MUSE Analysis of Gas around Galaxies (MAGG) - IV: The gaseous environment of $z \sim 3 - 4$ Lyα emitting galaxies

Emma K. Lofthouse[1,2]⋆, Michele Fumagalli[1,3]†, Matteo Fossati[1,2], Rajeshwari Dutta[1,2], Marta Galbiati[1],

Fabrizio Arrigoni Battaia[4], Sebastiano Cantalupo[1], Lise Christensen[5], Ryan J. Cooke[6], Alessia Longobardi[1,2],

Michael T. Murphy[7], J. Xavier. Prochaska[8,9]

[1]*Dipartimento di Fisica "G. Occhialini", Università degli Studi di Milano Bicocca, Piazza della Scienza 3, 20126 Milano, Italy*
[2]*INAF - Osservatorio Astronomico di Brera, via Brera 28, 20121 Milano, Italy*
[3]*INAF - Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, 34143 Trieste, Italy*
[4]*Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str 1, D-85748 Garching bei München, Germany*
[5]*Cosmic Dawn Center (DAWN), Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2200 N Copenhagen, Denmark*
[6]*Centre for Extragalactic Astronomy, Durham University, South Road, Durham, DH1 3LE, UK*
[7]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia*
[8]*Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064, USA*
[9]*Kavli Institute for the Physics and Mathematics of the Universe, 5-1-5 Kashiwanoha, Kashiwa 277-8583, Japan*

3 October 2022

ABSTRACT

We study the link between galaxies and H I-selected absorption systems at $z \sim 3 - 4$ in the MUSE Analysis of Gas around Galaxies (MAGG) survey, an ESO large programme consisting of integral field spectroscopic observations of 28 quasar fields hosting 61 strong absorbers with $N_{\rm HI} \gtrsim 10^{16.5}$ cm^{-2}. We identify 127 Lyα emitting galaxies (LAEs) around the absorbers, corresponding to a detection rate of 82±16 per cent. The luminosity function of these LAEs is ≈ 5 times higher in normalization than the field population and we detect a significant clustering of galaxies with respect to the gas, confirming that high column density absorbers and LAEs trace each other. Between 30 and 40 per cent of the absorbers are associated with multiple LAEs, which lie preferentially along filaments. Galaxies in groups also exhibit a three times higher covering factor of optically-thick gas compared to isolated systems. No significant correlations are identified between the emission properties of LAEs and the absorption properties of optically-thick gas clouds, except for a weak preference of brighter and multiple galaxies to reside near broad absorbers. Based on the measured impact parameters and the covering factor, we conclude that the near totality of optically-thick gas in the Universe can be found in the outer circumgalactic medium (CGM) of LAEs or in the intergalactic medium (IGM) in proximity to these galaxies. Thus, LAEs act as tracers of larger scale structures within which both galaxies and optically-thick clouds are embedded. The patchy and inhomogeneous nature of the CGM and IGM explains the lack of correlations between absorption and emission properties. This implies that very large samples are needed to unveil the trends that encode the properties of the baryon cycle.

Key words: galaxies: evolution – galaxies: formation – galaxies: high-redshift – galaxies: halos – quasars: absorption lines

1 INTRODUCTION

Gas is a fundamental constituent in the life of galaxies. The very formation of galaxies hinges on the ability of gas to cool from the intergalactic medium (IGM) onto collapsing dark matter halos. Gas is the raw fuel that feeds and sustains the formation of stars, and it is a critical agent in the self-regulation mechanisms that operate inside and around galaxies, within the circumgalactic medium (CGM). Therefore, a complete mapping of the different gas phases inside and around galaxies becomes essential to fully describe the processes that shape their evolution.

A complete census of how gas is distributed and exchanged between

galaxies across cosmic time is, however, far from trivial to obtain. The multiphase nature of gas requires concerted observational efforts across the entire electromagnetic spectrum, but some phases remain particularly hard to detect especially at large cosmological distances due to the intrinsic diffuse nature of the gas and the weak electromagnetic signal associated.

Recent advancements in instrumentation have however enabled a breakthrough in our ability to map gas as a function of cosmic time. The Atacama Large Millimeter/submillimeter Array (ALMA; Wootten et al. 2009) has expanded our view of the cold atomic and molecular phase inside the interstellar medium (ISM) and the dense parts of the CGM, up to very high redshifts (e.g. Franco et al. 2019; Péroux et al. 2019; Neeleman et al. 2017, 2020). High-sensitivity integral field (IFU) spectrographs at 8 m class telescopes such as the Multi Unit Spectroscopic Explorer (MUSE, Bacon et al. 2010) and

---

⋆ E-mail: emmakatherine.lofthouse@unimib.it
† E-mail: michele.fumagalli@unimib.it

the Keck Cosmic Web Imager (KCWI, Morrissey et al. 2018) are starting to yield a glimpse of the low-density and ionized hydrogen inside and around galaxies up to $z \approx 6$, including gas in filaments connecting multiple $z \approx 3$ star-forming galaxies (e.g. Leclercq et al. 2017; Wisotzki et al. 2018; Kikuta et al. 2019; Umehata 2019; Bielby et al. 2020; Bacon et al. 2021; Muzahid et al. 2021). Very recently, pathfinders of the Square Kilometre Array (SKA) such as MeerKAT (Jonas 2009) are pushing the detectability of the more diffuse and widespread cold/warm neutral medium traced by atomic hydrogen, reaching for the first time cosmological distances. Stacking of 21cm signals have allowed detections out to $z \approx 1$ (Chowdhury et al. 2020). However, until the advent of the full SKA, these efforts remain limited to the current modest redshifts ($z \lesssim 1$), and the $z \gtrsim 2$ Universe must be explored by other means, such as absorption spectroscopy.

Indeed, thanks to large quasar surveys, and particularly the Sloan Digital Sky Survey (York et al. 2000), it has been possible to map in detail the distribution, kinematics, and chemical composition of fully neutral and partially ionised H I along the line of sight to high-redshift quasars. The detailed study of these intervening absorption line systems (ALSs), and in particular the damped Lyα absorbers (DLAs; $N_{HI} \geq 10^{20.3}$) and the Lyman limit systems (LLSs; $10^{17.2}\mathrm{cm}^{-2} \leq N_{HI} < 10^{20.3}\mathrm{cm}^{-2}$) has been for several years a fruitful tool for the study of the high-redshift IGM, CGM, and ISM from $z \gtrsim 2$ all the way up to the end of the reionization epoch (e.g. Péroux et al. 2003; Wolfe et al. 2005; Prochaska & Wolfe 2009; Rafelski et al. 2012; Fumagalli et al. 2013, 2016).

The most difficult aspect of this approach, however, has been to establish a direct connection between gas probed in absorption along the quasar pencil beam and the associated galaxies. Attempted searches in emission of galaxies near ALSs have been accumulating in the literature, but have traditionally resulted in a large number of non-detections. In the last decade, however, there has been an acceleration in successfully identifying associations between gas probed in absorption and galaxies seen in emission. At $z \approx 2-3$, surveys using multi-object spectroscopy, e.g. the Keck Baryonic Structure Survey (KBSS; Rudie et al. 2012; Trainor & Steidel 2012; Rakic et al. 2012) and the VLT Lyman break galaxy (LBG) Redshift Survey (VLRS; Bielby et al. 2011), have shown that star-forming galaxies reside inside gas overdensities, and have revealed that metal-enriched gas exhibits complex kinematics modulated by a combination of outflow and inflows (e.g. Steidel et al. 2010; Rubin et al. 2010; Crighton et al. 2011; Rudie et al. 2012; Turner et al. 2014; Tummuangpak et al. 2014; Bielby et al. 2017). These surveys, albeit powerful, require a pre-selection of the targeted galaxy population and are thus likely to miss a large population of continuum-faint galaxies that emit strong emission lines (e.g. Crighton et al. 2015).

More targeted approaches based on long-slit spectroscopy have indeed revealed several emission line galaxies at close separation from ALSs, and in particular DLAs (e.g. Møller et al. 2002, 2004; Fynbo et al. 2010; Srianand et al. 2016). Based on empirical scaling relations and model expectations (Møller et al. 2004; Ledoux et al. 2006; Krogager et al. 2020), most of these efforts have focused on metal-rich ALSs, leaving therefore mostly unexplored the population of galaxies in proximity of the more metal poor systems, which happen to be the more common ones (Fumagalli et al. 2016). The advent of integral field spectrographs, starting with the pioneering efforts in the near infrared using e.g. SINFONI (Eisenhauer et al. 2003; Bonnet et al. 2004) or OSIRIS (Larkin et al. 2006), has transformed our ability to search for galaxies in proximity to ALSs without the need for pre-selecting either the galaxies or the absorbers. Indeed, several successful examples are accumulating in the literature showing how sensitive searches can be conducted blindly around ALSs across a large range of column densities and metallicities, resulting in a high number of new detections (e.g. Peroux et al. 2011, 2012; Fumagalli et al. 2016b, 2017; Mackenzie et al. 2019; Bielby et al. 2020). In parallel, searches for galaxies near ALSs with sub-millimeter interferometers have been equally successful (e.g. Møller et al. 2018; Klitsch et al. 2018; Neeleman et al. 2017; Kanekar et al. 2018; Neeleman et al. 2020), underscoring the need of a multiwavelength approach to finally paint a complete view of the link between gas probed in absorption and galaxies detected in emission.

In this paper, we exploit the power of MUSE and present a comprehensive study of the link between gas and galaxies at $z \gtrsim 3$ using the MUSE Analysis of Gas Around Galaxies (MAGG), an ESO Large Programme (PID 197.A-0384, PI Fumagalli) designed to blindly map the galaxy population in 28 quasar fields with 61 high H I column density ALSs along the line of sight. With at least 5 hours of MUSE observations in each field complemented by high resolution and high signal-to-noise ratio (S/N) quasar spectroscopy, the MAGG survey is designed to yield the discovery of orders of magnitude more galaxies within a projected radius of ≈ 250 kpc from strong ALSs compared to previous surveys. Moreover, differently from most of previous studies, MAGG targets high H I column density ALSs ($\gtrsim 10^{16.5}$ cm^{-2}) which are most likely to trace the denser parts of the CGM and IGM at $z \gtrsim 3$ (Faucher-Giguère et al. 2011; Fumagalli et al. 2011; Van de Voort et al. 2012) without selecting on absorption properties other than the hydrogen column density. This survey is therefore constructed to provide the most complete and unbiased snapshot of the link between gas and galaxies when the Universe was ≈ 4 billion years old, just before the peak of the cosmic star formation history.

The design of the MUSE survey, together with a detailed description of the observations and data reduction of both MUSE and ancillary high-resolution quasar spectroscopy, has already been presented in the first paper of this series (Lofthouse et al. 2020). The MAGG data have also been used to study the link between lower redshift ($z \lesssim 1.8$) galaxies and metal absorption lines, and in particular Mg II (Dutta et al. 2020) also as a function of galaxy overdensity (Dutta et al. 2021), and to map the gas and galaxy environment of high-redshift quasars (Fossati et al. 2021). We now turn to the core objective of the MAGG survey, i.e. the study of the gas-galaxy link at $z \gtrsim 3$. Throughout this work, we use the Planck Collaboration et al. (2016) cosmology, with $\Omega_m = 0.307$ and $H_0 = 67.7$ km s^{-1} Mpc^{-1}. Unless otherwise stated, all distances are in proper units and we use the AB magnitude system.

2 DATA ACQUISITION AND REDUCTION

In this section we provide a brief overview of the acquisition and data reduction for the MUSE observations and for the high-resolution quasar spectroscopy. We also briefly summarise the steps taken to generate the galaxy catalogues. Additional details can be found in previous papers of this series, particularly Lofthouse et al. (2020) and Fossati et al. (2021). In this work, we focus in more detail on the analysis of absorption line systems, including the measurements of the H I and metal column densities and the procedures adopted to constrain the gas-phase metallicity of the ALSs.

2.1 MUSE observations and data reduction

The MAGG sample consists of 28 quasars, including five fields from the ESO archive such as data from Borisova et al. (2016) as part of GTO observations (see Lofthouse et al. 2020, their table 1 for full details including individual exposure times for each sightline and references to the ESO programmes). MUSE observations of the original MAGG fields have been carried out during ESO periods 97–103, integrating on-source for ≈ 4 h per field. Longer exposures are achieved for fields with available exposures in the ESO archives such as the GTO fields which have up to 10 hours observations. To improve the homogeneity of the sensitivity across the different MUSE spectrographs, each observing block included dithers and instrument rotations of 90 deg. Data have been acquired during clear nights at airmass < 1.6, and the resulting image quality is on average 0.6-0.7 arcsec full width at half maximum (FWHM). For the data reduction we follow consolidated techniques, see e.g. Lofthouse et al. (2020) and Fossati et al. (2021), that combine the ESO MUSE pipeline (Weilbacher et al. 2014) with post-processing tools, and in particular the CUBEXTRACTOR package, hereafter CUBEX (Cantalupo et al. 2019, and Cantalupo in prep.). Bias subtraction, flat-field correction and the wavelength and flux calibrations are first completed using the ESO pipeline. Cubes are sky-subtracted and reconstructed on a regular 3D grid, with the astrometry registered on the Gaia DR2 reference frame (Gaia Collaboration et al. 2018).

Using CUBEX, we then correct residual imperfections in the sky subtraction and illumination corrections. In particular, we use the CUBEFIX tool to balance the different response of spectrographs, stacks and slices in MUSE. The CUBESHARP tool is then used to suppress sky residuals arising from variations in the instrument line spread function. These tools are applied iteratively, to achieve optimal masking of the astrophysical sources in the field. In the end, we combine all the reduced exposures in final cubes using both mean and median statistics, and we further produce two coadds containing only half of all the exposures, which we use to validate the detection of sources. The wavelength solutions of the final cubes are all in air; however, when extracting spectra of continuum sources or of LAEs, we convert to vacuum wavelengths to be consistent with the quasar spectroscopy and hence the absorber redshifts (see next section).

While the pixel uncertainty is propagated during the various reduction steps, the uncertainty of the final cubes is found to not reproduce with sufficient accuracy the effective noise as measured on the data. We therefore bootstrap pixels in individual exposures to reconstruct an empirical noise estimate of each cube, as described in Fossati et al. (2019). To minimise small fluctuations in the bootstrap procedure arising from the small number of individual exposures, we then correct the pipeline variance cube with a wavelength-dependent function computed from the bootstraps.

2.2 Archival quasar spectroscopy

For the search and analysis of ALSs along the quasar sightlines, we collect archival high resolution and high S/N spectroscopy from the Ultraviolet and Visual Echelle Spectrograph (UVES; Dekker et al. 2000), High-Resolution Echelle Spectrometer (HIRES; Vogt et al. 1994) and the Magellan Inamori Kyocera Echelle instruments (MIKE; Bernstein et al. 2003). Where available, we also collect moderate resolution spectroscopy from ESI and X-Shooter, which is useful to extend our analysis in the near infrared (NIR) and to fill the gaps in wavelengths that are occasionally present in the high-
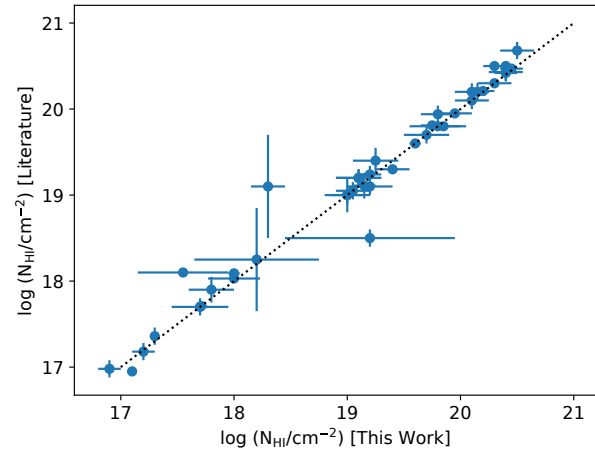
Figure 1. Comparison of the H I column densities measured in this work with the column densities reported in the literature (see text for details). A tight correlation is found over a wide range of column density.

resolution data. Details of the 62 individual spectra assembled for the 28 quasars are provided in Lofthouse et al. (2020).

All data have been reduced with optimised and instrument-specific pipelines, as detailed in Lofthouse et al. (2020). Although differences in the adopted algorithms and procedures are present, the reduction for all the instruments follow a similar set of procedures. First, basic calibrations (bias subtraction, flat-fielding, and where applicable dark subtraction) are applied to the 2D frames, and a vacuum wavelength solution is computed. After sky subtraction, the objects are extracted on an order-by-order basis and data are combined, also co-adding multiple exposures if present. If applicable, spectra are then flux-calibrated and continuum normalized. Observations from the same instruments but using different setups are further combined into a single spectrum by resampling the data onto a common wavelength solution. We do not attempt to combine data across different instruments, which we analyse separately as discussed in the following sections.

3 DATA ANALYSIS AND CATALOGUE PREPARATION

3.1 System selection and HI determination

The core objective of our analysis is to investigate the connection between gas and galaxies at $z \approx 3 - 4$, where MUSE is most effective at uncovering Lyα emission lines (due to wavelength coverage and the absence of sky lines) and where we can search for high-column density absorption line systems ($N_{HI} \gtrsim 10^{16.5}$ cm^{-2}) which are tracers of the denser IGM and of the CGM (Faucher-Giguère et al. 2011; Fumagalli et al. 2011; Van de Voort et al. 2012).

These absorption line systems are selected purely based on their H I column density, to ensure we assemble a representative sample that is not biased e.g., towards metal rich systems (see e.g., Berg et al. 2021). By scanning each spectrum, we identify the highest column density systems ($\log(N_{HI}/\text{cm}^{-2}) \gtrsim 19$) thanks to their distinctive Voigt profile in the Lyα and, when present, Lyβ transitions. Lower column density systems ($16.5 \lesssim \log(N_{HI}/\text{cm}^{-2}) \lesssim 19$) are instead recognised by their Lyman series and the flux decrement
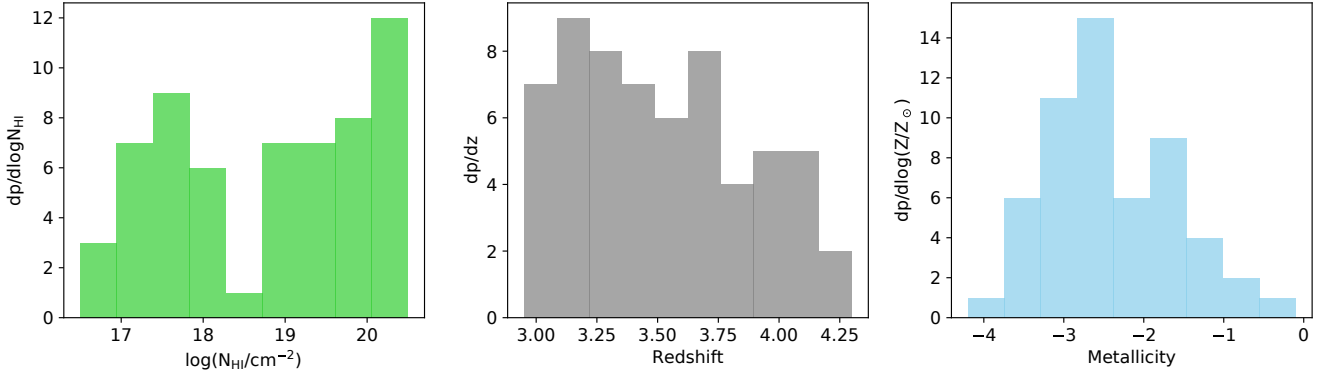
Figure 2. Summary of the statistical properties of the 61 ALSs identified along the 28 quasar sightlines. From left to right, we show the H I column density, the redshift and the metallicity distributions.

Table 1. Summary of the absorption line properties identified in the quasar spectra, including the N(H I)-weighted redshift, the total H I column density, the inferred metallicity, the interquartile range between 16% and 84% of the posterior metallicity distributions and the number of LAEs (see text for details). Note that the quoted error on the HI column density is the value input into our photoionisation model while the error on the metallicity is the interquartile range on the model output rather than the traditional 1σ error as these distributions are typically not Gaussian.

Quasar	z_{HI}	log N_{HI} (cm^{-2})	[M/H]	IQR	N_{LAE}
J010619+004823	4.23536	17.2 ± 0.1	-3.18	$^{+0.20}_{-0.18}$	1
	4.06780	17.1 ± 0.1	<-2.78		1
	3.72887	18.05 ± 0.60	>-0.70		1
	3.32100	19.2 ± 0.1	-2.61	$^{+0.47}_{-0.67}$	3
J012403+004432	3.76618	16.8 ± 0.1	-2.51	$^{+0.15}_{-0.16}$	2
	3.07777	20.20 ± 0.05	-2.09	$^{+0.05}_{-0.05}$	2
J013340+040059	4.11223	18.2 ± 0.3	-2.97	$^{+0.20}_{-0.16}$	5
	3.99526	19.8 ± 0.1	-2.56	$^{+0.04}_{-0.03}$	1
	3.77240	20.4 ± 0.1	-0.85	$^{+0.09}_{-0.07}$	1
	3.69164	20.5 ± 0.1	-2.56	$^{+0.03}_{-0.02}$	2
J013724−422417	3.66541	19.15 ± 0.15	-2.88	$^{+0.06}_{-0.06}$	2
	3.10030	19.75 ± 0.05	-1.69	$^{+0.04}_{-0.04}$	1
J015741−010629	3.38595	19.2 ± 0.1	-1.56	$^{+0.05}_{-0.05}$	8
J020944+051713	3.98786	17.6 ± 0.2	-2.00	$^{+0.15}_{-0.13}$	1
	3.86337	20.4 ± 0.1	-2.65	$^{+0.06}_{-0.05}$	0
	3.70700	19.2 ± 0.1	-2.96	$^{+0.13}_{-0.19}$	1
	3.66592	20.45 ± 0.05	-2.35	$^{+0.03}_{-0.03}$	1
J024401−013403	4.03100	17.40 ± 0.05	-3.69	$^{+0.15}_{-0.14}$	4
	3.96620	18.3 ± 0.6	-2.70	$^{+0.02}_{-0.02}$	3
J033413−161205	4.23449	17.55 ± 0.10	-3.03	$^{+0.11}_{-0.10}$	0
	4.14700	17.2 ± 0.1	-2.71	$^{+0.19}_{-0.21}$	3
	3.55655	20.7 ± 0.1	-1.09	$^{+0.01}_{-0.01}$	1
J033900−013317	3.11540	19.60 ± 0.05	-1.57	$^{+0.03}_{-0.03}$	1
	3.06222	20.4 ± 0.15	-1.05	$^{+0.05}_{-0.04}$	1
J094932+033531	3.92880	17.0 ± 0.1	-0.77	$^{+0.11}_{-0.11}$	2
	3.31137	19.80 ± 0.05	>-0.55		1
J095852+120245	3.22320	17.3 ± 0.1	-2.97	$^{+0.14}_{-0.13}$	2
	3.09622	17.2 ± 0.1	<-2.75		1
J102009+104002	3.05410	17.10 ± 0.05	-2.94	$^{+0.37}_{-0.22}$	0
J111008+024458	3.47555	17.55 ± 0.05	-2.85	$^{+0.15}_{-0.15}$	2
	3.44172	17.4 ± 0.1	<-2.75		0
J111113−080402	3.81137	18.2 ± 0.3	-3.19	$^{+0.14}_{-0.12}$	5
	3.60766	20.30 ± 0.05	-2.11	$^{+0.04}_{-0.04}$	5
	3.48185	19.95 ± 0.05	-2.24	$^{+0.04}_{-0.04}$	1

Table 1 – *continued*

Quasar	z_{HI}	log N_{HI} (cm^{-2})	[M/H]	IQR	N_{LAE}
J120917+113830	3.02224	19.2 ± 0.1	-0.36	$^{+0.11}_{-0.07}$	0
J123055−113909	3.31990	16.5 ± 0.1	-2.69	$^{+0.59}_{-0.45}$	4
J124957−015928	3.52694	17.71 ± 0.05	-3.61	$^{+0.47}_{-0.54}$	5
J133254+005250	3.42111	19.1 ± 0.1	-2.39	$^{+0.02}_{-0.02}$	1
	3.08390	19.25 ± 0.15	-2.49	$^{+0.09}_{-0.06}$	1
J142438+225600	3.38275	16.9 ± 0.1	-2.86	$^{+0.64}_{-0.65}$	1
J162116−004250	3.10494	19.7 ± 0.1	-1.27	$^{+0.07}_{-0.06}$	6
J193957−100241	3.57226	18.00 ± 0.05	-2.59	$^{+0.04}_{-0.04}$	0
J200324−325144	3.55212	18.00 ± 0.05	-1.72	$^{+0.04}_{-0.04}$	4
	3.18780	19.85 ± 0.05	-1.59	$^{+0.05}_{-0.05}$	3
	3.17243	19.80 ± 0.05	-3.33	$^{+0.03}_{-0.02}$	1
J205344−354655	3.17249	18.2 ± 0.6	-1.50	$^{+0.14}_{-0.20}$	4
	3.09420	19.05 ± 0.10	-1.07	$^{+0.07}_{-0.06}$	4
	2.98882	20.1 ± 0.1	-1.86	$^{+0.10}_{-0.10}$	0
J221527−161133	3.89620	17.40 ± 0.05	-3.73	$^{+0.19}_{-0.17}$	2
	3.70152	19.40 ± 0.05	-1.89	$^{+0.03}_{-0.03}$	0
	3.66185	20.1 ± 0.1	-2.18	$^{+0.07}_{-0.06}$	1
	3.46525	19.0 ± 0.2	-3.60	$^{+0.15}_{-0.15}$	3
	3.30557	19.0 ± 0.1	<-3.42		3
J230301−093930	3.31186	17.80 ± 0.15	-3.04	$^{+0.10}_{-0.09}$	1
J231543+145606	3.27319	20.30 ± 0.05	-2.48	$^{+0.03}_{-0.03}$	5
	3.13528	19.95 ± 0.05	-2.61	$^{+0.05}_{-0.05}$	3
J233446−090812	3.22578	17.7 ± 0.1	-2.81	$^{+0.10}_{-0.08}$	1
	3.05723	20.40 ± 0.05	-1.66	$^{+0.04}_{-0.03}$	7
J234913−371259	4.08985	18.80 ± 0.05	-3.48	$^{+0.11}_{-0.16}$	1
	3.69231	20.15 ± 0.10	-2.55	$^{+0.05}_{-0.05}$	2
	3.58203	19.1 ± 0.1	-4.17	$^{+0.08}_{-0.07}$	0

at the Lyman limit. Because of this selection, our sample of absorbers does not follow a statistical distribution of column densities (i.e., the $f(N_{HI}, z)$), as we can more easily identify systems with $\log(N_{HI}/\text{cm}^{-2}) \gtrsim 19$ at all wavelengths, while we are limited to the identification of $16.5 \lesssim \log(N_{HI}/\text{cm}^{-2}) \lesssim 19$ systems for which we have complete coverage of the Lyman limit (in other words, we have a longer search path for systems with damping wings). Moreover, MAGG sightlines are selected to preferentially contain LLSs at $z \gtrsim 3$, which naturally bias the incidence of LLSs above what expected in random sightlines.

In this work, we expand the initial analysis with the PYIGM pack-

age[1] presented in Lofthouse et al. (2020), moving to a custom-built tool that allows the simultaneous analysis of spectra from different instruments. Similarly to what is possible within the PYIGM package, this tool allows us to construct full models of the Lyman series plus Lyman limit of every absorber, by varying their column density, Doppler parameter and redshift. Each model is then convolved with the appropriate kernel to account for the resolution of each instrument. Within this tool, each model can then be compared to both the high-resolution and the moderate-resolution spectra that are available for each quasar (see table 2 in Lofthouse et al. 2020). By using the full set of data available, we revisit the search for high-column density absorbers of Lofthouse et al. (2020), and expand the catalogue originally presented to include 9 new absorbers (generally systems for which a refined fit of the Lyα line revealed a column density above the selection threshold), for a total of 61 systems in 28 sightlines.

For all of the systems, in accordance to well-consolidated practice in the literature (e.g. Fumagalli et al. 2013, 2014; Prochaska et al. 2015; Berg et al. 2021), we visually fit the continuum level and then systematically vary the redshift and column density to obtain the best model that describes the full Lyman series and the transmission at the Lyman limit. Within this model, we consider as the main component the one that dominates the optical depth at the Lyman limit and within the Lyman series. In most cases, a single component is sufficient to account for the absorption in high-order Lyman series lines. We then add a second component when clearly required by the data (that is, when high order lines separate in distinct components with column density of $\gtrsim 1/10$ that of the main component), following the same methodology. For these systems, we quote the total column density over all components and the resulting column-density weighted redshift. For the high HI column density systems, we rely primarily on the presence of damping wings in the Lyα absorption profile. After fitting the continuum level around the absorption profile, we adjust the model column density and Doppler parameter to find the best fit to the observed profile. The redshift and column density are, in most cases, well constrained by the unsaturated high-order lines in the Lyman series, the presence of damped wings in the Lyα lines, and/or the residual transmission at the Lyman limit. For the cases where high-order lines are saturated and there is no evidence of damped wings in the Lyα profile, we quote an $N_{\rm H\,I}$ range where the lower bound is set by the lack of transmission at the Lyman limit and the upper bound is set by the onset of the damped wings. The value we adopt is the mid-point of this interval and the error expresses the half-width of the interval (see also Prochaska et al. 2015).

A summary of all the systems identified with their column density is presented in Table 1. Fig. 1 shows a direct comparison of our H I determination with the one reported in the literature for the same systems (Zafar et al. 2013; Neeleman et al. 2013; Riemer-Sørensen et al. 2015; Prochaska et al. 2015; Berg et al. 2021; Lehner et al. 2021). Despite some scatter and a couple of outliers (which arise from the fact that some literature work relied on data at lower-resolution, or in a narrower wavelength range), we find excellent agreement. We note that some of the systems with $\log(N_{\rm HI}/{\rm cm}^{-2}) \lesssim 19$ have large errors. This is a result of the Lyman series lines being saturated limiting the ability to determine an accurate column density. In this case, the error bars represent the range of values allowed by the observations. The left and middle panels of Fig. 2 summarise the H I column density and the redshift distributions of our ALS sample respectively.

3.2 Models of individual ions and metallicity estimates

To identify the presence of metal absorption lines associated with each ALS we use the LINETOOLS package[2]. Given a quasar spectrum and the redshift of the absorber, the routine LT_XABSSYS simultaneously displays ± 1000 km s^{-1} windows around the expected position of potential metal lines (e.g., Si II, Mg II, Fe II, Al II, Al III, C IV, Si IV). From this, we select those which show detected absorption and determine the wavelength range that needs to be fit in the following analysis. We exclude lines which fall within the Lyα forest as the heavy blending with the hydrogen lines makes it difficult to identify the underlying metal components.

To fit the identified metal lines we utilise the Monte Carlo Absorption Line fitter code (MC-ALF, Longobardi et al., in prep; see also Fossati et al. 2019). For all the transitions in common for a given ion, in each absorber, MC-ALF uses a Bayesian statistical method to determine the posterior probability distribution of the models that describe the data. Specifically, the code uses the wavelength range specified using LT_XABSSYS as a prior on the redshift interval to consider. Within a given range for the number of components (based on visual inspection) that can be used to describe the absorption profile, MC-ALF returns the posterior of each model computed for a progressively increasing number of components together with the corresponding Akaike Information Criterion (AIC). We then select fits with an AIC within 5 of the lowest value and adopt the best model from this pool following visual inspection. Where the models cannot provide a good description of the data due to blending with lines not associated with the absorber under consideration, MC-ALF allows us to add in filler Voigt components to account for these additional sources of optical depth. These lines (shown in pink in Fig. 3) contribute to the overall fit (shown in red in Fig. 3) to the observed spectrum but are not included in the calculation of the metal column density as they arise from unassociated systems at different redshifts to the absorption system we are studying. The code further allows the continuum to vary by ± 2 per cent, to account for small offsets in the applied normalization. We also use a flat prior on the Doppler parameter between 1 and 35 km s^{-1} and a flat prior on the column density in the range $12 \leq \log(N/{\rm cm}^{-2}) \leq 16$. While the redshift and Doppler parameter of components are not constrained to be the same for different low/high ions, we tend to find that the different ions in the same system but with similar ionisation states have similar profiles with comparable number of components.

For each component of the adopted model, MC-ALF returns a column density. In cases where multiple components are required, the total column density is calculated as the sum of all these components. For ions where all of the available transitions are saturated as determined from visual inspection of the line profiles, it is not possible to determine an accurate measurement of the column density with this approach and we instead use the apparent optical depth method (AODM; Savage et al. 1991) to measure lower limits. Specifically, we adopt the most stringent lower limit on the column density by performing the AODM analysis of the weakest transition that is not blended. In cases where there is no detectable line within ± 1000 km s^{-1} of the absorber redshift, we instead determine a 2σ upper limit on the column density using the AODM. For this calculation we use a velocity window of ± 100 km s^{-1} centred on the redshift of the absorber as this is typical of the Δv_{90} (i.e. the velocity range which encompasses 90 per cent of the optical depth of the line) for our systems. As an example, the column densities for

---

Table 2. Column density of metal lines associated with the absorption system at $z = 3.69164$ along the J013340+040059 sightline. Tables showing the column densities of metals for all the other systems are included in the online material. The columns are as follows: (1) name of the ion; (2) rest frame wavelength of the line in Å; (3) observed equivalent width of line in Å; (4) instrument used to obtain the spectrum from which we calculate the values shown in this table; and (4) the adopted column density of the ion in $\log(N/\mathrm{cm}^{-2})$ from the analysis of all transitions as described in the text.

Ion	Wave (Å)	EW (Å)	Instrument	$\log(N / \mathrm{cm}^{-2})$
CIV	1548	0.413±0.008	UVES	13.56±0.07
	1550	0.255±0.005	UVES	
AlII	1670	0.437±0.007	UVES	12.47±0.02
AlIII	1854	-	X-shooter	<12.1
MgII	2796	2.2±0.02	X-shooter	>13.26
	2803	1.83±0.02	X-shooter	>13.26
FeII	1608	0.171±0.003	UVES	13.3±0.1
	2344	0.47±0.02	X-shooter	
	2374	0.19±0.01	X-shooter	
	2382	1.06±0.02	X-shooter	
	2600	1.0±0.02	X-shooter	
SiII	1526	1.62±0.01	UVES	14.3±0.3
	1808	0.0541±0.0007	UVES	
SiIV	1393	0.289±0.007	X-shooter	12.88±0.05
	1402	0.153±0.003	X-shooter	

the different metals along with our fits to the spectra are shown in Table 2 and Fig. 3, respectively, for the system at $z = 3.69164$ in the J013340+040059 sightline. Tables and profile models for all the remaining systems are provided in the online material.

3.2.1 Photoionisation modelling

The gas in the lower column density absorbers in our sample (LLSs and sub-DLAs) is not fully neutral. As a result, the measurements acquired in the previous section cannot be converted trivially into a value for metallicity as in DLAs. In these cases, we must apply ionization corrections when converting the observed column densities to metallicities. To do this, we follow the technique described in Fumagalli et al. (2016) (see also Crighton et al. 2015). We include a brief overview here and refer the reader to Fumagalli et al. (2016) for a more detailed description.

Using CLOUDY (Ferland et al. 2013, c17.01), we construct a grid of models using the assumptions of the minimal model in Fumagalli et al. (2016). Briefly, this consists of a 1D single-phase gas slab with a constant density (n_H), which is illuminated on one side by a redshift-dependent UV background radiation (Haardt & Madau 2012). The metallicity and density are given flat priors and allowed to vary within the ranges $-4.0 < \log Z/Z_\odot < -1.0$ and $-4.5 < \log (n_H/\mathrm{cm}^{-3}) < 0.0$, respectively. The redshift and HI column density are defined by Gaussian priors based on the observed values and errors except for where the LLSs have saturated Lyman series lines and it is not possible to accurately determine a value of the column density (see the systems in Table 1 and Fig. 1 with larger errors). For these cases we use a top-hat function for the prior where the limits are defined by the minimum and maximum values permissible by the observations.

We compare the observed column densities and limits for each absorber to the grid of models using EMCEE (Foreman-Mackey et al. 2013). This provides posterior probability distribution functions (PDFs) for the metallicity and density of each of our 61 absorbers. An example for the system shown in the previous section (at $z = 3.69164$ along the J013340+040059 sightline) is provided in Fig. 4 with all

the figures for the remaining systems provided as online-only material. The distribution of metallicity for the entire sample is further shown in the right panel of Fig. 2 and listed in Table 1. We note that the errors shown in this table for the H I column density are the values that are input to our model. On the other hand, the uncertainty on the metallicity are an output of the model. For these we use the 16 per cent to 84 per cent interquartile range (IQR) as the posterior metallicity distributions are typically not Gaussian. Therefore, while in some cases it seems that the metallicity error is smaller than for the H I column density from which it is derived, we confirm that this IQR is larger than the IQR of the corresponding output H I column density.

As 16 of the quasars in the MAGG sample are also in the HD-LLS sample studied by Prochaska et al. (2015), we can test for consistency both in the determination of the ion column densities and in the inferred metallicity. Regarding the ion column densities, we find good agreement with the values reported in Prochaska et al. (2015). Of the 22 absorption systems that we are able to compare, the MAGG metallicities of 19 agree within $< 2\sigma$ with the published values. The three other systems have some of the smallest errors in our sample resulting from strong constraints on both high and low ionisation absorption lines. We note that both studies resort to the same technique for inferring the metallicity, but our work relies on a new set of CLOUDY models.

Before proceeding with our analysis, we re-iterate some caveats of this modelling which have been already discussed in the literature (Fumagalli et al. 2016b; Berg et al. 2021). First, and foremost, this modelling treats the gas slab as a single phase, which is clearly not an ideal approximation given the multiphase nature of the CGM. While more advanced modelling has been introduced in the literature (e.g. Buie et al. 2020; Haislmaier et al. 2021), this approach has yet to scale to the wide range of redshift, metallicity, and density of interest to this work. We therefore resort to this simpler approach, keeping in mind that we are modelling in most cases a cold/cool metal phase (traced by low ions), which we attribute to all the neutral hydrogen measured in absorption. The inclusion of Si IV and C IV deserves a special note, as these ions straddle the separation between a cool and
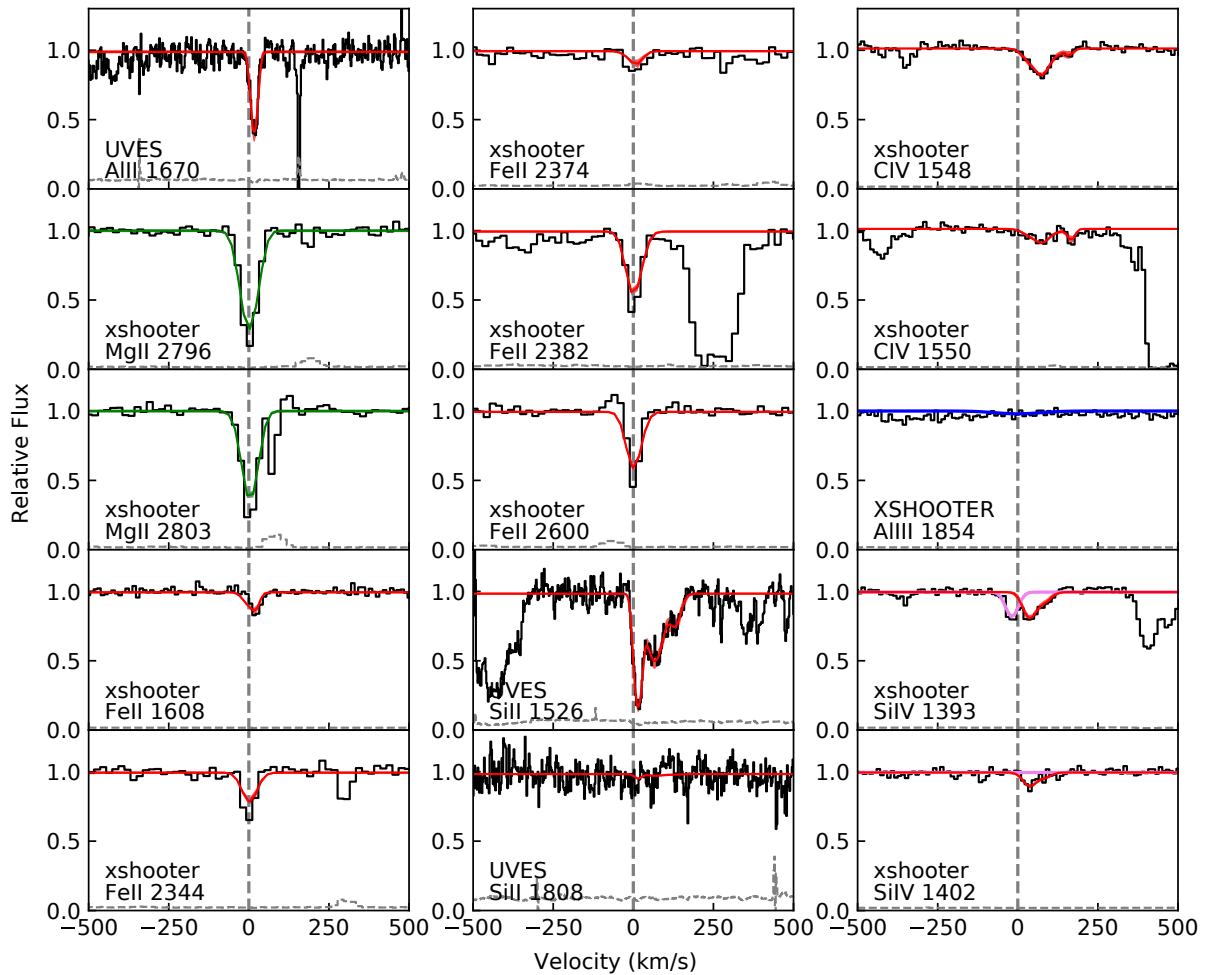
Figure 3. Example of fits to metal lines for a single system (J013340+040059, $z = 3.69164$). Models for all the remaining systems are provided in the online material. Fits from MC-ALF are shown in red for unsaturated line, in green for lines that are saturated and in blue for upper limits. Filler components, which arise from absorption due to unrelated systems that overlap with the metal associated with the absorption system, are shown in pink. The error spectra of the data are shown in each panel as the horizontal grey dashed line. The vertical dashed grey line marks the systemic velocity of the HI absorption.

photoionized phase and a warmer, possibly collisionally ionized, one. As shown in appendix 1 of Fumagalli et al. (2016) (see their figure A3), the inclusion of an additional C IV contribution produces an appreciable difference to the model results. At the current precision, however, this difference is comparable to $\approx 2\sigma$ of the recovered posteriors. We therefore choose to include these higher-ionization ions, mindful that an (unknown) contribution from a second phase is likely present.

As a second approximation, we do not apply a full model for dust depletion. While clear evidence of a small but non-negligible contribution of dust has emerged in the literature of absorption line systems, especially for DLAs (e.g. De Cia et al. 2016; Murphy & Bernet 2016), we currently have no detailed model to break the degeneracy between the intrinsic abundance pattern and the effect of dust depletion for LLSs. As shown in figure 9 of Fumagalli et al. (2016), models with and without dust depletion yield similar results for LLSs, with more scatter for $\log Z/Z_\odot > -1.5$ (as expected). Despite the good approximation, strictly speaking, the inferred metallicity is related only to the gas phase (see e.g. Berg et al. 2021).

Finally, we note that our operational definition of LLSs assumes that the detected metal abundance across the entire velocity range can be associated with the entire (or most dominant, see above) H I component. This approach smooths the fluctuations in the metallicity across components that are known to be present in these absorbers (e.g. Prochter et al. 2010; Fox et al. 2013), but is meant to capture the mean metallicity of the entire absorber and it is the only feasible measurement when lacking coverage of high-order Lyman series lines (Kacprzak et al. 2019).

3.3 Detection of galaxies in MUSE data

For each field, we search for any galaxies associated with one of the absorbers previously identified in the quasar spectra. We use two methods to detect galaxies. Firstly, we search for galaxies detected via their continuum emission in white light images constructed from the MUSE datacube. Secondly, we search for line emitters which are too faint to be identified in the continuum images. Specifically in this work, we look for Lyα emitters (LAEs) at the redshift of the
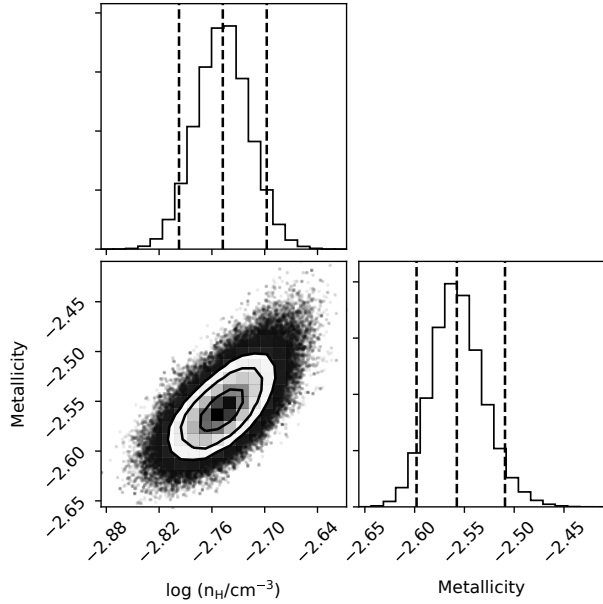
Figure 4. Example of the results from the photoionisation modelling to determine the metallicity of the absorption system. Here we show the posterior probabilities for the metallicity and the density for the system at $z = 3.69164$ in J013340+04005. Results for all the absorption systems in our sample are included in the online only material.
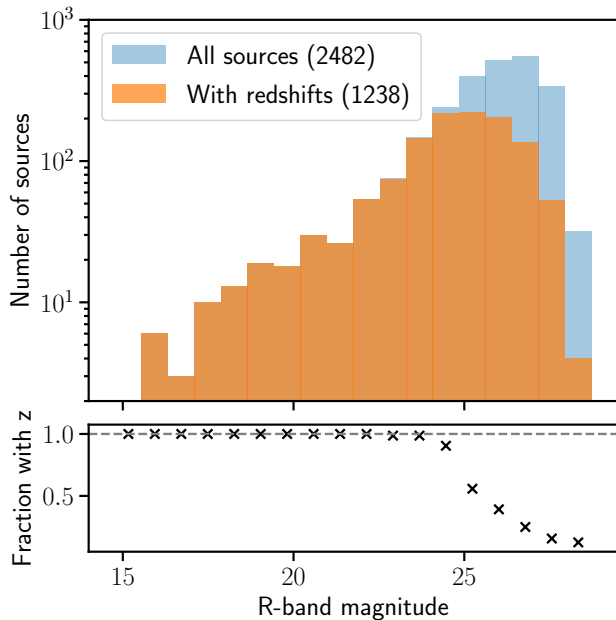
Figure 5. Number of detected sources in all 28 MAGG fields as a function of R-band magnitude. We are able to obtain reliable redshifts for 1238 sources for a spectroscopic survey that is $\approx$ 90 per cent complete down to $\approx$ 24.85 mag.
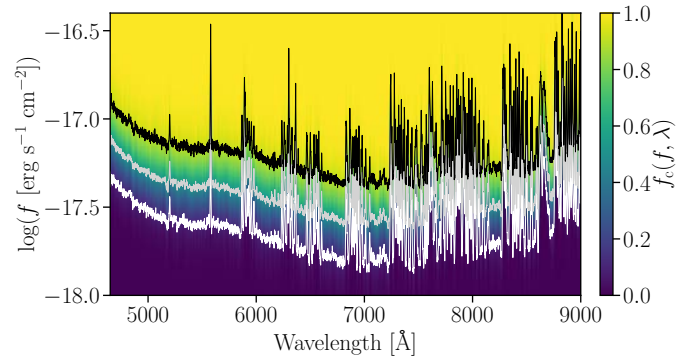
Figure 6. Recovery fraction of real LAEs as a function of wavelengths and flux in the MAGG sample. The black, grey, and white lines show, respectively, the 90, 50, and 10 per cent completeness limit.

absorbers. See Dutta et al. (2020) for a study using similar methods to detect [O II] emitters in the MAGG data at lower redshifts, namely $0.8 < z < 1.5$. We briefly describe the two methods used below and refer to Lofthouse et al. (2020) for a more detailed description of our methodology.

For the continuum source search, we use SExtractor (Bertin & Arnouts 1996) on the white light images created from the MUSE datacubes to identify all sources within the field of view (FOV) using the same parameters as in Lofthouse et al. (2020). Across all 28 fields, this results in a catalogue of 2482 sources. 1D spectra of the galaxies are extracted from the MUSE datacubes using the 2D segmentation map from SExtractor. We use these spectra and the MARZ tool (Hinton et al. 2016), specifically the customized branch by M. Fossati[3], to measure redshifts for each of these sources.

Fig. 5 shows the number of detected sources as a function of R-band magnitude. Of the 2482 sources identified by SExtractor we are able to determine reliable redshifts for 1238 sources with $\approx$ 90 per cent completeness down to $\approx$ 24.85 mag. These sources were assigned a confidence flag in MARZ of either 6 indicating a star or 3 and 4 indicating they showed multiple features which enabled a reliable estimate of the galaxy redshift according to the classification scheme followed in Bielby et al. (2019) and Lofthouse et al. (2020).

In addition to identifying continuum-detected sources, we search for LAEs associated with each absorber. In this work, we broadly define LAEs as high-redshift galaxies ($z \gtrsim 3$) displaying a significant Lyα emission, without imposing cuts in equivalent widths as traditionally found in the literature based on narrow band imaging. This additional search allows us to detect fainter galaxies with bright emission lines, which would otherwise be missed in the continuum search. To detect the LAEs, we first run the CubEx 3D detection algorithm on the full MUSE datacube, restricting the selection of sources to those with integrated $S/N > 5$. Here the S/N is based on the effective noise, as described in detail in section 3.2.5 of Lofthouse et al. (2020). We also refer the reader to this paper for the full description of the parameters used to reduce the detection of nonphysical sources, e.g. cosmic rays.

This produces a catalogue of potential sources across the full wavelength range of MUSE. For each absorber, we restrict the catalogue to a velocity window of ± 1000km s^{-1} and for each source we visually

---

[3] matteofox.github.io/Marz

inspect its spectrum, 3D segmentation map and images (constructed from mean, median and independent half cubes). This visual identification was performed by at least three authors (EKL,MFo,RD) and sources are only included where there is consensus that it is a real source. During the identification, we also classify sources into the following confidence groups (which expands the original classification presented by Lofthouse et al. 2020):

– 1: Highest confidence LAEs, showing e.g. bright and asymmetric lines and/or other absorption lines, with integrated $S/N > 7$;

– 1.5: Considered likely to be Lyα but less certain than confidence 1 sources, with integrated $S/N > 7$;

– 2: Likely to be Lyα but with $5 < S/N < 7$;

– 2.5: Tentative identification with $5 < S/N < 7$, but generally more noisy as visible on the images or affected by, e.g., sky residuals;

– 3: Poor quality sources, e.g. overlapping with the edge of MUSE cube and for which reliable photometry is not possible.

We exclude sources identified as confidence 3 in the following analysis.

In addition to the flux above a given S/N per voxel, directly provided by CubEx, we also estimate the total flux using a curve of growth method following Fossati et al. (2021) and, originally, Marino et al. (2018). In brief, we create a narrow-band image by collapsing the datacube within 15Å of the LAE redshift. We then compute the flux in circular apertures, centred on the CubEx coordinates of the LAEs, in steps of 0.2 arcsec with any nearby sources masked. The radius of the circular aperture is increased until the flux is no longer increasing by at least 2.5 per cent of the value at the previous radius. The flux at this final radius is taken to be the total Lyα flux. The median radius used is 7 pixels (1.4 arcsec) with a range between 3 pixels and up to 17 pixels for the largest sources. All LAEs and fluxes are visually inspected to ensure they are correctly centred and any unassociated sources of flux are masked. The Lyα flux for each source has been corrected for Milky Way extinction using extinction values for each of the MAGG fields derived from the dust map by Schlegel, Finkbeiner, & Davis (1998) and the extinction curve from Fitzpatrick (1999).

To establish the source redshifts we proceed according to the line shape. Specifically, LAEs fall into three categories. Firstly, the emitters which do not have any significant asymmetry in their profile. In this case we fit the line with a Gaussian profile and take the peak as the redshift of the LAE. In the second case, we have emitters with asymmetric line profiles, which we visually inspect to obtain the wavelength of the peak and then use this to determine the galaxy's redshift. Finally, a small number of LAEs (13) showed double-peaked lines with peak separations of around a few 100km s^{-1}. For these cases we take the wavelength of the red peak as the redshift of the galaxy, for consistency with what is measured for the single peak lines which are assumed to be the most redshifted component. See Fossati et al. (2021) for further details.

In some cases, emitters that are spatially close can be blended and are identified by CubEx as a single source. This is a particular problem in the regions that fall within the nebula around the quasar, as discussed in more detail in Fossati et al. (2021). Briefly, we separate these sources by generating a composite image of the Lyα emission on which we run SExtractor. A DEBLEND_CONT of 0.05 is used to balance the requirement to efficiently separate sources without over-shredding the signal into a high number of sources which are not real.

To further prevent over-separation of sources, we require each one to have an integrated $S/N > 5$ and to be composed of at least 9 pixels. For each source that satisfies this, we generate a 3D segmentation map from which we extract the spectrum.

As a last step, we measure the completeness of our LAE catalogue using the procedure described by Fossati et al. (2021), which consists of injecting LAEs from the MUSE Ultra Deep Survey (MUDF; Fossati et al. 2019) which is $> 5\times$ deeper than the MAGG observations, into the cubes. The recovery fraction as a function of flux and wavelengths for the MAGG cubes is shown in Fig. 6. Our LAE survey is 90 per cent complete for line fluxes $\approx 10^{-17.2}$ erg s^{-1} cm^{-2}, and ≈ 50 per cent complete for fluxes $\approx 10^{-17.4}$ erg s^{-1} cm^{-2}. The completeness for mock point sources is instead a factor ≈ 3 deeper, as shown by Fossati et al. (2021) and, originally, by Herenz et al. (2019). We note that only three sources are identified around absorbers in the continuum search, but these sources are subsequently also identified in the emitter search. In addition, four continuum sources that were detected by SExtractor but whose redshifts were not reliably determined in MARZ are subsequently confirmed at $z > 3$ as a result of the emitter search. Therefore, in the following analysis we will generally refer to LAEs as the galaxy population that is the subject of our study. The fact that our search yields a large and complete sample of LAEs confirms that MUSE is highly effective for the detection of emission line galaxies. This means, however, that our sample is incomplete with respect to dusty galaxies, passive galaxies, and/or active galaxies (such as some LBGs) showing only Lyα in absorption.

4 THE ENVIRONMENT OF OPTICALLY-THICK GAS CLOUDS

With the catalogues of galaxies and absorbers in hand, in this section we first explore the general properties of the environment of the optically-thick H I systems, before turning to a detailed investigation of the correlations between physical parameters of the absorbers and galaxies.

4.1 Detection rate of galaxies around absorbers

The emission line selection identified 127 LAEs across all confidence groups within ±1000 km s^{-1} of an absorber. The catalogue of these galaxies is included in the next paper in the MAGG series (Galbiati et al., in prep). Of these 55 (44 per cent) are flagged as confidence 1 (i.e. the highest confidence group), as shown in Fig. 7. A further 21 systems are assigned confidence 1.5, 36 systems are confidence 2, 3 are confidence 2.5 and only 12 in the lowest confidence group (confidence 3 which are not used in the analysis performed in this study). The majority of the absorption systems have only one or two detected galaxies with a decreasing frequency up to the highest number of detected LAEs of 8. There are only 11 absorbers (out of the full sample of 61) with no detected galaxies within the MUSE field of view and within the specified velocity range. This corresponds to a detection rate of 82±16 per cent when considering all LAEs and 57±12 per cent when considering only the highest confidence (confidence 1 and 1.5) LAEs. To check whether the presence of sky lines has impacted the non-detection of galaxies, we investigate where these 11 absorbers with no associated LAEs fall at wavelengths of low transmission as seen in Fig. 6. Only 3 of these absorbers are found at redshifts close to significant sky lines. The remaining 8 absorbers are all in regions with good sky transmission, and so we would have
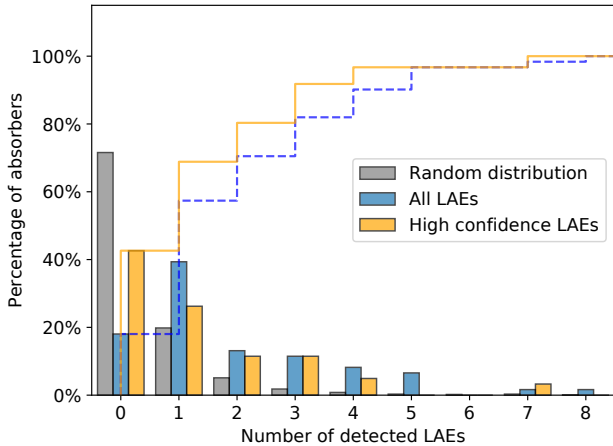
Figure 7. Number of LAEs detected within the MUSE FOV and at a velocity separation of < 1000 km s^{-1} around each absorber. In blue we show all detected LAEs while in orange we include only those that were classified as high confidence (confidence 1 and 1.5) which have S/N > 7. The solid orange and dashed blue lines show the cumulative number of absorbers for the all the galaxies and the high confidence sample respectively. For comparison, we also show the number of S/N > 7 LAEs expected from a random distribution in grey.

detected any galaxies above the typical flux limits quoted if they were present. Therefore we conclude that the lack of detections is an intrinsic properties and not a selection effect.

For comparison, we also show the results expected from a random distribution of LAEs. To calculate this we randomly choose a field and a redshift between 3 and 4.5. Given this position, we search for any LAEs within the MUSE FOV and ± 1000 km s^{-1} at $S/N > 7$. This process is repeated 1000 times and the resulting histogram is shown in grey. We find no associated LAEs in 72 ± 4 per cent of the cases, a single galaxy in 18 ± 1 per cent and a small number of multiple associations (less than 10 per cent). This is much lower than the percentage of multiple associations found around the absorbers (42 per cent).

While it has previously been common to consider a one-to-one association between absorbers and galaxies, the large FOV of MUSE, combined with its sensitivity, enables multiple galaxy detections, as seen here, of which many are plausibly associated to the absorbers. Indeed, such multiple detections are seen in other MUSE fields both at high redshift and low redshift (e.g. Fumagalli et al. 2016; Mackenzie et al. 2019; Dutta et al. 2020; Péroux et al. 2019). Multiple detections have also been seen around high-redshift DLAs using the Hubble Space Telescope (Møller et al. 1998) and in lower redshift studies focusing on Mg II absorption systems (e.g. Bergeron, Cristiani, & Shaver 1992; Whiting, Webster, & Francis 2006; Kacprzak, Murphy, & Churchill 2010).

4.2 Establishing the association between LAEs and ALSs

To investigate the environment of these absorbers on a more statistical ground, we compute a series of metrics to establish how closely related are LAEs to the ALSs in these fields.

4.2.1 Clustering in velocity

To identify LAEs with high completeness, we have adopted a generous velocity window of ± 1000 km s^{-1} to establish a link with ALS, which corresponds to a line-of-sight distance of ~ 3 Mpc at $z \sim 3$. We now evaluate more in detail how LAEs cluster in velocity space with respect to ALSs, to rule out the possibility of a substantial spurious association.

Fig. 8 (left) shows the velocity separation between the LAEs and the absorbers where the redshift is determined from the HI absorption. We find a clear peak in the number of detected LAEs at velocity separations around 250 km s^{-1} in both the full sample of LAEs (blue) and when restricted to only the highest confidence galaxies (orange). The bulk of the emitters are distributed within ≈ 500 km s^{-1} of this peak, with very few LAEs scattering at larger velocity separation.

The small offset relative to the velocity of the absorbers is an expected and known property due to scattering of the Lyα photons (e.g. Steidel et al. 2010; Verhamme et al. 2018). Similar offsets have been seen in previous works. For example, Muzahid et al. (2020) find an average velocity difference of 169 ± 10 km s^{-1} between their sample of 96 LAEs at $z \sim 3.3$ and nearby absorption systems while Shibuya et al. (2014) find that the Lyα line in LAEs is shifted by ~ 230 km s^{-1} relative to the nebular emission lines in the same galaxies. Similarly, when using the C IV absorption lines (where available) to determine the redshift of the absorber, we again find a peak around 250 km s^{-1} (Fig. 8, right). We find no significant offset between the velocity of the H I and the metals as traced by C IV.

This clustering around 250 km s^{-1} strongly indicates that, despite the large velocity window and possible contamination from projection effects, there is a clear physical association in velocity space between the absorbers and the galaxies detected in emission. We clarify at this point that, in the context of this work, we define an "association" as a link between absorbers and galaxies that are not arising at random, as clearly shown in Fig. 8. The grey distribution in this figures shows the results for a random arrangement of LAEs, unrelated to absorbers. This was calculated by randomly choosing a field and a redshift between 3 and 4.5 as with the random sample used in Fig. 7. Given this position, we search for any LAEs within the MUSE FOV at $S/N > 7$. This process is repeated 1000 times to obtain the distribution shown. This random sample shows a broadly uniform distribution across the full velocity range with no peak as observed around the absorbers. By comparing the random and observed distributions within ± 500 km s^{-1} of the peak, we infer an overdensity of ≈ 6 near LLSs.

In previous works, the focus has often been on identifying the "host galaxy" to the ALSs, more strictly defined as the galaxy within which the absorbing gas resides. In our analysis, for the reasons provided below, we do not attempt to make this strict identification and hence the definition of an association should be regarded in broader terms. This is especially true given that the projected distances between galaxies and ALSs (the impact parameter, see Fig. 9) is generally large, ranging from ≈ 18 kpc to ≈ 275 kpc, with a median value of 165 kpc.

4.2.2 Number density and luminosity function

Having established that LAEs are clustered in velocity with respect to the ALSs, we now consider in detail how the detected number density compares with expectation from a field population. Based on the Lyα luminosity function at $z \sim 3 - 4$, we would expect
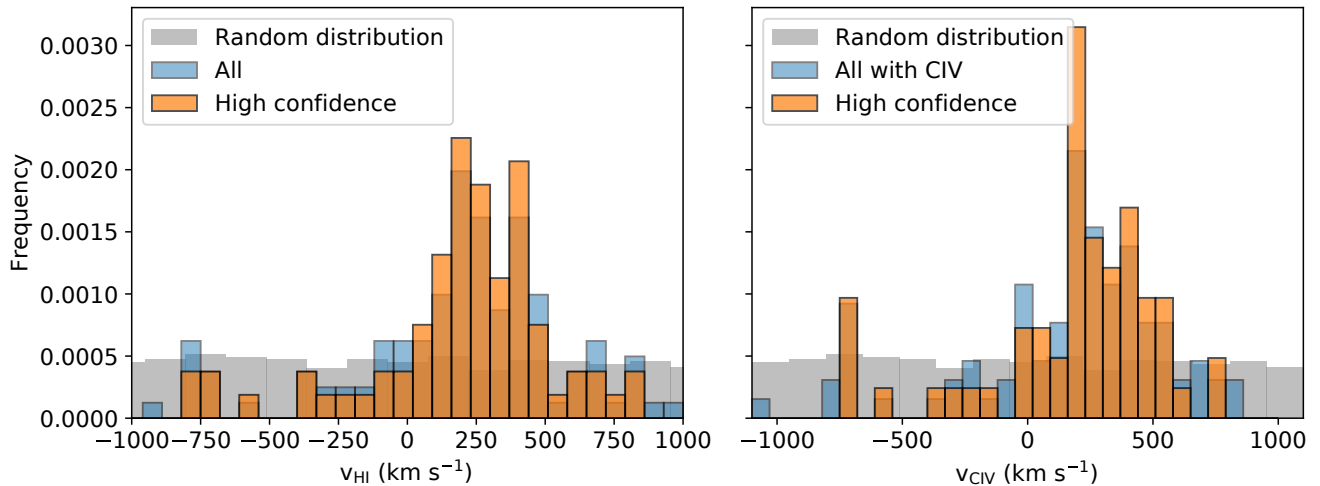
Figure 8. Velocity separation between the absorbers and the LAEs detected within the MUSE FOV. In blue we show all detected LAEs while in orange we include only those that were classified as high confidence (confidence 1 and 1.5). For comparison, we also show the results obtained from a random distribution of galaxies (in grey). The left figure uses the H I to determine the redshift of the absorber while the right figure shows the velocity separation when using the redshift determined from the C IV absorption lines. For all histograms, there is a peak in the velocity distribution at ~250 km s^{-1}.
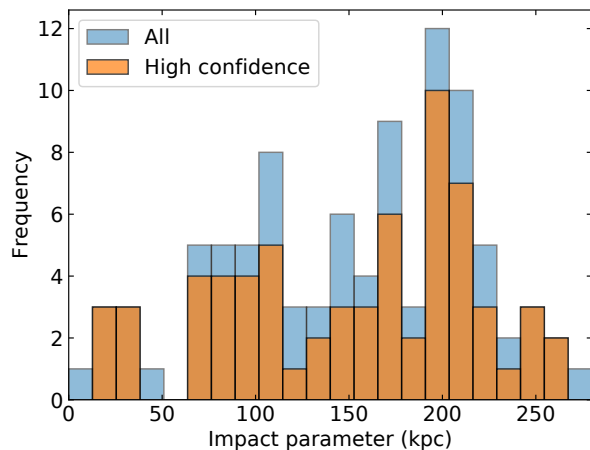
Figure 9. The distribution of impact parameters for all the identified LAEs within 1000 km s^{-1} of an absorber. In blue we show all detected LAEs while in orange we include only those that were classified as high confidence (confidence 1 and 1.5).

to find an average of between 0.3 and 1 source[4] at a luminosity of $\geq 4.5 \times 10^{41}$ erg s^{-1} within the ≈ 120 Mpc^3 comoving volume searched around each absorber (Grove et al. 2009; Cassata et al. 2011; Drake et al. 2017; Herenz et al. 2019). This luminosity is chosen as it is the value at which we are ≈ 50 per cent complete for extended sources based on the tests shown in figure 9 of Lofthouse et al. (2020). As an example, for an absorber at $z = 3.5$, a ± 1000 km s^{-1} search window equates to a line-of-sight distance of ~ 6.6 Mpc. This is assuming expansion purely due to the Hubble flow and $H_0 = $

[4] For a discussion on the differences among the various luminosity functions obtained with different techniques, see Herenz et al. (2019).

67.7 km s^{-1} Mpc^{-1} (Planck Collaboration et al. 2016), which gives 2000 km s^{-1} / (67.7 km s^{-1} Mpc^{-1} × (1 + 3.5)) = 6.56Mpc, or $\sim$ 30 cMpc (comoving Mpc). The MUSE FOV at this redshift is 0.45×0.45 Mpc giving a search area of $[0.45 \times (1+3.5)]^2 \approx 4$ cMpc. Combining this search area with the line-of-sight distance gives a total search volume of 30Mpc^2 × 4Mpc $\sim$ 120 cMpc^3.

Therefore, for the 61 absorbers studied, the random distribution of LAEs at this redshift would lead to a detection of $\approx 18 - 60$ LAEs within the full MAGG sample. We note that apart from two absorbers in J200324−325144 which have a separation of $\Delta z \approx 0.015$, there is no overlap in the search areas around any of the other absorbers. The identification of 127 LAEs in the MAGG sample is $\approx 2 - 7$ times greater than the number expected from a random distribution. This is in line with the inference from the clustering in velocity above, and it offers convincing evidence of an excess of galaxies, indicating that, on average, these LAEs are physically associated to optically-thick ALSs.

For a more formal comparison, we use the luminosity function (LF) to compare the environment around the absorbers to that of the field. The LF is defined as the number of galaxies at a given luminosity and redshift per unit volume. Given that the LF of LAEs is known to show no significant evolution over the redshifts studied (Herenz et al. 2019), we assume any variation due to redshift is negligible and bin all our LAEs together. To construct the LF for the galaxies we use the 1/$V_{\rm max}$ method as in Schmidt (1968) and Felten (1976) with a luminosity dependent selection function (Fan et al. 2001; Herenz et al. 2019). This can be defined as

$$\phi(\langle \tilde{L}_{\rm Ly\alpha} \rangle) = \frac{1}{\Delta \tilde{L}_{\rm Ly\alpha}} \sum_i \frac{1}{V_{\rm max,i}} \qquad (1)$$

where $\langle \tilde{L}_{\rm Ly\alpha} \rangle$ is the average Lyα luminosity in a given bin, $\Delta \tilde{L}_{\rm Ly\alpha}$ is the width of that bin and $V_{\rm max,i}$ is the volume of the survey which has been adjusted for the selection function (Johnston 2011). See Fossati et al. (2021) for more details on the methodology used here.

Motivated by the clustering in velocity space previously shown, we
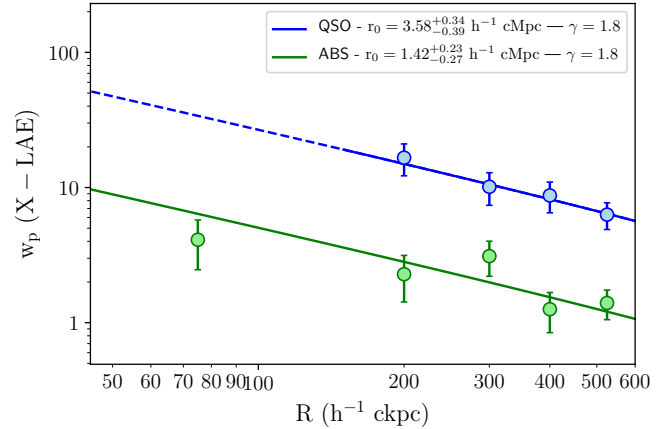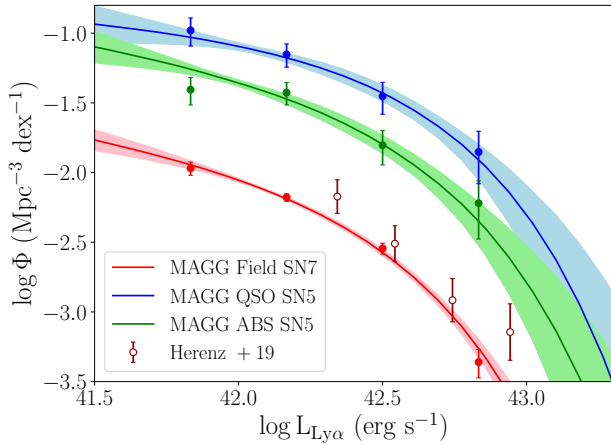
Figure 10. *Left:* Differential LAE luminosity functions for galaxies in different environments in the MAGG sample. In red we show field galaxies which are selected at $S/N > 7$ across the entire MUSE cubes. In green we show the LF for galaxies within ± 500 km s^{-1} of optically-thick H I absorbers and in blue are galaxies within ± 500 km s^{-1} of the quasars. Markers shown with error bars indicate the $1/V_{max}$ binned values while the solid lines are Schechter fits with a 1σ error shown by the shaded region. Results from Herenz et al. (2019) using their real source selection function are shown by the open red points. *Right:* Angular cross-correlation function between absorbers and LAEs (green) and quasars and LAEs (blue). Binned values are shown with 1σ Poisson error bars and the solid line shows the data fit to Equation 3.

restrict the LAEs to only those within ± 500 km s^{-1} of the peak of the Lyα emission as in Fig. 8 (i.e. within -245 and 745 km s^{-1} of the absorber redshift), given that galaxies beyond this are increasingly more likely to be chance rather than physical associations. We fit the LF with a parametric Schechter function (Schechter 1976) defined as

$$\Phi(L) = \ln(10)\Phi^* 10^{(\tilde{L}-\tilde{L}^*)(1+\alpha)} \exp(-10^{(\tilde{L}-\tilde{L}^*)}) \qquad (2)$$

We show the LF of galaxies around the absorbers in Fig. 10 (left). The points with errors are the $1/V_{max}$ binned values and the solid line shows the Schechter function fit. Also shown is the LF for galaxies detected around the quasars (blue) as in Fossati et al. (2021) using a search window of 500 km s^{-1} corrected by 245km s^{-1} due to the scattering of Lyα and with the same S/N cut of 5. The quasar redshift here is taken to be the peak of the Lyα emission. Given that the redshifts for both the QSO and the associated LAEs are both derived from Lyα, radiative transfer effects are expected to have similar impacts on the measured redshifts and therefore compensate for each other. Indeed, Fossati et al. (2021) discuss this effect and show that the average velocity offset between the LAEs and QSO is close to zero.

To compare with the field population, we include the LF by Herenz et al. (2019) and we also compute the LF for the $\approx 1,100$ LAEs found in the MAGG cubes up to a $S/N > 7$ (MAGG V; Galbiati et al., in prep), which are identified with the same procedure described here. The Herenz et al. (2019) LF is derived from observations of 237 LAEs over $2.9 < z < 6$ from the MUSE-Wide survey (Urrutia et al. 2019).

Compared to the two field LFs, we see an increase in the number of LAEs near absorbers, with a normalisation that is about a factor ≈ 5 higher than the field. This is in line with the excess noted above with simple counting arguments. Conversely, the shapes of the LF in the field and near absorbers are the same, suggesting that LAEs near optically-thick ALSs have comparable luminosity to the field. By comparing with the quasar sample in Fossati et al. (2021), we find again a similar shape but with a higher normalisation than the

absorber sample, demonstrating that the excess of LAEs near ALSs is bracketed by values in the field from below and values near quasars from above.

4.2.3 Projected cross-correlation function

An additional way to assess statistically an overdensity of LAEs around the absorption systems is to compute the cross-correlation function between the absorbers and LAEs. This can be thought of as the increased probability of detecting a galaxy in a given volume at a specified distance relative to the probability of detecting a similar galaxy at a random position. The quasar-LAE angular cross-correlation has been calculated for the MAGG sample by Fossati et al. (2021). We follow the same methodology here (see also Trainor & Steidel 2012) for the absorbers and compare our results. The reduced angular cross-correlation function within a given velocity window (in our case ± 500 km s^{-1} as motivated by the clustering seen in Fig. 8 and adjusted for the Lyα offset of ≈ 250 km s^{-1}) is defined as

$$w_P^{AL}(R) = (R_0^{AL}/R)^\gamma \times {}_2F_1(1/2, \gamma/2; 3/2; -z_w^2/R^2), \qquad (3)$$

where R is the impact parameter between the absorbers and the LAEs, $_2F_1$ is the Gaussian hypergeometric function which corrects the power law for the effects of truncations introduced by the use of a finite redshift window, and $z_w = (500$ km s$^{-1})(1 + z)H(z)^{-1}$ is the half-width of the redshift window in units of cMpc.

The results of the cross-correlation are shown in the right panel of Fig. 10. The green markers show the values binned in circular annuli centred on the absorbers up to 600 h^{-1} ckpc where this upper limit is defined by the MUSE FOV. The solid line shows the fit to Equation 3. Also shown are the results for the quasar-LAE cross-correlation (blue). The slope of both fits, γ, is fixed to 1.8 following the convention of other similar studies (see e.g. Diener et al. 2017; García-Vergara et al. 2019; Fossati et al. 2021) as a result of the small range of distances which can be probed within a single MUSE pointing. The resulting R_0 is insensitive to the precise choice of γ as
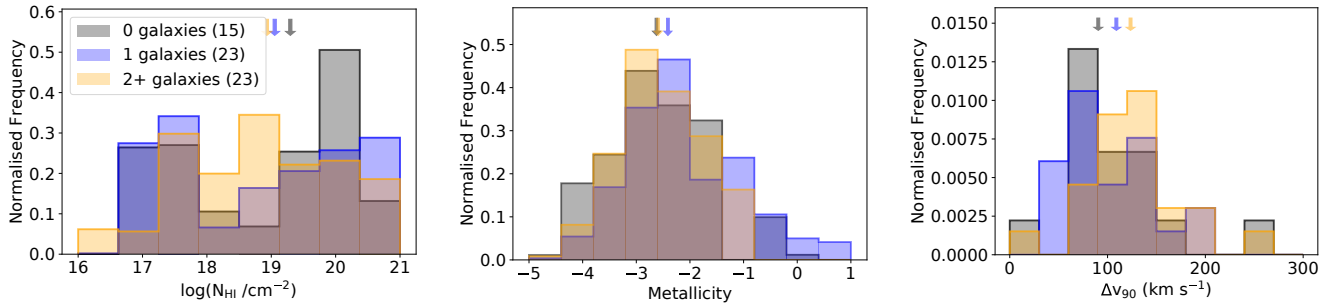
Figure 11. *Left:* Distributions of N_{HI} column density for absorbers where we detect: no galaxies within the MUSE FOV and within ± 500 km s^{-1} of the absorber redshift (grey); detect a single galaxy association (blue); and detect multiple galaxy associations (orange). The number of absorbers in each sample are indicated by the number in brackets in the legend. *Middle:* Metallicity distributions for the same three groups. *Right:* Δv_{90} distributions for each group. The median values of each distribution are indicated by the correspondingly coloured arrows at the top of each panel. For this analysis we use the full LAE sample (confidence < 3).

discussed in Diener et al. (2017). We report a clustering length $R_0 = 1.42^{+0.23}_{-0.27}$ h^{-1} cMpc for the absorber-LAE angular cross-correlation. This weak but positive signal indicates the presence of clustering on the small scales probed within the MUSE FOV, consistent with the results seen in the velocity distributions and the LAE luminosity functions. In comparison to the quasars, which have a $R_0 = 3.58^{+0.34}_{-0.39}$ h^{-1} cMpc, the clustering around the absorbers is weaker than around the quasars, and it is more similar to the value inferred for the autocorrelation of LAEs ($R_0 = 2.09^{+0.46}_{-0.47}$ h^{-1} cMpc; Galbiati et al. in prep.). In particular, we note that in the case of the ALSs, LAEs are distributed over a larger range of radii, diluting more the amplitude of the cross-correlation function on small scales compared to the larger contrast with respect to the field that emerges when studying the LF that marginalises over the angular separation.

Overall, from the evident clustering in velocity, the excess in the normalisation of the LF and a positive clustering from the projected cross-correlation, we conclude that there is strong statistical evidence that the environment around optically-thick absorption systems is richer in galaxies than the field, similarly (but less prominently) to what is found near the quasars themselves. The environment and clustering around hydrogen absorption systems has previously been investigated using LBGs. For example, using VLRS, Bielby et al. (2011) and Crighton et al. (2011) find an increase in the neutral hydrogen gas within few Mpc of galaxies. This is consistent with other LBG surveys, such as KBSS where a definite clustering between hydrogen and metal absorbers is found with respect to continuum-detected galaxies (e.g. Adelberger et al. 2003, 2005; Rudie et al. 2012). Our study extends this analysis to the lower masses probed by LAEs, using spectroscopic redshifts from MUSE that are more precise than those from narrow-band surveys and employing large enough samples to statistically study the cross-correlation between absorbers and such spectroscopically-selected galaxies.

4.3 Dependence of absorption properties on galaxy environment

Having established a correlation between LAEs and optically-thick H I selected ALSs, we now investigate whether there is a link between the richness of the galaxy environment (i.e. the number of LAEs identified next to ALSs) and the properties of the absorbers themselves, namely the H I column density or the metallicity.

Fig. 11 (left) shows the distribution of the HI column densities for absorbers split into three categories: absorbers where we do not detect any LAEs within the MUSE FOV and within ± 500 km s^{-1} of the peak of the velocity distribution seen in Fig. 8 (grey), absorbers where we detected a single galaxy within this volume (blue) and absorbers where we detect two or more galaxies (orange). While they are not shown in the figures due to low numbers, we specifically check the values for systems with a very high number of detected LAEs (5 or more) and find that they do not stand out as having particularly high or low values and they broadly follow the distributions of the 'two or more galaxies' sample shown. For this analysis, we consider LAEs with confidence class < 3. The median value of each distribution is indicated by the corresponding arrows at the top of the figure. All of the categories show broadly the same distribution in H I column density, spanning almost the entire range covered by our sample and with no clear peak in their distributions. We see no evidence of a global correlation between N_{HI} and the wider galaxy environment. Indeed, both the highest ($\log(N_{HI}/\text{cm}^{-2}) \approx 21.1$) and lowest ($\log(N_{HI}/\text{cm}^{-2}) \approx 16.5$) column density systems studied in MAGG show multiple galaxy associations.

To statistically test whether there is any difference in the N_{HI} distributions, we use the Kolmogorov-Smirnov (KS) test. The p-value, indicating the probability of the null hypothesis that the distributions are drawn from the same overall distributions, is > 0.7 for all the comparisons between the three distributions. In particular, comparing the distribution of N_{HI} for absorbers where we detect no galaxies and where we detect 2 or more galaxies results in a p-value of 0.84. Even when restricting the sample to only the highest confidence LAEs, the KS test gives a p-value of >0.5 between all distributions. From this, we conclude that statistically there is no dependence of the N_{HI} on the number of detected galaxies and it is therefore not possible to predict the number of LAEs, and hence the environment in which an absorber lies, purely from the H I column density of the system.

We investigate next whether the metallicity of the absorber can provide any hints to the environment. It may be expected that higher metallicity systems would have more associated galaxies (or be closer to a galaxy, which is explored in Section 5) as the cumulative effects of stellar feedback or the removal of enriched gas by environmental processes would provide a source of metals for the absorbing gas. Fig. 11 shows the distribution of metallicities for each of the same three categories: no galaxies, single galaxy, and multiple galaxy associations. The photoionisation modelling used to determine the metallicity of the absorbers from the observed column densities, as described in

Section 3.2.1 provides a PDF of the metallicity of each absorber. We combine these PDFs for all of the absorbers in each of the three categories to produce the histograms shown in the figure. As with the H I column density, we see no significant differences in the distributions for each of the three categories with each distribution spanning a wide range of metallicities. As with the $N_{\rm HI}$ distributions, we show the median values with a vertical arrow at the top of the figure.

All three of the distributions show the same shape and have very similar median metallicities. The group of absorbers with no associated galaxies have a median metallicity of $\log(Z/Z_\odot) \approx -2.62$, absorbers with a single detected galaxy have a median metallicity of $\log(Z/Z_\odot) \approx -2.41$ and the absorbers with 2 or more detected galaxies have a median metallicity of $\log(Z/Z_\odot) \approx -2.59$. This result still holds when we repeat the analysis using only the highest confidence LAEs in our sample (confidence < 2). We repeat the KS test for our metallicity distributions, again finding p-values that indicate that the distributions are consistent with each other, with probabilities over 0.5 for all the comparisons between each of the three distributions. The similarity in the distributions means that it would be impossible to predict the galaxy environment based purely on the metallicity of the absorber alone.

A final absorber property we investigate is the kinematics, in this case via the use of the Δv_{90} parameter. This parameter is the velocity range which encompasses 90 per cent of the optical depth of an unsaturated line (Prochaska & Wolfe 1998; Ledoux et al. 2006). In the right hand panel of Fig. 11 we show the distributions of the Δv_{90} parameter for the same three samples as in the previous plots ("no galaxies", "one galaxy" and "two or more galaxies"). These distributions are calculated using the average Δv_{90} value for all lines in a given absorber. However the analysis has been repeated using only Si II to investigate the kinematics of the low ionisation gas and separately with C IV for the higher ionisation gas. These distributions show the same results as the average Δv_{90}.

As with the H I column density and metallicity, the distributions for each of the three samples look broadly the same with the distributions of each one covering almost the full Δv_{90} range. The median values for the sample with no detections and a single detection are ≈ 90 km s^{-1} and ≈ 100 km s^{-1} respectively while the sample for two or more detections shows a slightly higher value of ≈ 125 km s^{-1}. Indeed the KS test between the 'no galaxies' and 'one galaxy' gives a p-value of 0.74 indicating a high likelihood that they are drawn from the same distribution. On the other hand, the KS test between the 'no galaxies' and 'two or more galaxies' samples results in a p-value of 0.07. This difference between the samples is more prominent when considering only the high confidence LAEs which gives a p-value of 0.002.

We conclude from this analysis that the absorption systems studied here are found within a wide range of environments and there does not appear to be any clear correlation between the global properties of absorbers such as their H I column density or metallicity and the richness of its galaxy environment. While there does appear to be a tendency for absorbers within rich environments with multiple galaxy detections to have higher Δv_{90}, the distribution spans the full range of values and hence it is not possible to determine the richness of an absorber's environment purely from the properties of the absorber itself.
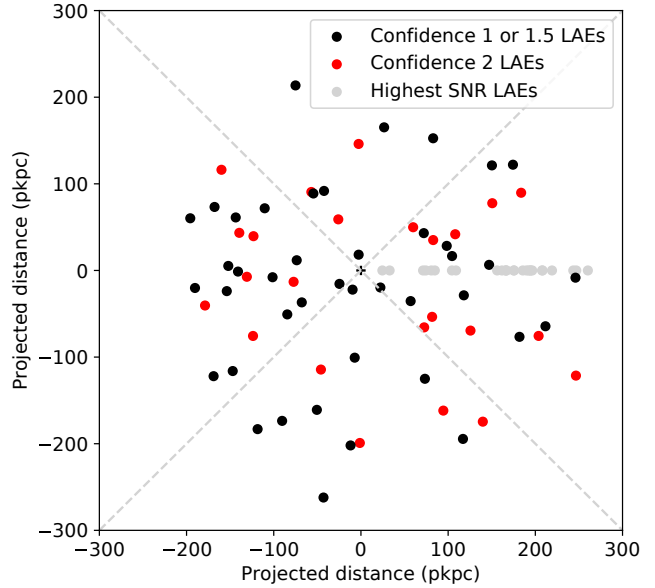

Figure 12. Spatial distribution of LAEs relative to the position of the absorption system (at the centre). For each field we rotate the positions of the LAEs such that the highest S/N LAE is along the positive x-axis (grey points). All other LAEs are identified as either high confidence (black; confidence 1 or 1.5) or lower confidence (red; confidence 2 or 2.5). Systems with only a single LAE are not included in this plot.

4.4 Spatial distribution and alignment of LAEs

One of the possible scenarios for the formation of LLSs is that they originate from gas within the cosmic filaments that connect galaxies. Various studies have attempted to probe such structures combining absorption spectroscopy and galaxies in emission as in MAGG (Møller & Fynbo 2001; Fumagalli et al. 2016b), and in recent years we have now begun to observe these filaments directly through the detection of their very faint emission (e.g. Umehata 2019; Bacon et al. 2021; Daddi et al. 2021). Indeed, for the first sightline studied in MAGG and presented in Lofthouse et al. (2020), we conclude that this is the most likely case for the extremely metal-poor LLS observed at $z \approx 3.53$ (see also Fumagalli et al. 2016b, Møller & Fynbo 2001, Mackenzie et al. 2019 for other examples). If this is the case for the majority of the absorption systems, we would expect to see a preferential alignment between the positions of the LLSs and the galaxies detected with MUSE.

To test for this effect, Fig. 12 shows the positions of all the detected LAEs in groups of two or more LAEs relative to the position of the absorption systems, which are at the centre of the figure. This analysis is therefore limited to 92/127 (72 per cent) LAEs out of the full sample. The positions have been rotated around this point such that the highest S/N LAE in each group is positioned along the positive x-axis (grey points). If the LLSs are lying within a filamentary structure we would expect an overdensity of galaxies along the horizontal axis (i.e. the axis between the absorber and the first aligned galaxy) relative to the number along the vertical axis.

Dividing the area into four sections, defined by cones with a 45 degree opening angle to the x-axis (see grey dashed lines in Fig. 12), we
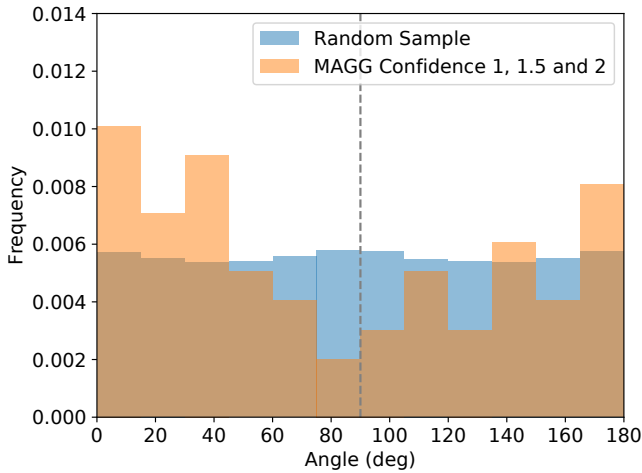
Figure 13. Number of galaxies at a given angle from the axis between the absorber and the closest LAE (orange; shown along the positive x-axis in Fig. 12). For comparison, we also show (in blue) the distribution of angles expected for a random distribution of galaxies.

find 79 galaxies (80 per cent) along the horizontal axes (i.e. left/right cones) compared to only 20 on the perpendicular axis (top/bottom cones). When removing the highest S/N LAEs from this calculation, as they are by definition along the x-axis, we still have 50 (71 per cent) galaxies in the left/right quadrants. We also test this result by restricting the calculation to only our highest confidence sources (confidence 1). Again, we still observe an excess along the horizontal axis with 29 galaxies in the left/right cones compared to only 11 in the top/bottom. Based on Poisson statistics the probability of obtaining this result with a random distribution of galaxies is ≈ 1 per cent.

To investigate this result further, we explore the distribution of angles between the LAEs and the axis defined by absorbers and the *closest* LAEs. Fig. 13 shows the number of galaxies at a given angle from this axis (orange). We also show the result that would be expected from a random distribution of galaxies (blue) which is broadly flat across all the possible angles. This random distribution was calculated by drawing coordinates at random in the range -250,+250 kpc in each direction. We then computed the angle distribution as normal and repeat the process 10,000 times.

For the LAEs near LLSs, we see an excess of galaxies near 0 degrees, i.e. along the same axis as the closest LAE, and at 180 degrees, i.e. in the opposite direction with respect to the quasar relative to the closest LAE but along the same axis. In contrast, there are significantly fewer galaxies around 90 degrees than would be expected from a random distribution. A KS test between the observed LAEs results and the random distribution gives a probability of 0.08 that the samples are drawn from the same overall distribution. From this, we conclude that there is evidence for a preferential alignment of LAEs around absorption systems at $z \sim 3 - 4$. In contrast, Fossati et al. (2021) performed a similar analysis for LAEs around quasars at $z \sim 3 - 4.5$ in MAGG and found no strong evidence for any spatial alignment of LAEs in these environments.
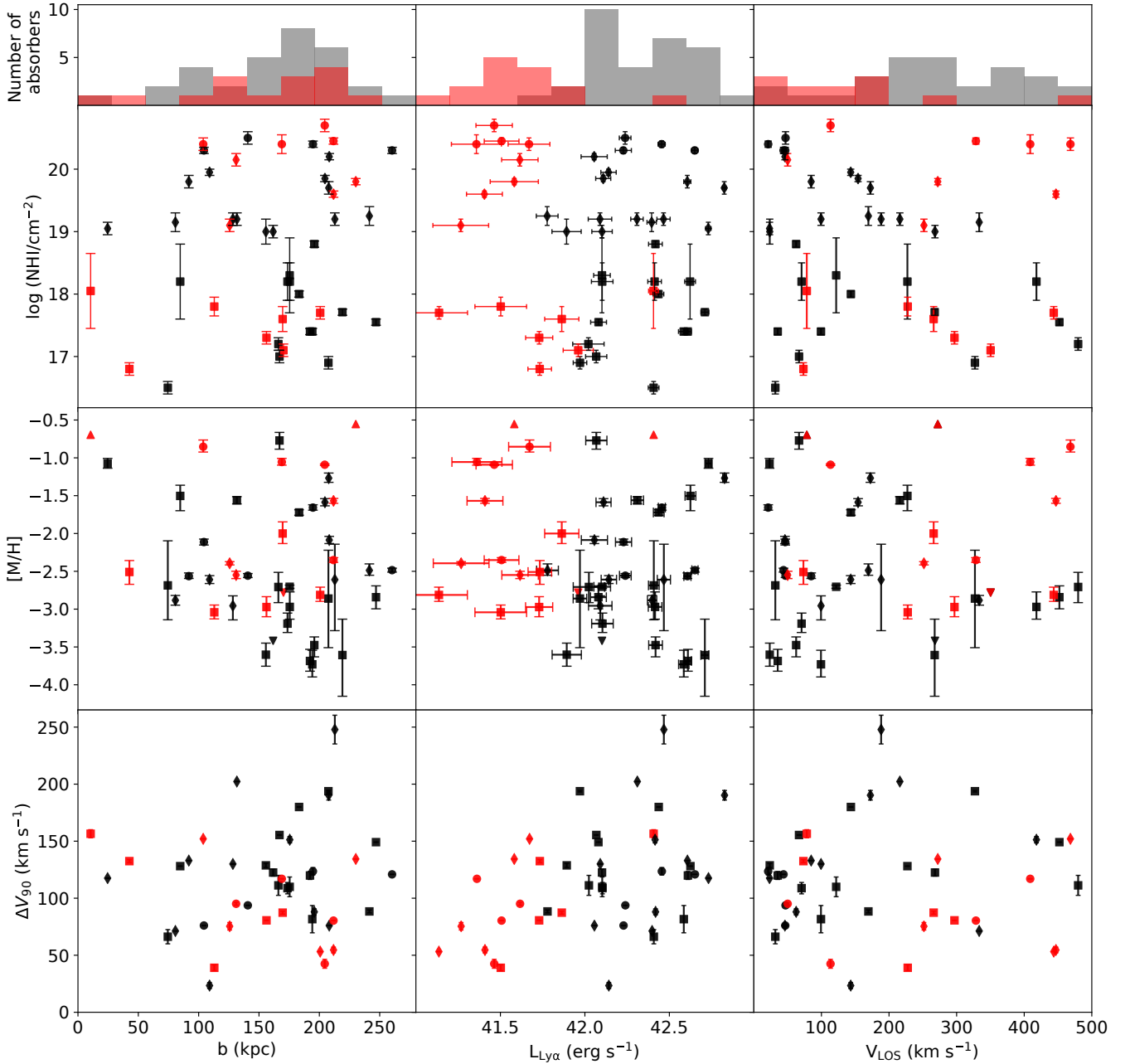
5 CORRELATIONS BETWEEN ABSORBERS AND LAEs

In the previous sections we have explored globally the relation between the properties of the optically-thick gas and galaxies. With the large sample of systems in MAGG, it is also possible to study the relationship between individual absorbers and galaxies in a statistical sense and investigate the presence of any correlations between absorption and emission properties.

As shown in Section 4, using MUSE observations we often detect multiple galaxies around a single absorber. This presents the issue of deciding which of the multiple galaxies should be used to compare properties with the absorber. The two widely-used options are to use the 'brightest' galaxy or the 'closest' galaxy. By choosing the brightest galaxy we are interested in investigating the relevance of what is assumed to be the highest star forming and most massive galaxy which is likely to be the central galaxy of the halo where gas is detected. Caveats to this approach are the use of Lyα which does not trace the entire star-forming population, it is an imperfect indicator of star formation, and the fact that there is non-negligible scatter in the underlying galaxy main sequence. Alternatively, by choosing the closest galaxy to the quasar line of sight, we focus on the galaxy that is not necessarily the central one (hence, we could select satellites) but the one that might have the largest influence given its proximity to the absorbing gas. As shown above, all the galaxies considered are highly clustered with the absorbers along the line of sight, which mitigates the impact of projection effects. In this case, however, we do not know the true location along the line of sight, although we can test for correlations with respect to the line of sight velocity. In the following, we show the results of using both of these options and discuss the impact of this choice on our conclusions.

We now proceed to examine, in turn, how the absorption properties (H I column density, metallicity, and velocity width) relate to properties defined by the galaxies in emission (relative distance between absorber and galaxy or impact paramater, velocity separation along the line of sight between absorber and galaxy, and Lyα luminosity). By definition, this experiment is only limited to the subset of 49/61 absorbers that are associated with at least one galaxy detected via Lyα emission. (as noted in Section 4.2.1, associations here are considered within the entire FOV and within ± 500 km s^{-1} of the peak of the LAE velocity distribution relative to the absorbing gas). Results are shown in Fig. 14-15 for the case of brightest and closest galaxies, respectively. Highest confidence sources are in black and the lower confidence sources are in red. Different symbols denote intervals of column densities for the absorbers (circles for DLAs, diamonds for sub-DLAs, and squares for LLSs). While we cannot investigate the correlations specifically for the handful of LAEs that are also continuum-detected sources due the limited statistics, we do check whether they show any tendencies to lie in certain regions of these plots. Beyond being at the higher end of the luminosity range, as expected, they appear to show no preferences in any of the properties studied here and fall within the scatter of the overall population.

5.1 Correlations with impact parameter

Considering the relationship between the H I column density of the absorber, N_{HI}, and the impact parameter, in the parameter space probed in MAGG which ranges from lower H I column density systems at $10^{16.8}$ cm^{-2} up to DLAs, we do not see any evidence of a correlation with absorbers. At all column densities, absorbers are found near galaxies and are scattered across impact parameters of $\approx 50 - 250$ kpc. This applies to both the results when we select

Figure 14. Correlations between the absorber properties and the properties of the brightest LAE. The top row shows the relationship between the absorber column density and impact parameter (left), Lyα luminosity (middle) and line of sight velocity separation (right). The middle row shows the absorber metallicity against the same three galaxy properties while the bottom row shows the Δv_{90}. In each plot we show the high confidence LAEs (i.e., confidence 1 and 1.5) in black and the lower confidence LAEs in red. Additionally we use different marker shapes to denote the absorber column densities with circles for DLAs, diamonds for sub-DLAs, and squares for LLSs. Upper and lower limits are shown by triangular markers. At the top of the figure we also show histograms of the number of absorbers within bins of the galaxy properties (impact parameter, luminosity and velocity separation). These histograms are coloured with grey denoting high confidence sources and red denoting the lower confidence sources.

the brightest LAEs in the FOV where multiple galaxies are detected (Fig. 14; top left) and where we select the closest galaxy to the absorber (Fig. 15; top left). To quantify this, we conduct a Kendall's Tau test on the two data sets (e.g. Brown et al. 1974; Isobe, Feigelson, & Nelson 1986). This provides a p-value indicating the probability of obtaining the observed data given that the two variables are uncor-

related. For the brightest LAE and closest LAE sample, we obtain p-values of 0.46 and 0.98 respectively indicating that it is highly likely that the $N_{\rm H\,I}$ is uncorrelated with impact parameter.

We examine next the comparison between the metallicity of the absorber and the impact parameter between the absorber and the
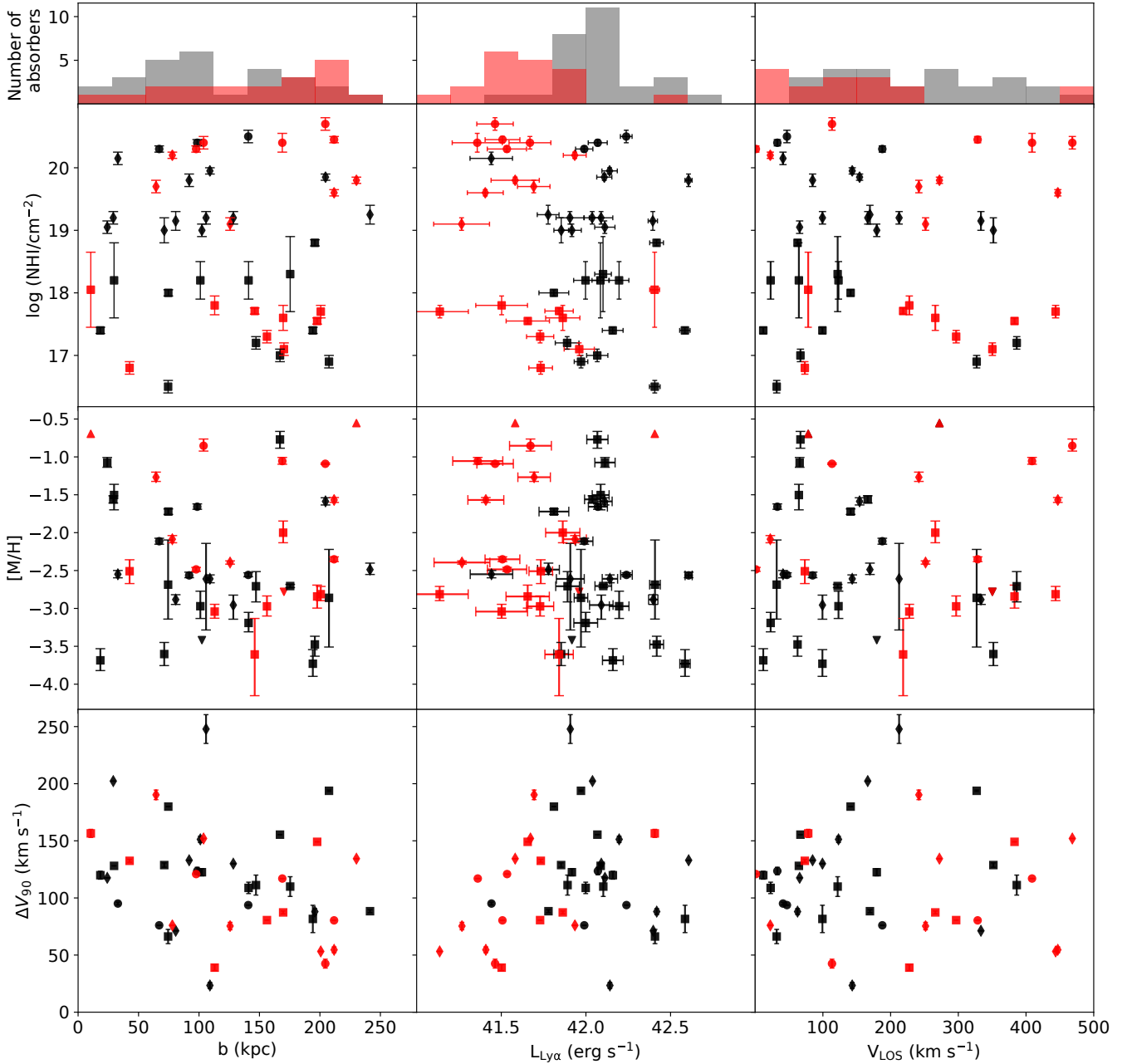
Figure 15. As Fig. 14 but selecting the closest LAE in impact parameter rather than the brightest.

brightest and closest galaxies. When selecting based on the brightest detected galaxy we see a potential absence of sources at very low metallicities and small impact parameters (middle left panel of Fig. 14). This hints at an admittedly weak trend where metallicity increases towards smaller separations, although with a large scatter. This trend, however, vanishes once we consider the galaxies at the closest separation (middle left panel of Fig. 15). Finally, considering the correlation between the velocity width of the absorbers (Δv_{90}) and the impact parameter, we observe no clear correlation between the two properties both when considering the brightest LAE (bottom left panel of Fig. 14) and the closest LAE (bottom left panel of

Fig. 15). We find both small and large velocity widths across the full range of impact parameters up to ≈ 250 kpc.

5.2 Correlations with Lyα luminosity

Next, we move on to investigating the correlations between the absorber properties and the Lyα luminosity of the selected LAE.

The H I column density of the absorbers is shown as a function of LAE luminosity in the top middle panels of Fig. 14 and Fig. 15. As for the impact parameter results, we see no clear correlation between the

H I column density and the galaxy luminosity, both for the brightest and for the closest LAEs. This result remains when we restrict the sample to only the highest confidence LAEs (confidence 1 and 1.5; black) which are typically found at the largest luminosities.

The centre plots of each figure show the relationship between the absorber metallicity and galaxy Lyα luminosity. Similarly to the results for the metallicity versus impact parameter, when selecting the brightest LAE we see a hint of a negative correlation between the two parameters with an absence of systems at low metallicities and low luminosities. However a Kendall's τ test on this data produces a $p-$value of 0.45 which is not sufficiently low to reject the null hypothesis and confirm the presence of a relation. When using the closest LAE instead, we see similar results but a Kendall's τ test on this sample now produces a lower probability (0.15) of there being no correlation between the two parameters.

When comparing the Lyα luminosity with the average Δv_{90} of the absorber's metal lines, as shown in the bottom middle panels of Fig 14, a positive correlation emerges such that we do not see fainter galaxies around absorbers with high Δv_{90} values. This result remains the same when we use both the brightest and the closest LAE. This may be the result of the brightest LAEs tracing the most massive haloes where the velocity dispersion is highest and hence the gas exhibits higher velocity, or a reflection of higher velocities in the CGM of galaxies with higher star formation, as expected for instance due to the presence of outflows.

5.3 Correlations with line of sight velocity

Finally, we focus on the correlation between absorption properties and the line of sight velocity difference between gas and galaxies, where the velocity separation has been systematically adjusted by 245 km s^{-1} to account for the shift in the Lyα due to scattering (see Fig. 8).

Similar to the impact parameter and Lyα luminosity results, we see no clear correlation between the H I column density of the absorption systems and the velocity separation either when considering the full galaxy sample (red) or when restricting only to the highest confidence LAEs (black). When using either brightest LAEs or the closest LAEs (top right panels of Fig. 14 and Fig. 15 respectively), we see absorbers across the full range of column density from LLSs to DLAs at both small and large velocity separations.

When we investigate the relationship between the velocity separation and the metallicity of the absorber, we do not observe any correlation, neither for the case of the brightest galaxy nor for the case of the nearest one (middle right panel of Fig. 14 and Fig. 15 respectively). Hence, while we did see some suggestion of a possible anti-correlation between the metallicity and the impact parameter, this trend is not confirmed by the line of sight velocity separation. Although the line-of-sight velocity is most likely dominated by peculiar motions and radiative transfer effects, we have also completed the exercise of combining the projected impact parameter with the equivalent distance along the line of sight (assuming Hubble flow velocities). Even when considering this three-dimensional distance, no obvious correlations emerge. This is not surprising given the results discussed above and the fact that the three-dimensional distance is, in this case, dominated by the LOS velocity given the small impact parameters we are probing with MUSE.

The last correlation we consider is that between the line width of the absorber (Δv_{90}) with respect to the relative velocity separation along
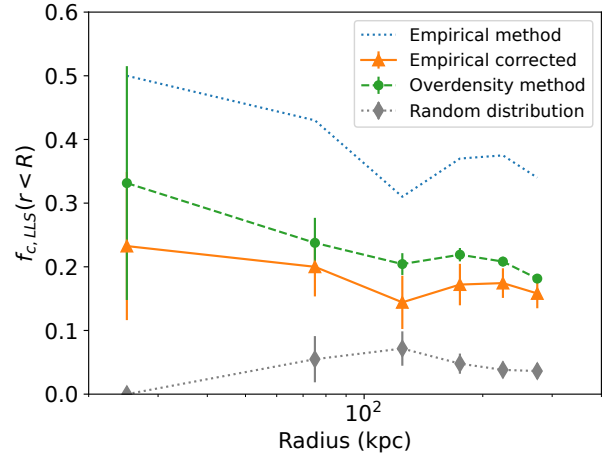
Figure 16. The cumulative covering fraction of H I absorption with $N_{HI} > 10^{17.2}$ cm^{-2} around LAEs detected at $S/N > 7$ as a function of impact parameter from the quasar line of sight. Empirical values from raw data are shown by the blue dotted line, while orange triangles mark the values corrected for the biased selection of LLSs in MAGG. The green circles show instead the covering fraction computed from the overdensity of LAEs, as described in the text. Both estimates are in good agreement. The results obtained for a random distribution of LAEs is also shown (grey diamonds).

the line of sight (bottom right panel of Fig. 14 and Fig. 15). We again do not observe any correlation either for the brightest LAE or for the closest LAE. Indeed, the Kendall's τ test gives probabilities that the two parameters are uncorrelated of ≈ 0.71 when choosing the brightest LAE and ≈ 0.85 for the closest LAE.

6 COVERING FRACTION OF OPTICALLY-THICK GAS AROUND LAEs

Having explored the link between absorbers and galaxies globally and in individual systems, we next move on to investigate, in a statistical sense, the distribution of gas around LAEs by measuring the covering fraction of neutral hydrogen. Empirically, the covering fraction can be defined as:

$$\tilde{f}_c(<R) = \frac{N_{det}(S/N \geq 7; N_{HI} > \bar{N}_{HI}; r < R)}{N_{tot}(S/N \geq 7; r < R)} \quad (4)$$

where N_{det} is the number of galaxies that show absorption above a given H I column density in the quasar spectrum within ± 500 km s^{-1} of their redshift (corrected by 245 km s^{-1} due to the scattering of Lyα) within a radius R; N_{tot} is instead the total number of galaxies in which we could potentially detect gas in absorption in the quasar spectrum if it was present (i.e. those lying in the path-length searchable for ALSs). To calculate this fraction, we need to extend the identification of LAEs beyond the regions occupied by ALSs. To do so, four authors (EKL, MiF, MF, MG) have visually inspected all sources extracted by CUBEX with a $S/N \geq 7$ across the entire MAGG cubes. We refer the readers to the forthcoming paper by Galbiati et al. (in prep) for more details on the full sample of LAEs across the entire MAGG survey. We exclude any systems within 3000 km s^{-1} of the quasar redshift from this analysis so that our results are not affected by the environment around the quasar itself. In this calculation, we further include only the highest redshift LLS of each sightline, as for

optically-thick systems without damping wings we cannot quantify the presence of additional ALSs at lower redshift due to the presence of the Lyman edge.

Given the selection of MAGG sightlines, which favours the presence of $z > 3$ LLSs, the incidence of optically-thick absorbers does not reflect the cosmological value, $\ell(z) \equiv dn/dz$. Indeed, considering the total pathlength to the first LLS in MAGG and the number of ALSs detected, we find $\ell_{\mathrm{MAGG}} \approx 4.3$, which is a factor ≈ 2.15 higher than the cosmological incidence measured at $z \approx 3.5$, equal to $\ell_{\mathrm{cosm}} \approx 2$ (Prochaska et al. 2010; Fumagalli et al. 2020). Due to this, the average pathlength to the first LLS in MAGG is ≈ 2.15 times shorter than what found in an unbiased survey. In turn, this shorter pathlength boosts the empirical estimate of the covering factor by the same amount, as it decreases the number of LAEs uncorrelated to LLSs by the ratio of the true to measured pathlength. Therefore, Equation 4 needs to be corrected by a factor $\xi = \ell_{\mathrm{MAGG}}/\ell_{\mathrm{cosm}}$ to yield an accurate estimate of the true covering fraction, $f_c = \tilde{f}_c/\xi$. In Figure 16, we show the cumulative covering fraction as a function of impact parameter for the $N_{\mathrm{HI}} \geq 10^{17.2}$ cm^{-2} systems, before (blue dotted line) and after (orange triangles) the correction factor ξ is applied. The error bars show the 1σ Wilson score confidence intervals. We note that a non-negligible uncertainty, up to a factor ≈ 2, exists in the literature for ℓ_{cosm} at $z \approx 3.5$ (see a discussion in Fumagalli et al. 2020). Here, as noted, we assume a value of $\ell_{\mathrm{cosm}} \approx 2$ which brackets current estimates; however, values both below and above ≈ 2 have been reported in the literature. As such, the covering fraction is subject to this uncertainty, and readers wishing to make a different assumption on the incidence of LLSs can simply rescale the values presented here by the factor ξ defined above.

Using MAGG data, we can derive a second estimate of the covering fraction based on the overdensity of LAEs near LLSs compared to the field, following the formalism outlined next. Starting from Equation 4, we can write

$$N_{\mathrm{det}}(< R) = \delta(< R)\rho_{\mathrm{LAE}}A_{\mathrm{FOV}}\Delta z N_{\mathrm{LLS}} , \qquad (5)$$

where $\delta(< R)$ is the overdensity of LAEs compared to the mean density ρ_{LAE} within an impact parameter R, A_{FOV} is the FOV area, Δz is the search window for LAEs near LLSs (± 500 km s^{-1}), and N_{LLS} is the number of LLSs detected. Moreover,

$$N_{\mathrm{tot}}(< R) = \rho_{\mathrm{LAE}}A_{\mathrm{FOV}}\Delta z_{\mathrm{tot}} , \qquad (6)$$

where Δz_{tot} is the total pathlength to the first LLS in MAGG. Combining, we obtain that

$$\tilde{f}_c(< R) = \frac{\delta(< R)\Delta z N_{\mathrm{LLS}}}{\Delta z_{\mathrm{tot}}} = \delta(< R)\Delta z \ell_{\mathrm{MAGG}} . \qquad (7)$$

Hence, the unbiased estimator becomes

$$f_c(< R) = \delta(< R)\Delta z \ell_{\mathrm{cosm}} , \qquad (8)$$

where we have replaced the incidence of LLSs from MAGG with the cosmological value. Hence, the covering fraction can be evaluated simply by measuring $\delta(< R)$ from MAGG data. For this, we integrate the distribution of LAEs detected within ± 500 km s^{-1} of LLSs, dividing by the random expectation in bins of radius (see Fig. 7). The resulting covering fraction based on the overdensity measure is shown by green circles in Fig. 16. The agreement with the empirical estimate is quite good, also considering the uncertainties.

The covering fraction profile for LLSs shows a mildly decreasing trend with impact parameter, reaching a maximum covering fraction of $f_c \approx 0.3$ for < 50 kpc, dropping down to just below $f_c \approx 0.2$ at ≈ 300 kpc. The covering fraction also appears to have an inflection
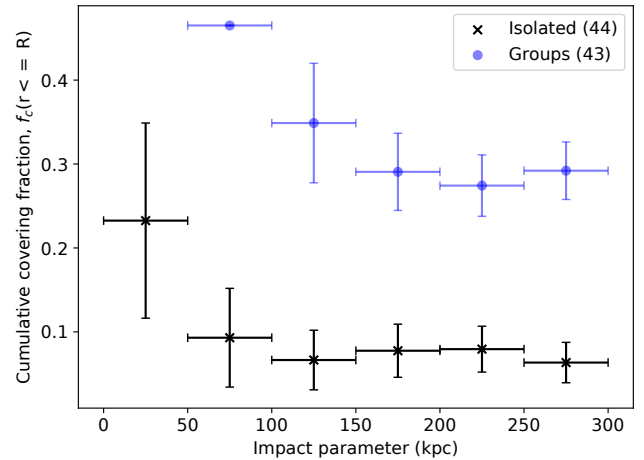
Figure 17. The cumulative corrected covering fraction as a function of impact parameter for $N_{\mathrm{HI}} \geq 10^{17.2}$ cm^{-2} absorbers around LAEs detected at $S/N > 7$. The sample is split into isolated LAEs (black crosses) and LAEs in groups of 2 or more (blue circles).

point near ≈ 100 kpc, a feature that is visible both in the empirical estimate and in the one based on overdensity. As discussed below, we interpret this dip as a transition between the signal arising from the CGM of individual LAEs to the larger scale environment. For comparison, we include the results expected from a random distribution of LAEs (grey). This was obtained in the same manner as the $N_{\mathrm{HI}} > 10^{17.2}$cm^{-2} sample but the redshifts of the LAEs were randomly shuffled before performing the calculation. For this distribution we find lower covering fractions than for the actual sample.

In Fig. 17, we repeat the measure of the covering fraction for $N_{\mathrm{HI}} > 10^{17.2}$ cm^{-2} absorbers, but splitting the sample into 'isolated' LAEs and LAEs in 'groups'. LAEs within groups are defined as those that have at least one other LAE within ± 500 km s^{-1}. Using this selection we see a clear difference between the covering fractions for isolated and group galaxies, with the latter having a higher covering fraction at all separations. For example, at separations < 300 kpc the covering fraction for isolated galaxies is only $f_c = 0.08 \pm 0.02$ while for groups it is $f_c = 0.29 \pm 0.03$. At smaller impact parameters the covering fractions increase for both samples with f_c reaching 0.23 ± 0.11 for isolated galaxies at impact parameters < 50 kpc. For groups of galaxies, the covering fraction approaches 50 per cent at < 100 kpc.

With the covering fraction in hand, we can also revisit the argument by Gallego et al. (2021), who presented a method to constrain the amplitude of the ultraviolet background (UVB) background at $z > 3$ starting from a measurement of the Lyα fluorescence in optically-thick clouds. In their original analysis, however, the LLS covering fraction was not known, hence they relied on results from numerical simulations. With our new measurement, we can independently verify their assumptions on the covering fractions. For this, we repeat the analysis of the covering fraction starting from a measurement of the overdensity using the HI column density cut of $N_{\mathrm{HI}} > 10^{17.5}$cm^{-2} and the same radial binning as in Gallego et al. (2021), focusing in particular on the interval $8 - 20$ arcsec ($60 - 150$ kpc), where we have the most constraining power from MAGG data. In this interval, we find a covering factor of ≈ 17 per cent, again assuming $\ell_{\mathrm{cosm}} \approx 2$. In this bin, Gallego et al. (2021) estimated with simula-

tions $f_c = 0.083 \pm 0.002$ for $z \approx 3$ and $f_c = 0.106 \pm 0.004$ for $z \approx 4$, which is a factor $\lesssim 2$ smaller than estimated in MAGG. This difference is, however, not particularly significant at the current precision with which ℓ_{cosm} is measured. At face value, our determination of the covering fraction combined with the fluorescence measure in Gallego et al. (2021) would yield an estimate of the UVB in good agreement with the Haardt & Madau (2001) UVB model, which is slightly lower than what is inferred by Gallego et al. (2021) using simulations.

7 DISCUSSION

Through a dedicated survey of galaxies around optically-thick absorbers, this analysis of the MAGG data provides the most complete determination of the UV-selected galaxy population surrounding partially neutral gas regions at $z \sim 3-4$, including UV-bright LBGs and, as a particularly novel aspect, LAEs with $L_{Ly\alpha} \gtrsim 10^{41}$ erg s^{-1}. Our in-depth study of the environment of LLSs reveals that: i) a fraction of LLSs ranging from 57 to 82 per cent (depending on the confidence of detection) is near one or more galaxies, the near-totality of which is an LAE; ii) there is solid evidence of a physical association between the LAEs and the LLSs based on the clustering in velocity and space, and on the normalization of the luminosity function; iii) the bulk of these associations is found at somewhat large distances from the absorbing gas, with impact parameters ranging from 50 to 250 kpc, and along filaments connecting LAEs and LLSs; iv) there is no evident correlation between the LLS absorption and the LAE emission properties, except when considering the absorber line widths and the Lyα luminosity. Building on these pieces of evidence, and in comparison with literature studies, we can now tackle a series of questions relating to the nature and origin of high-redshift LLSs and to what extent LAE-LLS associations provide information on the role of the baryon cycle at high redshift.

7.1 A rising detection rate of galaxies around H I absorbers

Strong links between LLSs or DLAs and galaxies at $z \gtrsim 2$ have been documented in the literature for some time now. The pioneering searches of galaxies associated with DLAs using traditional spectrographs (Fynbo et al. 2010; Krogager et al. 2017) have been some of the most successful efforts to build a statistical sample of galaxies associated to DLAs before the deployment of MUSE. This sample counted however only ≈ 10 associations. The advent of large-format IFUs and especially MUSE has accelerated the rate of discovery near both DLAs and LLSs, following in the footsteps of programmes with smaller IFUs such as SINFONI (Peroux et al. 2012). Indeed, early works based on MUSE often detected galaxies within the local environment of absorbers (e.g. Fumagalli et al. 2016; Mackenzie et al. 2019; Lofthouse et al. 2020), an effort that culminated in the full sample of MAGG where we find a high detection rate of galaxies around absorbers of ≈ 82 per cent. This is higher than previous surveys with SINFONI, e.g., by Peroux et al. (2012) who quote detection rates ranging from ≈ 14 to ≈ 37 per cent. The main factor behind this increase is the larger impact parameter probed by MUSE around absorbers compared with what is possible in SINFONI (hundreds of kpc versus tens of kpc). Indeed, the majority, ≈ 79 per cent, of the LAEs in the MAGG sample are found at impact parameters above 100 kpc. Another factor is the increased sensitivity to faint emission in the optical with MUSE compared to near-IR detections with SINFONI which improves the detection efficiency. Additionally, Lyα

is intrinsically brighter than Hα emission in systems with low-dust obscuration. When comparing to results from similar studies using MUSE, and hence similar search volumes albeit with smaller samples of absorbers, we find similar results such as Mackenzie et al. (2019) who achieve a detection rate of ~80 per cent from a study of six DLAs.

Despite the fact that the detection of galaxies near absorbers is now becoming the norm, there are still examples both in MAGG and in the literature of systems with no detectable galaxies within hundreds of kpc from the absorbing gas (e.g. Fumagalli et al. 2016). An obvious reason is related to the intrinsic bias of MUSE towards less dusty galaxies and galaxies that are line emitters compared to UV bright galaxies with no emission lines. Both ALMA surveys and targeted spectroscopic searches sensitive to absorption lines indeed provide examples of these types of associations. Moreover, especially in LLSs, the bright glare of the background quasar still affects the detectability of faint sources near $\approx 10-15$ kpc where PSF residuals are non-negligible.

Nevertheless, comparing the variation of the rate of non detections in confidence > 2 versus confidence ≤ 2 systems (≈ 18 per cent in all LAEs and ≈ 40 per cent for high-confidence ones; Fig. 7), it is clear that the fraction of detections rapidly rises with increasing sensitivity, as expected given the shape of the luminosity function. Deeper surveys have therefore the potential to reach detection rates approaching 100 per cent. Moreover, as argued in the next section, there is evidence that a significant percentage if not the entirety of the optically-thick gas observed at $z \approx 3-4$ can be associated with LAEs of comparable luminosity to the ones observed by MAGG within a radius of ≈ 150 kpc. This implies that, despite the completeness issues noted above, LAEs have a sufficient number density to be very often if not always located in proximity to optically-thick gas resulting in associations. We therefore predict that, in virtue of the clustering of other types of galaxies (e.g. more massive systems) with LAEs, there should be a significant overlap in associations between, e.g. dusty or UV bright galaxies and LAEs. Indeed, already within MAGG and as already reported by other studies, we see instances of multiple detections near a single LLS, with up to eight LAEs detected around one absorber. A combination of multi-wavelength and deep surveys of fields hosting LLSs is therefore likely to reveal a progressively higher number of multiple associations composed of systems which are detected at different wavelengths.

This rapid increase of detections made possible by MUSE and the recognition that optically-thick absorbers could trace either the CGM of galaxies clustered to other systems or even large-scale structures shared among multiple galaxies (see the next section) prompts us to move away from the traditional one–to–one associations between galaxies and absorbers, and instead consider these absorbers as residing in more complex galaxy environments. The direct consequence is that the interpretation of signatures of the baryon cycle (e.g. inflows or outflows) encoded in ALSs is made more complex, discouraging straightforward conclusions which could lead to incorrect or conflicting inferences from data. To fully exploit the constraining power of gas-galaxy associations for the baryon cycle it is therefore necessary to resort to more statistical analysis, requiring detailed numerical models or even fully cosmological hydrodynamic simulations (see Mackenzie et al. 2019, for an example). Thankfully, the increasing number of galaxy-absorber associations makes this approach more realistic compared to what is possible in previous surveys.

7.2 The nature of optically-thick gas

The nature of optically-thick gas giving rise to high-redshift LLSs has been long studied in the literature, particularly through the statistical analysis of LLSs and numerical simulations. Globally, the incidence (e.g. Sargent, Steidel, & Boksenberg 1989; Prochaska et al. 2010; Fumagalli et al. 2020) and the detailed absorption properties of LLSs such as density and metallicity (e.g. Steidel 1990; Prochaska 1999; Fumagalli et al. 2013), place these optically-thick absorbers in overdensities comparable to what is found in the virial regions of galaxies. Such a conclusion finds general agreement with the results of numerical simulations, which place a non-negligible fraction of LLSs in close proximity to (or even within) the halo of high-redshift galaxies (e.g. Faucher-Giguère et al. 2011; Fumagalli et al. 2011; Van de Voort et al. 2012). However, the picture is complicated by the fact that, as the neutral fraction of the IGM increases with redshift, progressively more systems are expected to arise from intergalactic regions (McQuinn, Oh, & Faucher-Giguère 2011; Fumagalli et al. 2013). Leveraging the blind survey of galaxies made possible by MAGG, we can revisit the question of where LLSs reside with respect to galaxies.

Being H I-selected systems, we consider in turn three scenarios for the origin of LLSs: the ISM, the CGM, and the IGM. Starting from the ISM, we note that the majority of the associations lie at projected impact parameters $\gtrsim 50$ kpc, larger than the expected extent of the gas "disks" at $z > 3$ (Rafelski et al. 2016) also considering some of the most extreme cases known (Neeleman et al. 2020; Cicone et al. 2015). A question arises, however, on possible selection biases due to the presence of the quasar that hampers the visibility of the innermost ≈ 1 arcsec (or $\approx 7 - 8$ kpc). We exclude this possibility on three grounds. First, Fig. 9 does not suggest the presence of a truncated distribution with detections piling up close to the putative limit at which we would start missing sources. Second, when stacking spectra extracted from circular apertures of $1 - 2$ arcsec at the quasar locations in the rest-frame of the absorbers, we do not observe evidence of Lyα emission which would arise from a large population of bright LAEs on top of quasars. Third, we inspect individual narrow-band images extracted from the cube at the redshift of the absorbers (also accounting for the typical shift of Lyα) as well as stepping through each wavelength slice of the cube around the wavelengths of the absorbers searching for emission near the quasar. Through this visual inspection we do not find any additional sources that were missed by our CubEx search. We also note that there is minimal (if any) contribution of the quasar residual in the absorption trough of DLAs. Such systems close to quasars have been previously identified in connection with extremely strong H I absorption systems (Ranjan et al. 2020), which are however absent in MAGG due to the intrinsic rarity of systems with $\log(N_{\rm HI}/{\rm cm}^{-2}) > 21.6$.

Disentangling the contribution of the CGM and IGM to the origin of LLSs near LAEs is less trivial, due to the fact there is no well defined transition marking the boundary between these two gas components. Following a similar argument to the one proposed for the ISM, however, we can rule out the inner CGM as the main contributor. Indeed, the clustering analysis of spectroscopic samples with comparable luminosity suggests a halo mass for LAEs that approaches $\approx 10^{11}$ M$_\odot$ (Herrero Alonso et al. 2021). The virial radius of such a halo at $z \approx 3$ is ≈ 35 kpc, implying that the bulk of the associations arise from beyond twice the virial radius. This seems to disfavour the inner CGM of LAEs as a primary source of optically-thick gas, but still leaves circumgalactic regions extending at $3 - 4$ times the virial radius as plausible sites where LLSs can be found. At these distances, the di-

vide between the CGM and IGM blurs. As argued in Section 4.4, there is also evidence of filaments connecting multiple LAEs and the LLSs. Møller & Fynbo (2001) provided one of the earliest detections of such structures by mapping out a string of galaxies which covered almost 5 Mpc (see also Fumagalli et al. 2016b; Mackenzie et al. 2019). In more recent years, we are now also able to see these filaments directly through Lyα emission. While these observations are difficult due to the intrinsic faintness of such emission, progress was made by observing regions where there are sources of ionizing photons which locally enhance the Lyα line. For example, Cantalupo et al. (2014) observed a 460 kpc cosmic web filament in Lyα which has been illuminated by emission from a quasar at $z \sim 2.3$. In Umehata (2019), the authors observed SSA22, a protocluster at $z = 3.1$, and reported the detection of a 1.3 Mpc filament with several star-forming galaxies embedded in it. Bacon et al. (2021) used the MUSE Extremely deep field, a 140h MUSE program in the Hubble Ultra Deep Field, to detect extended Lyα emission from the cosmic web in proximity to AGNs beyond the extreme environments probed in previous studies. Therefore, combining our analysis and mounting evidence in the literature of filaments connecting galaxies, we propose that a significant fraction of optically-thick gas could indeed arise from gas inside these IGM structures. Finally, the fraction of LLSs without associated LAEs within ≈ 250 kpc is sufficiently low to exclude an abundant population of LLSs arising in IGM filaments far from galaxies. Furthermore, no dependence on redshift is found for the fraction of LLSs associated with galaxies. Therefore we exclude the idea of an increased contribution from IGM regions as redshift increases, although with large uncertainties due to the limited sample.

We must also acknowledge that – despite the very complete survey enabled by MUSE – our selection is still biased in favour of line emitters, so galaxies lacking prominent Lyα emission are underrepresented in our study. Hence, we cannot derive an absolute statement on the origin of LLSs with these data alone. However, while the fraction of star-forming galaxies emitting LAEs is not particularly high ($\lesssim 20$ per cent for comparable samples; see Kerutt et al. 2022), there should be no preference in selecting LAEs at large separations with respect to the position of LLSs. Therefore, we expect that the distribution of impact parameters is representative of the underlying LAE population. Therefore, relative to these sources, we can affirm that LLSs originate primarily either from the outer CGM (defined here as $3 - 4$ $R_{\rm vir}$) or in the IGM but in close proximity to galaxies.

The search for galaxies near LLSs in MAGG provides a direct way to address independently the question on the incidence of optically-thick gas within extended Lyα halos surrounding galaxies. Deep long-slit and IFU observations have indeed frequently uncovered extended halos near Lyα emitters with luminosity comparable to the one selected by our survey (Rauch et al. 2008; Wisotzki et al. 2018). Thanks to these studies, it was possible to compare the redshift incidence of Lyα emitting regions with the incidence of optically-thick gas derived from absorption spectroscopy. Based on the number of LAEs and the extent of their Lyα halos, these authors found that LLSs have a comparable incidence to halos above a surface brightness of 10^{37} erg s^{-1} kpc^{-2} out to $\approx 50 - 70$ kpc (10^{38} erg s^{-1} kpc^{-2} out to $\approx 20 - 30$ kpc for DLAs). At face value, this correspondence would imply that a large fraction of LLSs (DLAs) arise from within ≈ 2 (≈ 1) times the virial radius of LAEs with $L_{\rm Ly\alpha} \gtrsim 10^{41}$ erg s^{-1}. By performing the opposite experiment (i.e., by searching for galaxies near LLSs), our MAGG survey places stringent constraints on this inference and, based on the findings discussed above, rules out a unique correspondence between the innermost part of halos and the distribution of optically-thick gas. Indeed, we have shown that LLSs
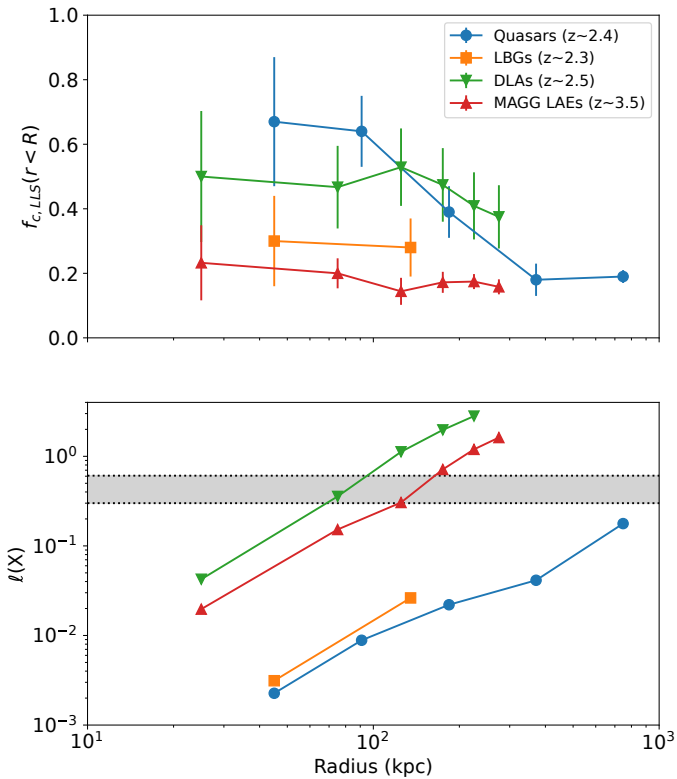
Figure 18. Top: Comparison between the covering fraction of optically-thick hydrogen around LAEs (red upward triangles) from this work, quasars (blue circles) from Prochaska et al. (2013), LBGs (orange squares) from Rudie et al. (2012), and DLAs (green downward triangles) from Rubin et al. (2015). Bottom: The contribution to the cosmic incidence of $\tau \gtrsim 2$ LLSs from the optically-thick gas detected around different populations (colour-coded as above) as a function of radius. The LLS incidence measured in quasar absorption line surveys is shown by the grey band, where the lower limit is the $z \approx 2$ value from O'Meara et al. (2013) and the upper limit is the $z \approx 4$ value from Prochaska et al. (2010). Due to their high number density, LAEs account for the near totality of LLSs by ≈ 150 kpc, and are thus overlapping with the DLA population.

are rarely found at such close proximity from LAEs. We conclude that comparisons between emission and absorption properties need to account for a more complicated distribution of optically-thick gas that extends to larger radii with covering factors much lower than unity.

We can further extend this investigation into the nature of optically-thick gas by combining additional tracers known from the literature, such as LBGs (e.g. Rudie et al. 2012), DLAs (e.g. Rubin et al. 2015), and quasars (Prochaska et al. 2013). This comparison in shown in Fig. 18. For covering fractions that are presented in the literature as differential values (i.e., in annuli of impact parameter where $R_1 < r < R_2$), we have converted the values to cumulative covering fractions (i.e., $r < R$) in order to directly compare them with the cumulative values used in MAGG. Considering the covering fraction, quasars are by far the population with the most extended optically-thick gas

in their surroundings, with elevated values up to ≈ 100 kpc which drop to lower but still considerable values up to 1 Mpc. Prochaska et al. (2013) attribute this rapid decline to a boundary in the quasar CGM and also note that, due the number of ambiguous cases in which they could not disentangle between optically-thick and optically-thin hydrogen based on partial coverage of the Lyman series, these values are believed to be lower limits. LAEs from our study exhibit a similar behaviour, but with a clear shift to smaller radii and with intrinsically lower values. Similarly to the interpretation of the covering fraction for quasars, we attribute the enhanced values at $\lesssim 60 - 70$ kpc to the contribution of the inner CGM of LAEs (within twice their virial radius), with a further contribution at larger distances arsing from the outer CGM and the IGM connecting LAEs, as already noted above. LBGs appear to show, in comparison to LAEs, a similar covering fraction but with a flatter radial dependence. It should be noted however that the survey by Rudie et al. (2012) has only a very limited number of LLSs, and hence this determination is the one that is subject to the largest uncertainties.

DLAs show instead an elevated covering fraction, but a peculiar radial dependence compared to the other populations. However, similar to the quasar case, this measurement suffers from ambiguous determinations of optically-thick versus optically-thin systems and hence these covering fractions should be, too, considered upper limits. Moreover, the measurement of the covering fraction near DLAs is conceptually different from the other three cases presented in this figure. DLAs arise from gas that is believed to be distributed around galaxies, and therefore the radial distance from a DLA is different from the one measured from galaxies. More explicitly, assuming for the sake of argument that DLAs are identically associated to LAEs, one should not expect the same measurement of covering fraction. Indeed, while the LAEs are expected always at the centre of the halo, DLAs are themselves distributed inside the halos and hence the covering fraction of DLAs is a convolution of "off-centre" measurements of the hydrogen distribution (see Fumagalli et al. 2014, for an explicit example).

Besides the direct comparisons of the covering fraction, insight into the nature of optically-thick gas and its relation to galaxies can be obtained by considering the contribution of each population to the total incidence of LLSs measured in quasar spectroscopic surveys (i.e., the number of LLSs per unit redshift $\ell(z)$). As commonly assumed in the literature (see e.g. Steidel et al. 2010; Prochaska et al. 2013; Fumagalli et al. 2013), knowing the covering fraction of optically-thick gas as a function of radius in a population of known number density $n_{\rm com}$, one can estimate the incidence of LLSs arising at a given radius from that population by computing the product of the effective cross section of optically-thick gas and the population number density

$$\ell(X, r < R) = \frac{c}{H_0} f_c(r < R) \pi R^2 n_{\rm com} . \qquad (9)$$

Here, the cross section is computed in physical units while the number density is in comoving units, resulting in a measure of the incidence in terms of the absorption distance

$$dX = dz \frac{(1+z)^2}{\sqrt{\Omega_{\rm m}(1+z)^3 + \Omega_\Lambda}} . \qquad (10)$$

The results are shown in the bottom panel of Fig. 18, and compared with the incidence of LLSs from absorption spectroscopy at $z \approx 2$ (O'Meara et al. 2013) and at $z \approx 4$ (Prochaska et al. 2010).

Despite their elevated covering fraction, due to their moderate number density ($n \approx 1.2 \times 10^{-4}$ Mpc^{-3} for halos with $M \gtrsim 10^{12.5}$ M$_\odot$),

quasars contribute to a minimal fraction of LLSs to within a few hundred kpc. However, when measured at large distances, their covering fraction rapidly rises to a substantial contribution of the LLS population, reaching ≈ 70 per cent at ≈ 1 Mpc. Similarly, the combination of somewhat modest covering fraction and number density ($n \approx 3.7 \times 10^{-4}$ Mpc^{-3} for halos with $M \gtrsim 10^{12}$ M$_\odot$) makes LBGs a comparably minor contribution of LLSs. The opposite behaviour is seen instead for both LAEs and DLAs: despite their covering fraction being almost a factor of $\approx 2 - 3.5$ lower than quasars at fixed radius, their number density is $n \approx 10^{-2}$ Mpc^{-3}, which is obtained for LAEs by integrating the luminosity function calculated in the MAGG survey (Galbiati et al., in prep) for $L_{\mathrm{Ly}\alpha} > 10^{41.5}$ erg s^{-1} and for DLAs integrating the halo mass function above 10^{11} M$_\odot$. In both cases, the high number density is sufficient to account for the entire LLS population already between ≈ 80 kpc and ≈ 150 kpc. The comparable number density and the fact that both DLAs and LAEs produce a similar contribution to LLSs strengthen the argument presented already by, e.g., Rauch et al. (2008) and Wisotzki et al. (2018) that LAEs account for the main DLA population[5].

Two further considerations are warranted at this point. First, it may appear surprising that either DLAs and LAEs overpredict the observed incidence of LLSs beyond $\approx 100 - 170$ kpc. In fact, there is reason to believe that these estimates are uncertain to at least a factor of two for several reasons: i) the minimum luminosity or halo mass function assumed for the calculation of the number density is subject to significant uncertainty especially for DLAs, but the resulting number density, n, is a steep function of this cutoff; ii) the determination of the abundance of LLSs is in itself uncertain (see a discussion in Fumagalli et al. 2020), with repercussions on the the precision with which we can estimate the covering fraction in MAGG; iii) the calculation in Equation 9 is rather crude as it attributes average binned values of covering fraction to all halos regardless of any underlying mass dependence. Furthermore and most importantly, this calculation breaks down at large radii due to the clustering of multiple LAEs near LLSs. If optically-thick gas arises from an astrophysical structure engulfing multiple LAEs, the number density of LAEs should be replaced by the number density of underlying structures giving rise to the absorption. Hence, the use of the number density of LAEs leads to an overestimate of the incidence. If DLAs are linked to LAEs as argued above, a similar consideration applies to the case of DLAs. The second consideration stems from the fact that LAEs (hence DLAs) account for all LLSs by $\approx 80-150$ kpc but that quasars themselves contribute with equal importance to the incidence out to ≈ 1 Mpc. This apparent contradiction can be reconciled by the fact that, at such large distances from quasars, one is likely to be probing the covering fraction of populations clustered to quasars themselves. For example, the analysis of MAGG data in proximity of quasars reveals a strong clustering of LAEs up to ≈ 0.5 Mpc (Fossati et al. 2021, see also Fig. 10). Hence, it is reasonable to believe that LAEs and quasars (at large distances) are not independent populations, and the two contributions shown in Fig. 18 should not be added together but rather regarded as arising from the same astrophysical structures.

In summary, based on the following three pieces of evidence revealed by the MAGG survey we can add new constraints to the nature of optically-thick gas in the Universe: i) LAEs strongly cluster near LLSs but with associations extending to beyond twice the typical

virial radius; ii) there is some degree of evidence of an alignment of multiple LAEs in filaments where LLSs are found; iii) LAEs easily account for the near totality of LLSs. In light of this, MAGG confirms previous claims that LAEs are indeed at the origin of the LLSs seen in absorption spectra, but extends this picture put forward by previous studies, shifting the importance from the inner CGM of individual LAEs to the filamentary cosmic structures that connect multiple LAEs and host a substantial fraction of these LLSs.

7.3 Environmental effects on the gas distribution

Following pioneering work based on multi-object spectroscopy (e.g., Bordoloi et al. 2011; Nielsen et al. 2018), surveys at high density enabled by MUSE have opened up a new window into exploring links between the density of galaxies and the properties of the surrounding gas. Several authors have reported at various redshifts a marked difference in the gas distribution near galaxies in pairs or groups versus those in isolation. The work by Fossati et al. (2019) is among the first studies with MUSE that reports an increase in the strength and extent of Mg II in $z \approx 1$ group galaxies, a trend later confirmed with MAGG by Dutta et al. (2020) and, more systematically and with a larger sample, by Dutta et al. (2021). At these redshifts, the authors suggest that environmental effects act as a perturber of the galaxy CGM within groups, displacing cool and enriched material at larger distances compared to more isolated galaxies. At higher redshift, excess in the optical depth of both hydrogen and C IV near LAEs in groups compared to more isolated ones at $z \approx 3$ has been reported by Muzahid et al. (2020). Similarly, Galbiati et al. (in prep.) have found using MAGG data an excess in the covering factor of C IV of $z \approx 3 - 4$ LAEs in groups compared to field galaxies.

Combining the information gathered by our analysis of the MAGG fields, we can further add to the study of Muzahid et al. (2020) and Galbiati et al. through a detailed exploration of the correlations between LAEs and LLSs. Our analysis reveals that $\approx 30 - 40\%$ of LLSs are associated with multiple galaxies (Fig. 7, depending on the confidence level considered), suggesting that a significant fraction of optically-thick clouds live near multiple LAEs, especially considering that our survey is incomplete at low luminosity and likely biased against obscured and non line-emitting systems. When examining closely whether the detailed absorption properties (column density, metallicity, kinematics; Fig. 11) vary systematically with galaxy environment, however, there does not appear to be any noticeable variation in the absorption properties as a function of the galaxy environment. Conversely, in line with what is observed by Muzahid et al. (2020) and Galbiati et al., we see a much elevated covering factor of optically-thick gas at all the impact parameters probed once considering galaxies that are in pairs or multiples compared to isolated ones (Fig. 17).

We therefore conclude that the galaxy environment plays an active role in increasing the distribution of optically-thick (and metal enriched) gas near LAEs, although there is no individual connection between the line-of-sight properties measured in absorption and the environment. These pieces of observational evidence can find an explanation in light of the above discussion on the origin of optically-thick gas near LAEs. In the previous section, we argued that the majority of LLSs arises from the outer CGM and IGM along filamentary structures connecting LAEs. In this scenario, it is reasonable to expect an enhanced cross-section of optically-thick gas in the presence of multiple LAEs which trace regions at the intersection of multiple filaments and more massive nodes. This scenario would

---

[5] In Fig. 18, we used the empirical estimate of covering fraction from MAGG. Assuming the one based on overdensity would lead to an even greater contribution to LLSs at all radii.
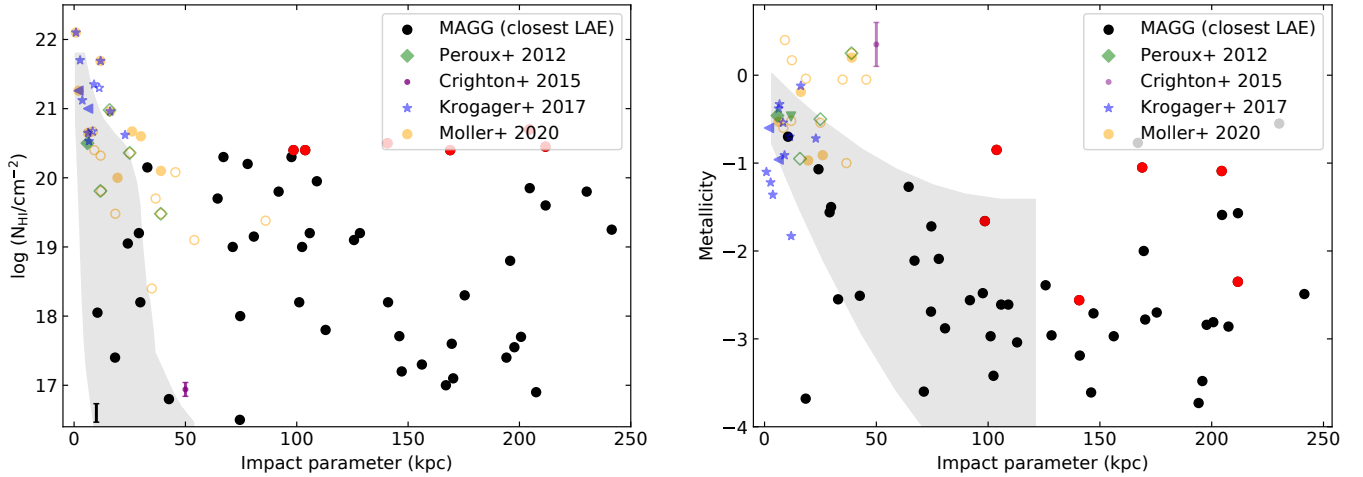
Figure 19. Comparisons between the MAGG sample and literature values. *Left:* H I column density of absorbers versus the impact parameter to the closest galaxy. We include both observational and theoretical results from the literature for comparison. The typical error on the MAGG results is shown in the lower left corner. *Right:* Comparison of the absorber metallicity as a function of impact parameter. For both plots, we show the MAGG results as dots (red for DLAs, black for LLSs). Other literature results are shown with the green diamonds (DLA sample from Peroux et al. (2012)), blue stars (Krogager et al. 2017), purple dot (Crighton et al. 2015) and orange circles (Møller & Christensen 2020). Absorbers at $z < 2$ are shown with open markers while those at $z \gtrsim 2$ are shown with filled markers. The grey shaded region shows the extent of H I systems drawn from Rahmati & Schaye (2014) at $z \sim 3$ (left panel) and from the FOGGIE simulation (Lehner et al. 2021) at $2.2 < z < 3.6$ (right panel).

also explain the more enhanced C IV distribution, provided that the gas within these structures is enriched, e.g., by outflows from galaxies as argued in some simulations (e.g., Shen et al. 2012). Due to the local nature of quasar absorption spectroscopy and the fact that the hydrogen and metal distribution are expected to be non uniform (see below for additional discussion), there is no reason to expect a tight correlation between column density, metallicity, kinematics and the number of LAEs detected.

This hypothesis for the origin of an enhanced gas distribution of LAEs relies primarily on the large-scale structures within which galaxies resides, and thus differs from the hypothesis put forward at lower redshift to explain the cool gas distribution in groups, where environmental effects such gravitational and hydrodynamic interactions are deemed responsible for pushing cool gas at larger distances. At present, there is only limited possibility to accurately test this hypothesis with data, and therefore future efforts should focus on assessing the plausibility of this interpretation within cosmological simulations (e.g., Nelson et al. 2021) and with new deep observations that can reveal, e.g., perturbations in the gas and stellar content.

7.4 A patchy gas distribution near LAEs

The study of scaling relations involving, on the one side, absorption properties derived from ALSs and, on the other side, emission properties from associated galaxies offers the potential to probe the baryon processes relevant for galaxy evolution on the gas surrounding galaxies inside the CGM. To this aim, particular attention has been given to the variation of hydrogen column density and metallicity as a function of distance from galaxies. Besides being quantities that can be easily computed also at high redshift, their relevance stems from the fact that both the hydrogen and metallicity distributions contain the signatures induced by accretion and feedback processes, as often suggested by simulations (e.g. Hummels et al. 2013).

Two complementary approaches have been followed in the literature

to pursue this investigation. One is the so-called 'galactocentric' approach, in which samples of galaxies are selected along the line of sight of quasars and then their CGM is studied in absorption (e.g. Steidel et al. 2010; Rudie et al. 2012). The other approach, which we follow primarily here, starts instead with samples selected based on a set of absorption properties and then proceeds to identify associated galaxies. With respect to H I selected ALSs, at $z \gtrsim 2$, the latter approach has been primarily exploited for DLAs in particular using X-Shooter observations at multiple orientations (Fynbo et al. 2010) or the SINFONI IFU (Peroux et al. 2012). Before our study however, only very small samples of lower column density systems have been examined. Thus, MAGG provides the largest sample for this kind of analysis to date.

In Fig. 19, we focus on the two main scaling relations explored at high redshift: the dependence of H I column density and metallicity as a function of projected distance from galaxies. We consider both the main DLA samples available in the literature and the LLS sample from MAGG. Two main features stand out. First, literature samples almost exclusively occupy the innermost 50 kpc, while MAGG data span a larger range of impact parameters. As already noted, this effects arises from technical reasons with X-Shooter and SINFONI probing most effectively or exclusively the inner $\lesssim 10$ arcsec near the quasars. The second feature is that, combined, the various samples suggest a generally decreasing trend of both hydrogen column density and metallicity as a function of distance from galaxies up to $\approx 30 - 60$ kpc, at which point the relations flatten and are dominated by a very large scatter, with column densities spanning a range of over three decades and metallicities covering over two orders of magnitude at fixed impact parameter. A large scatter is also evident at small impact parameters, where H I column densities and metallicities are on average more elevated but with data points covering almost 2 dex in hydrogen column density and a decade in metallicity also for impact parameters of $\lesssim 50$ kpc.

The existence of such a large scatter is not surprising in light of our current appreciation of the structure of the CGM and IGM. Obser-

vations against multiple lines of sight reveal at all redshifts large variations in the hydrogen and metal content on distances of few kiloparsecs, or even tens of parsecs (e.g. Rauch, Sargent, & Barlow 2001; Lopez et al. 2018; De Cia et al. 2021; Bordoloi et al. 2022). Moreover, according to the picture painted above, LAEs in groups or along filaments extending for hundreds of kiloparsecs are likely to experience even larger scatter not only due to the patchy distribution of the CGM/IGM, but also to the superposition of different structures along the line of sight (e.g. Bordoloi et al. 2011; Dutta et al. 2020). Nevertheless, a general systematic trend is still expected provided that sufficient samples are available to unveil robust mean values over distributions of large intrinsic scatter. This is in fact seen in numerical simulations, where mean trends emerge both in the neutral hydrogen and metallicity profiles albeit with dispersions of several orders of magnitude. While systematic comparisons with simulations are not our main aim, for illustration purposes we offer two examples in Fig. 19, where we show (to the left) the expected H I radial profile reconstructed by Rahmati & Schaye (2014) using the EAGLE simulation (Schaye et al. 2015) and the metallicity distribution (to the right) predicted by the FOGGIE simulations (Peeples et al. 2019) as shown by Lehner et al. (2021). In both cases, a decreasing trend with radius is predicted with a scatter in the distributions that spans several orders of magnitudes and that increases with distance from the galaxies.

An important caveat on Fig.19 applies to this discussion when combining DLAs samples from the literature with MAGG observations. The MAGG sample has been purely selected based on H I column density, with no information on metallicity and with the aim to cover a broad range of column densities. DLA samples, especially from Krogager et al. (2017) and collaborators, not only target systematically $\log(N_{\rm HI}/{\rm cm}^{-2}) \geq 10^{20.3}$ but focus primarily on metal rich absorbers for which modelling suggests a higher probably of identifying galaxy counterparts. Hence, the suggestion of a marked trend both in column density and metallicity in the data with radius is, at this point, a possible artefact of the different selection effects with literature surveys probing galaxies only at projected distances $\lesssim 50$ kpc around higher column density and higher metallicity absorbers. A second point to note is that, differently from NIR surveys, MAGG targets only Lyα at $z > 3$. While this means that we are potentially more sensitive to lower mass galaxies and in general reach lower flux limits, dust attenuation may become a non-negligible bias for the most metal rich absorbers in our sample and galaxies at closer separations may exist but go undetected in our survey.

In summary, while MAGG has drastically increased the number of galaxy detections near optically-thick absorbers, it has uncovered very large scatters in the distribution of hydrogen and metals due to the patchy nature of the CGM and the presence of large scale structures within which multiple LAEs are embedded. This intrinsic property of the data thus prevents us from deriving firm constraints on the impact of accretion and feedback on the CGM. Going forward, future efforts should increase the sample size to identify reliable trends within distributions that are characterised by a large dispersion. Moreover, multi-wavelength observations of at least UV and optical lines should identify and minimise possible systematic effects linked to dust near metal rich absorbers. Finally, although we admit this is not a trivial effort, larger scales reaching at least $\approx 1-2$ Mpc should be explored to investigate more the clustered nature of LAEs in large scale structures.

8 SUMMARY AND CONCLUSIONS

In this paper, we have presented the study of the link between $z \sim 3-4$ optically-thick ALSs and galaxies from the full MUSE Analysis of Gas around Galaxies (MAGG) sample. The MAGG sample covers 28 quasar fields with medium-depth IFU observations using VLT MUSE for at least 4h on-source. These quasar sightlines were selected to include at least one H I absorption system with $N_{\rm HI} \gtrsim 10^{16.5}$ cm^{-2} from archival high-resolution quasar spectroscopy. In total, the full sample includes 61 H I optically-thick absorbers. With MUSE, we identified galaxies via Lyα emission by searching a velocity window of 1000 km s^{-1} around these absorbers in the (500 kpc)^2 FOV of MUSE, and we studied the connection between gas and galaxies across a wide range of absorption properties with no pre-selection on parameters such as the metallicity.

The key results emerging from our analysis are:

(i) We identified 127 galaxies around 61 high column density H I absorbers at $2.9 < z < 4.2$. For 10 absorbers we detected no galaxies within the volume searched (impact parameters $\lesssim 300$ kpc and velocity window of ± 1000 km s^{-1}). This resulted in a detection rate of 82 per cent (50/61). The highest number of detected galaxies around a single absorber was 8 while most of the absorption systems had a single galaxy detection.

(ii) The H I absorbers were found to be clearly correlated with galaxy overdensities. Specifically, galaxies were found to cluster in velocity space with the majority of them lying within 500km s^{-1} of the absorbers' redshift, once accounting for a typical velocity offset of ≈ 250 km s^{-1} due to the resonant Lyα emission. The luminosity function for galaxies around absorbers has an $\approx 5\times$ higher normalisation than field galaxies, further supporting the physical connection between absorbers and LAEs. This enhancement is, however, less strong than the one observed for galaxies around the quasars themselves (Fossati et al. 2021). We also showed that the transverse cross-correlation between galaxies and absorbers has a positive correlation, indicating the presence of clustering on the scales probed within the MUSE FOV. We reported a clustering length of $R_0 = 1.43^{+0.26}_{-0.29}h^{-1}$ cMpc.

(iii) We observed no differences in the distributions of either the H I column density or metallicity for samples of absorbers for which there are no, 1 or ≥ 2 associated galaxies, indicating that the galaxy environment does not alter significantly the absorption properties.

(iv) When multiple LAEs are detected near an LLS, we observed a preferential alignment along the galaxy-absorber axis with an excess of galaxies along this direction relative to random distribution, suggestive of LAEs and optically-thick absorbers lying in filamentary structures.

(v) No strong correlations were observed in individual systems between the ALS properties (H I column density, metallicity, metal ion velocity width) and the galaxy emission properties (Lyα emission, projected distance between galaxy and absorber, and line of sight velocity separation). Only a mild correlation between the ALS velocity width (Δv_{90}) and Lyα emission was found, suggesting that the brightest LAEs live in regions of higher velocity dispersion, possibly due to effects induced by their halo mass or outflows. These results hold when considering either the brightest galaxy detected around each absorber or when selecting the closest in terms of impact parameter.

(vi) Once considering the full sample of > 1000 LAEs identified

in the MUSE cubes, we calculated the covering fraction of optically-thick gas around galaxies. We found an elevated covering factor on scales of $\lesssim 50$ kpc, typical of the inner CGM, within $1 - 2\,R_{\rm vir}$. We also observed an inflection in the radial dependence of the covering factor around ≈ 100 kpc, which we interpret as a transition between the contribution of the CGM of individual LAEs and the contributions from the IGM/CGM of other galaxies embedded in the same structures. When considering galaxies that reside in pairs or groups, we found an elevated covering factor (up to 3×) compared to what is seen around isolated galaxies.

The empirical findings described above led us to the following conclusions on the nature of optically-thick gas at $z \gtrsim 3$ and the use of galaxy-absorber pairs to constrain the baryon cycle:

(i) The sensitivity of MUSE and the ability to uniformly probe impact parameters up to ≈ 250 kpc from the quasars now routinely leads to very high detection rates of galaxies near optically-thick absorbers, with > 80 per cent detection rate at the moderate depth of our survey regardless of the absorption properties. IFU surveys thus have the potential to reach 100 per cent detection rates in deeper searches. The majority of the identified galaxies (≈ 80 per cent) lie at large impact parameters ($\gtrsim 100$ kpc) with respect to the absorbing gas, with $\approx 30 - 40$ per cent of LLSs being associated with more than one galaxy. Our survey thus reveals the full impact of the clustered nature of galaxies within cosmic structures containing hydrogen, forcing us to move from the study of individual galaxy counterparts of absorbers to a more statistical analysis of the galaxy environment near ALSs.

(ii) Based on the recovered statistics of the distribution of galaxies near optically-thick gas, we exclude the ISM and inner CGM ($\lesssim 1 - 2\,R_{\rm vir}$) of the identified galaxies as the origin of the observed LLSs. We favour instead the outer CGM (defined here as $3 - 4\,R_{\rm vir}$) and filaments of the IGM within which galaxies are embedded as the sites where LLSs reside. Furthermore, in line with conclusions from studies of Lyα in emission, we identify LAEs as the most likely contributors to the cosmic number density of optically-thick gas. Indeed, given the elevated number density and covering factor of optically-thick gas around LAEs, virtually all LLSs can be identified within 150 kpc of an LAE. Moreover, the clustered nature of LAEs with respect to themselves and with other galaxy populations (e.g. LBGs or quasars) imply that the covering factors measured at large distances in these other galaxy populations are not independent and are already accounted for in the budget attributed to LAEs. In other words, by virtue of their larger number density, LAEs contribute either directly with their CGM to the cosmic budget of LLSs or act as tracers of the larger scale structures which host both LLSs and other galaxies.

(iii) The emerging picture of LAEs embedded within filaments hosting optically-thick gas, combined with the literature evidence that the CGM/IGM is patchy and locally inhomogeneous, explains the lack of significant correlations between absorption and emission properties, and the very large scatter ($\gtrsim 2 - 3$ dex) observed in the hydrogen and metal distribution around LAEs on scales of $\approx 50 - 200$ kpc. In turn, this implies that the signature of baryon processes (i.e., inflows and outflows) around galaxies are hard to discern in moderate-size samples. Larger samples are thus required to unveil systematic trends inside distributions with intrinsically high dispersion. Furthermore, it is not particularly surprising that the galaxy environment does not have a clear influence on the detailed properties of ALSs, but the elevated covering factor in groups can be

explained as enhanced hydrogen cross-sections in massive nodes where multiple filaments intersect or overlap.

In conclusion, the MAGG survey has provided the most detailed view of the connection between optically-thick gas and galaxies at $z \gtrsim 3$, enabling for the first time a systematic study of the clustering of line emitting galaxies with respect to LLSs, and of their individual absorption and emission properties. This analysis has led to a clearer understanding of how hydrogen is distributed around star-forming galaxies at high redshift, painting a view similar to the one proposed by cosmological simulations according to which galaxies are clustered inside filaments that contain substantial amount of neutral hydrogen. From a theoretical point of view, the next steps are to provide a more rigorous and detailed assessment of the level of agreement that exists between cosmological simulations and the MAGG observations. Empirically, extending observations to rest-frame optical and infrared emission will become a crucial element to complete this picture with dusty or non star-forming galaxies that are underrepresented in MAGG. Campaigns with the *James Webb Space Telescope* and ALMA should make this possible. Furthermore, studies like the one proposed here should be continued to further grow the samples and unveil the systematic trends between galaxy properties in emission and gas properties in absorption, which are critically needed for novel constraints on the baryon processes operating near and around galaxies.

9 DATA AVAILABILITY

The VLT data used in this work are available from the European Southern Observatory archive (https://archive.eso.org/) either as raw data or phase 3 data products. Spectroscopy obtained at the Keck telescopes is available via the Keck Observatory Archive (KOA, https://www2.keck.hawaii.edu/koa/public/koa.php). Cubextractor can be obtained upon request by contacting Sebastiano Cantalupo, and other codes used in this paper have been made available at http://www.michelefumagalli.com/codes.html.

10 ONLINE SUPPORTING INFORMATION

The fits of the metal absorption lines for all of the systems in this work are included as online material along with associated tables listing the properties for each ion including equivalent widths and column densities. We also provide the full figures from the photoionisation modelling used to determine the metallicity of each absorption systems.

REFERENCES

Adelberger K. L., Steidel C. C., Shapley A. E., Pettini M., 2003, ApJ, 584, 45. doi:10.1086/345660

Adelberger, K. L., Shapley, A. E., Steidel, C. C., et al. 2005, ApJ, 629, 636

Adelberger K. L., Steidel C. C., Pettini M., Shapley A. E., Reddy N. A., Erb D. K., 2005, ApJ, 619, 697. doi:10.1086/426580

Astropy Collaboration, Price-Whelan A. M., Sipőcz B. M., Günther H. M., Lim P. L., Crawford S. M., Conseil S., et al., 2018, AJ, 156, 123. doi:10.3847/1538-3881/aabc4f

Astropy Collaboration, Robitaille T. P., Tollerud E. J., Greenfield P., Droettboom M., Bray E., Aldcroft T., et al., 2013, A&A, 558, A33. doi:10.1051/0004-6361/201322068

Augustin, R., Péroux, C., Møller, P., et al. 2018, 478, 3120

Bacon, R., Accardo, M., Adjali, L., et al. 2010, Proc. SPIE, 7735, 773508

Bacon, R., Mary, D., Garel, T., et al. 2021, A&A, 647, A107

Berg, T., Fumagalli, M., D'Odorico, V. et al. 2021, MNRAS, 502, 4009

Bergeron J., Cristiani S., Shaver P. A., 1992, A&A, 257, 417

Bernstein, R., Shectman, S. A., Gunnels, S. M., Mochnacki, S., & Athey, A. E. 2003, Proc. SPIE, 4841, 1694

Bertin, E. and Arnouts, S. 1996, A&AS, 117, 393

Bielby, R. M., Shanks, T., Weilbacher, P. M., et al. 2011, MNRAS, 414, 2

Bielby, R. M., Tummuangpak, P., Shanks, T. et al. 2016, MNRAS, 456, 4061

Bielby, R., Crighton, N. H. M., Fumagalli, M., et al. 2017, MNRAS, 468, 1373

Bielby, R., Shanks, T., Crighton, N. H. M., et al. 2017, MNRAS, 471, 2174

Bielby, R. M., Stott, J. P., Cullen, F. 2019, MNRAS, 486, 21

Bielby, R. M., Fumagalli, M., Fossati, M., et al. 2020, MNRAS, 493, 5336

Bonnet, H., Abuter, R., Baker, A., et al. 2004, The ESO Messenger 117, 17

Booth, R. S., de Blok, W. J. G., Jonas, J. L., et al. 2009, arXiv:0910.2935

Bordoloi R., O'Meara J. M., Sharon K., Rigby J. R., Cooke J., Shaban A., Matuszewski M., et al., 2022, arXiv, arXiv:2205.08554

Bordoloi R., Lilly S. J., Knobel C., Bolzonella M., Kampczyk P., Carollo C. M., Iovino A., et al., 2011, ApJ, 743, 10. doi:10.1088/0004-637X/743/1/10

Borisova E., Cantalupo S., Lilly S. J., Marino R. A., Gallego S. G., Bacon R., Blaizot J., et al., 2016, ApJ, 831, 39. doi:10.3847/0004-637X/831/1/39

Brown B. W. M., Hollander M., Korwar R. M., 1974, in Proschan F., Serfling R. J., eds, Reliability and Biometry. SIAM, Philedelphia, PA, p. 327

Buie, E. II, Gray, W. J., Scannapieco, E., et al. 2020, ApJ, 896, 136

Cassata, P., Le Fèvre, O., Garilli, B. et al. 2011, A&A, 535, A143

Cantalupo S., Arrigoni-Battaia F., Prochaska J. X., Hennawi J. F., Madau P., 2014, Natur, 506, 63. doi:10.1038/nature12898

Cantalupo, S., Pezzulli, G., Lilly, S. J., et al. 2019, MNRAS, 483, 5188

Corlies L., Peeples M. S., Tumlinson J., O'Shea B. W., Lehner N., Howk J. C., O'Meara J. M., et al., 2020, ApJ, 896, 125. doi:10.3847/1538-4357/ab9310

Chen H.-W., Lanzetta K. M., Webb J. K., Barcons X., 2001, ApJ, 559, 654. doi:10.1086/322414

Chowdhury A., Kanekar N., Chengalur J. N., Sethi S., Dwarakanath K. S., 2020, Natur, 586, 369. doi:10.1038/s41586-020-2794-7

Christensen, L., Wisotzki, L. Roth, M. M. et al. 2007, A&A, 468, 587

Cicone C., Maiolino R., Gallerani S., Neri R., Ferrara A., Sturm E., Fiore F., et al., 2015, A&A, 574, A14. doi:10.1051/0004-6361/201424980

Crighton, N. H. M., Bielby, R., Shanks, T., 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., et al., 2015, MNRAS, 446, 18

Daddi E., Valentino F., Rich R. M., Neill J. D., Gronke M., O'Sullivan D., Elbaz D., et al., 2021, A&A, 649, A78. doi:10.1051/0004-6361/202038700

De Cia A., Ledoux C., Mattsson L., Petitjean P., Srianand R., Gavignaud I., Jenkins E. B., 2016, A&A, 596, A97. doi:10.1051/0004-6361/201527895

De Cia A., Jenkins E. B., Fox A. J., Ledoux C., Ramburuth-Hurt T., Konstantopoulou C., Petitjean P., et al., 2021, Natur, 597, 206. doi:10.1038/s41586-021-03780-0

Dekker, H., D'Odorico, S., Kaufer, A., Delabre, B., & Kotzlowski, H. 2000, Proc. SPIE, 4008, 534

Diener C., Wisotzki L., Schmidt K. B., Herenz E. C., Urrutia T., Garel T., Kerutt J., et al., 2017, MNRAS, 471, 3186

Drake, A. B., Garel, T., Wisotzki, L., et al. 2017, A&A, 608, A6.

Dutta, R., Fumagalli, M., Fossati, M., et al. 2020, MNRAS, 499, 5022

Dutta R., Fumagalli M., Fossati M., Bielby R. M., Stott J. P., Lofthouse E. K., Cantalupo S., et al., 2021, arXiv, arXiv:2109.10927

Eisenhauer, F., Abuter, R., Bickert, K., et al. 2003, SPIE, 4841, 1548

Fan X., Narayanan V. K., Lupton R. H., Strauss M. A., Knapp G. R., Becker R. H., White R. L., et al., 2001, AJ, 122, 2833. doi:10.1086/324111

Faucher-Giguère, C. and Kereš, D. 2011, MNRAS, 412, L118

Felten J. E., 1976, ApJ, 207, 700. doi:10.1086/154538

Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, Rev. Mex. Astron. Astrofis., 49, 137

Fitzpatrick E. L., 1999, PASP, 111, 63. doi:10.1086/316293

Foreman-Mackey, D., Hogg, D., David, W., et al. 2013, PASP, 125, 306

Fossati, M., Fumagalli, M., Lofthouse, E. K., et al. 2021, MNRAS, 503, 3044

Fossati, M., Fumagalli, M., Lofthouse, E. K., et al. 2019, MNRAS, 490, 1451

Fox A. J., Lehner N., Tumlinson J., Howk J. C., Tripp T. M., Prochaska J. X., O'Meara J. M., et al., 2013, ApJ, 778, 187. doi:10.1088/0004-637X/778/2/187

Franco, M., Elbaz, D., Béthermin, M., et al. 2019, A&A, 620, A152

Fumagalli, M., Prochaska, J. X., Kasen, D., et al. 2011, MNRAS, 418, 1796

Fumagalli, M., O'Meara, J. M., Prochaska, J. X., et al. 2013, ApJ, 775, 78

Fumagalli, M., O'Meara, J. M., Prochaska, J. X., et al. 2014, MNRAS, 444, 1282

Fumagalli, M., O'Meara & J. M., Prochaska, J. X. 2016, MNRAS, 455, 4100

Fumagalli, M., Cantalupo, S., Dekel, A., et al. 2016, MNRAS, 462, 1978

Fumagalli M., Mackenzie R., Trayford J., Theuns T., Cantalupo S., Christensen L., Fynbo J. P. U., et al., 2017, MNRAS, 471, 3686. doi:10.1093/mnras/stx1896

Fumagalli M., Fotopoulou S., Thomson L., 2020, MNRAS, 498, 1951. doi:10.1093/mnras/staa2388

Fynbo, J. P. U., Prochaska, J. X., Sommer-Larsen, J., et al. 2008, ApJ, 683, 321

Fynbo J. P. U., Laursen P., Ledoux C., Møller P., Durgapal A. K., Goldoni P., Gullberg B., et al., 2010, MNRAS, 408, 2128. doi:10.1111/j.1365-2966.2010.17294.x

Grove, L. F., Fynbo, J. P. U., Ledoux, C., et al. 2009 A&A, 496, 689

Gallego S. G., Cantalupo S., Sarpas S., Duboeuf B., Lilly S., Pezzulli G., Marino R. A., et al., 2021, MNRAS, 504, 16. doi:10.1093/mnras/stab796

García-Vergara C., Hennawi J. F., Barrientos L. F., Arrigoni Battaia F., 2019, ApJ, 886, 79. doi:10.3847/1538-4357/ab4d52

Gardner, J. P., Katz, N., Hernquist, L., et al. 2001, ApJ, 559, 131

Gawiser E., Francke H., Lai K., Schawinski K., Gronwall C., Ciardullo R., Quadri R., et al., 2007, ApJ, 671, 278. doi:10.1086/522955

Glidden A., Cooper T. J., Cooksey K. L., Simcoe R. A., O'Meara J. M., 2016, ApJ, 833, 270. doi:10.3847/1538-4357/833/2/270

Haardt, F. & Madau, P. 2012, ApJ, 746, 125

Haislmaier K. J., Tripp T. M., Katz N., Prochaska J. X., Burchett J. N., O'Meara J. M., Werk J. K., 2021, MNRAS, 502, 4993. doi:10.1093/mnras/staa3544

Hamanowicz, A., Péroux, C., Zwaan, M. A., et al. 2020, MNRAS, 492, 2347

Herenz, E. C., Wisotzki, L., Saust, R., et al. 2019, A&A, 621, A107.

Herrero Alonso Y., Krumpe M., Wisotzki L., Miyaji T., Garel T., Schmidt K. B., Diener C., et al., 2021, A&A, 653, A136. doi:10.1051/0004-6361/202141226

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, arXiv:1804.09365

Haardt F., Madau P., 2001, cghr.conf, 64

Hinton, S. R., Davis, T. M., Lidman, C., et al. 2016, Astronomy and Computing, 15, 61

Hummels C. B., Bryan G. L., Smith B. D., Turk M. J., 2013, MNRAS, 430, 1548. doi:10.1093/mnras/sts702

Isobe T., Feigelson E. D., Nelson P. I., 1986, ApJ, 306, 490. doi:10.1086/164359

Johnston R., 2011, A&ARv, 19, 41. doi:10.1007/s00159-011-0041-9

Jonas J. L., 2009, IEEEP, 97, 1522. doi:10.1109/JPROC.2009.2020713

Jorgenson R. A., Murphy M. T., Thompson R., 2013, MNRAS, 435, 482. doi:10.1093/mnras/stt1309

Kanekar N., Prochaska J. X., Christensen L., Rhodin N. H. P., Neeleman M., Zwaan M. A., Møller P., et al., 2018, ApJL, 856, L23. doi:10.3847/2041-8213/aab6ab

Kacprzak G. G., Murphy M. T., Churchill C. W., 2010, MNRAS, 406, 445. doi:10.1111/j.1365-2966.2010.16667.x

Kacprzak G. G., Pointon S. K., Nielsen N. M., Churchill C. W., Muzahid S., Charlton J. C., 2019, ApJ, 886, 91. doi:10.3847/1538-4357/ab4c3c

Kerutt J., Wisotzki L., Verhamme A., Schmidt K. B., Leclercq F., Herenz E. C., Urrutia T., et al., 2022, A&A, 659, A183. doi:10.1051/0004-6361/202141900

Kikuta S., Matsuda Y., Cen R., Steidel C. C., Yagi M., Hayashino T., Imanishi M., et al., 2019, PASJ, 71, L2. doi:10.1093/pasj/psz055

Klitsch, A., Péroux, C., Zwaan, M. A., et al. 2018, MNRAS, 475, 492

Krogager, J. K., Fynbo, J. P. U., Møller, P., et al. 2012, MNRAS, 424, L1

Krogager J.-K., Møller P., Fynbo J. P. U., Noterdaeme P., 2017, MNRAS, 469, 2959. doi:10.1093/mnras/stx1011

Krogager J.-K., Fynbo J. P. U., Møller P., Noterdaeme P., Heintz K. E., Pettini M., 2019, MNRAS, 486, 4377. doi:10.1093/mnras/stz1120

Krogager J.-K., Møller P., Christensen L. B., Noterdaeme P., Fynbo J. P. U., Freudling W., 2020, MNRAS, 495, 3014. doi:10.1093/mnras/staa1414

Larkin, J., Barczys, M., Krabbe, A., et al. 2006, SPIE, 6269, 62691A

Leclercq, F., Bacon, R.,Wisotzki, L., et al. 2017, å, 608, A8

Ledoux, C., Petitjean, P., Fynbo, J. P. U., et al. 2006, A&A, 457, 71

Lehner N., Howk J. C., Tripp T. M., Tumlinson J., Prochaska J. X., O'Meara J. M., Thom C., et al., 2013, ApJ, 770, 138. doi:10.1088/0004-637X/770/2/138

Lehner N., Kopenhafer C., O'Meara J., Howk J. C., Fumagalli M., Prochaska J., Acharyya A., et al., 2021, arXiv, arXiv:2112.03304

Lofthouse, E. K., Fumagalli, M., Fossati, M., et al. 2020, MNRAS, 491, 2057

Lopez S., Tejos N., Ledoux C., Barrientos L. F., Sharon K., Rigby J. R., Gladders M. D., et al., 2018, Natur, 554, 493. doi:10.1038/nature25436

Mackenzie, R., Fumagalli, M., Theuns, T., et al. 2019, MNRAS, 487, 5070

McQuinn M., Oh S. P., Faucher-Giguère C.-A., 2011, ApJ, 743, 82. doi:10.1088/0004-637X/743/1/82

Marino, R. A., Cantalupo, S., Lilly, S. J., et al. 2018, ApJ, 859, 53

Møller, P. and Warren, S. J. 1998, MNRAS, 299, 661

Møller P., Fynbo J. U., 2001, A&A, 372, L57. doi:10.1051/0004-6361:20010606

Møller, P., Fynbo, J. P. U. and Fall, S. M. 2004, A&A, 422, L33

Møller, P., Warren, S. J., Fall, S. M., et al. 2002, ApJ, 574, 51

Møller, P., Christensen, L., Zwaan, M. A. 2018, MNRAS, 474, 4039

Møller P., Christensen L., 2020, MNRAS, 492, 4805. doi:10.1093/mnras/staa128

Monier, E. M., Turnshek, D. A. and Rao, S. 2009, MNRAS, 397, 943

Morrissey, P., Matuszewski, M., Martin, D. C., et al. 2018, ApJ, 864, 93

Murphy M. T., Bernet M. L., 2016, MNRAS, 455, 1043. doi:10.1093/mnras/stv2420

Muzahid, S., Schaye, J., Marino, R. A., et al. 2020, MNRAS, 496, 1013

Muzahid, S., Schaye, J., Cantalupo, S., et al. 2021, arXiv:2105.05260

Narayanan, A., Sameer, Muzahid, S., et al. 2021, arXiv:2105.04841

Neeleman, M., Wolfe, A. M., Prochaska, J. X., et al. 2013, ApJ, 769, 54

Neeleman M., Kanekar N., Prochaska J. X., Rafelski M., Carilli C. L., Wolfe A. M., 2017, Sci, 355, 1285. doi:10.1126/science.aal1737

Neeleman M., Prochaska J. X., Kanekar N., Rafelski M., 2020, Natur, 581, 269. doi:10.1038/s41586-020-2276-y

Nelson D., Byrohl C., Peroux C., Rubin K. H. R., Burchett J. N., 2021, MNRAS, 507, 4445. doi:10.1093/mnras/stab2177

Nielsen N. M., Kacprzak G. G., Pointon S. K., Churchill C. W., Murphy M. T., 2018, ApJ, 869, 153. doi:10.3847/1538-4357/aaedbd

Noterdaeme, P., Laursen, P., Petitjean, P., et al. 2012, A&A, 540, A63

O'Meara J. M., Prochaska J. X., Worseck G., Chen H.-W., Madau P., 2013, ApJ, 765, 137. doi:10.1088/0004-637X/765/2/137

Ouchi M., Shimasaku K., Furusawa H., Saito T., Yoshida M., Akiyama M., Ono Y., et al., 2010, ApJ, 723, 869. doi:10.1088/0004-637X/723/1/869

Peeples M. S., Corlies L., Tumlinson J., O'Shea B. W., Lehner N., O'Meara J. M., Howk J. C., et al., 2019, ApJ, 873, 129. doi:10.3847/1538-4357/ab0654

Péroux C., Dessauges-Zavadsky M., D'Odorico S., Kim T.-S., McMahon R. G., 2003, MNRAS, 345, 480. doi:10.1046/j.1365-8711.2003.06952.x

Péroux, C., Bouché, N., Kulkarni, V. P. et al. 2011, MNRAS, 410, 2237

Péroux, C., Bouché, N., Kulkarni, V. P. et al. 2012, MNRAS, 419, 3060

Péroux, C., Zwaan, M. A., Klitsch, A., et al. 2019, MNRAS, 485, 1595

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13

Pontzen, A., Governato, F., Pettini, M., et al. 2008, MNRAS, 390, 1349

Prochaska J. X., Wolfe A. M., 1998, ApJ, 507, 113. doi:10.1086/306325

Prochaska J. X., 1999, ApJL, 511, L71. doi:10.1086/311849

Prochaska J. X., Hennawi J. F., Lee K.-G., Cantalupo S., Bovy J., Djorgovski S. G., Ellison S. L., et al., 2013, ApJ, 776, 136. doi:10.1088/0004-637X/776/2/136

Prochaska J. X., Wolfe A. M., 2009, ApJ, 696, 1543. doi:10.1088/0004-637X/696/2/1543

Prochaska, J. X., O'Meara, J. M., Fumagalli, M., et al. 2015, ApJS, 221, 2

Prochaska J. X., O'Meara J. M., Worseck G., 2010, ApJ, 718, 392. doi:10.1088/0004-637X/718/1/392

Prochter G. E., Prochaska J. X., O'Meara J. M., Burles S., Bernstein R. A., 2010, ApJ, 708, 1221. doi:10.1088/0004-637X/708/2/1221

Quiret S., Péroux C., Zafar T., Kulkarni V. P., Jenkins E. B., Milliard B., Rahmani H., et al., 2016, MNRAS, 458, 4074. doi:10.1093/mnras/stw524

Ranjan A., Noterdaeme P., Krogager J.-K., Petitjean P., Srianand R., Balashev S. A., Gupta N., et al., 2020, A&A, 633, A125. doi:10.1051/0004-6361/201936078

Rafelski, M., Wolfe, A. M., Prochaska, J. X, et al. 2012, ApJ, 755, 89

Rafelski M., Gardner J. P., Fumagalli M., Neeleman M., Teplitz H. I., Grogin N., Koekemoer A. M., et al., 2016, ApJ, 825, 87. doi:10.3847/0004-637X/825/2/87

Rakic, O., Schaye, J., Steidel, C. C., et al. 2012, ApJ, 751, 94

Rahmati A., Schaye J., 2014, MNRAS, 438, 529. doi:10.1093/mnras/stt2235

Rao, S. M., Belfort-Mihalyi, M., Turnshek, D. A., et al. 2011, MNRAS, 416, 1215

Rauch, M., Haehnelt, M., Bunker, A., et al. 2008, ApJ, 681, 856

Rauch M., Sargent W. L. W., Barlow T. A., 2001, ApJ, 554, 823. doi:10.1086/321402

Riemer-Sørensen, S., Webb, J. K., Crighton, N., et al. 2015, MNRAS, 447, 2925

Rubin, K. H. R., Weiner, B. J., Koo, D. C., et al. 2010, ApJ, 719, 1503

Rubin K. H. R., Hennawi J. F., Prochaska J. X., Simcoe R. A., Myers A., Lau M. W., 2015, ApJ, 808, 38. doi:10.1088/0004-637X/808/1/38

Rudie, G. C., Steidel, C. C., Trainor, R. F., et al. 2012, ApJ, 750, 67

Sargent W. L. W., Steidel C. C., Boksenberg A., 1989, ApJS, 69, 703. doi:10.1086/191326

Savage, B. D. and Sembach K. R. 1991, ApJ, 379, 245

Schaye J., Crain R. A., Bower R. G., Furlong M., Schaller M., Theuns T., Dalla Vecchia C., et al., 2015, MNRAS, 446, 521. doi:10.1093/mnras/stu2058

Schechter P., 1976, ApJ, 203, 297. doi:10.1086/154079

Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525. doi:10.1086/305772

Schmidt M., 1968, ApJ, 151, 393. doi:10.1086/149446

Shen S., Madau P., Aguirre A., Guedes J., Mayer L., Wadsley J., 2012, ApJ, 760, 50. doi:10.1088/0004-637X/760/1/50

Shibuya, T., Masami, O., Kimihiko, N., et al. 2014, ApJ, 788, 74

Srianand, R., Hussain, T., Noterdaeme, P., et al. 2016, MNRAS, 460, 634

Steidel, C. C., Erb, D. K., Shapley, A. E., et al 2010, ApJ, 717, 289

Steidel C. C., 1990, ApJS, 74, 37. doi:10.1086/191493

Trainor, R. F. and Steidel, C. C., 2012, ApJ, 752, 39

Tummuangpak, P., Bielby, R. M., Shanks, T., et al. 2014, MNRAS, 442, 2094

Turner, M. L., Schaye, J., Steidel, C. C., Rudie, G. C., & Strom, A. L. 2014, MNRAS, 445, 794

Umehata, H., Fumagalli, M., Smail, I., et al., 2019, Science, 366, 97

Urrutia T., Wisotzki L., Kerutt J., Schmidt K. B., Herenz E. C., Klar J., Saust R., et al., 2019, A&A, 624, A141. doi:10.1051/0004-6361/201834656

van de Voort, F., Schaye, J., Altay, G., et al. 2012, MNRAS, 421, 2809

Verhamme, A., Garel, T., Ventou, E., et al. 2018, MNRAS, 478, L60

Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, Proc. SPIE, 2198, 362

Weatherley, S. J., Warren, S. J., Møller, P., et al. 2005, MNRAS, 358, 985

Weilbacher, P. M., Streicher, O., Urrutia, T., et al. 2014, Astronomical Data Analysis Software and Systems XXIII, 485, 451

Whiting M. T., Webster R. L., Francis P. J., 2006, MNRAS, 368, 341. doi:10.1111/j.1365-2966.2006.10101.x

Wisotzki, L., Bacon, R., Blaizot, J., et al. 2016, å, 587, 1

Wisotzki, L., Bacon, R., Brinchmann, J., et al. 2018, Nature, 562, 229

Wolfe, A. M., Gawiser, E., Prochaska, J. X. 2005, ARA&A, 43, 861

Wootten, A. and Thompson, A. R. 2009, IEEE Proceedings, 97, 1463

York, D. G., Adelman, J., Anderson, J. E. Jr., et al. 2000, AJ, 120, 1579

Zafar, T., Popping, A., & Péroux, C. 2013, A&A, 556, A140

This paper has been typeset from a TeX/LaTeX file prepared by the author.