAE
5	07:47:47.998	11:53:52.28	>26.6	–	1.25 ± 0.17	0.25 ± 0.03	0.6582	–	1	[O II][b]
6	07:47:48.188	11:54:07.18	>26.9	–	3.22 ± 0.37	1.98 ± 0.22	1.0443	–	1	[O II]
7	07:47:48.278	11:53:59.98	>27.0	>-21.8	1.82 ± 0.20	45.82 ± 5.12	4.8309	149	1	LAE
8	07:47:48.355	11:53:29.96	27.0 ± 0.5	-21.8 ± 0.6	16.44 ± 0.74	483.64 ± 21.92	5.1655	446	1	LAE
9	07:47:48.777	11:54:23.19	26.7 ± 0.6	-21.2 ± 0.6	16.76 ± 0.92	284.30 ± 15.59	4.0876	631	1	LAE
10	07:47:49.283	11:53:32.31	>26.7	–	37.49 ± 1.37	20.97 ± 0.77	1.0054	–	1	[O II]
11	07:47:49.375	11:54:02.98	25.6 ± 0.2	–	2.31 ± 0.26	0.46 ± 0.05	0.6586	–	1	[O II]
12	07:47:49.415	11:54:16.36	>27.0	>-21.1	11.13 ± 0.69	212.12 ± 13.23	4.2946	766	1	LAE
13	07:47:49.460	11:54:02.94	25.6 ± 0.2	-22.2 ± 0.4	1.09 ± 0.16	26.20 ± 3.94	4.7473	136	2	LAE[c]
14	07:47:49.915	11:54:15.94	>26.9	>-22.0	14.47 ± 0.64	407.48 ± 17.97	5.0703	513	1	LAE
15	07:47:49.919	11:53:49.74	>27.2	>-21.8	3.53 ± 0.32	91.09 ± 8.19	4.8830	249	1	LAE
16	07:47:50.196	11:54:18.30	26.3 ± 0.2	-21.6 ± 0.2	4.37 ± 0.41	75.53 ± 7.02	4.1194	231	1	LAE[d]
17	07:47:50.435	11:54:12.43	>26.9	>-21.6	0.80 ± 0.13	19.42 ± 3.08	4.7516	120	2	LAE
18	07:47:50.801	11:53:32.26	>26.9	–	1.64 ± 0.25	0.34 ± 0.05	0.6697	–	2	[O II]
19	07:47:50.805	11:54:18.27	>26.9	>-21.8	5.34 ± 0.43	133.47 ± 10.67	4.8186	211	1	LAE
20	07:47:51.207	11:53:49.61	>26.9	–	3.44 ± 0.38	0.71 ± 0.08	0.6698	–	1	[O II]
21	07:47:51.808	11:53:41.18	26.0 ± 0.3	-22.4 ± 0.4	1.35 ± 0.22	33.72 ± 5.39	4.8161	202	1	LAE

Note. [a]Overlaps with emitter 4 in projection, but appears at different redshift. [b]Classification uncertain. [c]Overlaps with continuum source 108 in projection, but appears at different redshift. [d] Overlaps in part with continuum source 150 in projection, but appears at different redshift.

PSF and line spread function with FWHM of 2.5 Å, and extended sources with exponential profiles of scale-length of 2, 3, and 4 kpc convolved with PSF and line spread function with FWHM of 2.5 Å. By analysing the mock cubes with CUBEX as done for the real data, we find the completeness function shown in Fig. 7 for $S/N > 6$. Our analysis is 50 per cent complete at $\approx 10^{-18}$ erg s^{-1} cm^{-2} for point sources and $\approx 3.1 \times 10^{-18}$ erg s^{-1} cm^{-2} for extended sources with 3 kpc scale-length (at 90 per cent completeness, the limits for point and extended sources become $\approx 2.4 \times 10^{-18}$ erg s^{-1} cm^{-2} and $\approx 7.7 \times 10^{-18}$ erg s^{-1} cm^{-2} for a 3 kpc scale-length).

4 PROPERTIES OF STRONG ABSORPTION LINES

4.1 Identification and measurement of column densities

We perform a continuum fit of the X-Shooter spectrum using a combination of the LT_CONTINUUMFIT code, contained within the LINETOOLS package,[2] and a template quasar continuum. The LT_CONTINUUMFIT code follows the steps outlined in Crighton et al. (2011), Bielby et al. (2017b), fitting an initial cubic spline form to the data. However, due to the frequency of Ly α absorbers at $z \sim 4-5$, the procedure underestimates the unabsorbed intrinsic continuum below $\lambda_{rest} = 1215.67$ Å and so we use the SDSS $0.04 \lesssim z \lesssim 4.78$ quasar composite spectrum calculated by Vanden Berk et al. (2001) as a guide to correcting the initial cubic-spline fit.

We first normalize and redshift the Vanden Berk et al. (2001) composite to fit the observed quasar continuum at $\lambda_{rest} \approx 1200-1400$ Å. To allow for a tilt in the template, we apply a power-law factor to the composite at $\lambda_{rest} < 1215$ Å of the form $(1215.67 - \lambda_{rest})^{\alpha}$, finding a slope of $\alpha = 0.045$ by normalizing to the peak flux at the Ly β emission wavelength in the quasar spectrum. We then use this template as a guide when fitting the continuum at $\lambda_{rest} < 1215.67$ Å using LT_CONTINUUMFIT.

Using this estimated continuum model, we make an initial census of strong Ly α absorption (in the range bounded by the intrinsic quasar Ly α and Ly β emission) and metal absorption systems along the sightline using PYIGM_GUESSES (part of the PYIGM suite of codes[3]). We include all systems where we observe either damping wings in the H I absorption or detected metal line absorption at a given redshift. This results in a sample of strong absorption systems, encompassing primarily LLSs and DLAs, all of which show associated C IV absorption features. The SDSS J074749.17+115352.4 sightline has a high incidence of metal and strong Ly α absorption system lines in the redshift range $4 < z < z_{qso}$, which we list in Table 3 (note that below $z \sim 4$, C IV begins to fall in the Ly α forest for our quasar sample, whilst $z \sim 4$ Ly α is obscured by $z \sim 5$ Ly β absorption). The strongest of these is a clear DLA at $z \approx 5.14$. All but one of the other metal-line detected systems lie at lower redshifts than the DLA, meaning that for these we have no measurement of their Lyman limit strengths, or in general any useful measurement of the Lyman series beyond that of Ly β.

For each strong absorption system, we perform a fit to the region of the Ly α forest over a velocity range depending on the strength of the absorber, using the ALIS line fitting software package.[4] In the first iteration, we keep the redshift, column density, and velocity broadening of each Ly α absorption line as free parameters, whilst keeping the continuum fixed to that estimated previously using LT_CONTINUUMFIT. Taking this initial fit to the Ly α regime, we then produce the resultant Ly β profile over the same velocity range. Comparing this Ly β profile to the data, we add appropriate lower-

[2]https://github.com/linetools

[3]https://github.com/pyigm

[4]https://github.com/rcooke-ast/ALIS

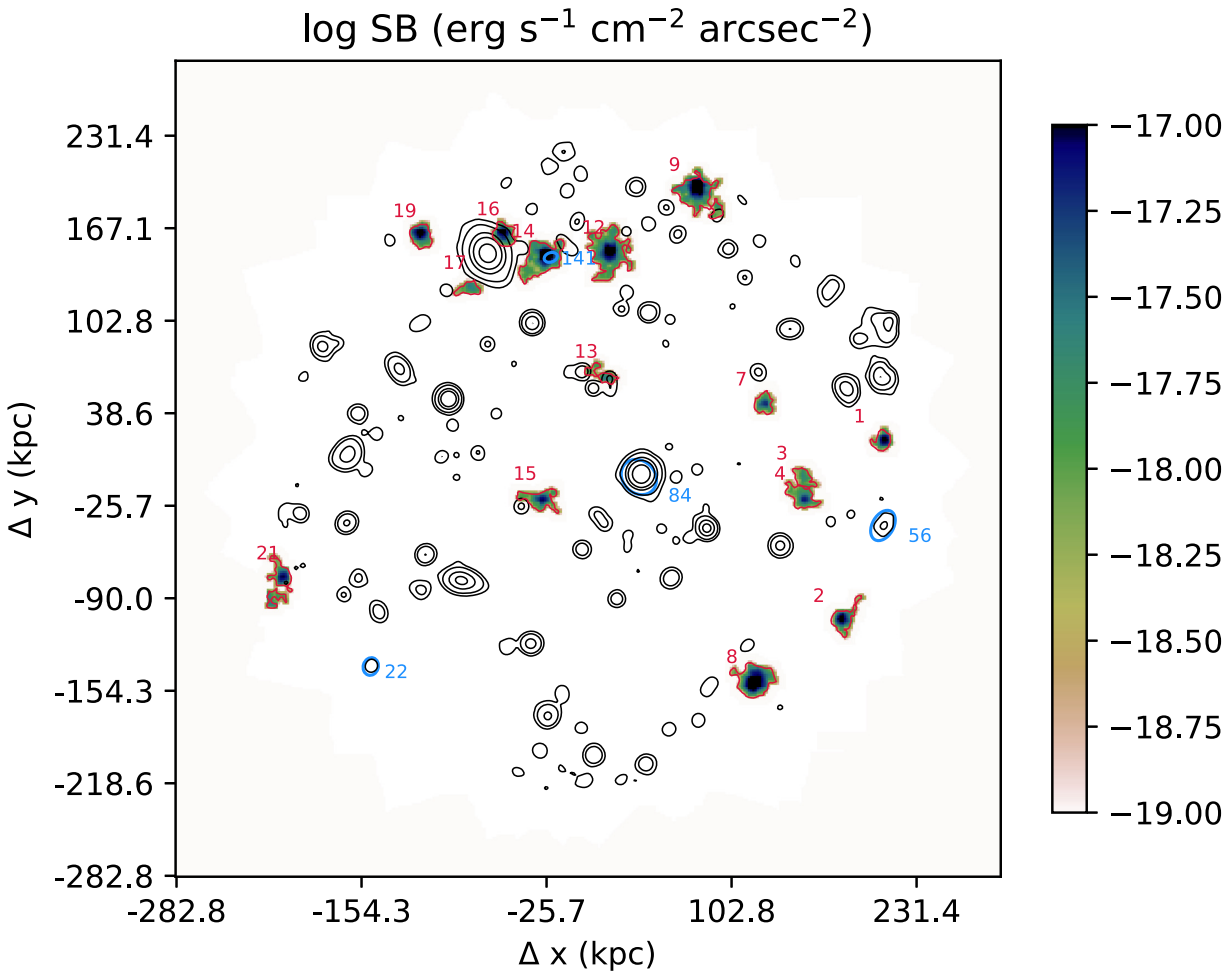

Figure 5. Map of the Ly α emission detected in the MUSE cube at $z > 4$ (excluding the quasar itself). The solid contours show continuum emission, whilst the colour scale gives the measured Ly α surface brightness. Objects detected in Ly α emission at $z > 4$ are numbered and outlined in red. Continuum sources at $z > 4$ are numbered and circled in blue.

redshift Ly α absorption systems to complete the model spectrum in the primary Ly β regime. We then perform a second iteration of the fitting using ALIS to fit the primary Ly α and Ly β absorption as well as the secondary Ly α absorption simultaneously.

As a final step, we perform a further iteration, allowing for addition of further absorbers in both regimes where necessary. Taking this configuration, we then estimate the effect of uncertainties on the continuum level on the fitting profiles. In order to do so, we re-run ALIS using the final model fit, but allowing the continuum level to be a free parameter, re-calculating the parameters for the absorption systems. Any change in the parameters for the strong absorber of interest is then folded into the ALIS fitting estimated uncertainties on each parameter.

The two lowest redshift ($z \approx 4.03$ and $z \approx 4.12$) absorbers have no associated coverage of Ly β (or any higher Lyman series orders), due to the presence, and absorption below the Lyman limit, of the $z \approx 5.14$ DLA. These are therefore purely constrained in H I by Ly α and thus have large associated uncertainties on the H I column density measurements. In contrast to these, the one absorber at $z \approx 5.16$ has available coverage up to Ly η, with all but the Ly η line being saturated. From the detected flux at Ly η observed wavelengths, we constrain this absorption line to most likely be a partial LLS, with $N(\mathrm{HI}) = 10^{16.25 \pm 0.25}$ cm^{-2}. This is the lowest column density absorber in our sample, with the highest being the $z = 5.14$ absorber with $N(\mathrm{HI}) = 10^{21.21 \pm 0.05}$ cm^{-2}, which is relatively well constrained via the damping wings of the absorber.

Our fits to the H I and metal line absorption profiles for the DLA system at $z = 5.1448$ are shown in Fig. 8, whilst we provide plots of the rest of the $z > 4$ absorption systems in the Appendix (Figs A1, A2, A3, and A4). The fitted parameters and the estimated uncertainties are instead given in Table 3, where we report only the integrated column densities (i.e. across all components) for each reported species.

Whilst the above provides a sample of absorbers that is effectively agnostic of the galaxy positions, we are interested in the absorption properties of the whole galaxy sample. We therefore make a secondary galaxy guided identification of absorption systems. We therefore identify the most proximate candidate strong H I absorption feature within ± 1000 km s^{-1} of each galaxy with no already identified strong absorption feature and no other galaxy within 1000 km s^{-1} at a closer impact parameter (i.e. where we have multiple galaxies at a given redshift, we only search for absorption within ± 1000 km s^{-1} of the closest to the sightline). To identify absorption systems, we use the same method as for the primary set of strong systems. We first identify absorption systems within ± 1000 km s^{-1} of the redshift of interest using PYIGM_GUESSES

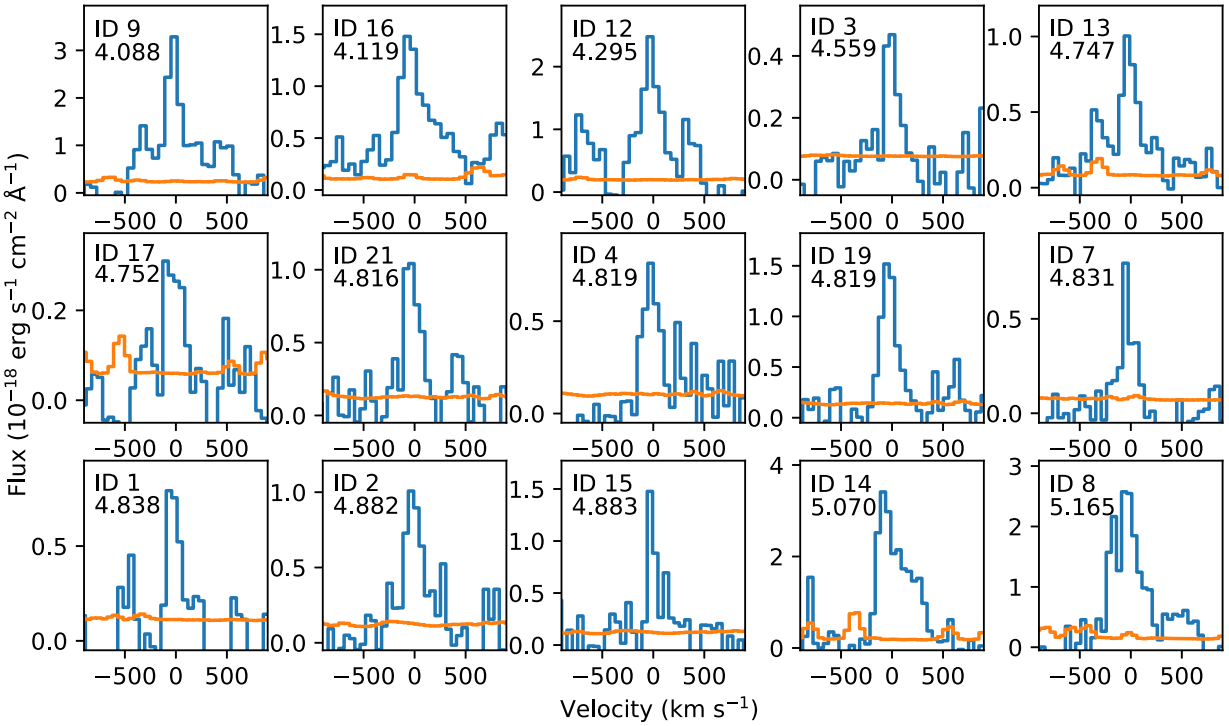

Figure 6. Emission line spectra centred on detected Ly α emission for the 15 $z > 4$ line-emission detected sources (in order of increasing redshift, see also Table 2). The blue histogram shows the detected flux profile and the orange the associated noise spectrum.

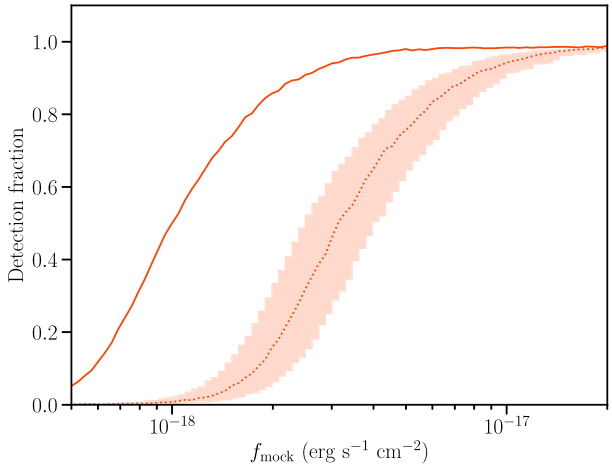

Figure 7. Completeness limits for extraction of line emitters at $S/N \geq 6$. The red solid curve refers to point sources matched to the MUSE PSF. In contrast, the shaded region shows extended sources with exponential profiles of scale-lengths of 2–4 kpc, with the dashed curve showing the 3 kpc scale length result.

and then perform an iterative fitting process using ALIS (adding lines where appropriate to improve the fit). We then identify the nearest strong absorber (i.e. N(H I) $\gtrsim 17$) to the galaxy redshift. The resulting column density estimates and absorber redshifts are listed below our primary sample in Table 3. Given these systems were not identified in the primary sample (i.e. either through metal lines or clear strong absorption), these secondary systems are all at the lower column density end of our strong absorber range and are poorly constrained, lying as they do in the flat section of the curve of growth.

4.2 Absorption line system metallicities

We now estimate the metallicities of the detected absorption line systems (including only those systems with at least one clearly detected metal line). Whilst the DLA is likely dense enough for partial ionization not to be an issue in determining metallicities, the LLSs are at least partially photoionized (e.g. Fumagalli et al. 2013), complicating the estimation of metallicities from the observed ion column densities. Using the identified H I and metal line features, we therefore constrain the metallicity of each LLS using a grid of photoionization models created with the CLOUDY code (Ferland et al. 2017). We follow the method outlined in Fumagalli et al. (2016a), creating a grid of models given by the parameter ranges provided in Table 4 using CLOUDY. Using the affine invariant MCMC ensemble sampler EMCEE (Foreman-Mackey et al. 2018), we sample the full parameter space, constraining the posterior probability distribution function (PDF) for the metallicity of each absorber taking the column density constraints given in Table 3 as priors (see also Crighton et al. 2015).

For the DLA, we calculate the metallicity directly from the measured column densities. Following Rafelski et al. (2012), we use sulphur as first preference for calculating the DLA metallicity finding [S/H] $= -1.70 \pm 0.11$. We also compute the measured Fe abundance, finding [Fe/H] $= -1.91 \pm 0.10$. This translates to a metallicity of [M/H] $= -1.61 \pm 0.19$ if we apply a correction of $+0.3$ dex to correct from Fe abundance to α-element metallicity, consistent with the estimate from sulphur (with the caveat that this conversion is estimated from lower redshift samples than our own, e.g. Petitjean, Ledoux & Srianand 2008; Rafelski et al. 2012). Whilst the estimates are consistent, we use the Sulphur derived metallicity as the total DLA metallicity ([M/H]) in the analysis that follows.

The resultant best estimates of the metallicities, plotted versus redshift and N(H I), are shown in Fig. 9, with the error bars showing

Table 3. Properties of absorption line systems identified in the sightline of SDSS J074749.17+115352.4.

Redshift	log N(H I) (cm^{-2})	log N(C II) (cm^{-2})	log N(C IV) (cm^{-2})	log N(Mg I) (cm^{-2})	log N(Mg II) (cm^{-2})	log N(Al II) (cm^{-2})	log N(Al III) (cm^{-2})	log N(Si II) (cm^{-2})	log N(Si IV) (cm^{-2})	log N(Fe II)
5.1612	16.25±0.25	–	13.76±0.02	–	<12.91	–	<12.92	–	<12.60	<12.81
5.1448	21.21±0.05	>14.94	13.63±0.03	11.95 ± 0.06	>14.00	>13.00	–	15.87±0.08	13.18±0.04	14.77±0.03
4.8804	18.70±0.70	<13.40	14.14±0.07	–	<12.78	–	<12.84	<13.52	13.18±0.04	–
4.8004	19.00±0.80	–	13.54±0.02	–	<12.46	<11.98	<12.91	<14.64	<12.94	<12.94
4.6917	17.40±1.00	–	13.18±0.04	–	<12.33	<12.24	–	–	12.62±0.05	<12.98
4.6651	19.03±0.40	–	13.49±0.03	–	12.76±0.03	<12.59	<12.77	<13.19	<12.80	<12.96
4.6169	19.92±0.17	–	13.60±0.04	–	13.02±0.04	<16.08	<12.85	<13.18	12.98±0.04	<13.11
4.1220	17.00±1.80	–	13.74±0.02	–	>13.81	13.25±0.05	<12.62	14.20±0.01	–	13.44±0.03
4.0302	17.90±1.10	–	>14.06	–	13.47±0.08	<12.38	<12.94	<13.15	–	12.73±0.22

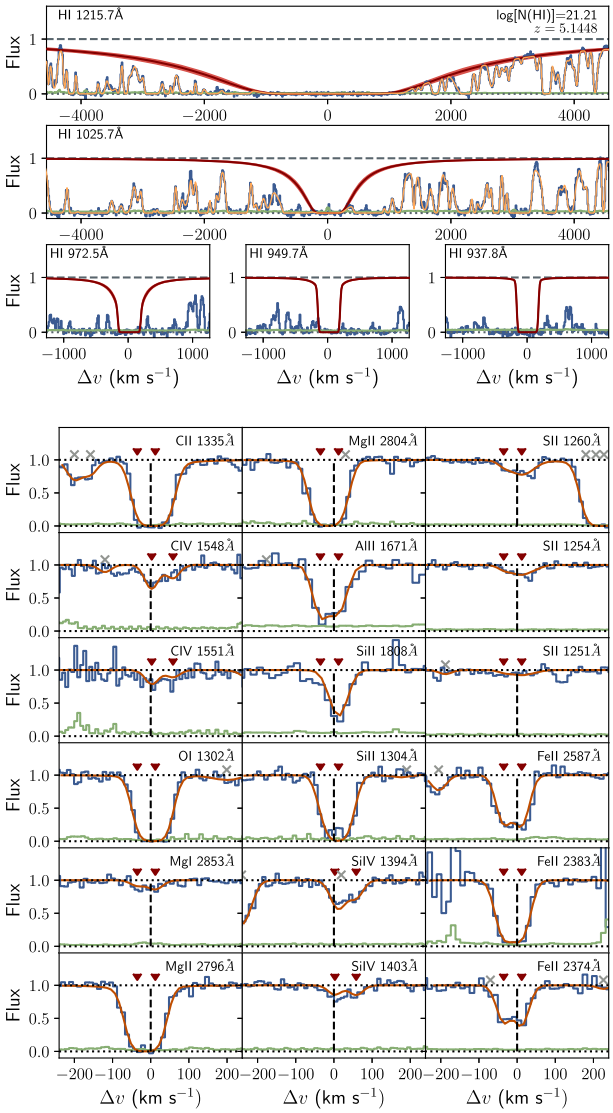

Figure 8. Model fits to the H I absorption and associated metal lines in the X-Shooter spectrum for the DLA system at $z = 5.1448$. The data are shown by the blue stepped line in each panel, with the orange curve showing the overall fit within the wavelength range and, in the case of the Lyman series panels, the red curves show the individual strong H I system (with the width of the curves showing the extent given by the best fitting ranges quoted in Table 3). The upturned triangles in each metal absorption panel identifies centroids of individual absorption components with the given velocity range of the DLA. The grey crosses indicate absorption features identified as not being at the same redshift (within the given velocity range) of the DLA.

Table 4. Parameter ranges and steps adopted to create the photoionization model grid used to constrain the metallicities of the strong absorbers.

Parameter	Min	Max	Step
Redshift	4.1	5.3	0.3
N(H I)	16.0	21.0	0.25
[M/H]	−4.5	0.00	0.25
$n(H)$	−4.25	0.00	0.25

the 10th and 90th percentile range in the MCMC-derived PDF. The circle, diamonds, and × symbols identify absorbers as DLAs, LLSs, and pLLSs (partial Lyman Limit System) respectively (although we note the uncertainties on the pLLS and LLS column density estimates mean the classification are themselves uncertain in some cases). For reference, we also plot the properties of the DLA samples presented in Rafelski et al. (2012, open circles) and Rafelski et al. (2014, open pentagons), and the mean results from the LLS sample of Fumagalli et al. (2016b, open squares). In addition, the hatched region in the left-hand panel shows the mean metallicity as a function of redshift estimated from the C IV mean density as reported in Simcoe et al. (2011). The nine absorbers identified in the SDSS J074749.17+115352.4 line of sight cover a wide range of metallicity (≈ 3 dex), enabling us to study a diverse range of enrichment histories.

5 SIMULATION DATA

To provide additional context to our observations, we incorporate predictions for absorption line column densities from the EAGLE suite of cosmological hydrodynamical ΛCDM simulations (Crain et al. 2015; Schaye et al. 2015). The simulations were run with a modified version of the smoothed particle hydrodynamics (SPH) code GADGET3, incorporating state-of-the-art numerical techniques and subgrid models used to capture physical processes important to galaxy formation and evolution. These include radiative gas cooling, star formation, mass-loss from stars, metal enrichment, energy feedback from star formation, and AGN and gas accretion on to, and mergers of, supermassive black holes. The efficiency of stellar feedback and the mass accretion on to black holes is calibrated to match the present-day stellar mass function of galaxies (subject to the additional constraint that the galaxy sizes need to be realistic), and the efficiency of AGN feedback is calibrated to match the observed relation between stellar mass and black hole mass.

To compare with observations, we select a representative sample of simulated galaxies from the L100N1504 simulation cube, which consists of a cube of side length 100 (comoving) Mpc and resolution 1504^3 dark matter particles. Based on LAE clustering

Table 5. Galaxy absorber associations and derived absorber metallicities. The first sample comprises absorbers purely detected via the presence of absorption by hydrogen and metal species in the quasar sightline. The second sample lists the absorbers identified via a search for saturated H I absorption features within $|\Delta v| < 1000\,\mathrm{km\,s^{-1}}$.

Absorber redshift	N(HI) log[cm^{-2}]	[M/H]	Closest galaxy	ρ (kpc)	Δv (km s^{-1})	$N_{\rm gal}$ ($<$1000 km s^{-1})
5.1612	16.25 ± 0.25	$-2.93^{+0.25}_{-0.33}$	Em-8	162.4	+209.3	1
5.1448	21.21 ± 0.05	-1.70 ± 0.11	Em-8	162.4	+1009.4	1
4.8804	18.70 ± 0.70	<-3.65	Em-15	74.2	+134.1	2
4.8004	19.0 ± 0.80	<-3.99	Em-4	118.0	+963.4	3
4.6917	17.40 ± 1.00	<-2.99	–	–	–	0
4.6651	19.03 ± 0.40	$-2.93^{+0.26}_{-0.35}$	–	–	–	0
4.6169	19.92 ± 0.17	$-2.75^{+0.18}_{-0.20}$	Cn-22	241.8	+430.5	1
4.1220	17.00 ± 1.80	$-1.28^{+0.08}_{-0.09}$	Em-16	210.9	−152.1	1
4.0302	17.90 ± 1.10	$-2.21^{+0.16}_{-0.14}$	–	–	–	0
5.0672	16.9 ± 1.1	–	Em-14/Cn-141	165.6	+107.5	1
4.7437	16.9 ± 1.4	–	Em-13	74.4	+190.1	2
4.5474	17.1 ± 1.2	–	Em-3	120.2	+649.1	1
4.2884	17.2 ± 1.1	–	Em-12	166.8	+352.7	1
4.0862	16.9 ± 2.1	–	Cn-56	193.5	−186.6	2

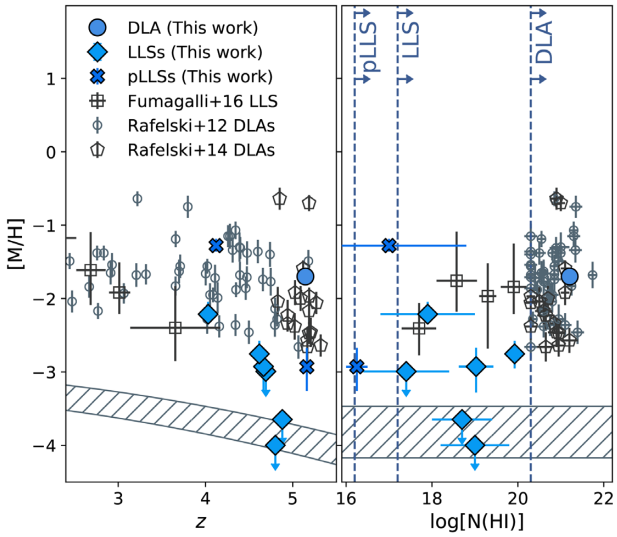

Figure 9. Inferred absorber metallicities as a function of redshift (left-hand panel) and H I column density (right-hand panel). The blue circles, triangles and squares show our own sample split into DLA, LLS, and pLLS subsets. Literature measurements are shown from Rafelski et al. (2012) and Rafelski et al. (2014) for DLAs (open circles and triangles, respectively). The hatched region shows the mean metallicity of the IGM based on the C IV measurements by Simcoe et al. (2011).

analyses (Ouchi et al. 2010), we select central galaxies of haloes in the mass range $10^{11.2-11.8}\,\mathrm{M_\odot}$ and with star formation rates of $>0.3\,\mathrm{M_\odot\,yr^{-1}}$. This selection, effectively matching to observed clustering of LAEs, gives galaxy space densities from EAGLE of $\rho = 1.04 \times 10^{-3}\,\mathrm{Mpc^{-3}}$. This space density is consistent with observed number densities of LAEs at $z \sim 4-5$ (e.g. Drake et al. 2017a), given the flux limits of our data (see Section 3.2).

Calculating accurate simulated H I column densities from the simulations requires that the main ionizing processes that shape the distribution of neutral hydrogen are accounted for. After the collisional ionization (which is dominant at high temperatures), photoionization by the metagalactic UVB radiation is the main contributor to the bulk of hydrogen ionization on cosmic scales, particularly at $z \gtrsim 1$ (e.g. Rahmati et al. 2013a). Rahmati et al. (2015) show that radiation from local sources is important at small scales and high column densities, reducing the covering factors of LLSs and DLAs by only $\approx$ 10 per cent at $\approx R_{\rm vir}$, but by up to $\approx$ 60 per cent at $\approx 0.1 R_{\rm vir}$ for high (N(H I) $\gtrsim 10^{22}$ cm^{-2}) column density systems (see also Rahmati et al. 2013b; Shen et al. 2013; Rahmati & Schaye 2014). The computationally expensive treatment of ionization by local sources is not included in the simulation volume used in this work and so we note that covering fractions of H I may be overpredicted in the simulated volumes by $\approx$ 10–20 per cent for LLSs in our sample and $\approx$ 20–50 per cent for DLAs (albeit within $\lesssim R_{\rm vir}$, equivalent to $\approx$30–40 kpc as discussed later). On the other hand, the H I covering fraction may be underrepresented based on resolution effects of the simulations themselves as recently demonstrated in zoom-in simulations (Peeples et al. 2019; Rhodin et al. 2019; van de Voort et al. 2019).

Following Rahmati et al. (2013b), Rahmati et al. (2015), Oppenheimer et al. (2016), we calculate column densities using SPH interpolation and projecting the ion content of desired regions on to a 2D grid (with 1 pkpc resolution). This projection into 2D was performed through a box of $\pm$600 kpc around each galaxy, equivalent to $\approx 20\,R_{\rm vir}$ from each galaxy along the line of sight. We project maps for each galaxy in the x, y, and z directions that are 600 pkpc across with the depth of 1200 kpc. SPH particles are smoothed onto a grid using the SPH kernel function. We then calculate the median column density profiles and percentile ranges from the resultant sample, using the three projections for each galaxy selected.

6 THE CIRCUMGALACTIC MEDIUM AT $z \approx 5$

6.1 Galaxy-absorber associations

6.1.1 Distribution of neutral hydrogen around $z = 4-5$ LAEs

The observations reveal a wide range of galaxy environments around $z > 4$ strong absorption systems (see Fig. 10 and Table 5 for an overview). Two of the strong absorbers show galaxies with Ly α emission detected within the field of view at $|\Delta v_{-200}| < 100\,\mathrm{km\,s^{-1}}$, with one of these showing two LAEs within this velocity range.

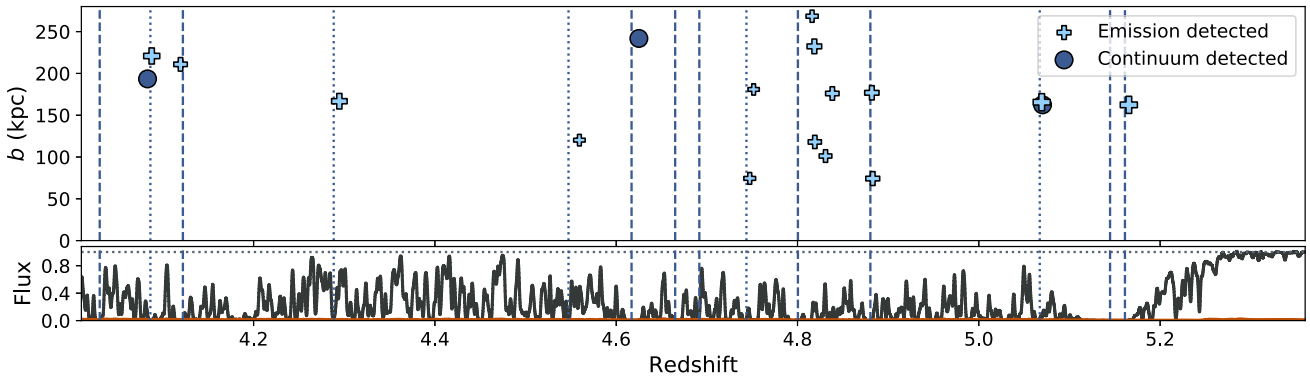

Figure 10. Redshift distribution of galaxies and absorbers. The top panel shows emission (plus symbols) and continuum (circle symbols) detected galaxies as a function of redshift and impact parameter (note the galaxy at $z \sim 5.07$ is detected by both the emission and continuum algorithms). The lower panel shows the flux normalized quasar spectrum. The dashed lines covering both panels highlight redshift positions of strong absorption lines with associated C IV absorption. The dotted lines indicate the positions of the strongest identified absorption features in the proximity of galaxies with no strong absorbers identified in the first pass of the quasar spectrum.

Three more of the absorption systems are seen to lie within $|\Delta v_{-200}| \approx 1000\,\mathrm{km\,s^{-1}}$ of single LAEs, whilst a further absorption system lies at $|\Delta v_{-200}| \approx 1000{-}2000\,\mathrm{km\,s^{-1}}$ from five LAEs tracing a galaxy overdensity at $z \approx 4.83$. The remaining three systems show no detected galaxies within $|\Delta v_{-200}| \approx 2000\,\mathrm{km\,s^{-1}}$. In summary, six of the nine strong H I absorption systems are found to lie within $|\Delta v_{-200}| \approx 1000\,\mathrm{km\,s^{-1}}$ and ≈250 kpc of at least one detected Ly α emitter at the depth of our search.

Comparing to targeted and small-scale IFU searches for galaxies around such absorption systems, where detection rates of ≈ 10 per cent are common at $z \gtrsim 1$ on scales of $\lesssim$100 kpc (e.g. Péroux et al. 2012), the MUSE observations offer a more comprehensive overall picture, with at least one galaxy detected for 66 per cent of the absorbers (up to scales of ≈300 kpc). Such high incidence of galaxies identifications per absorber are comparable with similar MUSE searches at $z \approx 3$ (Mackenzie et al. 2019).

From clustering analyses (e.g. Gawiser et al. 2007; Ouchi et al. 2010; Bielby et al. 2016), LAE samples across redshifts of $3 \lesssim z \lesssim 7$ are inferred to inhabit dark matter haloes with masses of $M_h \approx 10^{11.5\pm0.3}\,\mathrm{M_\odot}$. We derive an estimated virial radius from this halo mass as:

$$R_{\rm vir} = \left(\frac{3M_{\rm h}}{4\pi\rho_{\rm crit}(z)\Delta_{\rm c}}\right)^{1/3}, \quad (1)$$

where $\rho_{\rm crit}$ is the critical density and $\Delta_{\rm c} \approx 18\pi^2$ (Bryan & Norman 1998). This gives $R_{\rm vir} \approx 30{-}50$ kpc at $z = 4{-}5$. Translating this into a Hubble flow velocity, such a separation would equate to $\Delta v \approx 20\,\mathrm{km\,s^{-1}}$. Indeed, as the Hubble parameter at $z \approx 4{-}5$ is $H \approx 500\,\mathrm{Mpc/(km\,s^{-1})}$, the vertical axis in Fig. 11 extends to ≈2 Mpc. Following Bryan & Norman (1998), a halo mass of $\approx 10^{11.5\pm0.3}\,\mathrm{M_\odot}$ at $z \approx 4{-}5$ corresponds to a velocity dispersion of ≈140 km s^{-1}. The concentric dashed ellipses in Fig. 11 show scales of 1, 2, 4, and 8× the typical virial radius and velocity dispersion for a $M_h = 10^{11.5}\,\mathrm{M_\odot}$ halo. Given the assumed halo mass and associated virial radius, all of the proximate galaxies lie $\gtrsim 2R_{\rm vir}$ from the sightline absorbers.

Despite the high number of sources identified near strong absorbers, it is important to maintain the awareness that the detected LAE galaxies are not a comprehensive census of $z \sim 4{-}5$ galaxies within the field. Beside the incompleteness related to the sensitivity of our data, MUSE observations have also limited sensitivity to sources with moderate dust obscuration or where the galaxy Ly α emission is significantly absorbed by the ambient gas. For example, Shapley et al. (2003) report 60 per cent of LBGs at $z \sim 3$ to show Ly α primarily in absorption, which are more difficult to identify in fainter objects due to limited *S/N* in the continuum. It is therefore possible that undetected sources lie in closer proximity to the sightline for some absorbers, and likewise that the three absorbers with no galaxy counterpart may in fact still arise from objects that are undetected. Nevertheless, the current observations provide a valuable census of relatively-bright LAEs up to scales of ≈6−8 $R_{\rm vir}$.

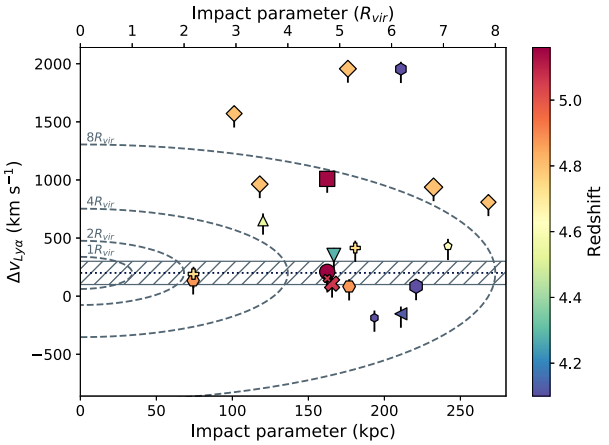

Figure 11. Associations between LAEs and strong absorbers along the sightline to SDSS J074749.17+115352.4. Points are colour coded by redshift with sizes proportional to the log($L_{\rm line}$) of each galaxy. Galaxies clustered around the same absorber share the same symbol. The dotted line and shaded region show $\Delta v = 200 \pm 100\,\mathrm{km\,s^{-1}}$, i.e. the typical offset of Ly α emission from intrinsic galaxy redshift. The top axis scale is in units of $R_{\rm vir}$ is calculated assuming a $M_{\rm h} = 10^{11.5}\,\mathrm{M_\odot}$ halo at $z = 4.8$.

In Fig. 12, we show the H I column density ranges for the absorbers as a function of galaxy impact parameters for all galaxies within $|\Delta v_{-200}| < 1000\,\mathrm{km\,s^{-1}}$ of an absorber. The differing symbols again represent the different absorbers, are colour coded by redshift and are assigned consistently with Fig. 11. Galaxies lying within $|\Delta v_{-200}| < 500\,\mathrm{km\,s^{-1}}$ of an absorber are plotted with a ring around their primary marker, whilst those with additional galaxies at the same redshift are given chevrons (with the number of chevrons

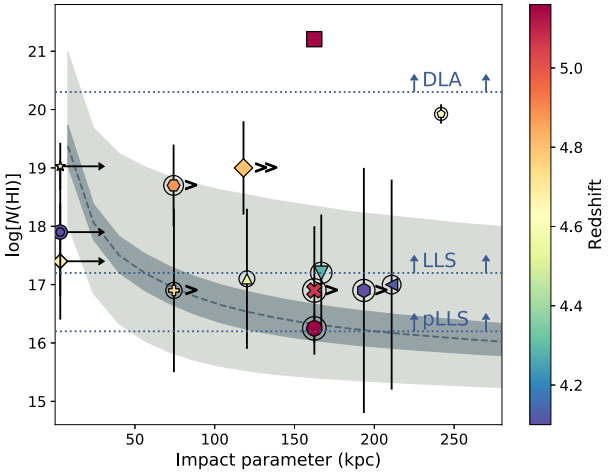

Figure 12. Neutral hydrogen column densities as a function of impact parameter from the nearest detected galaxy within $|\Delta v_{-200}| < 1000\,\mathrm{km\,s^{-1}}$. Symbols are consistent with those given in Fig. 11. Absorber-galaxy pairs lying within $|\Delta v_{-200}| < 500\,\mathrm{km\,s^{-1}}$ are circled. Where multiple galaxies are detected at the same redshift as a given absorber, chevrons are added to the symbol denoting the number of additional galaxies. Limits to the right mark the column density of the absorbers for which we do not identify galaxies. The dashed curve and shaded regions illustrate the median and [32 per cent, 68 per cent] and [10 per cent, 90 per cent] column density ranges predicted by the EAGLE simulations.

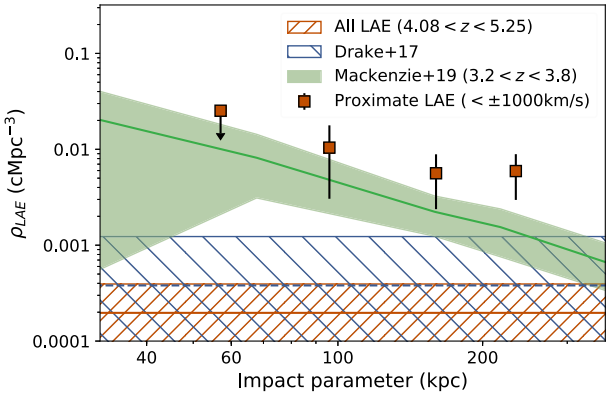

Figure 13. LAE space density within $|\Delta v_{-200}| = \pm 1000\,\mathrm{km\,s^{-1}}$ of identified strong absorption systems, as a function of impact parameter (square points). The light green band shows the overdensity measured around lower redshift DLAs by Mackenzie et al. (2019). The hatched regions show the mean spatial density of all galaxies detected over the redshift range $4.08 < z < 5.16$ in our MUSE field and the space density of LAEs at $z = 4.5$ from the luminosity function of Drake et al. (2017b), given our observational Ly α flux limits.

denoting the number of additional galaxies within $1000\,\mathrm{km\,s^{-1}}$ of the absorber). The dotted horizontal lines denote the observational criteria for the different categories of strong H I absorbers. The dashed curve and shaded regions show the median, [32 per cent, 68 per cent] and [10 per cent, 90 per cent] column density ranges predicted as a function of impact parameter around the galaxy sample extracted from the EAGLE simulation volume.

As would be expected given the galaxy incompleteness and the patchy nature of the CGM, this analysis shows a large scatter in column density versus impact parameter. Four of the detected galaxies show column densities broadly consistent with the range predicted by the simulation, with the weakest two absorbers aligning with galaxies at large impact parameters ($b \gtrsim 150$ kpc). The comparison with simulations further reinforce the idea that galaxies may exist at smaller impact parameters in any of these four cases, as well as the two cases where we see enhanced column densities at large impact parameters compared to the EAGLE predictions. As a slight aside, it is also interesting to note that the EAGLE predicted distribution appears to trace lower redshift results well (e.g. Chen et al. 2001; Keeney et al. 2018).

Whilst inevitably there will be galaxies at the studied redshifts that lie below both our continuum and line emission detection thresholds, the MUSE detected galaxies still provide us with a valuable tracer population for the overdensity of objects in the proximity of strong absorbers. To this end, we now estimate the average galaxy environment around our sample of absorption systems with $N(\mathrm{HI}) > 10^{16.2}\,\mathrm{cm^{-2}}$, and compare it with previous studies and the general field population. Specifically, we compute the space density of galaxies detected in our sample within $|\Delta v_{-200}| < 1000\,\mathrm{km\,s^{-1}}$ of the identified absorption line systems as a function of impact parameter (i.e. within cylindrical volumes assuming the velocity offsets transform to a distance via the Hubble flow), treating the upper end of this velocity range as a physical spatial separation along the line of sight. Note that for this analysis, we only use the nine absorption systems that were detected without prior knowledge of the galaxies identified in the MUSE data. The result is shown in Fig. 13 and is given by the square points. The hatched regions show the mean spatial density of galaxies over the redshift range $4.08 < z < 5.16$ in our MUSE field and the space density of LAEs at $z = 4.5$, given our observational Ly α flux limits, calculated from the luminosity functions of Drake et al. (2017a,b). A clear and significant overdensity of galaxies is evident on average around the absorption systems in our field, establishing a physical connection between the gas probed in absorption and the distribution of galaxies within the adopted window. Mackenzie et al. (2019) presented a similar analysis showing the overdensity of LAEs around DLAs at $z \approx 3.2-3.8$, shown here as the filled green region. Our own result at this higher redshift shows a comparable clustering of LAEs (in agreement within 1σ level) around absorption systems, including lower column density ones.

6.1.2 *Metal absorption around $z \sim 4-5$ LAEs*

C IV absorption in quasar sightlines has been shown to have a significant correlation with galaxies via clustering analyses at $z \sim 3$ (e.g. Adelberger et al. 2003; Crighton et al. 2011; Turner et al. 2014), whilst lower redshift studies (i.e. $z \sim 0.01$) have shown that C IV systems are more common in low galaxy-density regions (e.g. Burchett et al. 2016).

With these new observations, we can now investigate the link between C IV and galaxies at $z > 4$. In the top panel of Fig. 14, we show C IV column density versus galaxy impact parameters. Where a C IV detection is made, we show all galaxies within $|\Delta v_{-200}| = \pm 1000\,\mathrm{km\,s^{-1}}$. In this figure, we also include upper limits on the column density (grey crosses) computed with $|\Delta v_{-200}| = \pm 250\,\mathrm{km\,s^{-1}}$ from galaxies within the field of view with no detected absorber. Finally, we mark the three absorbers where we fail to detect any galaxies by lower limits with respect to the impact parameter. As in Fig. 12, the dashed and shaded regions show the median, [32 per cent, 68 per cent] and [10 per cent, 90 per cent] column density ranges predicted around the galaxy sample extracted from the EAGLE simulation volume. The central panel shows our equivalent to the top panel, but for Mg II, whilst the lower panel

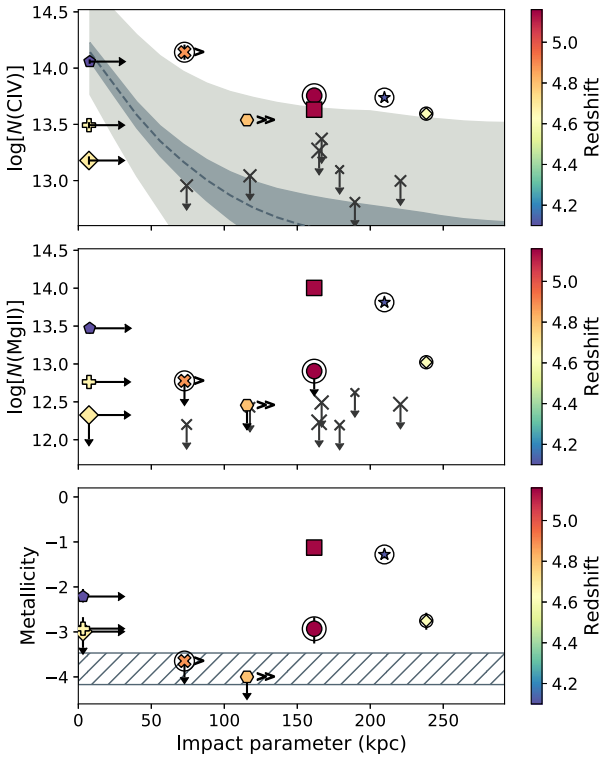

Figure 14. Column density of C IV and Mg II and inferred metallicity as a function of impact parameter to galaxies identified within $|\Delta v_{-200}| < 1000\,\mathrm{km\,s^{-1}}$. Galaxies lying within $\Delta v = \pm 500\,\mathrm{km\,s^{-1}}$ of absorbers are circled. Each detection is colour coded by redshift as given by the colour bar, with the symbols assigned as in Fig. 11. Where multiple galaxies are detected at the same redshift as a given absorber, chevrons are added to the symbol denoting the number of additional galaxies. Upper limits (grey × symbols, 3σ upper limits) are estimated within $\pm 250\,\mathrm{km\,s^{-2}}$ of high-confidence galaxies at $4 < z < 5.2$ with no detected C IV. Absorbers with no detected proximate galaxy are instead shown by lower limits in impact parameter.

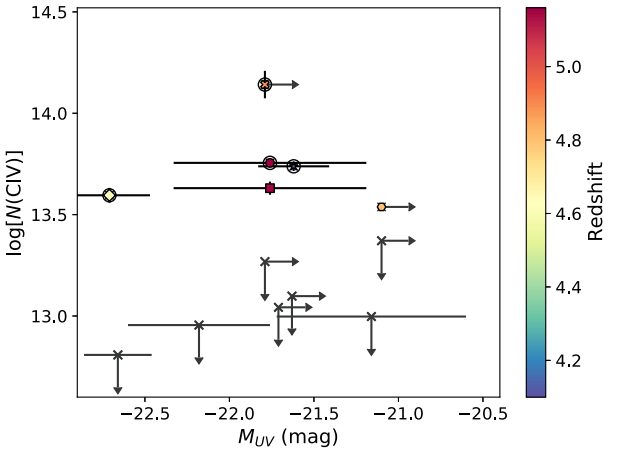

Figure 15. Column density of C IV as a function of the absolute UV magnitude of the nearest galaxy identified within $|\Delta v_{-200}| < 1000\,\mathrm{km\,s^{-1}}$. Galaxies lying within $\Delta v = \pm 500\,\mathrm{km\,s^{-1}}$ of absorbers are circled. Each detection is colour coded by redshift as given by the colour bar, with the symbols assigned as in Fig. 11. Where multiple galaxies are detected at the same redshift as a given absorber, chevrons are added to the symbol denoting the number of additional galaxies. Upper limits (grey × symbols, 3σ upper limits) are estimated within $\pm 250\,\mathrm{km\,s^{-2}}$ of high-confidence galaxies at $4 < z < 5.2$ with no detected C IV.

shows the estimated metallicities derived from the MCMC analysis described previously. The hatched region in the lower panel marks the range in metallicities of the CGM at $4 < z < 5$ estimated from Simcoe et al. (2011) as in Fig. 9.

We find that of the 17 galaxies detected in the MUSE field of view between $4 \lesssim z \lesssim 5.3$, 8 lie within $1000\,\mathrm{km\,s^{-1}}$ of a detected C IV absorber (with a further two galaxies within $2000\,\mathrm{km\,s^{-1}}$ of C IV absorption). The remaining galaxies show upper limits on C IV absorption. Fig. 14 highlights that the three systems with the highest C IV column densities are all within $500\,\mathrm{km\,s^{-1}}$ of a detected galaxy. Mindful of the small sample size and selection biases described above, this finding is at least consistent with the conclusion of previous work (Keating et al. 2020; Meyer et al. 2019a), according to which galaxies at these redshifts are expected to reside within ionized regions, and thus cluster more with high-ionisation lines rather than neutral species. We note, however, that this correspondence is not unique, and that gas with different ionization stages is still found close to galaxies at these redshifts. Indeed, our sample contains also galaxies in proximity to absorbers showing high column density of both C IV and Mg II, but with low column density of H I. We note how there seems to be a lack of clear correlation between metallicity and proximity to galaxies both in projection and in velocity space. Evidence of inhomogeneous enrichment is indeed building up at lower redshifts (e.g. Fumagalli, O'Meara & Prochaska 2011a; Mackenzie et al. 2019; Lofthouse et al. 2020; Fossati et al. 2019), and it is thus not surprising to find hints of similar inhomogeneity at these higher redshifts, where the time available for enriched pockets to grow is limited.

When compared to the EAGLE simulated data, we find in our observations a significant fraction of upper limits at impact parameters of > 100 kpc, consistent with the predictions from simulation. However, we also find several absorbers with high column density of C IV, in excess to the EAGLE predictions for a given impact parameter. While this may reflect again incompleteness in our search towards low impact parameters, a deficit of strong C IV near simulated galaxies has been documented in the literature before (Finlator et al. 2016), and may be common feature of simulations. Adding to this line of enquiry, we show in Fig. 15 the C IV absorber column densities as a function of the UV absolute magnitude of the nearest galaxy. There is no clear correlation between the UV brightness of the galaxies and the C IV column density, although the strongest absorber does lie in the proximity of one of the faintest galaxies (which also happens to be the closest galaxy to the quasar sightline).

Further, we plot in Fig. 16 the distributions of closest galaxy properties with the blue diagonal hatched histograms showing galaxies with proximate C IV absorption and the red vertical hatched histograms showing galaxies with no proximate C IV absorption. We see no clear difference between the two samples for impact parameter, UV brightness M_{UV}, Ly α flux and galaxy environment (N_{gal}).

Inverting the question, we now investigate the absorption line system properties as a function of the nearest galaxy Ly α luminosity. We show the proximate absorber H I column density, C IV column density, and metallicity as a function of galaxy Ly α luminosity in the top, middle, and bottom panels of Fig. 17, respectively.

We see little indication of any clear correlations between the absorber system properties and the proximate galaxy Ly α luminosity. The galaxies are found in the proximity of a range of absorbers at

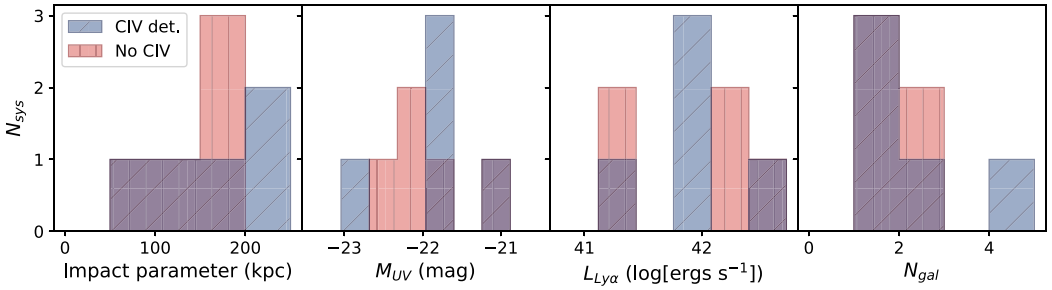

Figure 16. Properties of closest detected galaxies (including both emission and continuum detected) split by whether proximate C IV absorption is detected (blue diagonal hatched histograms) or not detected (red vertical hatched histograms).

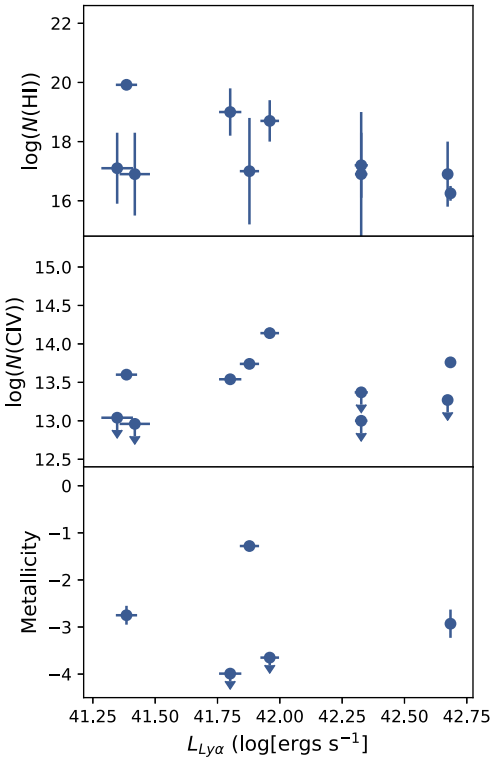

Figure 17. Absorption line properties as a function of nearest galaxy Ly α luminosity, showing the H I column density (top panel), the C IV column density (middle panel), and the absorber metallicity (bottom panel).

N(HI) $\gtrsim 10^{16}$ cm^{-2}, covering a wide range in metallicities with no clear correlation with galaxy Ly α emission. Of course the sample probes a wide range of impact parameter between galaxy and absorber, which as suggested by the EAGLE predictions in Figs 12 and 14 will contribute to scatter in the results. In addition, the simulations predict a significant scatter in the absorber column densities for a given galaxy halo, such that the column density of gas any given sightline may pass through is modulated by the patchy covering factor at any specific location. Our observations seem to support this in so much as the absorber properties show little correlation with the measured galaxy properties in this small sample.

Our own results resolving LAEs around strong absorption systems coincident with C IV absorption are complemented by two concurrent studies: Finlator et al. (2020) and Díaz et al. (2020). Both studies see a clear trend for C IV absorption to be coincident with the galaxy population at scales of ∼100–200 kpc, as we have also shown here. Further Díaz et al. (2020) also report a preference for the galaxies in the proximity of C IV absorption to be towards the faint end of the Ly α luminosity function, something that we do not observe in our own sample.

6.2 The gas environment of $z > 4$ groups

Looking in more detail at the two incidences of multiple galaxies at the same redshift as a given absorption line system (those highlighted by × and hexagonal symbols), we find in both cases prominent H I absorption giving rise to LLSs, which themselves are found to have low metallicity ($\lesssim 10^{-3}$).

Considering the group of five emitters detected at $z \approx 4.82$, these lie along an axis on the plane of the sky from East to West. This can be seen in Fig. 5, where the $z \sim 4.81$ structure is traced by LAE sources 1, 4, and 7 clustered together closely in the East to 19 and 21 scattered in the West of the image. The galaxy pair at $z = 4.88$ (LAEs 2 and 15) lies along a similarly projected axis on the sky. Should the five $z \approx 4.82$ galaxies lie in a galaxy group or protocluster environment, a halo mass of $M_{\rm h} \approx 10^{12.5-13.5}$ M$_\odot$ would perhaps be reasonable (e.g. Jenkins et al. 2001; Springel et al. 2005), which would give a virial radius of $r_{\rm vir} \approx 80$–160 kpc. In projection then at least, the strong absorber may well lie within $\approx 1-2R_{\rm vir}$ of a group environment traced by these galaxies.

Similar to the C IV case of Burchett et al. (2013) at low redshift, for example, the coincidence of multiple galaxies with metal-poor gas may be a case of intergalactic gas accreting onto the galaxy group, or it may be tracing the CGM of an undetected galaxy in the observed group environment. Indeed, that these two relatively pristine absorption systems align with apparently dense galaxy environments also finds commonality with recent studies of the galaxy environment of metal-poor LLSs at $z \sim 3$ (Fumagalli et al. 2016b; Lofthouse et al. 2020). As in the two cases here, the $z \sim 3$ metal-poor environments show galaxy overdensities at 80 kpc $\lesssim b \lesssim$ 300 kpc, within $|\Delta v_{-200}| \lesssim 1000$ km s^{-1}. Similar to these previous works, we argue that the low metallicities that we measure for these systems suggests they are not outflows in nature, but more likely tracing cold gas within the cosmic web of gas surrounding the galaxy overdensity. Indeed, as with the previous examples at $z \approx 3.1$, the presence of an overdensity strongly favours this gas ultimately being accreted onto the galaxy population with time, further fuelling star formation in this region. Interestingly, these systems would appear to fit with recent simulated results (Mandelker et al. 2019), which suggest there should be a population of metal-poor LLS in the IGM. Such systems are predicted from the merger and collapse of sheets of

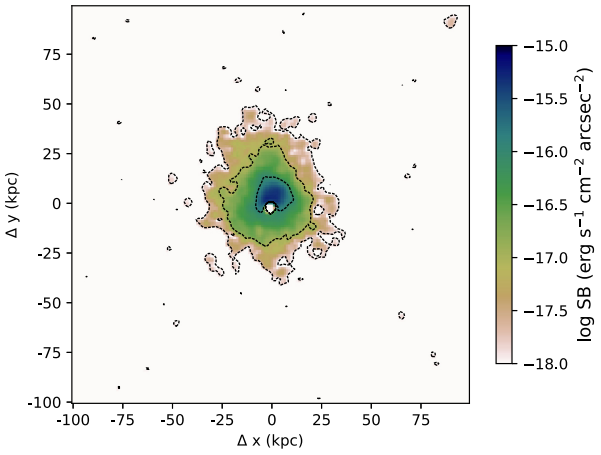

Figure 18. Optimally extracted narrow-band image of the diffuse Ly α emission at the redshift of the quasar SDSS J074749.17+115352.4. The dotted contours are at surface brightness levels of $10^{-18,\,-17,\,-16}$ erg s^{-1} cm^{-2} arcsec^{-2}.

gas, forming pressure confined cool $\lesssim$ kiloparsec-scale gas clouds with densities of $n \gtrsim 10^{-3}$ cm^{-3}.

6.3 Extended emission associated to the quasar

Complementing the investigation of the CGM in absorption, the MUSE data also provides a probe of the CGM via Ly α emission around the background quasar, exploiting the presence of the bright central active galactic nucleus which illuminates the surrounding gas (e.g. Christensen et al. 2006; Goto et al. 2009; Cantalupo et al. 2014; Martin et al. 2014; Hennawi et al. 2015; Borisova et al. 2016; Fumagalli et al. 2016b; Arrigoni Battaia et al. 2019; Farina et al. 2019). Whilst the LAEs and LBGs predominantly probe halo masses of $\sim 10^{11}$ M$_\odot$ (e.g. Adelberger et al. 2003; Ouchi et al. 2010; Bielby et al. 2016, 2013), quasars are more often found in higher mass haloes (e.g. Ross et al. 2009; Chehade et al. 2016; Stott et al., in preparation) offering a probe of the gas in haloes at the higher end of the halo mass function.

For this analysis, we first prepare the MUSE data cube by subtracting the quasar PSF and the continuum of other sources using the CUBEX tools, as described in detail in Borisova et al. (2016) and Arrigoni Battaia et al. (2019). After smoothing the cube with a Gaussian filter of 3 pixels in radius, we search for extended Ly α emission at the quasar redshift, running CUBEX to detect connected pixels with $S/N \geq 3$. We repeat this procedure on the mean- and median-combined cube, finding consistent results.

With this analysis, we clearly identify an extended nebula with size of ≈ 50 kpc at a Ly α surface brightness level of $\approx 10^{-18}$ erg s^{-1} cm^{-2} arcsec^{-2} and a peak surface brightness of $>10^{-16}$ erg s^{-1} cm^{-2} arcsec^{-2}. The nebula, shown in Fig. 18 in an optimally extracted narrow-band image, appears roughly symmetric at the current surface brightness limit of 3.8×10^{-18} erg s^{-1} cm^{-2} arcsec^{-2} integrated in a 25 Å window, with a hint of elongation in the North direction. In Fig. 19, we compare the nebula's surface brightness profile (extracted from a narrow-band image of 25 Å in width and centred at the peak line emission of the nebula) with the average profiles of other nebulae detected with MUSE at $z > 3$ (Borisova et al. 2016; Arrigoni Battaia et al. 2019; Farina et al. 2019) and with the Keck Cosmic Web Imager (KCWI) at $z \sim 2$ (Cai et al. 2019). Once corrected for the cosmological surface brightness dimming, the nebula around SDSS J074749.17+115352.4 is found to be broadly consistent with the ranges measured for such nebulae at $z > 3$, but with somewhat enhanced surface brightnesses at scales of $r \lesssim 80$ comoving kpc. The surface brightness profile of SDSS J074749.17+115352.4 is also comparable to the profiles reported around $z \approx 6$ quasars in Drake et al. (2019) and Farina et al. (2019).

Although only a single object, SDSS J074749.17+115352.4 aligns with the mild redshift evolution noted in Cai et al. (2019) between $z \lesssim 3$ and $z \gtrsim 3$. Based on the argument in Arrigoni Battaia et al. (2019), we could be witnessing an increase in the amount of cold gas at $z \approx 5$ in this host galaxy possibly due to intense gas accretion. It should be noted, however, that SDSS J074749.17+115352.4 is among the brightest quasars known at $z \approx 5$ (Wang et al. 2016). Indeed, after re-normalizing the X-Shooter spectrum using a curve-of-growth analysis of the quasar photometry in the reconstructed i -band from MUSE data, we measure a quasar luminosity of $M_{1450} = -28.069 \pm 0.003$ mag. This is in excellent agreement with the value reported by Wang et al. (2016), $M_{1450} = -28.04$ mag, thus ruling out significant variability over the time-scale of our observations. Compared to the characteristic luminosity of $z \approx 5$ quasars, $M^*_{1459} = -26.98 \pm 0.23$ (Yang et al. 2016), SDSS J074749.17+115352.4 is clearly among the brightest quasars at these redshifts, and is at the bright end of the luminosity distribution probed at $z \approx 3$ (e.g. $-28.29 \leq M_{1450} \leq -25.65$ in Arrigoni Battaia et al. 2019). Therefore, while no significant trend has been observed thus far between the quasar luminosity and the brightness of the Ly α nebulae (e.g. Christensen et al. 2006; Arrigoni Battaia et al. 2019), it remains possible that the surface brightness enhancement may be linked to the quasar luminosity.

The apparent steeper surface brightness profile (compared to the average nebulae profiles) shown in Fig. 19 would appear suggestive of a more compact nebula. In the right-hand panel of Fig. 19, we investigate whether this difference can be attributed to a scaling proportional to the typical size of the host halo. For this, we re-scale the profiles assuming the virial radius at a characteristic halo mass of $\log M_{\rm halo}/M_\odot \approx 12.5$, independent of redshift (see discussion in Arrigoni Battaia et al. 2019), which is $R_{\rm vir} \approx 80$ kpc at $z \approx 5.26$. Admittedly, there is very large uncertainty on the evolution of the typical halo mass of quasar hosts beyond $z > 3$ (Timlin et al. 2018), with uncertainties over a factor of 2−3. However, unless the virial radius is $\lesssim 50$ kpc, corresponding to haloes with $\log M_{\rm halo}/M_\odot \lesssim 12$ at this redshift (open squares in Fig. 19), the nebula around SDSS J074749.17+115352.4 suggests an intrinsic difference, e.g. due to the underlying density profile. Again, we note the quasar's high intrinsic luminosity, which would suggest a larger, rather than smaller, host halo. Indeed, given the high luminosity of SDSS J074749.17+115352.4 a virial radius of $R_{\rm vir} \approx 100$ kpc would not be unreasonable.

7 SUMMARY AND CONCLUSIONS

We have presented a survey of galaxies in a single field observed with the VLT MUSE IFU, in conjunction with VLT X-Shooter moderate resolution spectroscopy of a background quasar at $z \approx 5.26$ exhibiting multiple strong absorption lines. Our analysis shows the following key results.

(i) Based on 9 C IV detected, N(HI) $\gtrsim 10^{16}$ cm^{-2}, absorbers we find a galaxy detection fraction of 67 per cent (6/9) for galaxies within $b < 250$ kpc and $|\Delta v_{-200}| < 1000$ km s^{-1}. Applying a more

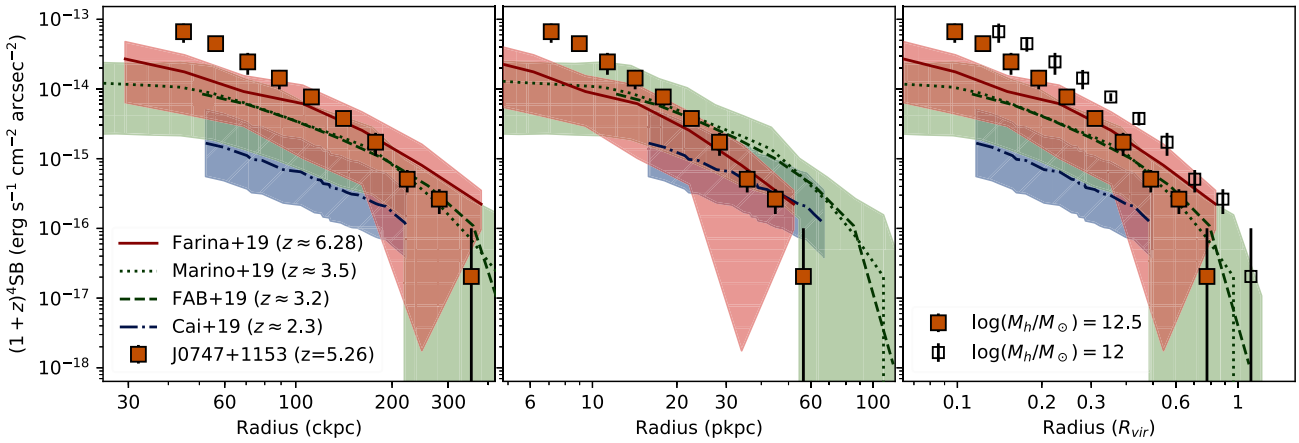

Figure 19. Surface brightness profiles of Ly α nebulae, as a function of comoving radius (left-hand panel), proper radius (centre), and units of the virial radius of the inferred host halo mass (assuming $M_h = 10^{12.5}$ M$_\odot$, right-hand panel). The results based on SDSS J074749.17+115352.4 in this work are shown by the filled black squares. Also shown are the average profiles from Cai et al. (2019) at $z \approx 2.3$ (blue dot-dashed line and shaded region – which shows the 10–90th percentile range), Arrigoni Battaia et al. (2019) at $z \approx 3.17$ (dashed line), Borisova et al. (2016) at $z \approx 3.5$ (green dotted line and shaded region – which shows the 10–90th percentile range taken from Marino et al. 2019); and Farina et al. (2019) at $z \approx 6.3$ (solid red curve and shaded region). The open black squares in the right-hand panel show the profile for SDSS J074749.17+115352.4 renormalized assuming a virial radius of 50 kpc (i.e. $M_h = 10^{12}$ M$_\odot$) instead of 80 kpc. The profile for SDSS J074749.17+115352.4 appears to be brighter and steeper compared to lower redshifts systems.

stringent velocity offset cut-off $|\Delta v_{-200}| < 500\,\mathrm{km\,s^{-1}}$, we find a detection fraction of 44 per cent (4/9). This is based on a Ly α flux limit of $3.1 \times 10^{18}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$ (50 per cent completeness for extended sources).

(ii) Taking the systems together as a whole, we find that galaxies are correlated with the strong absorption systems, showing a statistically significant overdensity of galaxy numbers when compared to the background population at $z \sim 4-5$ and given the flux limit of our observations.

(iii) We detect two low-metallicity absorption systems apparently associated (at ≈100 kpc separation) with candidate group environments, with 2 or more detected LAE galaxies within the MUSE field of view.

(iv) We detect extended Ly α emission around the $z = 5.26$ quasar, reaching extents of ≈50 kpc at a surface brightness limit of $3.8 \times 10^{-18}\,\mathrm{erg\,s^{-1}\,cm^{-2}\,arcsec^{-2}}$. After scaling for surface brightness dimming, we find that this nebula is centrally brighter, with a steeper radial profile, than the average for nebulae studied at $z \sim 3$, hinting at a mild redshift evolution in such nebulae.

Overall, we find a picture that, whilst strong H I systems are correlated with LAE galaxies at $z \sim 5$, this coverage is patchy. Our results have shown that, on average, the galaxy environment is overdense around LLSs and DLAs at $z \sim 4$–5, similar to Mackenzie et al. (2019) and Lofthouse et al. (2020) at $z \approx 3$–4. The individual systems lie across a wide range of galaxy environments however, from groups of LAEs to no detected LAEs at all (although of course there will inevitably be fainter galaxies below our detection limits in most cases, c.f. Rahmati & Schaye 2014). Ultimately, we find that H I gas is often associated with complex structures as opposed to, or as well as, single galaxies.

All the detected LAEs in our sample are at impact parameters of $b \gtrsim 2R_{\rm vir}$, lending support to the analysis of Keating et al. (2020), suggestive of Ly α emission from galaxies in close proximity to strong absorption being absorbed, potentially due to patchy reionization. Whilst Keating et al. (2020) and Becker et al. (2018) measure underdensities of LAEs at impact parameters of up to ≈3 Mpc in simulations and observations respectively, we find a lack of LAEs only at $\lesssim$60 kpc around the LLS and DLA systems presented here. This smaller on-sky scale may indeed be expected if the strong absorbers in our sample are tracing multiple smaller islands of non-reionized material compared to the extended trough observed by Becker et al. (2015, 2018). The statistical uncertainties on our sample however cannot rule out a continuation of the overdensity of LAEs detected at 60 kpc $\lesssim b \lesssim$ 300 kpc to these smaller scales at $b \lesssim 60$ kpc from strong H I absorbers. Complementing the analyses of the neutral H I gas distribution, our analysis of metals and system metallicity suggest an inhomogeneously enriched medium. Both observations and simulation predictions show a large scatter in C IV column density around $z \sim 5$ galaxies, whilst no correlation is evident between metallicity and galaxy proximity.

These data, alongside similar recent studies at lower redshift (Bielby et al. 2017a; Chen et al. 2019; Fossati et al. 2019; Lofthouse et al. 2020; Mackenzie et al. 2019), highlight the value of deep IFU observations in performing blind identification of galaxies in the proximity of quasar sightline data. Going forward the available VLT X-Shooter sample of bright $z \sim 5$ quasars present an indispensable base for developing these studies with further VLT/MUSE IFU observations. In conjunction with this, the simulations have shown that if we are to properly understand the galaxy–absorber connection, higher signal-to-noise sightline spectra are required beyond the existing data in order to probe to lower column density metal absorption features. In summary, these data lay the ground for a survey comprising deep sightline spectra combined with extensive MUSE IFU data and ultimately supported by NIR spectroscopic observations, either with *JWST* or the E-ELT, to probe more fully the component of the galaxy population not obscured in Ly α emission.

ACKNOWLEDGEMENTS

Based on observations made with ESO Telescopes at the La Silla Paranal Observatory under programme ID 0102.A−0261. RMB and MF acknowledge support by the Science and Technology Facilities Council [grant number ST/P000541/1]. This project has received funding from the European Research Council (ERC) under

the European Union's Horizon 2020 research and innovation programme (grant agreement No. 757535). MR acknowledges support by Space Telescope Science Institute's Director's Research Funds. SC gratefully acknowledges support from Swiss National Science Foundation grant PP00P2-163824. The Cosmic Dawn center is funded by the Danish National Research Foundation (DNRF). SL was funded by project FONDECYT 1191232. For access to the codes and advanced data products used in this work, please contact the authors or visit http://www.michelefumagalli.com/codes.html. Raw data are publicly available via the ESO Science Archive Facility.

REFERENCES

Adelberger K. L., Steidel C. C., Shapley A. E., Pettini M., 2003, ApJ, 584, 45
Arrigoni Battaia F., Hennawi J. F., Prochaska J. X., Oñorbe J., Farina E. P., Cantalupo S., Lusso E., 2019, MNRAS, 482, 3162
Bacon R. et al., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III. SPIE, Bellingham, p. 773508
Barnes L. A., Haehnelt M. G., 2014, MNRAS, 440, 2313
Barnes L. A., Haehnelt M. G., Tescari E., Viel M., 2011, MNRAS, 416, 1723
Becker G. D., Sargent W. L. W., Rauch M., Carswell R. F., 2012, ApJ, 744, 91
Becker G. D., Bolton J. S., Madau P., Pettini M., Ryan-Weber E. V., Venemans B. P., 2015, MNRAS, 447, 3402
Becker G. D., Davies F. B., Furlanetto S. R., Malkan M. A., Boera E., Douglass C., 2018, ApJ, 863, 92
Becker G. D. et al., 2019, ApJ, 883, 163
Bergeron J., Boisse P., 1986, A&A, 168, 6
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bielby R. et al., 2013, MNRAS, 430, 425
Bielby R., Crighton N. H. M., Fumagalli M., Morris S. L., Stott J. P., Tejos N., Cantalupo S., 2017a, MNRAS, 468, 1373
Bielby R. M. et al., 2016, MNRAS, 456, 4061
Bielby R. M. et al., 2017b, MNRAS, 471, 2174
Bielby R. M. et al., 2019, MNRAS, 486, 21
Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2846
Bigiel F. et al., 2011, ApJ, 730, L13
Bird S., Haehnelt M., Neeleman M., Genel S., Vogelsberger M., Hernquist L., 2015, MNRAS, 447, 1834
Borisova E. et al., 2016, ApJ, 831, 39
Bouché N., Lowenthal J. D., Charlton J. C., Bershady M. A., Churchill C. W., Steidel C. C., 2001, ApJ, 550, 585
Bouché N., Murphy M. T., Péroux C., Davies R., Eisenhauer F., Förster Schreiber N. M., Tacconi L., 2007, ApJ, 669, L5
Bouché N. et al., 2010, ApJ, 718, 1001
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Bunker A. J., Warren S. J., Clements D. L., Williger G. M., Hewett P. C., 1999, MNRAS, 309, 875
Burchett J. N., Tripp T. M., Werk J. K., Howk J. C., Prochaska J. X., Ford A. a. B., Davé R., 2013, ApJ, 779, L17
Burchett J. N. et al., 2016, ApJ, 832, 124
Cai Z., Fan X., Dave R., Finlator K., Oppenheimer B., 2017, ApJ, 849, L18
Cai Z. et al., 2019, ApJS, 245, 23
Cantalupo S., Arrigoni-Battaia F., Prochaska J. X., Hennawi J. F., Madau P., 2014, Nature, 506, 63
Cantalupo S. et al., 2019, MNRAS, 483, 5188
Chehade B. et al., 2016, MNRAS, 459, 1179
Chen H.-W., Lanzetta K. M., Webb J. K., Barcons X., 2001, ApJ, 559, 654
Chen H.-W., Boettcher E., Johnson S. D., Zahedy F. S., Rudie G. C., Cooksey K. L., Rauch M., Mulchaey J. S., 2019, ApJ, 878, L33
Christensen L., Jahnke K., Wisotzki L., Sánchez S. F., 2006, A&A, 459, 717
Crain R. A. et al., 2015, MNRAS, 450, 1937
Crighton N. H. M. et al., 2011, MNRAS, 414, 28
Crighton N. H. M., Hennawi J. F., Simcoe R. A., Cooksey K. L., Murphy M. T., Fumagalli M., Prochaska J. X., Shanks T., 2015, MNRAS, 446, 18
D'Odorico V. et al., 2018, ApJ, 863, L29
Daddi E. et al., 2007, ApJ, 670, 156
Davé R., Finlator K., Oppenheimer B. D., 2012, MNRAS, 421, 98
Dekel A., Mandelker N., 2014, MNRAS, 444, 2071
Dekel A., Sari R., Ceverino D., 2009, ApJ, 703, 785
Díaz C. G., Koyama Y., Ryan-Weber E. V., Cooke J., Ouchi M., Shimasaku K., Nakata F., 2014, MNRAS, 442, 946
Díaz C. G., Ryan-Weber E. V., Cooke J., Koyama Y., Ouchi M., 2015, MNRAS, 448, 1240
Díaz C. G., Ryan-Weber E., Karman W., Caputi K., Salvadori S., Crighton N., Ouchi M., Vanzella E., 2020, preprint (arXiv:2001.04453)
Djorgovski S. G., Pahre M. A., Bechtold J., Elston R., 1996, Nature, 382, 234
Drake A. B. et al., 2017a, MNRAS, 471, 267
Drake A. B. et al., 2017b, A&A, 608, A6
Drake A. B., Farina E. P., Neeleman M., Walter F., Venemans B., Banados E., Mazzucchelli C., Decarli R., 2019, ApJ, 881, 131
Elbaz D. et al., 2007, A&A, 468, 33
Farina E. P. et al., 2019, ApJ, 887, 196
Faucher-Giguère C.-A., Kereš D., 2011, MNRAS, 412, L118
Faucher-Giguère C.-A., Feldmann R., Quataert E., Kereš D., Hopkins P. F., Murray N., 2016, MNRAS, 461, L32
Ferland G. J. et al., 2017, Rev. Mex. Astron. Astrofis., 53, 385
Finlator K., Oppenheimer B. D., Davé R., Zackrisson E., Thompson R., Huang S., 2016, MNRAS, 459, 2299
Finlator K., Doughty C., Cai Z., Díaz G., 2020, MNRAS, 493, 3223
Foreman-Mackey D. et al., 2018, Dfm/Emcee: Emcee V3.0Rc2. doi:10.5281/zenodo.1436565
Fossati M. et al., 2019, MNRAS, 490, 1451
Fumagalli M., O'Meara J. M., Prochaska J. X., 2011a, Science, 334, 1245
Fumagalli M., Prochaska J. X., Kasen D., Dekel A., Ceverino D., Primack J. R., 2011b, MNRAS, 418, 1796
Fumagalli M., O'Meara J. M., Prochaska J. X., Worseck G., 2013, ApJ, 775, 78
Fumagalli M., O'Meara J. M., Prochaska J. X., 2016a, MNRAS, 455, 4100
Fumagalli M., Cantalupo S., Dekel A., Morris S. L., O'Meara J. M., Prochaska J. X., Theuns T., 2016b, MNRAS, 462, 1978
Fumagalli M., Haardt F., Theuns T., Morris S. L., Cantalupo S., Madau P., Fossati M., 2017a, MNRAS, 467, 4802
Fumagalli M. et al., 2017b, MNRAS, 471, 3686
Fynbo J. U., Burud I., Møller P., 2000, A&A, 358, 88
Fynbo J. P. U. et al., 2013, MNRAS, 436, 361
Gawiser E. et al., 2007, ApJ, 671, 278
Goto T., Utsumi Y., Furusawa H., Miyazaki S., Komiyama Y., 2009, MNRAS, 400, 843
Grove L. F., Fynbo J. P. U., Ledoux C., Limousin M., Møller P., Nilsson K. K., Thomsen B., 2009, A&A, 497, 689
Hamanowicz A. et al., 2020, MNRAS, 492, 2347
Hartoog O. E. et al., 2015, A&A, 580, A139
Heckman T. M., Armus L., Miley G. K., 1990, ApJS, 74, 833
Hennawi J. F., Prochaska J. X., Cantalupo S., Arrigoni-Battaia F., 2015, Science, 348, 779
Hinton S. R., Davis T. M., Lidman C., Glazebrook K., Lewis G. F., 2016, Astron.Comput., 15, 61
Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372

Kacprzak G. G., Murphy M. T., Churchill C. W., 2010, MNRAS, 406, 445
Keating L. C., Weinberger L. H., Kulkarni G., Haehnelt M. G., Chardin J., Aubert D., 2020, MNRAS, 491, 1736
Keeney B. A. et al., 2018, ApJS, 237, 11
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Klitsch A., Péroux C., Zwaan M. A., Smail I., Oteo I., Biggs A. D., Popping G., Swinbank A. M., 2018, MNRAS, 475, 492
Klitsch A. et al., 2019, MNRAS, 482, L65
Krogager J. K., Møller P., Fynbo J. P. U., Noterdaeme P., 2017, MNRAS, 469, 2959
Kulkarni G., Keating L. C., Haehnelt M. G., Bosman S. E. I., Puchwein E., Chardin J., Aubert D., 2019, MNRAS, 485, L24
Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2782
Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, ApJ, 772, 119
Lofthouse E. K. et al., 2020, MNRAS, 491, 2057
López S. et al., 2016, A&A, 594, A91
Lowenthal J. D., Hogan C. J., Green R. F., Woodgate B., Caulet A., Brown L., Bechtold J., 1995, ApJ, 451, 484
Mackenzie R. et al., 2019, MNRAS, 487, 5070
Mandelker N., van den Bosch F. C., Springel V., van de Voort F., 2019, ApJ, 881, L20
Marino R. A. et al., 2019, ApJ, 880, 47
Martin C. L., 2005, ApJ, 621, 227
Martin D. C., Chang D., Matuszewski M., Morrissey P., Rahman S., Moore A., Steidel C. C., 2014, ApJ, 786, 106
Meyer R. A., Bosman S. E. I., Kakiichi K., Ellis R. S., 2019a, MNRAS, 483, 19
Meyer R. A., Bosman S. E. I., Ellis R. S., 2019b, MNRAS, 487, 3305
Møller P., Warren S. J., 1993, A&A, 270, 43
Møller P., Warren S. J., 1998, MNRAS, 299, 661
Nasir F., D'Aloisio A., 2019, preprint (arXiv:1910.03570)
Neeleman M., Kanekar N., Prochaska J. X., Rafelski M., Carilli C. L., Wolfe A. M., 2017, Science, 355, 1285
Neeleman M., Kanekar N., Prochaska J. X., Christensen L., Dessauges-Zavadsky M., Fynbo J. P. U., Møller P., Zwaan M. A., 2018, ApJ, 856, L12
Neeleman M., Kanekar N., Prochaska J. X., Rafelski M. A., Carilli C. L., 2019, ApJ, 870, L19
Noeske K. G. et al., 2007, ApJ, 660, L43
Oppenheimer B. D. et al., 2016, MNRAS, 460, 2157
Ouchi M. et al., 2010, ApJ, 723, 869
Peeples M. S. et al., 2019, ApJ, 873, 129
Péroux C., Bouché N., Kulkarni V. P., York D. G., Vladilo G., 2011, MNRAS, 410, 2251
Péroux C., Bouché N., Kulkarni V. P., York D. G., Vladilo G., 2012, MNRAS, 419, 3060
Péroux C. et al., 2016, MNRAS, 457, 903
Péroux C. et al., 2017, MNRAS, 464, 2053
Péroux C. et al., 2019, MNRAS, 485, 1595
Petitjean P., Ledoux C., Srianand R., 2008, A&A, 480, 349
Pettini M., Shapley A. E., Steidel C. C., Cuby J.-G., Dickinson M., Moorwood A. F. M., Adelberger K. L., Giavalisco M., 2001, ApJ, 554, 981
Planck Collaboration et al., 2016, A&A, 594, A13
Prochaska J. X., Wolfe A. M., 2009, ApJ, 696, 1543
Prochaska J. X., O'Meara J. M., Worseck G., 2010, ApJ, 718, 392
Rafelski M., Wolfe A. M., Prochaska J. X., Neeleman M., Mendez A. J., 2012, ApJ, 755, 89
Rafelski M., Neeleman M., Fumagalli M., Wolfe A. M., Prochaska J. X., 2014, ApJ, 782, L29
Rahmati A., Schaye J., 2014, MNRAS, 438, 529
Rahmati A., Pawlik A. H., Raičević M., Schaye J., 2013a, MNRAS, 430, 2427
Rahmati A., Schaye J., Pawlik A. H., Raičević M., 2013b, MNRAS, 431, 2261
Rahmati A., Schaye J., Bower R. G., Crain R. A., Furlong M., Schaller M., Theuns T., 2015, MNRAS, 452, 2034
Rhodin N. H. P., Agertz O., Christensen L., Renaud F., Fynbo J. P. U., 2019, MNRAS, 488, 3634
Ross N. P. et al., 2009, ApJ, 697, 1634
Salim S. et al., 2007, ApJS, 173, 267
Schaye J. et al., 2015, MNRAS, 446, 521
Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, ApJ, 588, 65
Shen S., Madau P., Guedes J., Mayer L., Prochaska J. X., Wadsley J., 2013, ApJ, 765, 89
Simcoe R. A. et al., 2011, ApJ, 743, 21
Springel V. et al., 2005, Nature, 435, 629
Steidel C. C., Pettini M., Dickinson M., Persson S. E., 1994, AJ, 108, 2046
Ströbele S. et al., 2012, in Ellerbroek B. L., Marchetti E., Véran J. P., eds, Proc. SPIE Conf. Ser. Vol. 8447, Adaptive Optics Systems III. SPIE, Bellingham, p. 844737
Tacchella S., Dekel A., Carollo C. M., Ceverino D., DeGraf C., Lapiner S., Mand elker N., Primack Joel R., 2016, MNRAS, 457, 2790
Tacconi L. J. et al., 2013, ApJ, 768, 74
Timlin J. D. et al., 2018, ApJ, 859, 20
Turner M. L., Schaye J., Steidel C. C., Rudie G. C., Strom A. L., 2014, MNRAS, 445, 794
van de Voort F., Springel V., Mandelker N., van den Bosch F. C., Pakmor R., 2019, MNRAS, 482, L85
Vanden Berk D. E. et al., 2001, AJ, 122, 549
Vernet J. et al., 2011, A&A, 536, A105
Wang F. et al., 2016, ApJ, 819, 24
Weilbacher P. M., Streicher O., Urrutia T., Pécontal-Rousset A., Jarno A., Bacon R., 2014, in Manset N., Forshay P., eds, ASP Conf. Ser. Vol. 485, Astronomical Data Analysis Software and Systems XXIII. Astron. Soc. Pac., San Francisco, p. 451
Yang J. et al., 2016, ApJ, 829, 33
Yang J. et al., 2017, AJ, 153, 184

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Table 1. Spectroscopic redshifts (including quality flags) and coordinates for continuum-detected sources.

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

APPENDIX A: ABSORPTION LINE PROFILES

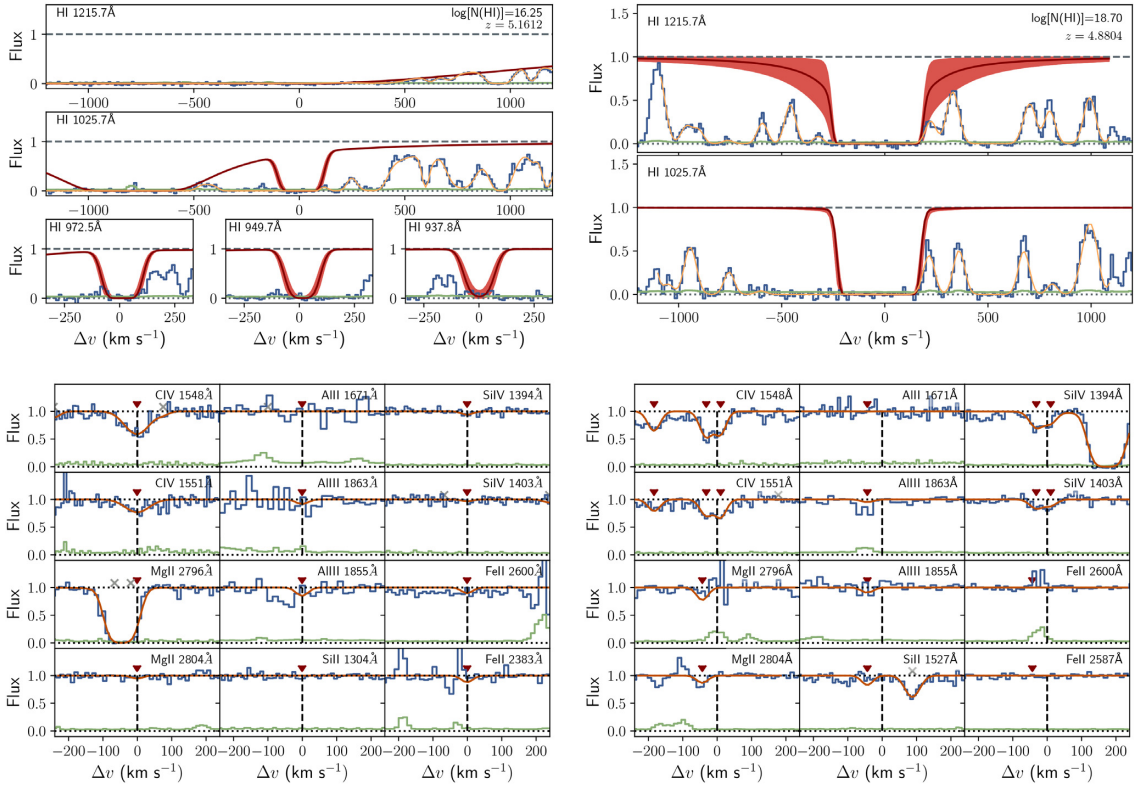

Figure A1. As in Fig. 8, but for systems at $z = 5.16$ (left-hand panel) and $z = 4.88$ (right-hand panel).

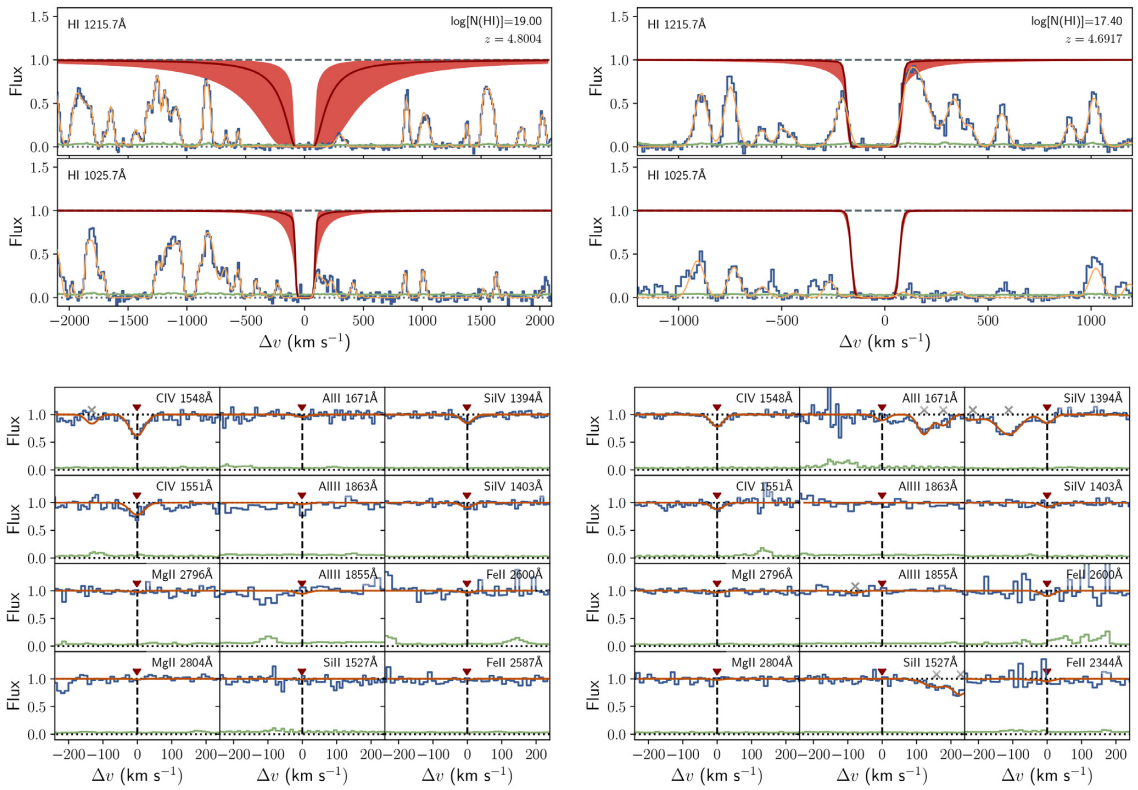

Figure A2. As in Fig. 8, but for systems at $z = 4.80$ (left-hand panel) and $z = 4.69$ (right-hand panel).

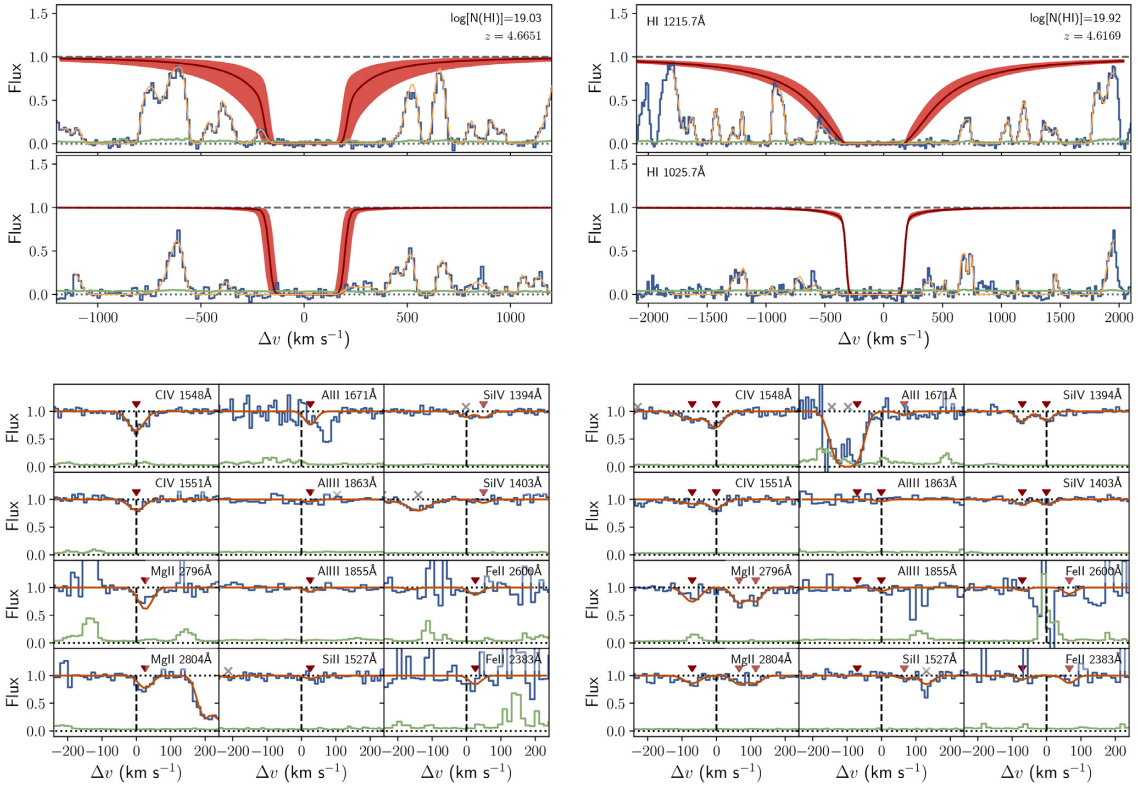

Figure A3. As in Fig. 8, but for systems at $z = 4.67$ and $z = 4.62$.

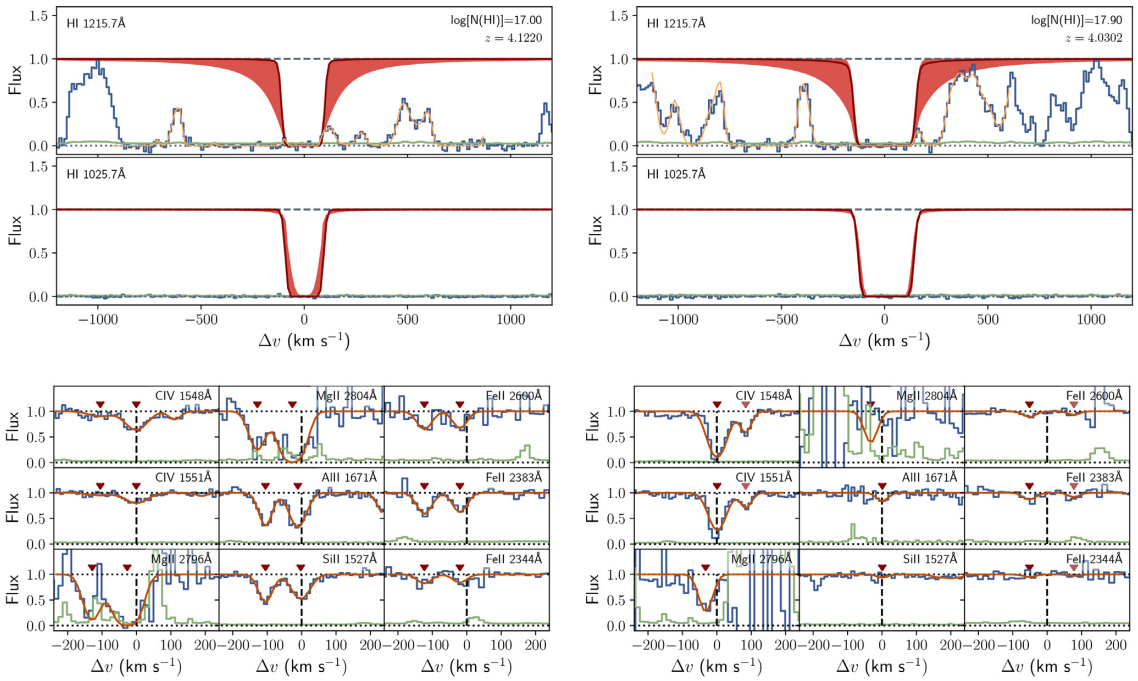

Figure A4. As in Fig. 8, but for systems at $z = 4.12$ and $z = 4.03$.

[1]*Centre for Extragalactic Astronomy, Physics Deptartment, Durham University, South Road, Durham DH1 3LE, UK*
[2]*Institute for Data Science, Durham University, South Road, Durham DH1 3LE, UK*
[3]*Institute for Computational Cosmology, Physics Deptartment, Durham University, South Road, Durham DH1 3LE, UK*
[4]*Dipartimento di Fisica G. Occhialini, Università degli Studi di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy*
[5]*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA*
[6]*Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA*

[7]*CASA, Department of Astrophysical and Planetary Sciences, University of Colorado, 389 UCB, Boulder, CO 80309, USA*
[8]*Department of Physics, ETH Zurich, Wolgang-Pauli-Strasse 27, CH-8093 Zurich, Switzerland*
[9]*Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark*
[10]*The Cosmic Dawn Center (DAWN), Vibenshuset, Lyngbyvej 2, DK-2100 Copenhagen, Denmark*
[11]*Niels Bohr Institute, Lyngbyvej 2, DK-2100 Copenhagen, Denmark*
[12]*Departamento de Astronomía, Universidad de Chile, Casilla 36-D Santiago, Chile*
[13]*INAF, Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, I-34143 Trieste, Italy*
[14]*Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy*
[15]*European Southern Observatory (ESO), Karl-Schwarzschild-Str.2, D-85748 Garching b. München, Germany*
[16]*Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, F-13388 Marseille, France*

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