



Universiteit  
Leiden  
The Netherlands

## The relationship between gas and galaxies at $z < 1$ using the Q0107 quasar triplet

Beckett, A.; Morris, S.L.; Fumagalli, M.; Bielby, R.; Tejos, N.; Schaye, J.; ... ; Cantalupo, S.

### Citation

Beckett, A., Morris, S. L., Fumagalli, M., Bielby, R., Tejos, N., Schaye, J., ... Cantalupo, S. (2021). The relationship between gas and galaxies at  $z < 1$  using the Q0107 quasar triplet. *Monthly Notices Of The Royal Astronomical Society*, 506(2), 2574-2602.  
doi:10.1093/mnras/stab1630

Version: Submitted Manuscript (under Review)

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/3276006>

**Note:** To cite this publication please use the final published version (if applicable).

# The relationship between gas and galaxies at $z < 1$ using the Q0107 quasar triplet

Alexander Beckett<sup>1</sup>★, Simon L. Morris<sup>1</sup>, Michele Fumagalli<sup>3,2,1</sup>, Rich Bielby<sup>1</sup>,  
Nicolas Tejos<sup>4</sup>, Joop Schaye<sup>5</sup>, Buell Jannuzi<sup>6</sup>, Sebastiano Cantalupo<sup>3,7</sup>

<sup>1</sup>Centre for Extragalactic Astronomy, Durham University, South Road, Durham DH1 3LE, UK

<sup>2</sup>Institute for Computational Cosmology, Durham University, South Road, Durham, DH1 3LE, UK

<sup>3</sup>Dipartimento di Fisica G. Occhialini, Università degli Studi di Milano Bicocca, Piazza della Scienza 3, 20126 Milano, Italy

<sup>4</sup>Instituto de Física, Pontificia Universidad Católica de Valparaíso, Casilla 4059, Valparaíso, Chile

<sup>5</sup>Leiden Observatory, Leiden University, PO Box 9513, 2300, RA Leiden, The Netherlands

<sup>6</sup>Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA

<sup>7</sup>Department of Physics, ETH Zurich, Wolfgang-Pauli Strasse 27, Zurich, Switzerland

Last updated 20XX Jan 1; in original form 20XX Jan 1

## ABSTRACT

We study the distribution and dynamics of the circum- and intergalactic medium using a dense galaxy survey covering the field around the Q0107 system, a unique  $z \approx 1$  projected quasar triplet. With full Ly $\alpha$  coverage along all three lines-of-sight from  $z=0.18$  to  $z=0.73$ , more than 1200 galaxy spectra, and two MUSE fields, we examine the structure of the gas around galaxies on 100-1000 kpc scales. We search for H I absorption systems occurring at the same redshift (within 500 km s<sup>-1</sup>) in multiple sightlines, finding with > 99.9% significance that these systems are more frequent in the observed quasar spectra than in a randomly distributed population of absorbers. This is driven primarily by absorption with column densities  $N(\text{H I}) > 10^{14}$  cm<sup>-2</sup>, whilst multi-sightline absorbers with lower column densities are consistent with a random distribution. Star-forming galaxies are more likely to be associated with multi-sightline absorption than quiescent galaxies. HST imaging provides inclinations and position angles for a subset of these galaxies. We observe a bimodality in the position angle of detected galaxy-absorber pairs, again driven mostly by high-column-density absorbers, with absorption preferentially along the major and minor axes of galaxies out to impact parameters of several hundred kpc. We find some evidence supporting a disk/outflow dichotomy, as H I absorbers near the projected major-axis of a galaxy show line-of-sight velocities that tend to align with the rotation of that galaxy, whilst minor-axis absorbers are twice as likely to exhibit O VI at the same redshift.

**Key words:** intergalactic medium – quasars: absorption lines – galaxies: formation – large-scale structure of Universe

## 1 INTRODUCTION

Galaxies follow a large-scale filamentary structure throughout the Universe, known as the cosmic web (Bond et al. 1996), formed by the gravitational accretion of gas towards the potential wells of dark matter around initial overdensities, as has been modelled in dark-matter-only simulations for decades (e.g. Klypin & Shandarin 1983; Springel et al. 2005). Once stars and galaxies form, this accretion from the intergalactic medium (IGM) towards galaxies becomes affected by complex baryonic physics, including stellar and AGN feedback (e.g. van de Voort et al. 2011; Nelson et al. 2019; Mitchell et al. 2020b). Simulating these effects on a range of scales from sub-parsec-scale supernovae to megaparsec-scale gas flows around

clusters necessitates sub-grid models which require constraints from observations.

Such constraints are necessary for studies of galaxy evolution, as exchanges of material between galaxies and their environments play an important role in regulating star formation (e.g. Kereš et al. 2005; Schaye et al. 2010; Davé et al. 2012; Lehnert et al. 2013; Somerville et al. 2015; Salcido et al. 2020). Gas outside of galaxies is believed to contain a substantial fraction of the baryons in the Universe (e.g. Fukugita et al. 1998; Behroozi et al. 2010; Werk et al. 2014), so any census of baryons used to constrain cosmological parameters must consider the state of the gas around galaxies (usually by modelling the ionization state of the gas based on absorption spectra, e.g. Shull et al. 2012, although recent methods using fast radio bursts account for all ionized baryons, e.g. Macquart et al. 2020). The dynamics of galaxies are also strongly linked to the state

★ Contact e-mail: alexander.beckett@durham.ac.uk

of the surrounding gas through the transfer of angular momentum (e.g. Pichon et al. 2011; Stewart et al. 2017; DeFelippis et al. 2020). Therefore, observational insights into the distribution of baryons around galaxies can not only inform our understanding of the gas itself, but also stellar processes within galaxies and their effects on galaxy formation and evolution.

Outside the local environment, most observations of the gas around galaxies are made by identifying absorption features in the spectrum of background sources, usually quasars. These features can probe neutral gas, most commonly using H I (e.g. Morris et al. 1991; Lanzetta et al. 1995; Chen et al. 1998; Adelberger et al. 2003; Tumlinson et al. 2013; Rakic et al. 2013; Heckman et al. 2017; Chen et al. 2018), especially since the Hubble Space Telescope has allowed the Lyman- $\alpha$  transition to be observed at low redshifts. Other studies use low ions such as Mg II to probe cool ( $\sim 10^4$  K) gas (e.g. Bergeron 1986; Bergeron & Boissé 1991; Bouché et al. 2006; Nielsen et al. 2013; Schroetter et al. 2016; Ho et al. 2017), or search for highly-ionized material (e.g. Bergeron et al. 1994; Tripp & Savage 2000; Cen et al. 2001; Tumlinson et al. 2011; Turner et al. 2014; Finn et al. 2016; Werk et al. 2016; Nicastro et al. 2018; Bielby et al. 2019) through transitions from ions such as O VI to O VIII. This variety of observed ions found in sightlines probing the gas near to galaxies suggests that it has a complex, multi-phase structure (e.g. Veilleux et al. 2005; Werk et al. 2013; Mathes et al. 2014; Péroux et al. 2019; Chen et al. 2020a).

Galaxy-scale outflows are observed in emission (e.g. Bland & Tully 1988; Finley et al. 2017; Burchett et al. 2021) and absorption (e.g. Grimes et al. 2009; Turner et al. 2015; Lan & Mo 2018; Schroetter et al. 2019), their multi-phase nature indicated by a range of diagnostics from low ions (tracing cool gas) (e.g. Concas et al. 2019), to X-ray-emitting hot gas (e.g. Lehnert et al. 1999). These are consistent with stellar-feedback, in which supernovae and stellar winds drive material out of the galaxy (e.g. Chevalier & Clegg 1985; Heckman et al. 1990) to distances of several tens of kiloparsecs. Cosmological simulations also find these winds, and can produce the multi-phase biconical outflows observed despite isotropic injection of energy and momentum (e.g. Nelson et al. 2019; Mitchell et al. 2020a). Hopkins et al. (2021) also find that the addition of cosmic rays to the outflow-driving mechanism may allow such flows to reach megaparsec scales.

There are several lines of evidence suggesting that substantial gas accretion occurs onto galaxies from the surrounding medium, including the metallicities of dwarf stars (e.g. Casuso & Beckman 2004), the short depletion timescales of star-forming galaxies (e.g. Freundlich et al. 2013; Scoville et al. 2017), and the declining H I density of the Universe over time (e.g. Neeleman et al. 2016). Evidence for this accretion is found in ‘down-the-barrel’ observations (using the host galaxy as the background source) where absorption line profiles indicate gas flows towards the galaxy (e.g. Martin et al. 2012; Rubin et al. 2012). Observations in Mg II have long suggested that the gas shows strong rotation, often with an in-falling velocity component (e.g. Charlton & Churchill 1998; Steidel et al. 2002). When this is compared to the rotation curves of galaxies, often using integral field units such as MUSE, most find that absorbing gas close to the major axis of a galaxy preferentially shows co-rotation in both Mg II (e.g. Ho et al. 2017; Martin et al. 2019; Zabl et al. 2019) and H I (e.g. French & Wakker 2020), particularly within 100 kpc. However, this co-rotation is not always apparent when extending to weaker absorbers or larger impact parameters (e.g. Dutta et al. 2020).

This ‘galactic fountain’ model, consisting of minor axis outflows and major axis co-rotating accretion, can explain the bimodal-

ity in position angle found in the MEGAFLOW survey (Schroetter et al. 2016; Zabl et al. 2019; Schroetter et al. 2019). It is also supported by cosmological simulations, for example the FIRE simulations (Hafen et al. 2019, 2020) produce galaxies around which the gas around galaxies is a mixture of material ejected from the galaxy interstellar medium (ISM) in winds, and material accreted from the IGM. Much of this material then accretes onto the central galaxy. However, this model would predict generally lower metallicities along the major axis, which has not been detected in recent H I observations (Pointon et al. 2019; Kacprzak et al. 2019), and Mg II observations often do not find a significant bimodality in position angle (Dutta et al. 2020; Huang et al. 2021). How frequently these structures form, and to what extent, therefore remains uncertain.

Whilst a galactic fountain can reproduce many observations of the gas close to galaxies (the circumgalactic medium, or CGM; see Tumlinson et al. 2017 for a recent review), studies of the correlation between gas and galaxies show transverse correlation lengths of 2–3 Mpc at  $z < 1$  (Tejos et al. 2014; Finn et al. 2016). On these scales the interactions between galaxies are expected to have a significant effect on the structure of the CGM/IGM (e.g. Fossati et al. 2019b; Dutta et al. 2020). Intra-group material can be observed in absorption, that can not easily be assigned to an individual galaxy (e.g. Péroux et al. 2017; Bielby et al. 2017; Chen et al. 2020a). We expect interactions in groups and clusters to build up pressure-confined tidal debris, visible across a range of column densities (Morris & van den Bergh 1994). This is clearly visible in H I maps of the Magellanic Stream (e.g. Mathewson et al. 1974; Nidever et al. 2008) and the M81/M82 group (e.g. Croxall et al. 2009; Sorgho et al. 2019), as well as H $\alpha$  and occasionally [O III] emission in dense environments (e.g. Fumagalli et al. 2014; Johnson et al. 2018; Fossati et al. 2019a).

Absorption, especially in Ly $\alpha$ , is our most sensitive probe of the CGM and IGM, but it is in most cases limited to a single pencil-beam sightline through the CGM of any individual galaxy. This makes it difficult to directly constrain the size scales and structures in the CGM, although some studies have stacked large samples together attempting to infer these properties (e.g. Chen 2012; Turner et al. 2014).

Additional information can be extracted using multiple sightlines probing the CGM of an individual galaxy. Bowen et al. (2016) take advantage of the large angle subtended by relatively nearby haloes, such that the halo is pierced by a number of QSO sightlines, finding the gas surrounding NGC 1097 to have a disk-like structure with rotating and in-falling velocity components. Similarly Keeney et al. (2013) find the absorption in three sightlines around a nearby galaxy to be consistent with a ‘galactic fountain’. Gravitational lensing produces multi-sightline systems, through multiple images (e.g. Smette et al. 1992; Rauch et al. 2002; Chen et al. 2014; Zahedy et al. 2016), and extended arcs that can be considered as multiple closely-spaced lines-of-sight (e.g. Lopez et al. 2018, 2020). It is also possible to use bright background galaxies as sources (which may also be extended), as an alternative or in addition to quasar sightlines (e.g. Adelberger et al. 2005; Steidel et al. 2010; Cooke & O’Meara 2015; Rubin et al. 2018; Fossati et al. 2019b; Zabl et al. 2020).

This study focuses on the Q0107 system, a quasar triplet at  $z \approx 1$ : LBQS 0107-025A, LBQS 0107-025B, and LBQS 0107-0232, hereafter denoted A, B, and C. Table 1 summarizes some of the main parameters of this system. This allows multiple sightlines to be probed through the CGM/IGM around galaxies in this field, separated by hundreds of kpc.

As the only known, bright, low-redshift quasar triplet, this

**Table 1.** Co-ordinates, redshifts and R-band magnitudes of the three quasars, taken from [Crighton et al. \(2010\)](#)

Object	RA (J2000)	Dec (J2000)	Redshift	R-mag
Q0107-025 A	01:10:13.14	-2:19:52.9	0.960	18.1
Q0107-025 B	01:10:16.25	-2:18:51.0	0.956	17.4
Q0107-0232 (C)	01:10:14.43	-2:16:57.6	0.726	18.4

system has been the focus of many studies. [Dinshaw et al. \(1997\)](#) observed A and B, finding 5 absorption systems that cover both sightlines, and 6 limited to a single sightline. Using a maximum likelihood analysis, they concluded that their data was best explained by randomly inclined disks approximately 1 Mpc in radius. [Young et al. \(2001\)](#) complemented this with analysis of the likelihood of individual multi-sightline absorption systems, and found that matches involving stronger absorption features tended to have smaller velocity separations. Coincidences of absorption between the sightlines also occur more frequently among high-column-density absorbers, as found by [Petry et al. \(2006\)](#).

A later study by [Crighton et al. \(2010\)](#) (hereafter C10) used improved QSO spectra, including QSO-C, in addition to galaxy data from CFHT-MOS, to extend these results. They observed a highly-significant excess of absorption systems covering all three sightlines over an ensemble of randoms, providing clear evidence that gas and galaxies are associated on scales of hundreds of kpc. Additionally, galaxies and groups of galaxies could be associated with multiple absorbers, allowing the structure of the gas to be analysed in individual systems (although we defer an updated analysis of these systems to a later paper, focusing here on the statistical properties of our samples).

Ionization modelling was used by [Muzahid \(2014\)](#) to study one example in this field at  $z \sim 0.22$ , using the presence of O VI in sightlines A and B to estimate the radius of the CGM as 330 kpc, and therefore detect both the warm and cool CGM of an  $L_*$  galaxy.

This field was also included in a study of galaxy-absorber cross-correlations covering six independent fields by [Tejos et al. \(2014\)](#) (hereafter T14) and another using 50 fields by [Finn et al. \(2016\)](#) (F16). Their galaxy catalogue forms the basis for the galaxy data used in this work.

In this paper we present an updated analysis on the Q0107 triplet, using a much larger sample of galaxies extending to fainter magnitudes, in addition to Hubble Space Telescope imaging providing improved morphologies, and MUSE fields providing kinematics on a subsample of galaxies close to the A and B sightlines. We use this data to examine the CGM/IGM on large scales, where the improved imaging allows us to constrain the extent of the ‘galactic fountain’ and the larger galaxy samples allow us to study the presence of absorption features covering multiple sightlines around galaxies of different properties.

In Section 2 we describe the quasar spectra and galaxy survey used to produce our catalogues of absorption features and galaxies. Section 3 discusses our test for correlated absorption between the three sightlines. Section 4 gives results from studying the relationship between absorption properties and the position angles of nearby galaxies, whilst in Section 5 we use the MUSE data to identify co- and counter-rotation among material close to the major axis of galaxies. In Section 6 we discuss the consequences of our results and how future work can further progress our understanding.

We use the cosmology from [Planck Collaboration et al. \(2020\)](#)

throughout, with  $\Omega_m = 0.315$  and  $H_0 = 67.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$ , and quote physical sizes and distances unless otherwise stated.

## 2 DATA

Our dataset consists of HST/COS and FOS spectra of the three quasars and galaxy surveys from the VIMOS, DEIMOS, GMOS and CFHT-MOS instruments, supplemented by HST R-band imaging and two MUSE fields. In this section we describe the reduction processes for each of these observations and the compilation of the results into our final catalogues of galaxies and absorbers, which are then used in Sections 3 - 5.

### 2.1 IGM data

The UV spectra of the quasars were taken by the Cosmic Origins Spectrograph (COS, [Green et al. 2012](#)) and Faint Object Spectrograph (FOS) on the Hubble Space Telescope. The COS spectra were observed in 2010-11 (program G011585, PI: Neil Crighton), using the G130M and G160M gratings, with a FWHM of 0.07–0.09 Å, and a signal-to-noise ratio (SNR) per pixel of 7–9. These complement longer-wavelength FOS data described in [Young et al. \(2001\)](#). The observations are detailed in Table 2. The G130M and G270H gratings were not used for QSO-C, due to the Lyman Limit of a sub-damped Lyman $\alpha$  system in the sightline and the lower redshift of this quasar respectively. These observations cover Ly $\alpha$  in COS from  $z = 0$  (or the Lyman Limit in the case of QSO-C) to  $z \approx 0.45$ , and in FOS from 0.45 to the redshift of the quasar.

We use the line catalogue from T14 (also used by F16). They provide a more detailed description of the reduction process with further references, but we summarise it here. Individual exposures from COS were downloaded from the Space Telescope Science Institute (STScI) archive and reduced using CALCOS v2.18.5.

T14 performed their own background smoothing procedure masking out portions of the spectra affected by strong geocoronal emission lines (namely the Ly $\alpha$  and O I 1302,1306 Å lines) and pixels with bad data quality flags. The error array was calculated in the same way as in CALCOS, but using the new background estimation, interpolated across the masked-out regions. Each spectrum was then flux calibrated using sensitivity curves provided by STScI.

Co-alignment was performed by cross-correlating strong galactic absorption features. Pixels with bad data quality flags were then excluded, whilst pixels with warning flags were halved in weight, before fluxes were re-binned to have a constant spacing equal to the dispersion of the grating. Co-addition was weighted by exposure time, and was followed by re-binning on a linear scale sufficiently narrow to ensure Nyquist sampling<sup>1</sup> across the entire wavelength range (0.0395 Å per pixel).

Individual exposures from FOS were downloaded from the STScI archive and reduced using the standard CALFOS pipeline. Wavelength corrections given by [Petry et al. \(2006\)](#) were applied to each individual exposure. The shortest wavelength region of the FOS G190H settings overlap with the longest wavelength COS settings, and T14 confirmed that the wavelength scales in these overlapping regions were consistent between the two instruments. All

<sup>1</sup> Usually 2 pixels per resolution element, although this is not well-defined for the non-Gaussian line-spread function of COS. Here the FWHM is used to estimate the required sampling.

**Table 2.** Summary of the QSO spectra used to generate the absorption-line catalogue. Columns are: (1) The QSO featured in the spectrum, (2) the spectrograph used, (3) the grating used, (4) the wavelength range for which this instrument and grating provides the spectrum, (5) the full width at half-maximum of the line-spread function of the spectrograph, (6) dispersion of the spectrum, (7) average signal-to-noise across the wavelength range given, (8) exposure time of the observations, (9) HST program ID of the observations.

QSO	Instrument	Grating	Wavelength Range (Å)	FWHM (Å)	Dispersion (Å/pix)	SNR (per pix)	Exposure Time (h)	Prog ID
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Q0107-025A	COS	G130M	1135-1460	0.07	0.01	9	7.8	11585
	COS	G160M	1460-1795	0.09	0.01	8	12.3	11585
	FOS	G190H	1795-2310	1.39	0.36	28	7.5	5320, 6592
Q0107-025B	FOS	G270H	2310-3277	1.97	0.51	32	2.4	6100
	COS	G130M	1135-1460	0.07	0.01	9	5.9	11585
	COS	G160M	1460-1795	0.09	0.01	7	5.9	11585
	FOS	G190H	1795-2310	1.39	0.36	28	1.8	5320, 6592
Q0107-0232 (C)	FOS	G270H	2310-3277	1.97	0.51	32	1.8	6100
	COS	G160M	1434-1795	0.09	0.01	7	23.2	11585
	FOS	G190H	1795-2310	1.39	0.36	18	9.1	11585

individual exposures were then combined together, resampling to a common wavelength scale of  $0.51 \text{ \AA}$  per pixel.

T14 then estimated the continuum of each spectrum using a semi-automated method. They split the spectrum into ‘chunks’ of  $\approx 12 \text{ \AA}$  (blueward of QSO Ly $\alpha$  emission, longer chunks at longer wavelengths), fit a straight line through the points within each chunk (iteratively removing outliers until convergence), then fit a cubic spline to give a smooth result. They checked the resulting continuum ‘by eye’ to ensure a reasonable fit (see their Figure 1).

VPFIT (Carswell & Webb 2014) was used to estimate redshifts, column densities and Doppler parameters for each identifiable absorption system. The H I systems were assigned a flag (‘a’, ‘b’ or ‘c’) based on the number of absorption lines observed and the signal-to-noise in the column density estimate:

- (a): At least two Lyman transitions observed with log(N) estimates at least 30 times their uncertainty
- (b): Only Ly $\alpha$  observed with log(N) estimate at least 30 times its uncertainty
- (c): log(N) estimate less than 30 times its uncertainty

Only those with ‘a’ and ‘b’ flags are included in our analysis. The distribution of H I column densities is consistent (using results from Keeney et al. 2012, as discussed in Figure 5 and Section 4.5 of T14) with a  $3\sigma$  detection limit estimate of  $\sim 10^{13} \text{ cm}^{-2}$  in the COS spectra and  $\sim 10^{13.5} \text{ cm}^{-2}$  in the FOS spectra. Our catalogue should therefore be complete above this column density, with the exception of unresolved blended systems more likely to be found in the lower-resolution FOS spectra. We incorporate the differing detection limits of the COS and FOS gratings into our analysis in Section 3, and check for any resulting redshift bias throughout this work.

The catalogue contains 430 absorption systems, of which 272 are H I. Most of our discussion focuses on these H I absorbers, although metals are briefly discussed in Section 4.2.

## 2.2 Galaxies

The galaxy data used in this study comes from a number of different surveys. The catalogue is based on that used in T14 and F16, with spectra from VIMOS, DEIMOS, GMOS and CFHT-MOS observations (referred to as MOS data throughout). More recent observations using the Multi-Unit Spectroscopic Explorer (MUSE,

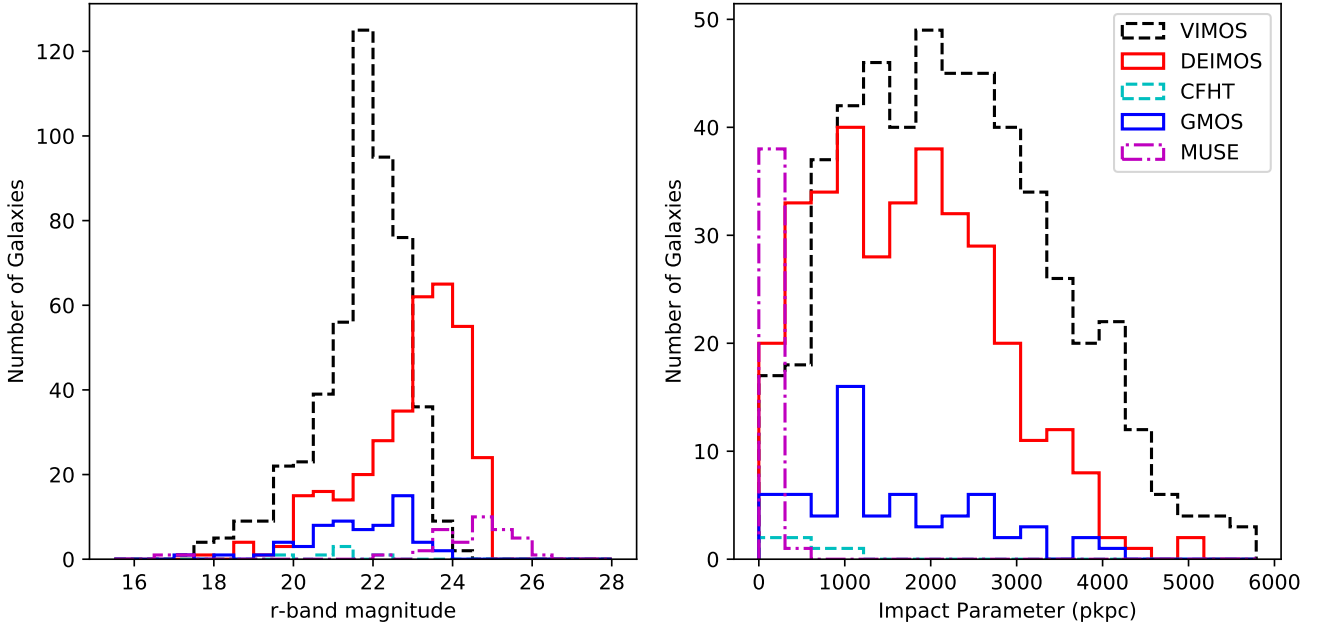
Bacon et al. 2010) on the VLT are added to this catalogue. Table 3 summarises the number of spectra taken in each of these observations. Additionally, HST imaging is used to determine position angles and inclinations of galaxies with identified redshifts. These observations and the associated data reduction are discussed below. Figure 1 shows the magnitude and impact parameter distributions of the MOS and MUSE surveys, whilst their projected extent on the sky is illustrated in Figure 2.

### 2.2.1 MUSE

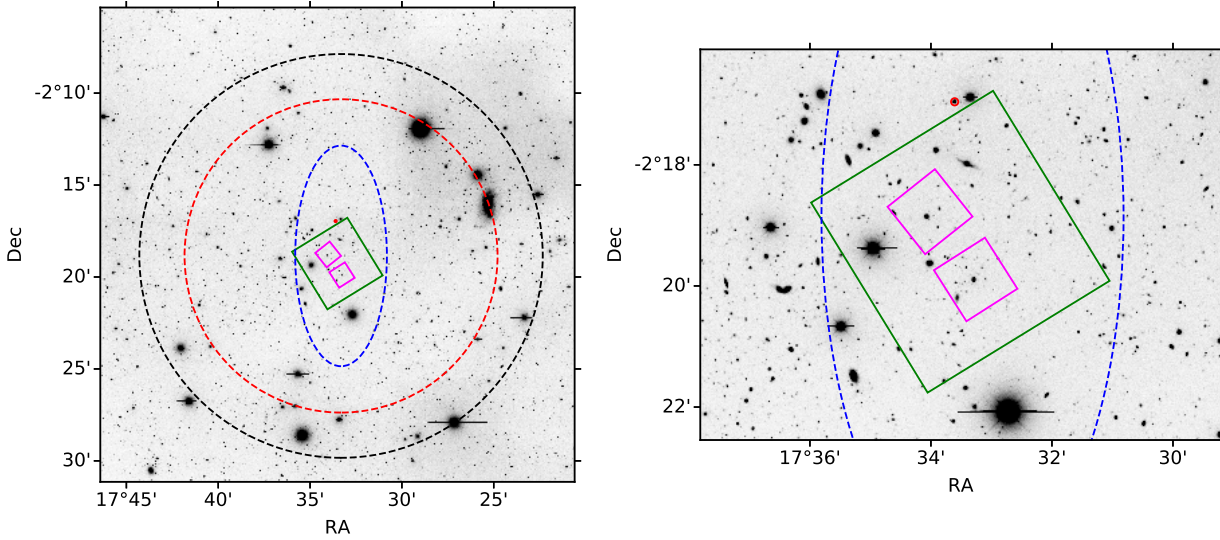
Information on the kinematics of galaxies close to the quasar lines-of-sight can be extracted from MUSE data. During 2014, the MUSE GTO team took eight exposures covering  $1' \times 1'$  fields of view around both QSO-A and QSO-B (program ID 094.A-0131, PI Schaye), totalling two hours for each quasar. This produces a 3D datacube, with a spectrum in each  $0.2'' \times 0.2''$  ‘spaxel’ from 4750 to 9350 Å, with a delivered seeing of  $0.96''$  for QSO-A, and  $0.82''$  for QSO-B. In the spectral direction, the datacube has a FWHM of  $\approx 2.7 \text{ \AA}$ .

The reduction of these data follows a similar process to Fumagalli et al. (2016, 2017); Fossati et al. (2019b); Lofthouse et al. (2020); Bielby et al. (2020). MUSE ESO pipeline routines were used to remove bias, apply flat-fielding, and calibrate astrometry and wavelength for each exposure. The ‘scibasic’ and ‘scipost’ pipeline routines combine the IFUs for each exposure, resampling using a drizzle algorithm onto a 3D grid, as well as correcting for telluric absorption using sky continuum and sky line models produced using the darkest pixels in the exposure. Exposures can be aligned using point sources to produce a reasonable combined datacube, but this process generally leaves sky line residuals as well as uneven illumination across the field (e.g. Bacon et al. 2017).

Further corrections were applied using the CUBEX package (S. Cantalupo, in prep). We use two main routines from this package (Cantalupo et al. 2019). CubeFix performs a renormalization on each IFU, stack, and ‘slice’ (similar to a single slit), making the background as flat as possible across the MUSE field-of-view and wavelength range, and removing the ‘chequered’ pattern that often afflicts images reduced solely using the ESO pipeline. CubeSharp provides a flux-conserving sky subtraction, using the empirical shape of sky lines to calculate a line-spread function. Within



**Figure 1.** The magnitude and impact parameter distributions of galaxies in our catalogue, divided by the instrument used to obtain the galaxy spectrum, illustrating the depth and area of each observing program. Objects observed by multiple MOS instruments are shown only in the instrument with the best redshift flag (these flags are described in Section 2.3.4), or the best resolution if flags are equal. The MUSE objects shown are all new detections not featuring in the MOS surveys. *Left:* The apparent magnitude distribution of the galaxy samples, using the SDSS r-band in bins of 0.5-magnitude width. *Right:* The impact parameter from each galaxy to the nearest of the three quasars.



**Figure 2.** The layout of the surveys used in this study. The background image was taken with the Mayall 4-metre Telescope at NSF's Kitt Peak National Observatory. The dashed lines enclose the approximate regions covered by the MOS surveys, with VIMOS, DEIMOS and GMOS shown in black, red and blue respectively. The solid green square shows the region covered by HST imaging, whilst the smaller magenta squares show the MUSE fields centered on QSOs A and B. QSO-C lies inside the small red circle just outside the northern edge of the HST field. The left panel shows the full extent of the galaxy surveys, and the right panel more clearly shows the region close to the lines-of-sight.

**Table 3.** Summary of the galaxy spectra used in the catalogue. Columns are: (1) instrument used; (2) number of spectra taken (including duplicates); (3) number of unique objects observed by that instrument; (4) number of objects for which that instrument provides the reference spectrum and redshift; (5) number of objects for which we have estimated a redshift using MARZ (used to estimate redshift uncertainties as described in Section 2.3); (6) approximate FWHM of the line-spread function in km/s; (7) the program ID of the observations. The instrument with the highest confidence flag is used in the catalogue; if multiple instruments have the same flag, then the best resolution is used.

Instrument (1)	N spectra (2)	N unique (3)	N cat (4)	N marz (5)	FWHM (km/s) (6)	Prog ID (7)
VIMOS	935	757	746	436	1500	086.A-0970, 087.A-0857
DEIMOS	642	543	487	286	60	A290D
GMOS	210	196	112	107	470	GS-2008B-Q-50
CFHT-MOS	30	29	20	9	400	
MUSE	140	140	59	67	120	094.A-0131

each spatial pixel, flux is then allowed to move between neighbouring spectral pixels to best match the LSF, allowing the sky lines to be more accurately removed whilst conserving total flux. CubeFix and CubeSharp are run twice on each exposure, the results of the first run allowing better masking of sources in the second run. This further ensures that fluxes are not overcorrected and preserves the source flux as well as possible.

A  $3\sigma$  clipping is then used, combining the exposures using mean statistics. This combined cube is then used to mask sources for a final run of CubeFix and CubeSharp. Using this package greatly reduces sky and flat-field residuals, although some residuals remain visible towards the red end of the MUSE spectra.

Objects in the MUSE fields were identified using SExtractor (Bertin & Arnouts 1996) on the white-light image. We produced ID spectra by summing the flux within the SExtractor aperture.

We then estimated redshifts using the MARZ software (Hinton et al. 2016), with additional galaxy templates provided by Matteo Fossati (described in Fossati et al. 2019b, created using Bruzual & Charlot 2003 stellar population models). MARZ identifies the five redshift/template combinations producing the best cross-correlation between the observed spectrum and the template. We then chose the most likely of these based on the features fitted. Objects identified by SExtractor that are much smaller than the point-spread function ( $\approx 1''$ , 5 pixels) can be excluded as artifacts, and other objects are best fit by a stellar template instead of a galaxy template. The remaining objects have been assigned a confidence flag between 1 (redshift unknown) and 4 (highly confident) based on the spectral features visible at the redshift given by the best-fit result from MARZ. 67 galaxies were assigned a flag  $> 1$ , and are therefore used in this study. The number of objects with each confidence flag is shown in Table 4.

## 2.2.2 MOS

The MOS galaxy data is the subset of the catalogue from T14 that covers the Q0107 field, consisting of spectra from CFHT-MOS, VIMOS, DEIMOS and GMOS. Many objects in the catalogue were observed multiple times, either by the same instrument or by multiple instruments. One example is shown in Figure 3. The data are described briefly here. T14 and references therein describe the data collection and reduction processes in more detail.

**Table 4.** Objects detected in the MUSE fields by the MARZ flag assigned. The flag 2-4 galaxy detections are those added to the galaxy sample used for this study.

Flag	Descriptor	N
4	highly confident	24
3	good	21
2	possible	22
1	unknown	39
6	artefacts/stars	32
	QSOs	2
2-4	galaxy detections	67
	total	140

## 2.2.3 CFHT-MOS

The multi-object spectrograph on the Canada-France-Hawaii Telescope (CFHT-MOS) (Le Fevre et al. 1994) was used for observing runs in 1995 and 1997 by Morris & Jannuzi (2006). The Q0107 field was observed on the 29th and 30th July 1995, producing 30 galaxy spectra (one object observed twice, so 29 objects). The observation and reduction are described in more detail in the above paper.

The observed spectra were bias-subtracted using IRAF, and bad pixels were interpolated over. Cosmic rays were removed by comparing multiple exposures using the same mask. Sky subtraction used adjacent regions of the slit, whilst wavelength calibration used an arc frame obtained whilst pointing towards the same region of sky (to minimize the effects of instrument flexure). Flux calibrations used a nearby standard star.

The number of CFHT galaxies observed is not sufficient to find a statistically significant offset between these and other observations, and no redshift confidence flags were provided.

## 2.2.4 VIMOS

The VIMOS data (LeFevre et al. 2003) used a low-resolution ( $R \approx 200$ ) grism, giving 935 spectra with coverage between 5500 and 9500 Å (programs 086.A-0970, PI:Crighton; and 087.A-0857, PI:Tejos). The data were reduced using VIPGI (Scodreggio et al. 2005). Wavelength corrections were made using both lamp frames and skylines, whilst flux calibration used a standard star. We note that these data were taken shortly before the VIMOS charge-coupled

devices (CCDs) were updated in August 2010, and are unfortunately affected by fringing effects at wavelengths  $\geq 7500 \text{ \AA}$ .

Redshifts were estimated using cross-correlation between the observed spectra and SDSS templates. These redshifts were then manually assigned a confidence flag (a: secure, b: possible, c: uncertain), where secure redshifts required at least three spectral features. The redshifts of all MOS objects in the Q0107 field were then adjusted to match the DEIMOS frame (as DEIMOS has the best resolution of the MOS instruments used), based on the objects observed by VIMOS and DEIMOS. The magnitude of this shift was  $\Delta z \approx 0.0008$ , or  $120\text{--}240 \text{ km s}^{-1}$  (see T14 for details).

### 2.2.5 DEIMOS

The DEIMOS (Faber et al. 2003) settings give a much better resolution ( $R \approx 5000$ ) and substantially deeper data but over a smaller field, covering the  $6400\text{--}9100 \text{ \AA}$  range for 642 objects (taken in 2007–08, program A290D, PIs: Bechtold and Jannuzi). Redshifts were obtained from the DEEP2 data reduction pipeline (Newman et al. 2013), which applied all necessary de-biasing, flat-fielding, wavelength and flux calibration, and heliocentric corrections. The DEIMOS redshift confidence was measured in the pipeline using four categories (1: not good enough, 2: possible, 3: good, 4: excellent), which were reassigned to match the three categories above (1 to c, 2 and 3 to b, 4 to a) when added to the catalogue.

### 2.2.6 GMOS

GMOS (Davies et al. 1997) was used in 2008 on this field (program GS-2008B-Q-50, PI: Crighton), with an intermediate resolving power ( $R \approx 640$ ) and a slightly bluer wavelength range of  $4450\text{--}8250 \text{ \AA}$ .

Each spectrum consists of three 1080s exposures, dithered in wavelength to allow removal of bad pixels. IRAF was used to calibrate fluxes and wavelengths, using arc frames and a standard star taken contemporaneously with the science exposures.

210 redshifts were estimated using the same method as for VIMOS data, although the shift needed to match the DEIMOS frame was smaller, only  $\Delta z \approx 0.0004$  or  $60\text{--}120 \text{ km s}^{-1}$ .

## 2.3 Combined galaxy catalogue

In order to combine spectra from the multiple instruments previously described into a single galaxy catalogue for this field, we need to match objects observed by multiple instruments in order to remove duplicates, as well as ensure that the galaxy properties we utilise in our analysis are measured consistently. We use photometry and astrometry from the Sloan Digital Sky Survey source catalogue (SDSS, Albareti et al. 2017) as an anchor, as the SDSS catalogue includes close to half of the MOS objects, and 14 of the MUSE galaxies.

### 2.3.1 Astrometry

In order to remove duplicates and select a single spectrum from which to derive the properties of each galaxy, we matched the coordinates of objects observed by multiple MOS instruments to those observed in SDSS. With the exception of CFHT, all instruments had sufficient cross-matches within  $1''$  to confirm that the astrometry is consistent between the MOS instruments, and required an offset of less than  $0.5''$  to match SDSS. The same process applied to

the MUSE astrometry revealed a similar  $\approx 1''$  offset for both fields, which were corrected separately.

The astrometry of both the MOS and MUSE catalogues were adjusted to match SDSS, to ensure objects appearing in both catalogues were correctly paired. 28 objects were found to match within  $1''$  after the correction to SDSS, and no additional matches within  $2''$ . Due to the larger number of objects, the offset between MOS and SDSS is more accurately determined than that for MUSE, so the corrected MOS coordinates are used for the objects appearing in both catalogues.

### 2.3.2 Magnitudes

R- and I- band magnitudes included in the MOS catalogue are systematically shifted to match the r- and i- band SDSS magnitudes, using the Lupton (2005) transformations<sup>2</sup>. Magnitudes for MUSE-only objects were estimated by integrating the spectrum through the SDSS r- and i- filters. The corrected MOS magnitudes were preferred for objects appearing in both MOS and MUSE catalogues, and the differences were approximately consistent with the uncertainties provided.

The survey depths are limited by the original target selections, at  $R \approx 23.5$ , 24 and 24.5 for VIMOS, GMOS and DEIMOS respectively. As shown in Figure 8 of T14, the success rate of assigning redshifts to objects in the Q0107 field is  $\geq 80\%$  to a depth of  $R=22$  in both VIMOS and DEIMOS, with the deeper DEIMOS data showing  $\geq 60\%$  success to  $R=24$ .

We do not detect any significant variation in the depth of the MOS data within the HST field, although the small number of MUSE galaxies extend to fainter objects. Any variation in depth across the field does not have a significant effect on our results, as we retain the same selection for our analysis in Section 3, and our Section 4 and 5 use the much smaller HST field.

### 2.3.3 Redshifts

Calculations of relative velocities between galaxies and absorbers require that there are no systematic shifts between their redshifts. We therefore compared redshifts to ensure that all of our galaxy samples are in the same frame as the absorption features.

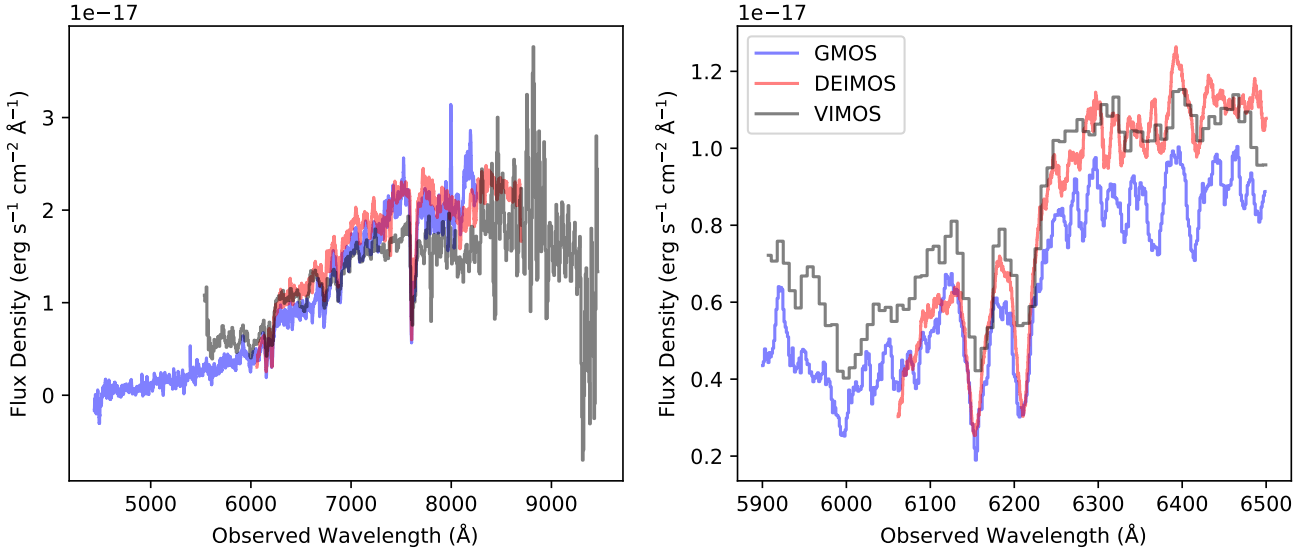
The separate MOS samples from T14 have already been corrected to a single frame, described in the section for each instrument. We therefore first check that this galaxy frame is the same as the frame in which the absorber redshifts are measured. One test of this is to calculate the difference in redshift between galaxy–absorber pairs. At the scales on which galaxies and absorbers are correlated, this should produce a signal that is symmetric about zero (velocity offsets e.g. outflows should average to zero over a large sample), as presented by Rakic et al. (2011). If the signal is offset, that suggests a shift is needed to bring all the redshifts into the same frame.

We find no such shift when pairing high column density H I absorbers ( $N(\text{H I}) > 10^{14} \text{ cm}^{-2}$ ) with galaxies within 2.5 Mpc of at least one sightline, both from the T14 catalogues. We fit a model consisting of a Gaussian plus constant offset to the velocity distribution of these pairs, these two components representing associated absorbers with small velocity differences, and unconnected absorbers with uniformly distributed velocity differences.

As most possible pairs lie in the uniform distribution, a velocity cut is needed to centre the fit close to zero, and avoid noise in this

<sup>2</sup> <https://www.sdss.org/dr16/algorithms/sdssubvritransform>





**Figure 3.** Example spectra of a galaxy observed in VIMOS, GMOS and DEIMOS. The left panel shows the full spectrum highlighting the different wavelength ranges covered, whilst the right panel centers on the Ca K & H absorption features. This galaxy is at  $z \approx 0.56$ , has luminosity close to  $L_{\star}$ , and is classified as non-star-forming. We smooth over a  $3 \text{ \AA}$  kernel to improve visibility.

distribution dominating over the Gaussian peak (i.e the Gaussian component of the fit jumps to a nearby ‘noise spike’ rather than the peak of associated galaxy–absorber pairs). The fit varies with the velocity used for this cut, but the centroid of the Gaussian component remains within  $\pm 10 \text{ km s}^{-1}$  for cuts smaller than  $5000 \text{ km s}^{-1}$ . Therefore the galaxies and absorbers provided by T14 can be taken to lie in the same frame. We similarly confirm that the separate samples for each instrument in the T14 catalogue are each consistent with this frame.

A small offset can be found if weaker absorbers are used or galaxies with a larger impact parameter are included, as well as if a larger velocity cut is used. These are less likely to be physically connected, so introduce noise that dominates over the peak. We therefore do not attempt to correct for any such spurious offset.

We then add the MUSE observations. The distribution of velocity differences between MOS and MUSE observations of duplicate objects showed that the MUSE objects required a further shift of  $30 \text{ km s}^{-1}$  in order to match the frame of the MOS catalogue.

We note that, although the redshifts of galaxies and absorbers are in the same frame, the overall redshift distributions are not similar, as seen in the left panel of Figure 4. Therefore, when comparing the galaxy–absorber associations involving different sub-samples, the difference in redshift must be considered in our analysis.

### 2.3.4 Redshift uncertainties

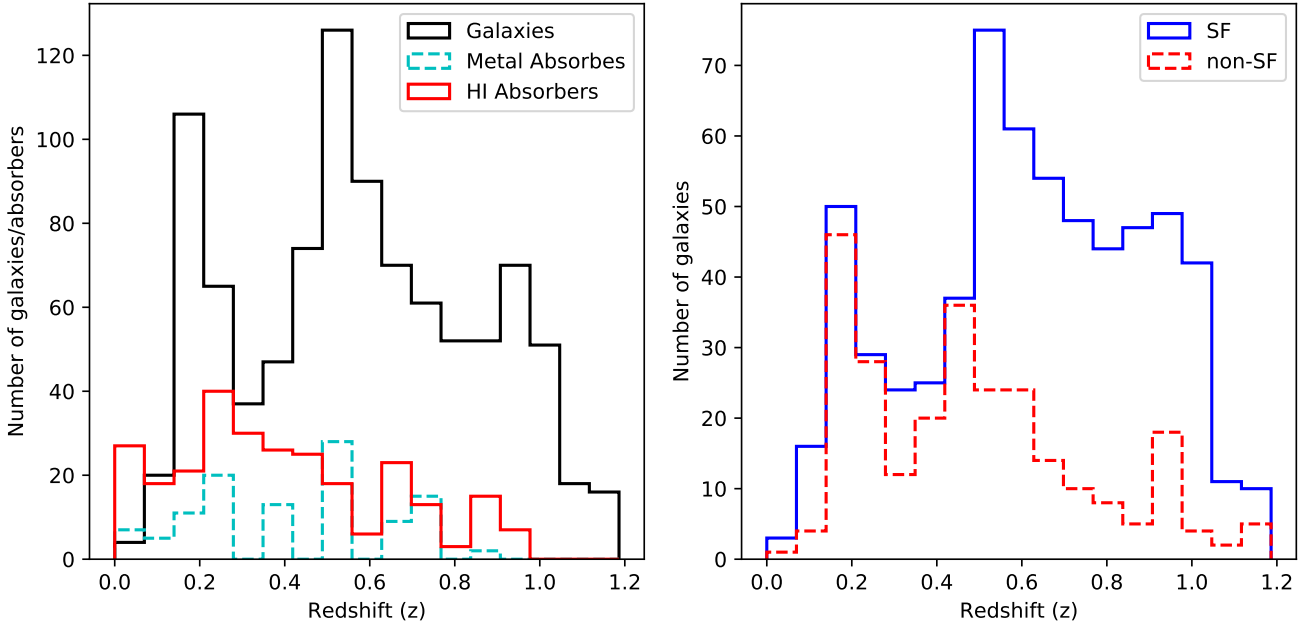
Redshift uncertainties for the MOS objects are based on those with multiple spectra. We attempted to fit all MOS spectra using the MARZ code used earlier to obtain redshifts from MUSE. For objects with multiple spectra from the same instrument, the redshift differences could be compared. The width of the distribution of velocity differences for each instrument provides an estimate of the uncertainty in the redshift of objects observed by that instrument. These are given in Table 5, where  $\sigma$  is adopted as the uncertainty for all objects in the T14 catalogue with the confidence flag shown

**Table 5.** Comparison of the measured velocity differences between objects measured twice by the same MOS instrument. We show the lowest-confidence flag assigned to spectra of each object from each instrument, and consider the distribution of velocity differences of each sample. Redshifts and flags shown are those assigned using the MARZ software (flags from 1 being ‘unknown’ to 4 being ‘highly confident’), and we also show the flags given by the T14 catalogue to which the MARZ flags are matched. A velocity cut of  $2000 \text{ km s}^{-1}$  was used to remove pairs of observations for which the velocity difference is due to a different identification of spectral features. The cut affects only a small number of ‘flag 2’ objects. The number removed by this cut is shown in brackets.  $\sigma$  is the standard deviation of the velocity differences, which can be adopted as the uncertainty on any individual redshift measurement within that sample.

Instruments	Flags	N	$\sigma$ (km/s)
VIMOS	3, 4 (a)	29	120
	2 (b)	69 (6)	190
DEIMOS	3, 4 (a)	23	26
	2 (b)	40 (3)	47
GMOS	3, 4 (a)	4	44
	2 (b)	5	48

in brackets. As redshifts with poorer confidence flags are generally identified using fewer features, we assign a larger redshift uncertainty to these b-flag galaxies. Whilst the GMOS sample size is very small, the relative values of the three instruments are consistent with the resolutions given in Table 3, so we adopt these uncertainties.

Our MUSE observations contain no duplicates, so this method cannot be used to estimate uncertainties on their redshifts. As MUSE has a higher resolution than GMOS, but lower than DEIMOS, we take the GMOS values as estimates of the velocity errors for MUSE galaxies with the same confidence flags.



**Figure 4.** The redshift distributions of galaxies and absorbers in our catalogue. *Left:* The redshift distributions of the galaxy and absorber samples. *Right:* The redshift distributions of the star-forming and non-star-forming galaxies in our sample.

In order to directly compare the flags assigned for objects with MOS and MARZ redshifts (a–c in the catalogue, 4–1 in MARZ), flag ‘a’ objects were numbered 3.5, ‘b’ to 2.5 and ‘c’ to 1.5. Redshifts quoted in our final catalogue are those with the highest flag.

Finally, the MARZ flags were reassigned to match those from the MOS catalogue: 1 to c, 2 to b and 3 & 4 to a. This produces our final catalogue of 1424 galaxies, of which 1026 have ‘a’ or ‘b’ flags and are used in the following analyses. Their locations in space and redshift are shown in Figure 6.

### 2.3.5 Spectral Classification

We divide our galaxy sample into ‘star-forming’ and ‘non-star-forming’ galaxies. We maintain the classifications used in T14 for the MOS galaxies (see their section 5.1) and apply similar criteria for dividing the MUSE galaxies. Namely, those galaxies best fit by a star-forming template (e.g. ‘late-type’, ‘starburst’ or ‘star-forming’ templates) are classified as star-forming, and those fit by a passive template (e.g. ‘passive’, ‘early-type’ or ‘absorption galaxy’ templates) are classified as non-star-forming.

These templates differ primarily due to the presence of emission lines, so this classification is denoting galaxies with measurable emission lines as star-forming, and those without as non-star-forming.

We also estimate the star-formation rates by directly fitting the  $H\alpha$  and  $[O\text{II}]$  emission lines<sup>3</sup>, and estimate stellar masses using the

relationship given by Johnson et al. (2015), finding that a cut at a specific star-formation rate of  $0.02\text{ Gyr}^{-1}$  correctly identifies  $\approx 75\%$  of both samples. Our stellar mass and SFR estimates are illustrated in Figure 5. However, these estimates have substantial measurement and systematic uncertainties, so we use the binary classification in our analysis.

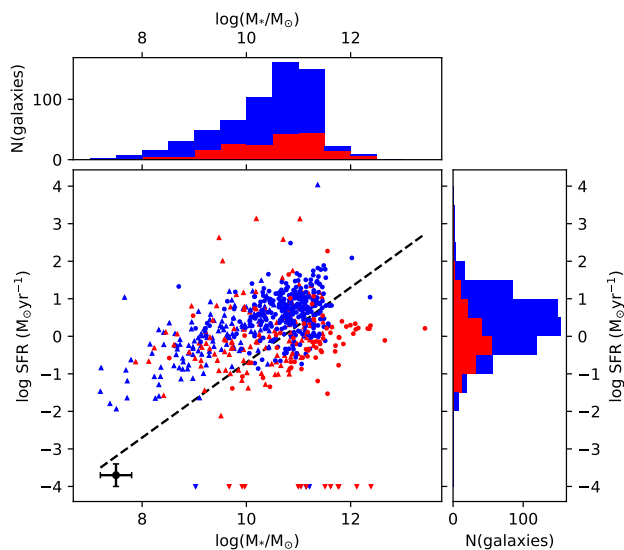
These subsamples of star-forming (SF) and non-star-forming (non-SF) galaxies show no substantial bias in their impact parameter distributions or mass distributions. It must be noted that there is a small excess of star-forming galaxies at the smallest impact parameters ( $< 200\text{ kpc}$ ), due to the ease of finding emission lines using the MUSE datacubes. The SF and non-SF samples also feature slightly larger low-mass and high-mass tails respectively in their mass distributions, but both of these biases affect a small number of galaxies. We confirm throughout that these have no substantial effect on our results by re-running our tests with samples excluding the MUSE galaxies and samples excluding the tails of the mass distribution, obtaining similar results.

However, there is a substantial bias in redshift, with SF galaxies preferentially featuring at higher redshifts than non-SF galaxies, as seen in the right panel of Figure 4. This is likely a combination of real redshift evolution (higher cosmic star-formation rates at higher redshift, e.g. Madau & Dickinson 2014) and observational effects (higher signal-to-noise is required to confirm the redshifts of galaxies without emission lines), and must be taken into account when comparing the CGM/IGM around galaxies in these samples.

## 2.4 HST imaging

In addition to the spectroscopic data, we also use high-resolution imaging of the field to constrain galaxy morphologies and orientations. We use publicly available Hubble Space Telescope imaging of this field (Program ID: 14660, PI Straka), obtained through ACS

<sup>3</sup> We use the Kennicutt (1998) and Kewley et al. (2004) calibrations to convert line luminosity to SFR for  $H\alpha$  and  $[O\text{II}]$  respectively. We assume 1 magnitude of extinction at  $H\alpha$  (Charlot et al. 2002), and use the Calzetti et al. (2000) curve to estimate extinction at  $[O\text{II}]$ . The predicted wavelength of at least one of these emission lines is available for  $\approx 90\%$  of galaxies with well determined redshifts.



**Figure 5.** The star-formation rate against stellar mass for our galaxy sample. Star-formation rates are calculated using the [OII] or  $H\alpha$  emission lines, shown in circles and upward-pointing triangles respectively. The downward-pointing triangles at the bottom of the figure show the masses of galaxies for which the line fit was unable to produce a result. (These are excluded from both projected histograms.) The blue points and histograms denote galaxies classified as ‘star-forming’ using the template fits, whilst the red points were classified as ‘non-star-forming’. The black point in the lower-left shows the median uncertainty in each axis, although the non-star-forming galaxies exhibit systematically larger uncertainty in SFR. The dotted black line is the  $0.02 \text{ Gyr}^{-1}$  cut in sSFR that best reproduces our template classification. Most of the objects lying substantially above the apparent ‘main sequence’ are due to the emission line being masked by sky lines or the fringes seen in the VIMOS spectra at red wavelengths.

(Ryon 2019) and the F814W filter. This consists of four exposures totalling 2171 seconds.

It must be noted that one of the exposures was affected by an unidentified bright object moving across the field, leaving an artefact in the final combined image. However, we only use the HST imaging to study the morphology of galaxies in this field, so this artifact does not substantially affect our results.

Galaxies were identified using SExtractor, then matched to the coordinates of objects in the MOS/MUSE combined catalogue. No systematic offset was found, so objects within  $1''$  were matched, as above.

We run GALFIT (Peng et al. 2002), which uses chi-squared minimization algorithms to produce a best-fitting 2D model of a galaxy. We initially fit a Sersic disk to every galaxy found in both our redshift catalogue and the HST image, using SExtractor results as initial guesses for the fit, and then introduce additional components where necessary to find a reasonable fit. This provides improved position angles and inclinations, taking full account of the point-spread function of the image and reducing the average uncertainty by a factor  $\approx 3$  relative to the position angles produced by SExtractor.

We again assign quality flags to the GALFIT results (1: good fit by eye and no clear structure remains in residuals, 2: good by eye, 3: possible, 4: clearly a poor fit), allowing poorly constrained results to be excluded. Flag 4 objects are excluded from all results. This returns 109 galaxies with position angles constrained by GALFIT and a counterpart in our spectroscopic survey, of which 72 also have

well-constrained redshifts (‘a’ or ‘b’ flags). We illustrate examples of this fitting in Appendix C.

### 3 COHERENCE BETWEEN SIGHTLINES

One unique test allowed by the configuration of multiple lines-of-sight is to compare the absorption seen across multiple sightlines at the same redshift. Both Dinshaw et al. (1997) and Crighton et al. (2010) (C10) attempted to estimate the scale size of absorbers using the numbers of coincidences (where absorption was identified at multiple redshifts) and anti-coincidences (where detection of absorption in one sightline was not matched in the other(s)). We can therefore use our larger sample to both review the results from these papers, demonstrating that absorption is often correlated on the 400–1200 kpc scales separating these sightlines, and to split the sample, allowing us to study how these coincidences are affected by the properties of the absorption and of nearby galaxies.

#### 3.1 Random absorbers

In order to test the significance of the correlations between sightlines, we must compare the observed distribution to the number expected if there were no physical connection between the observed gas and galaxies. For this reason, we have generated 5000 sets of randomly distributed absorbers, using a method similar to that used in T14, as follows:

- (i) Calculate the signal-to-noise per resolution element for each QSO spectrum.
- (ii) Convert this to a minimum rest-frame equivalent width for the absorption feature as a function of redshift. The detection limits for QSO-A are shown as an example in Figure 7.
- (iii) For each real absorption feature, find the allowed region in redshift space for which the EW of the progenitor is larger than the minimum, and is not covered by galactic absorption.
- (iv) Distribute absorbers randomly through the allowed region, giving the random absorber the same properties as the observed progenitor.

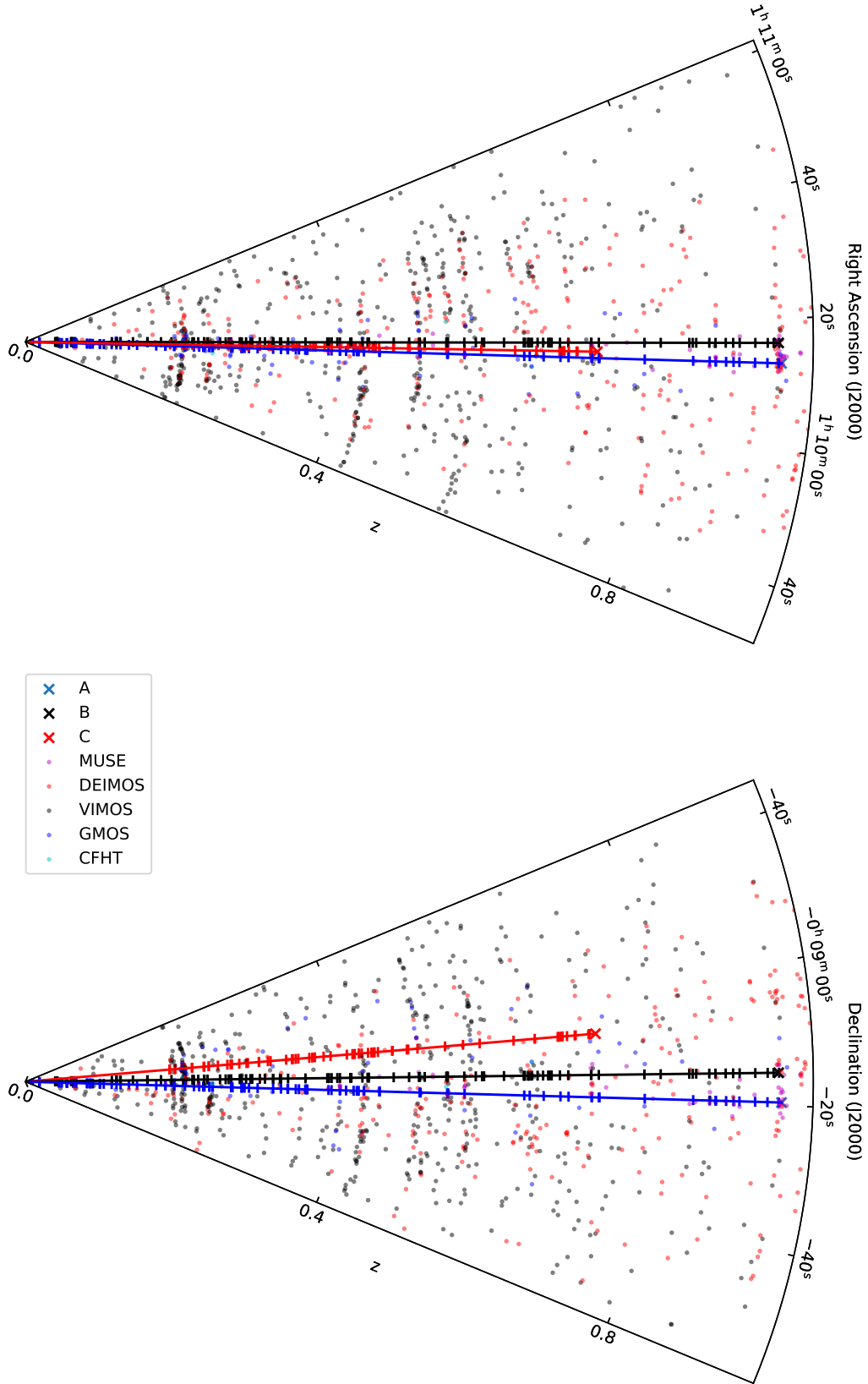
In order to maintain the approximate redshift/wavelength distribution of absorbers, we restrict random absorbers to the same grating as their observed progenitor. We describe this process in more detail in Appendix B.

#### 3.2 Absorber-Galaxy Groups

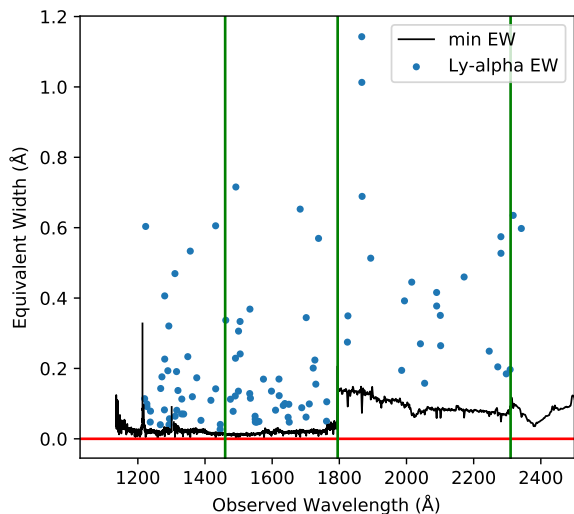
In order to associate galaxies and absorbers, we use two different grouping algorithms in this section, similar to those used in C10. The first is a nearest-neighbour algorithm, which simply takes the nearest absorber (in velocity) to each galaxy in each sightline.

The second method is a velocity-cut around each absorber/galaxy, in which we consider all absorbers/galaxies within that velocity window as associated with that absorber/galaxy.

Galaxies must be within 2.5 Mpc of at least one of the sightlines in order to be included, and have a redshift flag of ‘a’ or ‘b’. These restrictions remove the galaxy-absorber pairs with the largest separations (and are therefore the least likely to be physically connected), as well as those with poorly-determined redshifts.



**Figure 6.** The distribution of galaxies and H I absorption features in the sample, plotted in redshift and right-ascension (top) or declination (bottom). Solid lines show the lines-of-sight to the three quasars, with H I absorption shown as tick marks. Points are galaxies coloured by observed instrument. Note that the angles are massively expanded, so arc-like features likely correspond to galaxy groups.



**Figure 7.** The estimated minimum rest-frame equivalent width detectable in the COS and FOS spectrum of QSO-A. The calculation is described in the text, SNR is calculated using the continuum estimate and the instrumental noise, with the FWHM taken from Table 2 in T14 for each grating. A significance limit of 3 is used. The observed H I absorbers from the catalogue are scattered on the plot, assuming that the Ly- $\alpha$  line was observed. The vertical green lines divide the different gratings used in the spectrum, which therefore have different signal-to-noise and FWHM.

### 3.3 Results

In the series of Figures 8, 9 and 10, three plots are shown for each set of constraints applied to the samples, each with three panels. The top three panels show the velocity difference between each galaxy and its nearest-neighbour absorber in each sightline. The histogram shows the distribution of real velocity differences, with the black points giving the median number of galaxies in each bin in the random sets, and the cyan points showing the 99% level. Thus, the level of the histogram in relation to the points shows the excess of galaxy-absorber separations in that  $100 \text{ km s}^{-1}$  bin over the expectation if the absorber redshifts were randomly distributed. Note that, as the total number of galaxies is the same, any excess in bins with small velocity difference must be accompanied by a deficit in other bins.

The middle three panels show the number of galaxies around which at least 1, at least 2, or all 3 sightlines contain H I absorption within the velocity-cut. The histogram shows the distribution of the random sets, with the black vertical line giving the mean value. The red vertical line shows the number of galaxies found in the real system. Also given are the percentage of random sets in which the number of galaxies found with absorption in 1, 2 or 3 sightlines is greater than or equal to the number in the real Q0107 system, and the significance of the difference from the mean in units of the standard deviation of the random distribution. Therefore panels in which the red line lies to the right of the histogram show that there are more galaxy-absorber groups in the real Q0107 system than in the random distributions.

The final panels show the number of H I absorber groups for which absorption within the velocity window is found in precisely 1, 2, and 3 sightlines. The layout of the plot is as described above, with the percentage of random sets containing at least as many absorber

groups as the real system given alongside the significance of the difference. The uppermost of these three panels often shows the real system having fewer single-LOS absorber groups than the random distribution, as the real absorbers are more likely to form coherent structures across multiple lines-of-sight, and be more clustered along any one line-of-sight.

These figures are intended as a direct comparison with those in C10. However, with the much larger samples now available, it is possible to obtain results from subsamples of galaxies and absorbers. These include separating star-forming from non-star-forming galaxies, and dividing absorption systems into low- and high-column-density samples. This allows testing of numerous models of the links between the galaxies and surrounding gas, as described below.

#### 3.3.1 Full Sample

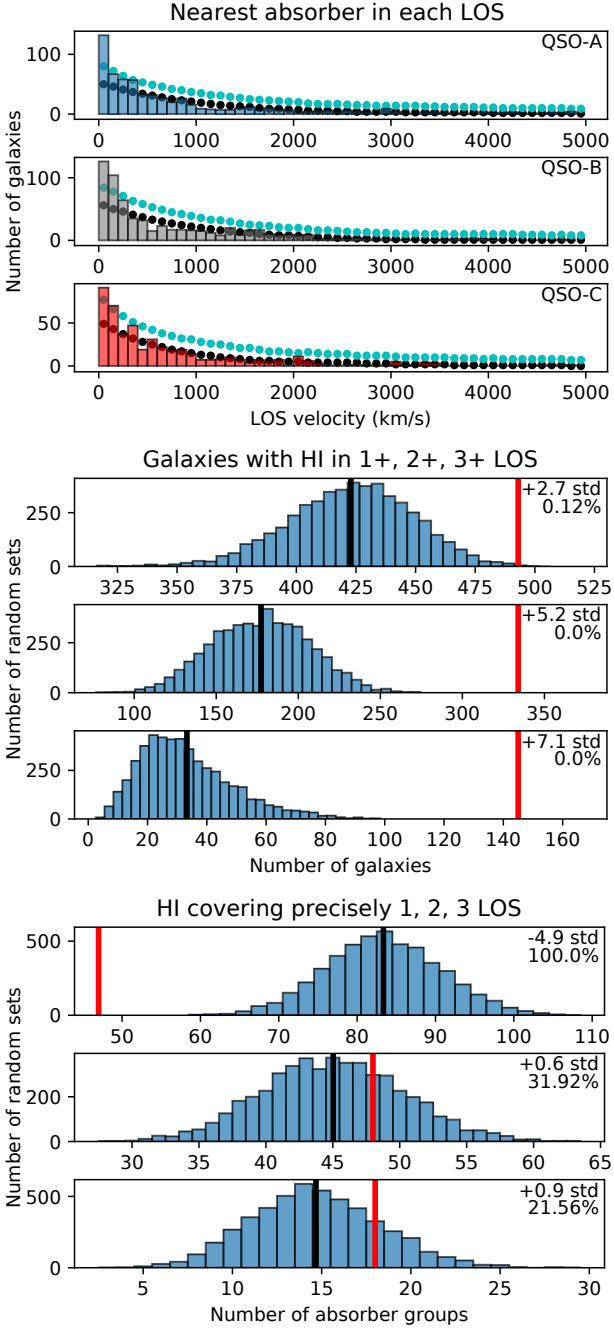
Figure 8 shows the results from the full sample, in which we consider all observed H I absorption systems, and all galaxies with ‘a’ and ‘b’ redshift flags within 2.5 Mpc of a sightline. The top panels show the velocity difference between each galaxy and its nearest-neighbour absorber in each sightline. For each sightline, the number of galaxies with absorption within  $100 \text{ km s}^{-1}$  (the innermost bin) is above the 99th percentile of random absorber sets, and remains consistently above the median out to  $400 \text{ km s}^{-1}$ . This suggests that most of the physical associations between galaxies and absorbers occur with smaller velocity differences. For most of this work, we use a velocity cut of  $500 \text{ km s}^{-1}$ , thus capturing most of the likely galaxy-absorber groups whilst minimizing the noise from unrelated pairs. This is also directly comparable to the grouping used by C10.

The middle panels shows the number of galaxies for which absorption systems are found in at least 1, at least 2, and all three sightlines within the  $500 \text{ km s}^{-1}$  window. In each case significantly more galaxies have absorption in the real Q0107 field than expected from the systems with randomly generated absorbers. Only six of the 5000 random distributions show as many galaxies with associated absorption, and no set of randoms has as many matches between a galaxy and multiple absorbers as the real Q0107. The significance of the excess also increases with the number of sightlines covered. This is similar to the results of P06 and C10 (their figure 16), in which there is a significant excess of galaxies associated with absorbers on these scales. The larger sample of absorbers and galaxies in this study has allowed a higher confidence level to be reached.

The lower panels show the number of absorber groups covering one, two and three sightlines respectively. As in C10 (the lowest panel of this plot is directly comparable to the middle panel of their figure 7) the number of triple-absorber groups is larger in the real system than in most sets of randoms. However, the excess is less significant here, with  $\approx 22\%$  of random sets showing more triples (as compared to  $\approx 11\%$  in the C10 results). This may be due to the improved sensitivity to low-column-density absorption. This observed absorption across all three sightlines within a small velocity window supports the idea that the gas is found in structures at least 500-1000 kpc in extent (the distances between the sightlines).

#### 3.3.2 Column Density

We then extend this test by cutting the absorber sample by column density, using a cut of  $N(\text{H I}) = 10^{14} \text{ cm}^{-2}$  identical to that used in T14, and perform this analysis on both the high- and low-column density samples. Figure 9 compares the results obtained from these two samples, with the left column giving results from the high-



**Figure 8.** Nearest-neighbour and velocity-window groups using the **full sample**, as described in Section 3.3.1. These compare the absorbers in the real Q0107 sightlines with the ensemble of randomly distributed absorber sets, showing line-of-sight velocity distances between galaxies and absorbers, number of galaxies with associated HI in one or more lines-of-sight, and number of absorber groups covering one or more lines-of-sight. Features are identical to those in Figure 9 and are described in detail there. *Top 3 panels:* The nearest absorber to each galaxy in each sightline. *Middle 3 panels:* The number of galaxies around which absorption is seen within  $500 \text{ km s}^{-1}$  in at least one, at least two and all three sightlines. *Lower 3 panels:* The number of absorber groups covering precisely 1, 2 and 3 sightlines within  $500 \text{ km s}^{-1}$ .

column-density sample, and the right column showing the low-column-density sample.

Using these constraints, the nearest-neighbour results (top) show a much clearer excess of galaxy-absorber pairs in the innermost velocity bins when only high-column-density absorbers are considered, substantially above the 99% level in all three sightlines. In the low-column-density case, the excess of small-velocity pairs is below the 99% level in two of the sightlines.

The excess of galaxies with associated absorption is more significant in the high-density sample (middle-left) than the low-column-density sample (middle-right). Indeed, there is no significant excess of galaxies with absorption in multiple sightlines in the low-density case. Similarly, the significance of the excess of real absorber triples is greater in the high-column-density case (bottom panels), and the low-column density observed coincidences are also consistent with the randomly generated sample.

These results confirm previous observations that high-column-density gas preferentially resides close to galaxies, in the CGM or intra-group medium, whereas the excess of low-column-density absorption around galaxies is less significant, suggesting these absorbers occur in the IGM. As only four absorber groups exist with high-column-density absorption in all three lines-of-sight, and these lie at the same redshift as more than 40 galaxies, these triplets correspond to galaxy overdensities.

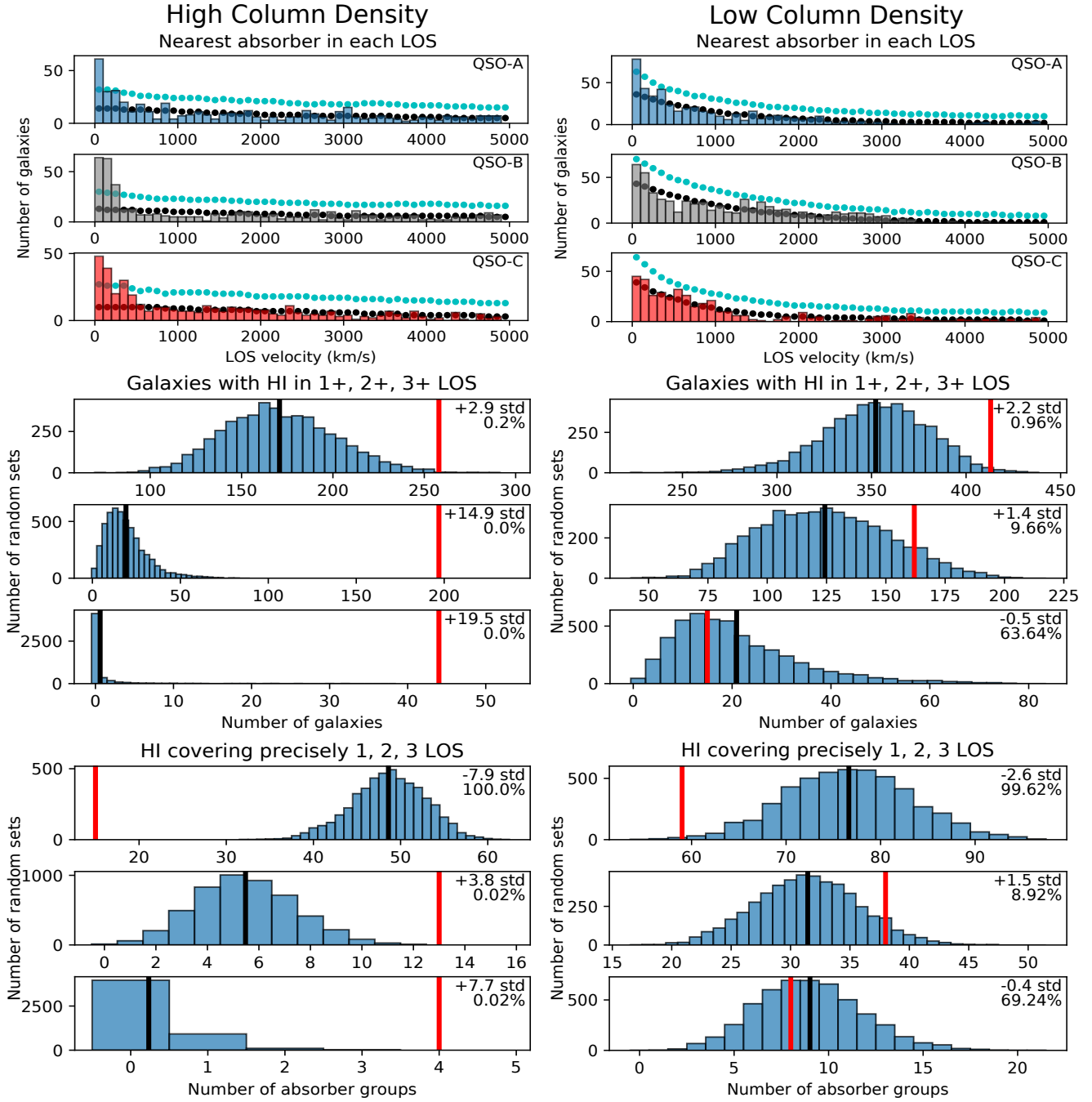
That low-column-density absorption is not found in multiple sightlines at a significantly higher frequency than in the random distribution may also indicate that these weak absorbers do not often form Mpc-scale structures.

These results are consistent with those in T14, in which the correlation function between weak absorbers and galaxies suggests that they do not trace the same dark matter distribution, and the low auto-correlation between low-column-density absorbers indicates that they rarely form large structures. Burchett et al. (2020) suggest that these column densities should be tracing the outer regions of filaments, as well as some overdensities in voids. This could lead to detection in two sightlines if the filament is aligned across two sightlines, but is unlikely to produce triplets, a possible explanation for the  $\sim 1.5\text{-}\sigma$  excess of low-column-density, two-sightline detections.

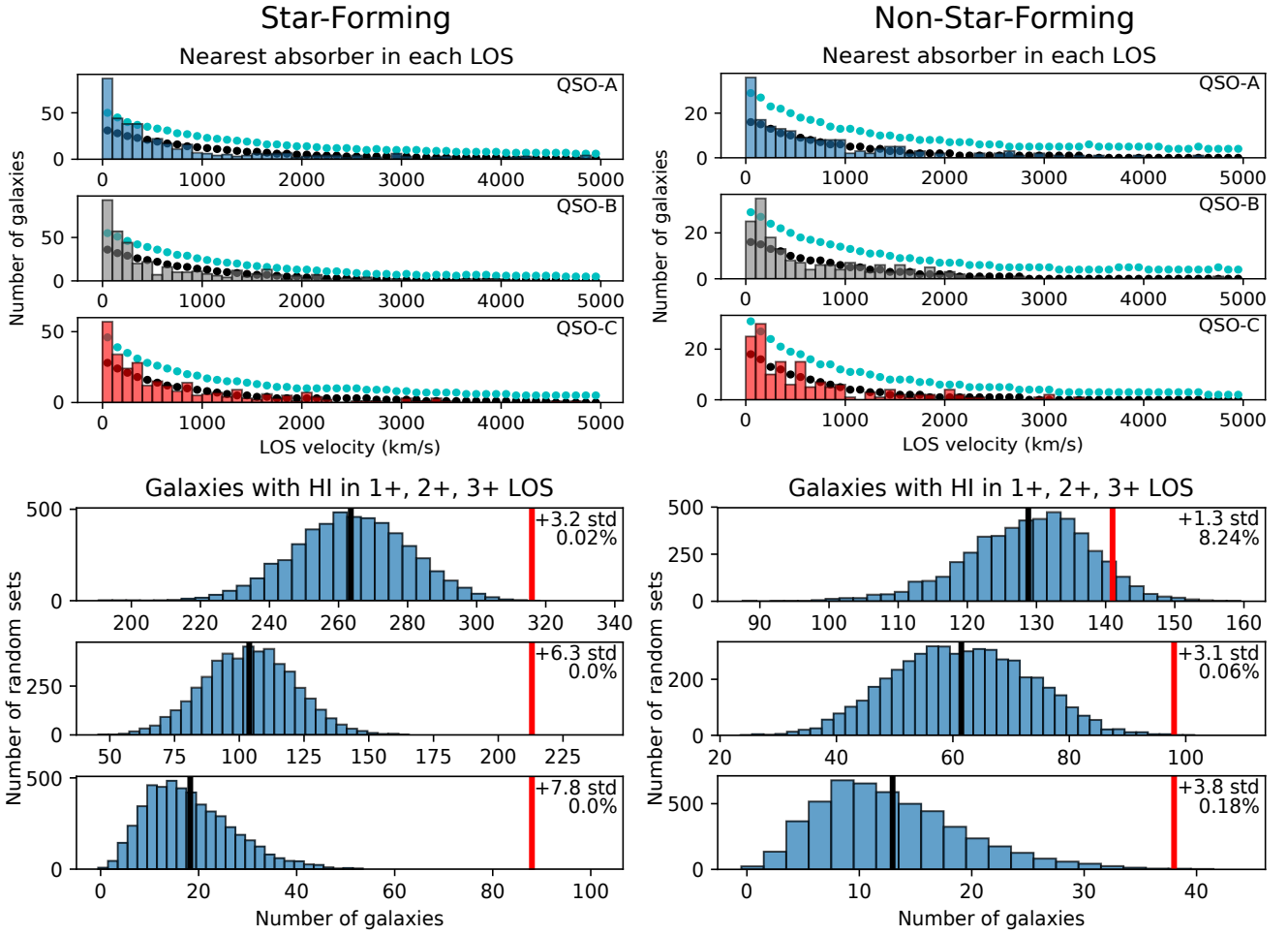
We also repeated this test after randomly discarding low-column-density absorbers until the sample sizes of high- and low-density absorption features were of equal size. There were no significant differences in the results, confirming that the greater excess seen in the high-column-density case is not merely an effect of the larger sample size. Due to the higher detection limit in the FOS spectra, there is also a difference in redshift between the low- and high-column-density absorbers, with low-column-density systems rarer at  $z \geq 0.48$  (median redshift of low- and high-column-density absorbers 0.32 and 0.52 respectively). However, the results are similar if we only include absorbers found in the COS spectra. Redshift evolution of the IGM is found to be slow at  $z \leq 1$  (e.g Kim et al. 2020), so no substantial difference is expected.

### 3.3.3 Star Formation

Another cut was made on the star-formation of galaxies in the sample (using the star-formation class for each galaxy defined in Section 2.3.5). Figure 10 shows these results. The excess of groups in the real system is more significant in the star-forming case (lower-left) than the non-star-forming case (lower-right). This can also be seen in the nearest-neighbour matching (upper-left and upper-right), in



**Figure 9.** Comparison between the real Q0107 system and the ensemble of randomly distributed H I absorber sets, using the **column density cut**, as described in Section 3.3.2. We show galaxy and absorber groups involving galaxies with ‘a’ or ‘b’ redshift flags within 2.5 Mpc of at least one line-of-sight, considering only high-column-density ( $N(\text{H I}) > 10^{14} \text{ cm}^{-2}$ ) absorbers in the left column, and lower-column-density absorbers in the right column. The features shown are identical to those in Figures 8 and 10. *Top 3 panels:* The nearest absorber to each galaxy in each sightline, in 100  $\text{km s}^{-1}$  bins. The histogram shows the velocity difference between each galaxy and the nearest absorber in the sightline given in the top-right of each panel, in 100  $\text{km s}^{-1}$  bins. The black and cyan points show the 50th and 99th percentiles of the random distributions respectively, illustrating the excess of galaxies within 300–500  $\text{km s}^{-1}$ . Note that each galaxy must appear once in each of the three panels, so an excess in the inner bins must be accompanied by a deficit in the bins with higher velocity separations. *Middle 3 panels:* The number of galaxies around which absorption is seen within 500  $\text{km s}^{-1}$  in at least one, at least two and all three sightlines. The histograms show the distribution of the random sets, with the black vertical line showing the mean, and the red vertical line shows the number observed in the real Q0107 system. In the top-right we state the number of galaxies with associated H I found in the real system relative to the mean of the distribution of randomized sets (in units of the standard deviation of the random distribution), and the percentage of random sets for which the number of galaxies with associated H I absorption is greater than or equal to the number in the observed system. *Lower 3 panels:* The number of absorber groups covering precisely 1, 2 and 3 sightlines within 500  $\text{km s}^{-1}$ , with features as in the middle three panels. The red vertical line again shows the number of absorber groups covering 1, 2 or 3 sightlines in the observed Q0107 field. This lies further to the left in the upper of these panels for more strongly clustered absorbers, as absorbers are more likely to be coincident with those in another line-of-sight.



**Figure 10.** Nearest-neighbour and velocity-window groups using the **star-formation cut**, as described in Section 3.3.3. These compare the the real Q0107 system with the ensemble of randomly distributed absorber sets, showing line-of-sight velocity distances between galaxies and absorbers, number of galaxies with associated H I in one or more lines-of-sight, and number of absorber groups covering one or more lines-of-sight. Features are identical to those in Figure 9 and are described in detail there. *Top 3 panels:* The nearest absorber to each galaxy in each sightline. *Lower 3 panels:* The number of galaxies around which absorption is seen within 500 km/s in at least one, at least two and all three sightlines. The absorber-only groups are omitted as these are not affected by the cut, so the results are the same as in the lower panel of Figure 8.

which the excess of galaxies in the innermost bins is greater in the star-forming sample.

These results show absorption is more likely to be detected around star-forming galaxies on scales of hundreds of kpc. This is similar to the results from (Chen & Mulchaey 2009), in which the correlation between Ly $\alpha$  absorbers and galaxies is stronger among emission galaxies than absorption galaxies. Our results are generally consistent with those in T14, despite the different approach. They do not find a significant difference in the correlation slope or length between the strong-H I/SF-galaxy and strong-H I/non-SF-galaxy cross-correlations. However, they show that the strong/SF results are consistent with linearly tracing the same dark matter distribution, whilst the strong/non-SF results are not, due to the greater auto-correlation between non-SF galaxies. Therefore the probability of finding a strong H I absorber near a non-SF galaxy is lower than that of finding a strong H I absorber near a SF galaxy, in agreement with our results.

Whilst the star-forming galaxies in our sample are more likely to be surrounded by detectable levels of H I gas, it is difficult to determine the cause of this. We note that excess absorption within

the virial radius is not sufficient to explain the excess of two- and three-sightline absorption, as the distance between the sightlines is larger than the virial radius for most of our galaxies.

There are suggestions from simulations that stellar-feedback-driven outflows may extend to and beyond the virial radius (e.g. DeFelippis et al. 2020; Mitchell et al. 2020a; Hopkins et al. 2021), so their presence around star-forming galaxies could lead to the observed absorption. If outflows make a substantial contribution to this excess absorption, we may expect to see a broader distribution of galaxy-absorber velocity offsets in the SF sample, arising from the line-of-sight component of the outflow velocity (see e.g. Chen et al. 2020b). This broader distribution is not seen in our data, which would suggest that outflows do not make a large contribution on these scales. However, there are potential redshift errors in the VIMOS galaxies in the same 100–200 km s $^{-1}$  regime as the likely outflow velocities (see Table 5), which could similarly broaden the distribution. We further discuss the presence of outflows in Section 4.

We consider the environments in which the SF and non-SF galaxies are likely to be found. Non-SF galaxies are more likely



to reside in groups (41% of non-SF and 32% of SF galaxies lie in groups of five or more galaxies, using the friends-of-friends algorithm described in Appendix A), as expected due to quenching when galaxies fall into larger haloes (e.g. [Wetzel et al. 2013](#)). As these groups will often correspond to larger overdensities in the cosmic web than single galaxies, we may expect increased incidence of H I absorber groups around non-SF galaxies, the opposite effect to that observed. However, the higher virial temperatures of group haloes would lead to a suppression of neutral hydrogen, possibly negating this effect.

We also consider the effects of biases in our sample selection (as discussed in Section 2.3.5). Firstly, to remove any effects of having a larger star-forming than non-star-forming sample, the analysis was rerun after randomly discarding star-forming galaxies until the sample sizes were identical, and the significance of the excess in galaxy-absorber groups was only marginally reduced in each case. Whilst there are small biases in the mass and impact parameter distributions of the two samples, a Kolmogorov-Smirnov (K-S) test cannot distinguish them at the 5% level, and they appear localised to the mass tails and the smallest impact parameters (i.e. the MUSE fields). Removing the mass tails and/or MUSE-only galaxies again slightly reduces the significance of the excess, but does not substantially affect the comparison.

The bias of SF galaxies towards higher redshifts does appear to have an effect, but this is found to be primarily due to the absorber selection function removing lower column density absorbers at higher redshifts (as seen in Figure 7). If only high-column-density absorbers are considered, no substantial difference between high- $z$  and low- $z$  coherence is found, but there remains a larger excess of absorption around star-forming galaxies than non-star-forming galaxies.

The difference between our SF and non-SF samples is therefore unlikely to have arisen from any sample biases, and instead does indicate that star-forming galaxies are more likely than quiescent galaxies to exhibit H I absorption covering single and multiple lines-of-sight.

## 4 GALAXY ORIENTATIONS

As described in Section 2.2 and shown in Figure 2, we have obtained MUSE data for the  $1' \times 1'$  fields around QSOs A and B, in addition to high-resolution HST imaging of a larger field (but not extending as far as QSO-C). We used GALFIT ([Peng et al. 2002](#)) to determine the position angles of galaxies in the HST field, as described in Section 2.4. This allows galaxy-quasar pairs in which absorption is detected (and those where it is not) to be shown as a function of position angle and impact parameter. We can then determine whether absorption lies preferentially along the major or minor axis of the galaxy, and identify possible co-rotating material and polar outflows.

The MEGAFLOW survey mostly covered impact parameters out to  $\sim 100$  kpc using Mg II, focusing their selection on sight-lines with high-equivalent-width absorbers in their SDSS spectra ([Schroetter et al. 2016](#)). They find a clear bimodality in position angle, with more absorbers found along both the major and minor axes, identified with discs and outflows respectively. It must be noted that they remove absorbers found with three or more nearby galaxies, and assign a single galaxy to the absorber in cases with two nearby galaxies.

Similarly, [Bordoloi et al. \(2011\)](#) and [Bordoloi et al. \(2014\)](#) look for Mg II around zCOSMOS galaxies, finding strong minor-

axis absorption attributed to outflows that is much stronger than the major-axis absorption, and an increasing outflow equivalent width with host galaxy SFR. They only find this for impact parameters  $\lesssim 50$  kpc, but also include group galaxies in their analysis, obtaining results consistent with a superposition of outflows from group members.

On the other hand, [Dutta et al. \(2020\)](#) find that their sample, primarily consisting of Mg II absorption at larger impact parameters, does not show such a bimodality (a similar result is found by [Huang et al. 2021](#)), nor can their measurements in group environments be fully explained by a superposition of absorption from individual group members. This suggests that, at least in Mg II, the disk/outflow dichotomy does not extend to scales much beyond  $\sim 50$  kpc.

This appears in contrast to some recent models. For example, [Hopkins et al. \(2021\)](#) find that their simulations (based on the FIRE simulations, with the addition of cosmic ray effects) can allow biconical outflows to reach megaparsec scales. Outflows in the EAGLE simulations are also seen to maintain their bi-directional structure to at least the virial radius of  $\sim 10^{12} M_{\odot}$  halos ([Mitchell et al. 2020a](#)), whilst Illustris also features hot outflowing material and cool co-rotating material along the minor and major axes out to close to the virial radius ([DeFelippis et al. 2020](#)).

Using H I absorption from lines-of-sight with no pre-selection allows an unbiased sample to extend to larger impact parameters than most Mg II studies. When considering galaxies in groups, we avoid artificially selecting the associated galaxy to each absorber, instead including all galaxy-absorber pairs within the  $500 \text{ km s}^{-1}$  cut. We test for this bimodality in position angle for the galaxy-absorber pairs in our sample using the Hartigan dip test ([Hartigan & Hartigan 1985](#)), which calculates the likelihood of observing the ‘dip’ in the sample histogram if the underlying distribution is unimodal. As in Section 3, we also compare the results when the sample is split into complementary subsamples such as high- and low-column-density absorbers, and star-forming and non-star-forming galaxies.

### 4.1 Hydrogen

The results from applying this test to the distribution of galaxy-absorber pairs involving H I absorption are summarized in Table 6. Where the sizes of the complementary samples (paired using horizontal lines in the table) are substantially different, we randomly discard detections from the larger sample until the sizes match in order to perform a fair comparison. We repeat with 100 random samples and take a median, giving the results in brackets.

#### 4.1.1 Full Sample

Figure 11 shows the position angle against impact parameter for the full sample of galaxies and absorbers, associating each galaxy with all absorbers within  $500 \text{ km s}^{-1}$ . Non-detections are shown on the scatter plot, but only the position angles and impact parameters of detected absorption are included in the histograms. There is a visible bimodality, and the dip test finds a significant result, returning a p-value of 0.025. This reproduces in H I the bimodality obtained using Mg II in [Zabl et al. \(2019\)](#) and [Martin et al. \(2019\)](#), although extending to larger impact parameters, suggesting that some fraction of our observed H I is tracing the same outflowing and accreting material. We note that the major axis absorption we observe covers position angles  $\lesssim 40^{\circ}$ . It is not simple to distinguish between a thin disk viewed at moderate inclination angles, and a thicker ‘wedge’ of material.

Some studies (e.g. Tempel & Libeskind 2013; Zhang et al. 2015) suggest that galaxy spins are preferentially aligned with or perpendicular to the surrounding large-scale structure. Observing the cosmic web around these galaxies, as traced by H I, could produce a bimodality. This alignment is a weak effect (in Tempel & Libeskind 2013, galaxies are a maximum of  $\approx 20\%$  more likely to be aligned over a random distribution), so unlikely to fully explain our stronger observed bimodality.

The overall position angle distributions of galaxies and galaxy-sightline pairs are consistent with uniform, with neither the dip test nor a K-S test against a uniform distribution showing a significant result. Therefore our bimodality is not due to an inherent alignment between galaxies in our sample. It is instead most likely due to the inflow/outflow dichotomy discussed above.

#### 4.1.2 Galaxy Groups

Cutting the sample to only include galaxies in groups of five or more decreases the significance of the bimodality to  $\approx 10\%$ , whereas no bimodality is seen among the non-group galaxies (using the friends-of-friends algorithm described in Appendix A).

The significance of the bimodality is expected to be reduced in galaxy groups. If the gas is primarily a superposition of outflows and/or accretion from individual galaxies, the signal would still be partially masked by other galaxies being paired with the same absorber. If the gas does not form these structures within most groups, and instead forms an intra-group medium not attributable to a single galaxy, this bimodality should not be visible.

The bimodality we see in the in-group sample is not significant at the 5% level, and the significance is further reduced when galaxy-absorber pairs are randomly discarded until the group and non-group samples contain the same number of pairs. Our results are consistent with the difference in the p-values shown in Table 6 being primarily due to sample size, as neither sample shows a significant bimodality at the 5% level, yet the combined sample does. A K-S test also fails to find a significant difference between the position angle distributions of galaxy-absorber pairs of these two samples.

No significant effect on the bimodality or the position angle distribution is found when using shorter linking lengths, when adjusting the minimum number of galaxies needed to constitute a group between three and five, or when splitting the non-group galaxies into those with no detected neighbours and those with 1-3 neighbours.

#### 4.1.3 Impact Parameters

When the sample is cut to absorber-galaxy pairs with an impact parameter of less than 500 kpc, the bimodality remains strong, a Hartigan dip test returning a significance of 3%, whereas the pairs with an impact parameter larger than 500 kpc exhibit no significant bimodality. The difference between the large- and small- impact parameter results may suggest that the model of minor axis outflow and major axis accretion can extend well beyond the  $\sim 100$  kpc observed in the MEGAFLOW results (Schroetter et al. 2016). We confirm that this bimodality is not entirely driven by galaxy-absorber pairs on small scales, obtaining a significant result from pairs with impact parameters of 200–500 kpc.

This is further illustrated in Figure 12, in which the galaxy-absorber pairs are binned by position angle and impact parameter. For the two innermost bins in impact parameter ( $< 500$  kpc), the  $30\text{--}60^\circ$  bin has a clear lack of absorption relative to the major axis and

**Table 6.** Summary of constraints applied to the sample and resulting bimodalities in the position angle distributions of galaxy–H I absorber pairs. Columns show: (1) The subsample used, (2) the number of possible galaxy–sightline pairs, accounting for the reduced redshift range for which QSO-C can be observed as well as the additional constraints, (3, 4) the number of galaxy–sightline pairs with and without observed H I absorption within  $\Delta v < 500 \text{ km s}^{-1}$ , and (5) the p-value obtained from a Hartigan dip test applied to the position angle distribution of galaxy–absorber pairs with detected absorption. Where the sizes of complementary samples are substantially different, we randomly discard detected galaxy-absorber pairs from the larger sample until the sizes match, repeating 100 times and showing the average result in brackets.

Constraint (1)	Pairs (2)	Det (3)	Non-Det (4)	P-value (5)
Full Sample	289	242	47	0.025
In-Group	190	173	17	0.091(0.281)
Non-Group	99	69	30	0.511
$r < 500$ kpc	154	137	17	0.031
$r > 500$ kpc	152	105	47	0.185
$r < 300$ kpc	93	81	12	0.030
$r > 300$ kpc	208	161	47	0.399(0.393)
$200 < r < 500$ kpc	101	84	17	0.033
High-N(H I)	245	128	117	0.009
Low-N(H I)	229	114	115	0.709
Star-forming	187	155	32	0.123(0.224)
Non-SF	83	73	10	0.284

minor axis bins. The third radial bin ( $\approx 500\text{--}750$  kpc) does not show this bimodality, suggesting that IGM absorption or other structures such as group material form the dominant component at this scale. Beyond this distance, the sample size is very small. We note that few conclusions can be drawn from the radial distribution, as this depends primarily on the geometry of the sightlines and redshift surveys. (The apparent ‘edge’ at  $\approx 750$  kpc is an artifact of this layout, as this is roughly the maximum distance from the A and B sightlines to the edge of the HST imaging. Most galaxy–absorber pairs beyond this distance involve absorption in sightline C.)

#### 4.1.4 Column Densities

Cutting to column densities  $N(\text{H I}) > 10^{14} \text{ cm}^{-2}$  improves the significance of the bimodality to better than 1%, as shown in Figure 13, whereas no clear bimodality is found for the low-column-density absorbers. These two subsamples show the strongest and weakest bimodalities according to the dip test results, suggesting that high-column-density gas is preferentially found in these putative inflows and outflows. This likely captures much of the same physics as the variation with impact parameter, as high-column-density absorbers are generally found closer to galaxies, for example in Chen (2012) and Keeney et al. (2017). Wilde et al. (2021) find that the probability of finding H I above our column density threshold around a galaxy drops to 50% at impact parameters of  $\approx 300$  kpc, similar to the extent of our bimodality.

As discussed in Section 3.3.2, the high- and low-column density samples have different redshift distributions, but this does not have a substantial impact on the results. We confirm that the bimodality is retained for the high-column-density absorbers in the

COS gratings, removing the FOS absorbers at higher redshifts where low-column-density absorbers are not detected ( $p=0.018$ ).

#### 4.1.5 Star Formation

We also consider the bimodality around star-forming and non-star-forming galaxies. Neither subsample shows a significant bimodality in position angle of absorption ( $p$ -values 0.12 and 0.28 respectively). This appears to be primarily due to the sample sizes. When star-forming galaxies are randomly discarded from the sample until the number of detections around the star-forming and non-star-forming galaxies are equal, the resulting  $p$ -values are similar between the two samples. A K-S test also fails to find a significant difference between the position angle distributions of absorbers around star-forming and non-star-forming galaxies.

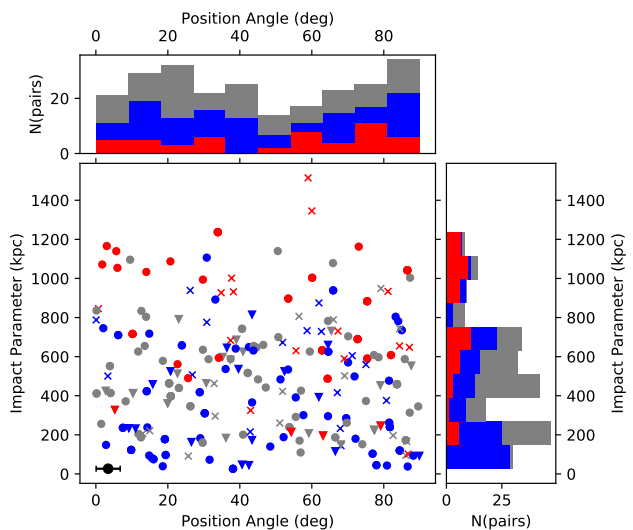
It is expected that more strongly star-forming galaxies are likely to have stronger outflows (e.g. Mitchell et al. 2020a). Our star-forming classification may be including many galaxies with SFRs too small to launch large-scale outflows, thus reducing the strength of the observed bimodality. We attempt to test this by applying a cut in sSFR (using the estimates described in Section 2.3.5) instead of our binary classification. This results in a bimodality significant at the 5% level in the strongly star-forming sample when using a threshold between  $\approx 0.05$  and  $0.1 \text{ Gyr}^{-1}$ . Higher thresholds leave the star-forming sample too small to obtain a significant result, and lower thresholds give similar results to our original classification.

This can be discussed in the context of the star-formation comparison in Section 3.3.3. Whilst our sSFR estimates have high uncertainties, our finding that more strongly star-forming galaxies show a stronger bimodality is likely an indication that stellar-feedback-driven outflows are a contributor to this bimodality we observe in the position angle distribution of H I absorbers around galaxies on scales  $\lesssim 300$  kpc. Our result that outflows on these scales are not evidenced by a bimodality around our more inclusive sample of star-forming galaxies, yet larger-scale coherent structures are observed (on the 500-1200 kpc scales probed by coincidences between the lines-of-sight), suggests that these larger-scale structures are not primarily a result of outflows, but are instead a consequence of the environment around these galaxies.

#### 4.1.6 Inclination

We note that when an inclination cut of  $i > 40^\circ$  (identical to that used in the MEGAFLOW survey, Zabl et al. 2019) is used to remove face-on galaxies for which the major and minor axes cannot easily be distinguished, the same subsamples show significant bimodalities at the 5% level as those in Table 6. For randomly oriented galaxies,  $\cos(i)$  is expected to be uniform, so the distribution of inclinations is not uniform, but is instead suppressed at low inclinations (face-on galaxies). This leads to only 8 of the 72 galaxies considered throughout this section having inclinations less than  $40^\circ$ , so it is unsurprising that the results do not change. We also note that a small number of galaxies, although fit well by the GALFIT modelling, have large uncertainties on their position angle (5 have position angle uncertainties  $> 20^\circ$ ). Excluding these galaxies also has no effect on which subsamples show bimodalities significant at the 5% level.

We briefly consider variation with inclination, by dividing the  $i > 40^\circ$  sample into two bins of 40-65 and 65-90 $^\circ$  (which contain 39 and 25 galaxies respectively). Neither sample shows a clear bimodality, and both are consistent with the overall distribution



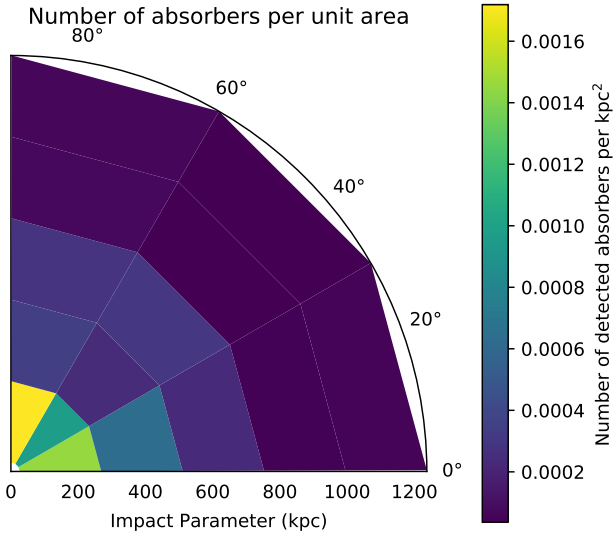
**Figure 11.** The position angle and impact parameter distributions of galaxy-absorber pairs involving H I absorption features around all galaxies in the sample for which position angles were measured. Position angles close to  $0^\circ$  represent absorption found close to the projected galaxy major axis, whilst those close to  $90^\circ$  show absorption found near the minor axis. Points are coloured by quasar sightline (A, B and C shown in blue, grey and red respectively), using triangles for galaxies in groups of five or more galaxies and circles for non-group galaxies, and non-detections are marked with crosses. Only the detected absorbers are included in the histograms. The bars on the black point in the lower-left are illustrative of the median error in each axis.

using the K-S test. Interestingly, the two bins are not consistent with each other ( $p \approx 0.04$ ). This appears to be mostly driven by galaxy-absorber pairs with impact parameters in the 250-500 kpc bin.

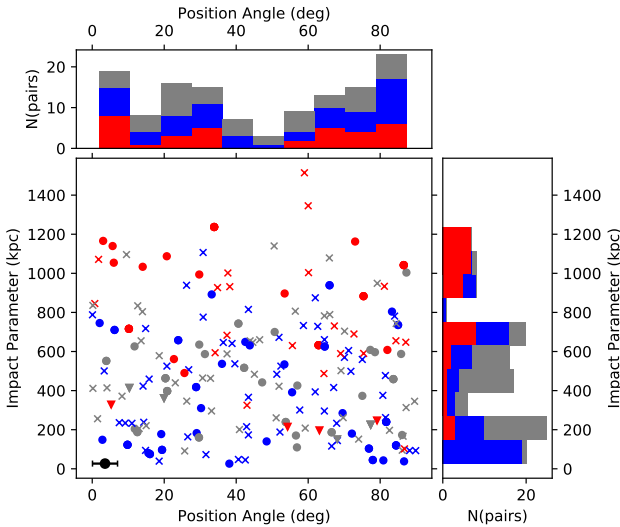
We find that 26 of the 39 pairs in this impact parameter range involving ‘intermediate-inclination’ galaxies lie in the 0-30 $^\circ$  major axis bin in position angle, whilst 14 of the 23 pairs involving  $i > 65^\circ$  galaxies lie in the 60-90 $^\circ$  minor axis bin in position angle. Whilst these sample sizes are relatively small, this perhaps supports the presence of a disk-like structure along the major axis with a small cross-section when viewed close to edge-on.

#### 4.1.7 Closest Galaxies

We note that if we reduce our sample to only the closest galaxy in impact parameter to each absorber (within the  $500 \text{ km s}^{-1}$  cut), the bimodality becomes somewhat stronger, with significance improving from  $\approx 2.5\%$  to  $\approx 1.0\%$ . This will capture the physical origin of the absorbing material in many cases, but does not appear to in all cases, as the galaxy-absorber pairs removed do themselves show some hint of bimodality (with a significance of  $\approx 7\%$ ). Splitting the sample of galaxy-absorber pairs involving only the closest galaxy produces results similar to those in Table 6, with no sub-sample crossing the 5% threshold. We do not take this approach through our earlier analysis as the cut appears to remove some physical associations and substantially reduces the sample size in some cases. This makes detecting a bimodality more difficult, especially in the in-group, SF and non-SF sub-samples.



**Figure 12.** The position angles and impact parameters of H I absorbers detected around galaxies, shown in bins of  $\sim 250 \text{ kpc} \times 30^\circ$ . In this representation the galaxy major axis lies along the x-axis. Colour scale shows number of detections per projected  $\text{kpc}^2$ . Note that the radial distribution is determined primarily by the geometry of the survey: with QSOs A and B lying near the centre of the HST field, and QSO-C outside the field, only absorbers in the line-of-sight to QSO-C can be found beyond 800 kpc.

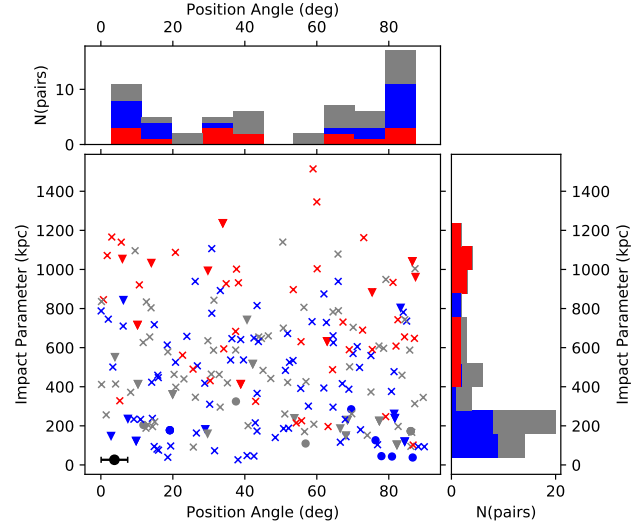


**Figure 13.** The position angle and impact parameter distributions of galaxy-absorber pairs involving H I absorption features with column densities  $N(\text{H I}) > 10^{14} \text{ cm}^{-2}$ . Features are as in Figure 11.

## 4.2 Metals

We briefly discuss here the presence of metals in our absorption-line sample. Lines in the spectra exist that are identified with numerous ions, including C I-C IV, Si II, Si III, O I-O IV and O VI. Only O VI forms a significant sample, with 34 lines identified across the three spectra.

Figure 14 shows the resulting distribution of position angles and impact parameters. The bimodality is again visible by eye, and



**Figure 14.** The position angles and impact parameters of detected O VI absorption features around all galaxies with measured position angles. Features shown are as in Figure 11.

statistically significant ( $p \sim 0.009$  from the Hartigan dip test). It must be noted that all non-group detections occur at impact parameters of less than 350 kpc (postage-stamp images of the 6 non-group galaxies with detected O VI absorption are shown in Appendix C). It may also be suggested that the minor-axis ‘peak’ is stronger relative to the major-axis than in the H I figures. This is in general agreement with the results from Kacprzak et al. (2015), in which a bimodality is also found around isolated galaxies, with higher O VI equivalent widths along the minor axis.

We also compare the column-density distributions of O VI detections in galaxy groups with those not in groups. A K-S test indicates that these distributions are likely different (98% confidence), with a substantial number of in-group absorbers observed with lower column density.

This is again consistent with the disk/outflow model at small scales, whereas all metal absorbers found at large impact parameter from galaxies occur in groups. This could be due to absorber-galaxy pairs involving an absorber associated with a different (possibly undetected) galaxy in the same group, virialized gas within the group halo, or due to tidal and intra-group material within the group. Whilst observations of this material in emission (e.g. Fossati et al. 2019a) allow us insight into the structure of this material on small scales, the lower column densities found on larger scales cannot easily be probed in this way. Morris & van den Bergh (1994) found that tidal absorbers extending to 1 Mpc radius would be consistent with observations. Unfortunately, this is approximately the limit of impact parameters for which position angles are available, so we cannot test whether this O VI extends to larger radii.

Oppenheimer et al. (2016) predict that the presence of O VI should be suppressed in group haloes due to a higher virial temperature than  $L_\star$  galaxy haloes, reducing the fraction of O VI in favour of higher ions. A similar suppression is found in Wijers et al. (2020) for high-mass haloes.

Alternatively, observations and radiative transfer modelling from Werk et al. (2014) suggest that the warm gas phase is not supported hydrostatically in  $L_\star$  galaxies, favouring a galactic ‘fountain’ over a virialized halo. The strong bimodality we observe in our data

favours the galactic fountain for non-group galaxies. That the scale of O VI in non-group galaxies is similar to the extent of the observed bimodality in H I also supports the existence of such fountains extending to  $\sim 300$  kpc. The apparent shift in the ratio between the major- and minor-axis peaks, towards the minor axis, further fits this picture, with warmer, higher-metallicity outflows exhibiting more O VI than cool accreting flows.

However, our sample contains relatively few isolated  $L_{\star}$  galaxies (see Figure 16), with most substantially smaller or lying in groups of galaxies. Either case moves the virial temperature of the halo away from the  $\sim 10^{5.5}$  K at which O VI is most prominent, making these virialized haloes harder to observe. Therefore, whilst a basic model without such haloes is consistent with the data, producing O VI from fountains and group interactions, we cannot rule out their presence. With our small number of isolated  $L_{\star}$  galaxies, we cannot detect any ‘peak’ in absorber number or column density around  $L_{\star}$  galaxies.

We also note that there are four redshifts at which O VI absorption occurs in two sightlines within  $500 \text{ km s}^{-1}$ , with no three-sightline O VI systems. In one of these systems, both sightlines featuring O VI lie within  $30^{\circ}$  and 200 kpc of the minor axis of an isolated  $1.5 L_{\star}$  galaxy. The other three systems exist in galaxy groups, where it is difficult to distinguish between tidal debris, virialized haloes, and warm outflows. More detailed ionization modelling of the absorption in two of these groups are described in Muzahid (2014) and Anshul et al. (2021).

### 4.3 Comparison between major and minor axes

We now compare the properties of absorption along the major- and minor- axes. We take two samples of galaxy-absorber pairs, with position angles  $< 40^{\circ}$  and  $> 50^{\circ}$  respectively. We remove face-on galaxies by requiring a galaxy inclination of  $> 40^{\circ}$  and impose a 500 kpc maximum impact parameter to focus on the regions in which the bimodality is significant<sup>4</sup>. This returns 67 galaxy–H I pairs in each sample.

A K-S test reveals no difference between the column density or Doppler width distributions of these two samples. We also compare the galaxy–absorber velocity offsets around moderately inclined and edge-on galaxies ( $40 - 65^{\circ}$  and  $> 65^{\circ}$  inclination respectively), finding no significant difference. If our observed position angle distribution is primarily due to major axis co-rotation and minor axis outflows, we may expect to see a wider velocity distribution along the minor axis of moderately inclined galaxies and the major axis of edge-on galaxies, due to the larger line-of-sight velocity components. This is not observed in our data, although the further split results in a small sample size.

However, 14 of the major-axis absorbers have associated O VI, whilst 28 of the minor-axis absorbers show this association. A similar ratio is found when using a 300 kpc maximum impact parameter. This difference in O VI incidence is found in Kacprzak et al. (2015) and Kacprzak et al. (2019), but is attributed to a higher H I column density rather than metallicity differences. We do not find any significant increase in H I column density along the minor-axis, so

<sup>4</sup> These cuts in inclination and azimuthal angle cover an identical region to the MEGAFLOW survey for the major axis sample (Zabl et al. 2019), whilst the slightly larger region along the minor axis is used both for symmetry and to produce a sample of identical size. The major axis cut is also identical to that shown in Figure 15, although we do not require observed kinematic data in this case.

the increased O VI found near the minor-axis may be due to more warm-hot material than in the absorbers found near the major axis.

## 5 KINEMATICS

We attempt to determine whether the major-axis population discussed above is indeed the same co-rotating material identified in Mg II (e.g Zabl et al. 2019), as well as found in simulations (e.g Huscher et al. 2021; DeFelippis et al. 2020). We compare the velocity offsets between galaxies and H I absorbers found near to their major axes with the stellar rotation of the galaxies determined from emission lines in the MUSE data.

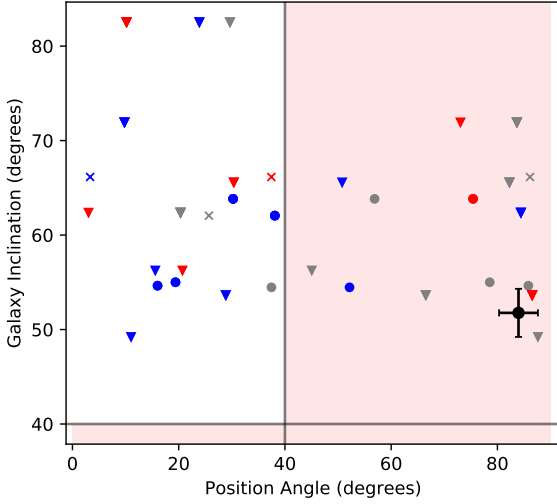
We use Astropy (Astropy Collaboration et al. 2018) to fit Gaussian line profiles to each spaxel within the aperture defined by SExtractor. In the case of [O II], we fit a double Gaussian with equal line width and redshift-dependent difference in central wavelength, and an intensity ratio between 0.1 and 10 to avoid fitting unphysically large or small values. We use kinematic maps from [O II] 3727, [O III] 5007 and H $\alpha$  6563, and generally consider the strongest line available for each galaxy.

We consider galaxies in which a velocity gradient is visible in the MUSE data. For comparison with Zabl et al. (2019) we only consider galaxies with inclinations above  $40^{\circ}$ , and position angles on the sky below  $40^{\circ}$ , although all galaxies with observed velocity gradients already fit this inclination constraint. A velocity window of  $500 \text{ km s}^{-1}$  is used, as we showed in Section 3.3.1 that this includes most of the associated absorption whilst reducing the noise from unassociated IGM absorbers. The position angles and inclinations of the resulting subsample are shown in Figure 15, whilst the locations of the galaxies in redshift-luminosity space are shown in Figure 16.

Using the position angle found from GALFIT as described in Section 2.4, we divide the galaxy into two regions, one each side of the projected minor axis. The median velocity of spaxels within each region then gives a redshifted and a blueshifted region. We then compare this to the velocity difference between the galaxy and absorber to identify possible co-rotation. As the subsample is small (22 galaxies with velocity gradients, of which 15 feature major-axis absorption), we also confirm by eye that this algorithm produces the correct result.

We first consider only galaxies not lying within groups of five or more galaxies, as any co-rotating structures are less likely to be disturbed by galaxy interactions in these cases. The magnitude and direction of the velocity offset relative to the rotation of the galaxy for each system is illustrated in Figure 17. All five of the galaxy–absorber pairs meeting the above constraints and not within groups show absorption offset in the same direction as the galaxy rotation. The probability of all five absorbers that are not consistent with zero (within  $1\sigma$ ) being offset in the same direction if the direction is random is 3%. A weighted mean of the velocity differences also gives a significant result ( $105 \pm 40 \text{ km s}^{-1}$ ). We note that where multiple absorption components are observed in the same spectrum within  $500 \text{ km s}^{-1}$  of one of these galaxies, each component is counted separately and shown separately in the figures.

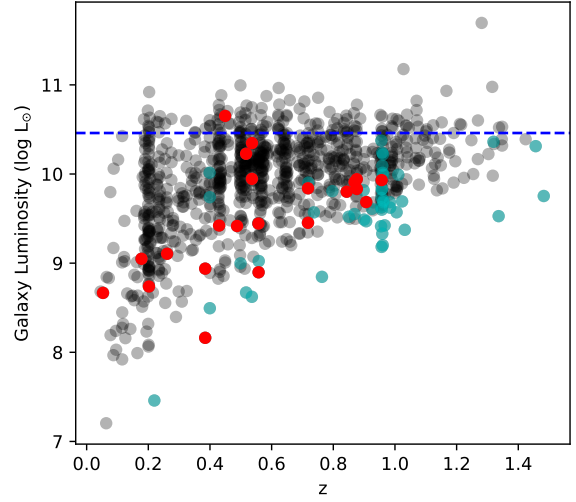
When galaxies in groups are also included, the major-axis H I absorption still shows a tendency to align with the galaxy rotation when using a weighted mean (Figure 18), with a  $\approx 2.5\sigma$  result of ( $74 \pm 27$ )  $\text{km s}^{-1}$ . The binomial calculation returns a p-value of 0.3 (the probability that at least 8 of the 13 absorber velocities not consistent with zero would be aligned with the galaxy rotation if there were no physical link between the galaxy and absorber).



**Figure 15.** The position angles and inclinations of absorption features around galaxies with observed velocity gradients in MUSE (using any emission line). This selection introduces the strong bias against face-on galaxies seen in the figure. We cut by inclination and position angle at  $40^\circ$  as described in the text, in order to identify candidate co-rotating material (unshaded region). Points are coloured by quasar sightline (A in blue, B in grey, C in red). Galaxies in groups are marked with triangles and non-detections are marked with crosses. The bars on the black point in the lower-right are illustrative of the median  $1-\sigma$  uncertainty in each axis.

Whilst the tendency for line-of-sight velocities of gas along the major axis to be aligned with the galaxy rotation would be expected given the rotating disk model, the sample size is too small to draw strong conclusions. Both of the non-group sightlines more than 100 kpc from the galaxy exhibit this alignment well beyond the virial radius of the galaxy, whereas studies such as MEGAFLOW (Zabl et al. 2019) and others (Bouché et al. 2016; Diamond-Stanic et al. 2016; Ho et al. 2017) only show co-rotation out to tens of kpc. In our small sample this can be explained by coincidence, but such alignment is thought to be a result of coherent accretion over Gyr timescales, due to the location of the galaxy within the cosmic web, transferring angular momentum from the gas to the central galaxy (e.g. Danovich et al. 2015; Stewart et al. 2017). DeFelippis et al. (2020) find, in Illustris TNG simulations, that this major-axis co-rotating material often extends in a much thicker ‘wedge’ out to  $\sim 0.75 R_{\text{vir}}$  for sub- $L_\star$  galaxies. However, Huscher et al. (2021) use the EAGLE simulations and study larger  $\approx L_\star$  galaxies, finding substantial rotation only to  $\approx 0.3R_{\text{vir}}$  in cool gas at  $z = 0$ . Any attempts to constrain the extent of this material through observations will require detections of absorption with impact parameters of 100–300 kpc from isolated galaxies, which our sample lacks. The sample of French & Wakker (2020) does contain H I absorption in this range, but they do not detect a significant preference for co-rotation beyond  $\approx 100$  kpc.

The less-clear alignment in galaxy groups is consistent with models suggesting interactions between galaxies are the origin of most strong H I absorbers in these environments (e.g. Morris & van den Bergh 1994), as the velocities of such tidal material would be more strongly affected by the relative velocities of the galaxies at large impact parameters. The velocity offsets between galaxies



**Figure 16.** The r-band luminosities of observed galaxies as a function of redshift. Red points illustrate objects in which a velocity gradient is observed in MUSE. Cyan points show other galaxies within the MUSE fields. The horizontal dashed line represents an estimate of  $L_\star$ , as given by Montero-Dorta & Prada (2009) for  $z \leq 0.2$ . ( $L_\star$  does vary over the given redshift range, but this is expected to be less than a factor of two, e.g. Gabasch et al. 2006, so is not shown.)

and absorbers are generally within the velocity dispersion of the observed galaxies.

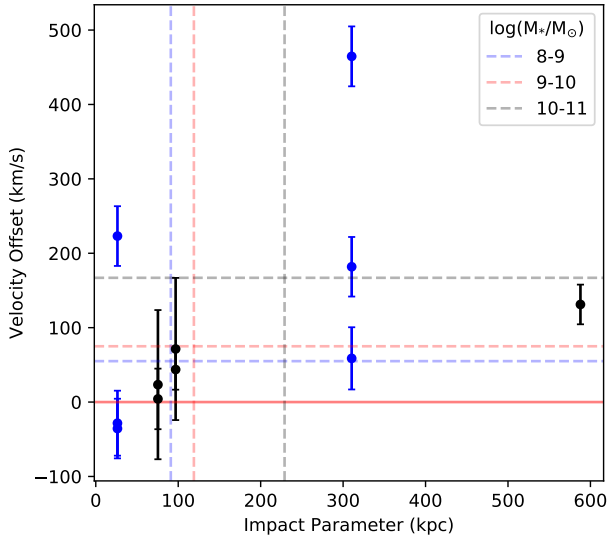
## 6 SUMMARY & CONCLUSIONS

In this study we take advantage of multiple observing campaigns, resulting in a deep, dense galaxy survey covering the field around a unique quasar triplet, to examine the geometry and extent of gas flows in the CGM and IGM on larger scales than most CGM studies.

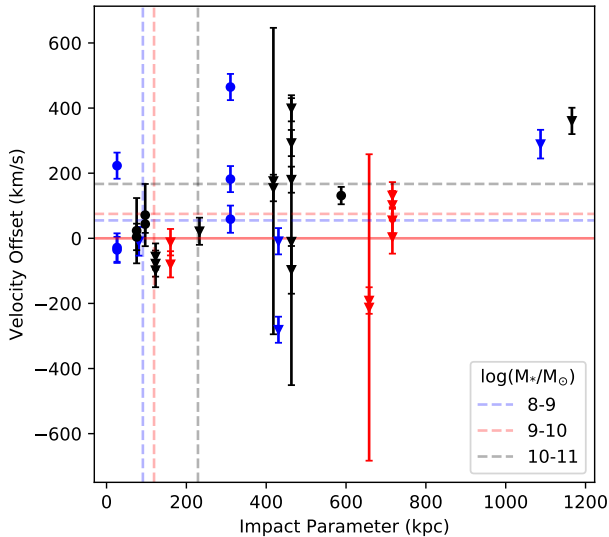
We find that:

(i) The H I absorbing gas is often found in structures several hundred kpc in extent, as observed in absorption lines visible in multiple lines-of-sight at the same redshift. These structures traced by H I over all column densities are more likely to match those traced by galaxies and high-column-density absorbers ( $N > 10^{14} \text{ cm}^{-2}$ ). We do not find significant evidence that H I absorbers with lower column densities form structures reaching this extent. It is not clear whether these are arising from smaller overdensities in galaxy voids, or forming filaments, most of which are not aligned such that they cover multiple sightlines. A larger number of background sources would be needed to confirm this through absorption measurements.

(ii) Galaxy-absorber pairs with H I absorption exhibit a significant bimodality in position angle for impact parameters up to  $\approx 300$  kpc, possibly evidence that outflows and extended disks reaching these scales are common. However, we do not find that H I absorbers near the minor axis have significantly higher Doppler widths than those on the major axis, nor do they show smaller velocity offsets around edge-on galaxies for which a putative outflow should not have a large line-of-sight velocity component. We therefore cannot rule out alternative hypotheses for this bimodality, such as preferential alignment between galaxies and large-scale structure.



**Figure 17.** The impact parameter and line-of-sight velocity difference between galaxies and absorbers in the major-axis subsample described in Section 5, showing only non-group galaxies. The sign of the velocity difference is chosen to reflect alignment (positive) between the angular momenta of the galaxy and gas, with the red solid line indicating zero offset. Uncertainties in velocity offset are the redshift measurement errors for the galaxy and absorber added in quadrature. Points are coloured by galaxy stellar mass, with galaxies of between  $10^8$  and  $10^9 M_{\odot}$  shown in blue,  $10^9$ - $10^{10} M_{\odot}$  in red, and  $10^{10}$ - $10^{11} M_{\odot}$  in black. The dashed horizontal and vertical lines show the median virial velocities and radii for galaxies within each bin across our full galaxy survey.



**Figure 18.** The impact parameter and line-of-sight velocity difference between galaxies and absorbers in the major-axis subsample described in Section 5, as in Figure 17 but including group galaxies. Galaxies in groups are shown with triangles, those not in groups are shown with circles. Points are coloured by galaxy stellar mass. Median virial velocities and radii do not consider other galaxies in the group that may be sharing the same halo.

(iii) H I absorbers found close to the minor axis of a galaxy are twice as likely to have associated O VI absorption as those found close to the major axis. The position angle distribution of O VI shows a clear bimodality, weighted towards the minor axis more than H I, supporting models in which O VI is primarily observed in galactic winds. O VI around isolated galaxies is far more limited in extent than H I, with no isolated galaxies featuring O VI absorption beyond 400 kpc. We also see O VI in galaxy groups, but cannot currently distinguish between virialized halo gas, tidal debris, and outflows/halo gas from other galaxies in the same group.

(iv) The line-of-sight velocities of H I absorption found near the major axis of a galaxy show a significant tendency to align with the galaxy rotation, suggesting a strong coupling between the angular momentum of the CGM and galaxy (as found by e.g. DeFelippis et al. 2020; Huscher et al. 2021).

(v) Star-forming galaxies are more likely to trace the same structures as H I seen in multiple sightlines. Our star-forming sample does not exhibit the bimodality in position angle that could be indicative of accreting and outflowing material, but this bimodality is apparent when using a more restrictive ‘strongly-star-forming’ sample. This suggests that this inflow/outflow dichotomy contributes to the overall bimodality seen on scales  $\lesssim 300$  kpc, although is not ubiquitous among star-forming galaxies, and that the excess absorption seen on larger scales is primarily due to environmental effects (although the redshift biases of our star-forming and non-star-forming samples may also contribute).

This study illustrates some of the ways in which multiple sightlines can be used to probe the structure of the gas around galaxies. In future work we will also study the absorption around individual galaxies and galaxy groups probed by these quasar lines-of-sight, where knowledge of absorption at multiple impact parameters and position angles around the same galaxy will enable us to test a range of models. Similar methods applied to mock lines-of-sight through simulations may provide further insights into the processes behind these results.

Whilst Q0107 is one of very few systems for which these techniques can currently be used, the prospects at higher redshift are much improved. The ELT will allow far deeper galaxy surveys, enabling high-redshift galaxies to be used as background sources (e.g. Japelj et al. 2019), and improving on current studies that have poor signal-to-noise (e.g. Lee et al. 2018) or stack observations to improve the sensitivity (e.g. Chen et al. 2020b). This progresses the field towards full tomography, in which numerous sightlines can probe a variety of scales throughout the same structures, and improve upon the usual pencil-beam observations of the IGM.

Similarly, observations of emission in H I 21cm rarely reach detections of material in the  $10^{13} - 10^{16} \text{ cm}^{-2}$  regime covered here, but the Square Kilometre Array will likely be able to reach sensitivities of  $10^{15} \text{ cm}^{-2}$  (Popping et al. 2015), sufficient to begin detecting the denser filaments of the cosmic web, and allowing study of the interaction between the CGM galactic fountains, the tidal material produced in galaxy groups, and the large-scale cosmic web.

## ACKNOWLEDGEMENTS

We thank the referee for useful comments that have improved the quality of this paper.

This work is based on observations made with the NASA/ESA Hubble Space Telescope, obtained from the data archive at the Space Telescope Science Institute. STScI is operated by the Association of

Universities for Research in Astronomy, Inc. under NASA contract NAS 5-26555. We also make use of observations collected at the European Southern Observatory under ESO programmes 086.A-0970, 087.A-0857 and 094.A-0131; at the W.M. Keck Observatory under programme A290D; and at the Gemini Observatory under programme GS-2008B-Q-50.

We thank Matteo Fossati for providing the MARZ templates used for estimating redshifts and their uncertainties. We also thank Jill Bechtold for leading the effort to obtain the Keck/DEIMOS data.

We thank the contributors to SCIPY<sup>5</sup>, MATPLOTLIB<sup>6</sup>, ASTROPY<sup>7</sup>, and the PYTHON programming language, the free and open-source community and the NASA Astrophysics Data system<sup>8</sup> for software and services.

This work also made use of the DiRAC system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility<sup>9</sup>.

AB acknowledges the support of a UK Science and Technology Facilities Council (STFC) PhD studentship through grant ST/S505365/1. SLM acknowledges the support of STFC grant ST/T000244/1.

SC gratefully acknowledges support from Swiss National Science Foundation grants PP00P2\_163824 and PP00P2\_190092, and from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme grant agreement No 864361. MF also acknowledges funding from the ERC under the Horizon 2020 programme (grant agreement No 757535). This work has been supported by Fondazione Cariplo, grant No 2018-2329.

## DATA AVAILABILITY

The raw data from the Hubble Space Telescope may be accessed from the MAST archive<sup>10</sup>, and that from the WM Keck Observatory from the Keck Observatory Archive<sup>11</sup>. That from the European Southern Observatory may be accessed from the ESO Archive<sup>12</sup>, and that from the Gemini Observatory may be accessed from the Gemini Science Archive<sup>13</sup>. The relevant program IDs are given in Tables 2 and 3 of this paper. The derived data generated in this research will be shared on reasonable request to the corresponding author.

## REFERENCES

- Adelberger K. L., Steidel C. C., Shapley A. E., Pettini M., 2003, *ApJ*, 584, 45
- Adelberger K. L., Shapley A. E., Steidel C. C., Pettini M., Erb D. K., Reddy N. A., 2005, *ApJ*, 629, 636
- Albareti F. D., et al., 2017, *ApJS*, 233, 25
- Anshul P., Narayanan A., Muzahid S., Beckett A., Morris S. L., 2021, *MNRAS*, 503, 3243
- Astropy Collaboration et al., 2018, *AJ*, 156, 123
- <sup>5</sup> <http://www.scipy.org/>
- <sup>6</sup> <https://matplotlib.org/>
- <sup>7</sup> <https://www.astropy.org/>
- <sup>8</sup> <https://ui.adsabs.harvard.edu/>
- <sup>9</sup> <http://www.dirac.ac.uk/>
- <sup>10</sup> <http://archive.stsci.edu/>
- <sup>11</sup> <https://koa.ipac.caltech.edu/cgi-bin/KOA/nph-KOALogin>
- <sup>12</sup> [http://archive.eso.org/eso/eso\\_archive\\_main.html](http://archive.eso.org/eso/eso_archive_main.html)
- <sup>13</sup> <http://www.cadc-ccda.hia-ihp.nrc-cnrc.gc.ca/en/gsa/>
- Bacon R., et al., 2010, *SPIE*, 7735, 773508
- Bacon R., et al., 2017, *A&A*, 608, A1
- Behroozi P. S., Conroy C., Wechsler R. H., 2010, *ApJ*, 717, 379
- Bergeron J., 1986, *Astronomy and Astrophysics*, 155, L8
- Bergeron J., Boissé P., 1991, *A&A*, 243, 344
- Bergeron J., et al., 1994, *ApJ*, 436, 33
- Bertin E., Arnouts S., 1996, *ApJS*, 117, 393
- Bielby R., Crighton N. H. M., Fumagalli M., Morris S. L., Stott J. P., Tejos N., Cantalupo S., 2017, *MNRAS*, 468, 1373
- Bielby R. M., et al., 2019, *MNRAS*, 486, 21
- Bielby R. M., et al., 2020, *MNRAS*, 493, 5336
- Bland J., Tully B., 1988, *Nature*, 334, 43
- Bond J. R., Kofman L., Pogosyan D., 1996, *Nature*, 380, 603
- Bordoloi R., et al., 2011, *ApJ*, 743, 10
- Bordoloi R., et al., 2014, *ApJ*, 794, 130
- Bouché N., Murphy M. T., Péroux C., Csabai I., Wild V., 2006, *MNRAS*, 371, 495
- Bouché N., et al., 2016, *ApJ*, 820, 121
- Bowen D. V., Chelouche D., Jenkins E. B., Tripp T. M., Pettini M., York D. G., Frye B. L., 2016, *ApJ*, 826, 50
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Burchett J. N., Elek O., Tejos N., Prochaska J. X., Tripp T. M., Bordoloi R., Forbes A. G., 2020, *ApJ*, 891, L35
- Burchett J. N., Rubin K. H. R., Prochaska J. X., Coil A. L., Rickards Vaught R., Hennawi J. F., 2021, *ApJ*, 909, 151
- Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, 533, 682
- Cantalupo S., et al., 2019, *MNRAS*, 483, 5188
- Carswell R. F., Webb J. K., 2014, *Astrophysics Source Code Library*, p. ascl:1408.015
- Casuso E., Beckman J. E., 2004, *A&A*, 419, 181
- Cen R., Tripp T. M., Ostriker J. P., Jenkins E. B., 2001, *ApJL*, 559, L5
- Charlot S., Kauffmann G., Longhetti M., Tresse L., White S. D. M., Maddox S. J., Fall S. M., 2002, *MNRAS*, 330, 876
- Charlton J. C., Churchill C. W., 1998, *ApJ*, 499, 181
- Chen H.-W., 2012, *MNRAS*, 427, 1238
- Chen H.-W., Mulchaey J. S., 2009, *ApJ*, 701, 1219
- Chen H.-W., Lanzetta K. M., Webb J. K., Barcons X., 1998, *ApJ*, 498, 77
- Chen H.-W., Gauthier J.-R., Sharon K., Johnson S. D., Nair P., Liang C. J., 2014, *MNRAS*, 438, 1435
- Chen H.-W., Zahedy F. S., Johnson S. D., Pierce R. M., Huang Y.-H., Weiner B. J., Gauthier J.-R., 2018, *MNRAS*, 479, 2547
- Chen H.-W., et al., 2020a, *MNRAS*, 497, 498
- Chen Y., et al., 2020b, *MNRAS*, 499, 1721
- Chevalier R. A., Clegg A. W., 1985, *Nature*, 317, 44
- Concas A., Popesso P., Brusa M., Mainieri V., Thomas D., 2019, *A&A*, 622, A188
- Cooke J., O'Meara J. M., 2015, *ApJL*, 812, L27
- Crighton N. H. M., Morris S. L., Bechtold J., Crain R. A., Jannuzi B. T., Shone A., Theuns T., 2010, *MNRAS*, 402, 1273
- Croxall K. V., van Zee L., Lee H., Skillman E. D., Lee J. C., Côté S., Kennicutt R. C., Miller B. W., 2009, *ApJ*, 705, 723
- Danovich M., Dekel A., Hahn O., Ceverino D., Primack J., 2015, *MNRAS*, 449, 2087
- Davé R., Finlator K., Oppenheimer B. D., 2012, *MNRAS*, 421, 98
- Davies R. L., et al., 1997, in *Optical Telescopes of Today and Tomorrow*, pp. 1099–1106, doi:10.1117/12.268996
- DeFelippis D., Genel S., Bryan G. L., Nelson D., Pillepich A., Hernquist L., 2020, *ApJ*, 895, 17
- Diamond-Stanic A. M., Coil A. L., Moustakas J., Tremonti C. A., Sell P. H., Mendez A. J., Hickox R. C., Rudnick G. H., 2016, *ApJ*, 824, 24
- Dinshaw N., Weymann R. J., Impey C. D., Foltz C. B., Morris S. L., Ake T., 1997, *ApJ*, 491, 45
- Draine B. T., 2011, *Physics of the Interstellar and Intergalactic Medium*. Princeton University Press
- Duarte M., Mamon G., 2014, *MNRAS*, 440, 1763
- Dutta R., et al., 2020, *MNRAS*, 499, 5022
- Faber S. M., et al., 2003, in *Instrument Design and Performance*



- for Optical/Infrared Ground-Based Telescopes. pp 1657–1669, doi:10.1117/12.460346
- Finley H., et al., 2017, *A&A*, 605, A118
- Finn C. W., et al., 2016, *MNRAS*, 460, 590
- Fossati M., Fumagalli M., Gavazzi G., Consolandi G., Boselli A., Yagi M., Sun M., Wilman D. J., 2019a, *MNRAS*, 484, 2212
- Fossati M., et al., 2019b, *MNRAS*, 490, 1451
- French D. M., Wakker B. P., 2020, *ApJ*, 897, 151
- Freundlich J., et al., 2013, *A&A*, 553, A130
- Fukugita M., Hogan C. J., Peebles P. J. E., 1998, *ApJ*, 503, 518
- Fumagalli M., Fossati M., Hau G. K. T., Gavazzi G., Bower R., Sun M., Boselli A., 2014, *MNRAS*, 445, 4335
- Fumagalli M., Cantalupo S., Dekel A., Morris S. L., O’Meara J. M., Prochaska J. X., Theuns T., 2016, *MNRAS*, 462, 1978
- Fumagalli M., Haardt F., Theuns T., Morris S. L., Cantalupo S., Madau P., Fossati M., 2017, *MNRAS*, 467, 4802
- Gabasch A., et al., 2006, *A&A*, 448, 101
- Green J. C., et al., 2012, *ApJ*, 744, 60
- Grimes J. P., et al., 2009, *ApJS*, 181, 272
- Hafen Z., et al., 2019, *MNRAS*, 488, 1248
- Hafen Z., et al., 2020, *MNRAS*, 494, 3581
- Hartigan J. A., Hartigan P. M., 1985, *Ann. Statist.*, 13, 70
- Heckman T. M., Armus L., Miley G. K., 1990, *ApJS*, 74, 833
- Heckman T., Borthakur S., Wild V., Schiminovich D., Bordoloi R., 2017, *ApJ*, 846, 151
- Hinton S. R., Davis T. M., Lidman C., Glazebrook K., Lewis G. F., 2016, *Astronomy and Computing*, 15, 61
- Ho S. H., Martin C. L., Kacprzak G. G., Churchill C. W., 2017, *ApJ*, 835, 267
- Hopkins P. F., Chan T. K., Ji S., Hummels C., Keres D., Quataert E., Faucher-Giguere C.-A., 2021, *MNRAS*, 501, 3640
- Huang Y.-H., Chen H.-W., Shectman S. A., Johnson S. D., Zahedy F. S., Helsby J. E., Gauthier J.-R., Thompson I. B., 2021, *MNRAS*, 502, 4743
- Huscher E., Oppenheimer B. D., Lonardi A., Crain R. A., Richings A. J., Schaye J., 2021, *MNRAS*, 500, 1476
- Japelj J., et al., 2019, *A&A*, 632, A94
- Johnson S. D., Chen H.-W., Mulchaey J. S., 2015, *MNRAS*, 449, 3263
- Johnson S. D., et al., 2018, *ApJL*, 869, L1
- Kacprzak G. G., Muzahid S., Churchill C. W., Nielsen N. M., Charlton J. C., 2015, *ApJ*, 815, 22
- Kacprzak G. G., Pointon S. K., Nielsen N. M., Churchill C. W., Muzahid S., Charlton J. C., 2019, *ApJ*, 886, 91
- Keeney B. A., Danforth C. W., Stocke J. T., France K., Green J. C., 2012, *Publ. Astron. Soc. Pac.*, 124, 830
- Keeney B. A., Stocke J. T., Rosenberg J. L., Danforth C. W., Ryan-Weber E. V., Shull J. M., Savage B. D., Green J. C., 2013, *ApJ*, 765, 27
- Keeney B. A., et al., 2017, *ApJS*, 230, 6
- Kennicutt J., 1998, *ARA&A*, 36, 189
- Kereš D., Katz N., Weinberg D. H., Davé R., 2005, *MNRAS*, 363, 2
- Kewley L. J., Geller M. J., Jansen R. A., 2004, *AJ*, 127, 2002
- Kim T.-S., et al., 2020, *MNRAS*
- Klypin A. A., Shandarin S. F., 1983, *MNRAS*, 204, 891
- Lan T.-W., Mo H., 2018, *ApJ*, 866, 36
- Lanzetta K. M., Bowen D. V., Tytler D., Webb J. K., 1995, *ApJ*, 442, 538
- Le Fevre O., Crampton D., Felenbok P., Monnet G., 1994, *A&A*, 282, 325
- LeFevre O., et al., 2003, in *Instrument Design and Performance for Optical/Infrared Ground-Based Telescopes*. International Society for Optics and Photonics, pp 1670–1681, doi:10.1117/12.460959
- Lee K.-G., et al., 2018, *ApJS*, 237, 31
- Lehnert M. D., Heckman T. M., Weaver K. A., 1999, *ApJ*, 523, 575
- Lehnert M. D., Le Tiran L., Nesvadba N. P. H., van Driel W., Boulanger F., Di Matteo P., 2013, *A&A*, 555, A72
- Lofthouse E. K., et al., 2020, *MNRAS*, 491, 2057
- Lopez S., et al., 2018, *Nature*, 554, 493
- Lopez S., et al., 2020, *MNRAS*, 491, 4442
- Macquart J.-P., et al., 2020, *Nature*, 581, 391
- Madau P., Dickinson M., 2014, *ARA&A*, 52, 415
- Martin C. L., Shapley A. E., Coil A. L., Kornei K. A., Bundy K., Weiner B. J., Noeske K. G., Schiminovich D., 2012, *ApJ*, 760, 127
- Martin C. L., Ho S. H., Kacprzak G. G., Churchill C. W., 2019, *ApJ*, 878, 84
- Mathes N. L., Churchill C. W., Kacprzak G. G., Nielsen N. M., Trujillo-Gomez S., Charlton J., Muzahid S., 2014, *ApJ*, 792, 128
- Mathewson D. S., Cleary M. N., Murray J. D., 1974, *ApJ*, 190, 291
- Mitchell P. D., Schaye J., Bower R. G., Crain R. A., 2020a, *MNRAS*, 494, 3971
- Mitchell P. D., Schaye J., Bower R. G., 2020b, *MNRAS*, 497, 4495
- Montero-Dorta A. D., Prada F., 2009, *MNRAS*, 399, 1106
- Morris S. L., Jannuzi B. T., 2006, *MNRAS*, 367, 1261
- Morris S. L., van den Bergh S., 1994, *ApJ*, 427, 696
- Morris S. L., Weymann R. J., Savage B. D., Gilliland R. L., 1991, *ApJL*, 377, L21
- Muzahid S., 2014, *ApJ*, 784, 5
- Neeleman M., Prochaska J. X., Ribaldo J., Lehner N., Howk J. C., Rafelski M., Kanekar N., 2016, *ApJ*, 818, 113
- Nelson D., et al., 2019, *MNRAS*, 490, 3234
- Newman J. A., et al., 2013, *ApJS*, 208, 5
- Nicastro F., et al., 2018, *Nature*, 558, 406
- Nidever D. L., Majewski S. R., Burton W. B., 2008, *ApJ*, 679, 432
- Nielsen N. M., Churchill C. W., Kacprzak G. G., Murphy M. T., 2013, *ApJ*, 776, 114
- Oppenheimer B. D., et al., 2016, *MNRAS*, 460, 2157
- Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, *AJ*, 124, 266
- Péroux C., et al., 2017, *MNRAS*, 464, 2053
- Péroux C., et al., 2019, *MNRAS*, 485, 1595
- Petry C. E., Impey C. D., Fenton J. L., Foltz C. B., 2006, *AJ*, 132, 2046
- Pichon C., Pogogyan D., Kimm T., Slyz A., Devriendt J., Dubois Y., 2011, *MNRAS*, 418, 2493
- Planck Collaboration et al., 2020, *A&A*, 641, A6
- Pointon S. K., Kacprzak G. G., Nielsen N. M., Muzahid S., Murphy M. T., Churchill C. W., Charlton J. C., 2019, *ApJ*, 883, 78
- Popping A., Meyer M., Staveley-Smith L., Obreschkow D., Jozsa G., Pisano D. J., 2015, *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, p. 132
- Rakic O., Schaye J., Steidel C. C., Rudie G. C., 2011, *MNRAS*, 414, 3265
- Rakic O., Schaye J., Steidel C. C., Booth C. M., Dalla Vecchia C., Rudie G. C., 2013, *MNRAS*, 433, 3103
- Rauch M., Sargent W. L. W., Barlow T. A., Simcoe R. A., 2002, *ApJ*, 576, 45
- Rubin K. H. R., Prochaska J. X., Koo D. C., Phillips A. C., 2012, *ApJL*, 747, L26
- Rubin K. H. R., Diamond-Stanic A. M., Coil A. L., Crighton N. H. M., Moustakas J., 2018, *ApJ*, 853, 95
- Ryon J. E., 2019, *Advanced Camera for Surveys HST Instrument Handbook*
- Salcido J., Bower R. G., Theuns T., 2020, *MNRAS*, 491, 5083
- Schaye J., et al., 2010, *MNRAS*, 402, 1536
- Schroetter I., et al., 2016, *ApJ*, 833, 39
- Schroetter I., et al., 2019, *MNRAS*, 490, 4368
- Scodreggio M., et al., 2005, *Publ. Astron. Soc. Pac.*, 117, 1284
- Scott J., Bechtold J., Morita M., Dobrzycki A., Kulkarni V., 2002, *ApJ*, 571, 665
- Scoville N., et al., 2017, *ApJ*, 837, 150
- Shull J. M., Smith B. D., Danforth C. W., 2012, *ApJ*, 759, 23
- Smette A., Surdej J., Shaver P. A., Foltz C. B., Chaffee F. H., Weymann R. J., Williams R. E., Magain P., 1992, *ApJ*, 389, 39
- Somerville R. S., Popping G., Trager S. C., 2015, *MNRAS*, 453, 4337
- Sorgho A., Foster T., Carignan C., Chemin L., 2019, *MNRAS*
- Springel V., et al., 2005, *Nature*, 435, 629
- Steidel C. C., Kollmeier J. A., Shapley A. E., Churchill C. W., Dickinson M., Pettini M., 2002, *ApJ*, 570, 526
- Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, *ApJ*, 717, 289
- Stewart K. R., et al., 2017, *ApJ*, 843, 47
- Tejos N., et al., 2014, *MNRAS*, 437, 2017
- Tempel E., Libeskind N. I., 2013, *ApJL*, 775, L42

- Tripp T. M., Savage B. D., 2000, *ApJ*, 542, 42  
 Tumlinson J., et al., 2011, *Science*, 334, 948  
 Tumlinson J., et al., 2013, *ApJ*, 777, 59  
 Tumlinson J., Peebles M. S., Werk J. K., 2017, *ARA&A*, 55, 389  
 Turner M. L., Schaye J., Steidel C. C., Rudie G. C., Strom A. L., 2014, *MNRAS*, 445, 794  
 Turner M. L., Schaye J., Steidel C. C., Rudie G. C., Strom A. L., 2015, *MNRAS*, 450, 2067  
 Veilleux S., Cecil G., Bland-Hawthorn J., 2005, *ARA&A*, 43, 769  
 Werk J. K., Prochaska J. X., Thom C., Tumlinson J., Tripp T. M., O’Meara J. M., Peebles M. S., 2013, *ApJS*, 204, 17  
 Werk J. K., et al., 2014, *ApJ*, 792, 8  
 Werk J. K., et al., 2016, *ApJ*, 833, 54  
 Wetzel A. R., Tinker J. L., Conroy C., van den Bosch F. C., 2013, *MNRAS*, 432, 336  
 Wijers N. A., Schaye J., Oppenheimer B. D., 2020, *MNRAS*, 498, 574  
 Wild V., et al., 2008, *MNRAS*, 388, 227  
 Wilde M. C., et al., 2021, *ApJ*, 912, 9  
 Young P. A., Impey C. D., Foltz C. B., 2001, *ApJ*, 549, 76  
 Zabl J., et al., 2019, *MNRAS*, 485, 1961  
 Zabl J., et al., 2020, *MNRAS*, 492, 4576  
 Zahedy F. S., Chen H.-W., Rauch M., Wilson M. L., Zabludoff A., 2016, *MNRAS*, 458, 2423  
 Zhang Y., Yang X., Wang H., Wang L., Luo W., Mo H. J., van den Bosch F. C., 2015, *ApJ*, 798, 17  
 van de Voort F., Schaye J., Booth C. M., Dalla Vecchia C., 2011, *MNRAS*, 415, 2782

## APPENDIX A: FRIENDS-OF-FRIENDS ALGORITHM

In this section we describe the friends-of-friends algorithm used to find galaxy groups, and the calculations required to determine the linking lengths used. We use the analysis and results from [Duarte & Mamon \(2014\)](#) to motivate our choices.

For physical linking lengths  $D_{\perp}$  and  $D_{\parallel}$ , any pair of galaxies at redshifts  $z_1$  and  $z_2$  (with  $d_c$  the comoving distance to that redshift), separated by an angle  $\theta$  on the sky, are ‘friends’ if both of the following conditions are satisfied:

$$\frac{d_c(z_1) + d_c(z_2)}{2} \theta \leq D_{\perp} \quad (\text{A1})$$

$$|d_c(z_2) - d_c(z_1)| \leq D_{\parallel} \quad (\text{A2})$$

We scale the constant linking lengths – taken from [Duarte & Mamon \(2014\)](#) and denoted  $b$  – by observed galaxy density,  $D = b n^{-\frac{1}{3}}$ , in order to account for the changing magnitude limit with redshift. In order to ensure that large groups do not bias the estimated density, we use large bins with a redshift width 0.2. Note that the angular diameter distance could be used in place of the comoving distance, and it could be argued that  $d_A$  should be used since galaxy groups are expected to be virialized. However, this also requires the line-of-sight linking length to be reduced by a factor of  $(1+z)$ , and most studies do not apply this.

The comoving volume is calculated by integrating a conic section with an half-opening angle of  $0.15^\circ$ , through comoving distance between the redshift limits. This is approximately the radius of the DEIMOS data. MUSE galaxies are not counted in this process, as the greater depth of the MUSE data would bias the selection of groups in and around the MUSE fields (by adding galaxies to groups that would be undetected if they lay outside of the MUSE fields, and potentially joining groups together that would otherwise remain separate). By excluding these galaxies, the depth of VIMOS,

DEIMOS and GMOS coverage throughout this aperture is close to constant.

$$V = \int_{d_c(z_{min})}^{d_c(z_{max})} \pi(\theta d_c)^2 d(d_c) \quad (\text{A3})$$

In each redshift bin, the number of observed galaxies is divided by the volume of the cone, to give the galaxy number density, which is converted to a linking length as above. Whilst the physical linking length is constant within a single bin, the angular separation and redshift difference are both affected by the galaxy redshift itself.

This algorithm is implemented by assigning each galaxy a group identifier. We then loop through each galaxy pair, and assign both galaxies the lower identifier if they are ‘friends’. This is repeated until the sum of all identifiers converges (usually 3 to 5 iterations). All galaxies with the same identifier form a single galaxy group. Absorbers and the MUSE galaxies are associated with galaxy groups if they are ‘friends’ with any galaxy in that group.

[Duarte & Mamon \(2014\)](#) list three pairs of linking lengths, from which we use the larger, completeness-optimizing, values ( $b_{\perp} = 0.2$ ,  $b_{\parallel} = 3$ ). This results in linking velocity differences of  $\approx 800 - 1500$  km/s and transverse linking lengths of  $\approx 800 - 1200$  pkpc. These linking lengths are large enough to allow a reasonable sample size of galaxy groups, as well as approximately matching the distance scales separating the three lines-of-sight. Such large linking lengths ensure that our non-group sample excludes any galaxies whose CGM is likely to be affected by ongoing group interactions.

We do repeat most analyses shown in the main text whilst using the ‘medium’ linking lengths from [Duarte & Mamon \(2014\)](#), and this does not reveal any significant differences. Using the shortest linking lengths, we do not find enough groups to draw any conclusions.

## APPENDIX B: GENERATING RANDOM ABSORBERS

In Section 3, we test the excess of galaxy-absorber associations within a series of different constraints. Here we describe the process of generating the random sets of absorbers used in these tests, in which any physical association between gas and galaxies is removed.

We generated 5000 sets of randomly distributed absorbers similar to the method used in T14, as follows:

- (i) Calculate the signal-to-noise per resolution element for QSO spectrum
- (ii) Convert this to minimum rest-frame equivalent width for the absorption feature as a function of redshift (see Figure 7)
- (iii) For each real absorption feature, find the allowed region in redshift space for which the EW of the progenitor is larger than the minimum, and is not covered by galactic absorption
- (iv) Distribute absorbers randomly through the allowed region, giving the random absorber the same properties as the observed progenitor

The signal-to-noise per resolution element is estimated by dividing the continuum value of the QSO spectrum by the uncertainty in the value of that pixel, and then convolving with the line-spread function of the instrument. In order to improve execution time we use a box car of the correct FWHM as an approximation to the line-spread function.

The minimum observable equivalent width for an unresolved transition observed at wavelength  $\lambda$  is:

$$W_{min}(\lambda) = S \frac{FWHM}{\langle SNR \rangle_\lambda} \quad (\text{B1})$$

where FWHM is the width of the line-spread of the instrument, S is the significance required of the detection in units of sigma (where a value of 3 best matches the distribution of observed absorbers), and  $\langle SNR \rangle_\lambda$  is the mean signal-to-noise per resolution element. If the resolution  $R = \frac{\lambda}{FWHM}$  is constant:

$$W_{min}(\lambda) = \frac{S\lambda}{R\langle SNR \rangle_\lambda} = \frac{S\lambda_0(1+z)}{R\langle SNR \rangle_\lambda} \quad (\text{B2})$$

where  $\lambda_0$  is the rest-frame wavelength of the transition. This gives a rest-frame equivalent width:

$$W_{min,r}(z) = \frac{S\lambda_0}{R\langle SNR \rangle_z} \quad (\text{B3})$$

We note that the lines in the COS spectra are usually resolved, with a median width  $b \approx 30 \text{ km s}^{-1}$ . For these lines a more representative equivalent width limit can be found using the prescription given in [Keeney et al. \(2012\)](#), in which the optimal integration window is found using a convolution of the COS line-spread-function and the absorption line profile, rather than just the resolution of the instrument. However, many of the absorption features are narrower than this, so we continue to use the COS FWHM as a limiting value. This only has a notable effect on absorbers that are both weak and narrow, so only four absorbers of the 184 across the three COS spectra would have their allowed redshift ranges reduced by more than 15% if the [Keeney et al. \(2012\)](#) prescription were used.

From the absorption line catalogue (taken directly from T14), the best fit column density and Doppler parameter are given for each absorber. This can be robustly converted to an equivalent width via optical depth using the approximation given in [Draine \(2011\)](#) (eqn 9.8, assuming negligible stimulated emission):

$$\tau_0 = \sqrt{\pi} \frac{e^2}{m_e c} \frac{n f_{osc} \lambda_0}{b} \quad (\text{B4})$$

where n and b are the column density and Doppler parameter of the absorber,  $\lambda_0$  the rest-frame wavelength, and  $f_{osc}$  the oscillator strength. Draine then provides an approximation for equivalent width:

$$W = \sqrt{\pi} \frac{b}{c} \frac{\tau_0}{1 + (\tau_0/2\sqrt{2})} \quad (\text{for } \tau_0 < 1.25393) \quad (\text{B5})$$

$$W = \sqrt{\frac{4b^2}{c^2} \ln\left(\frac{\tau_0}{\ln 2}\right) + \frac{b}{c} \frac{\gamma \lambda_0}{c} \frac{\tau_0 - 1.25393}{\sqrt{\pi}}} \quad (\text{for } \tau_0 > 1.25393) \quad (\text{B6})$$

$\gamma$  is the damping term of the Lorentzian component of the Voigt profile, which becomes non-negligible for high-column-density systems.

This value of equivalent width can then be compared with the minimum equivalent width detectable in the spectrum as a function of redshift, shown for QSO-A in Figure 7. We mask out the regions where the absorber would be undetectable due to insufficient equivalent width, as well as regions within 200 km/s of galactic absorption

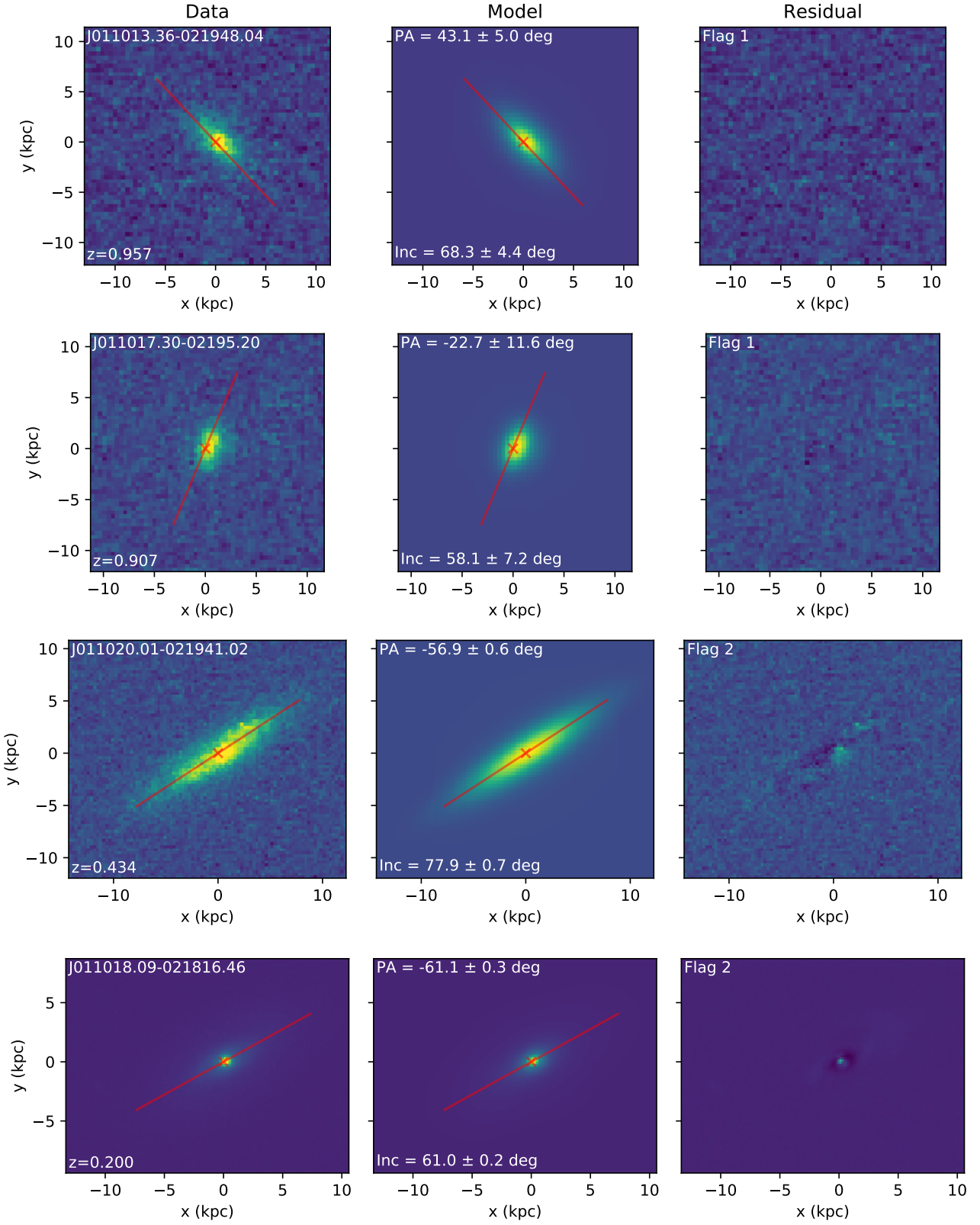
(as in T14, using C II, N V, O I, Si II, S II and Fe II). The region less than 3500 km/s blueward of the QSO is also masked, in order to remove most associated absorption. As found by [Wild et al. \(2008\)](#),  $\approx 40\%$  of C IV absorption out to 3000 km/s is associated with outflows from the QSO, although absorption from outflows can extend beyond 10,000 km/s. The velocity cut used is therefore a compromise between excluding most absorption associated with the QSO, whilst excluding minimal intervening IGM absorption. This cut is also much larger than the region within which the proximity effect is significant (e.g [Scott et al. 2002](#)). Absorbers in this set that have already been placed are also masked using a window of 35 km/s (approximately the median width of H I absorption systems) so that systems that would not be resolved due to blending are not generated by this process. The absorber is then randomly assigned a redshift within the unmasked region.

## APPENDIX C: GALFIT RESULTS

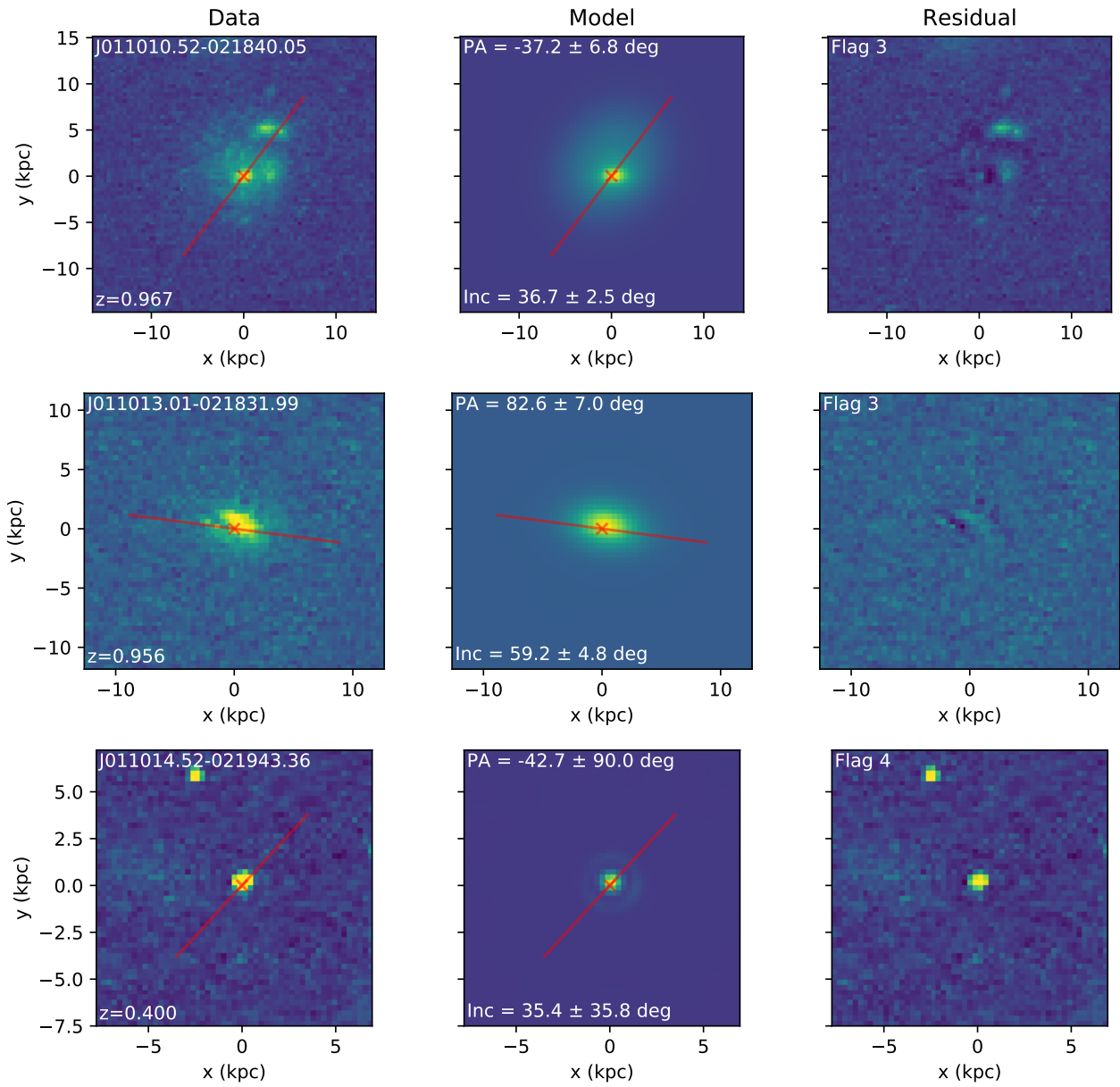
Here we include some examples of our GALFIT results, in order to illustrate the utility of the modelling and the use of the quality flags described in Section 2.4. As illustrated in the Figures C1 and C2, the quality flags are based on the goodness of fit, not on the uncertainty in position angle or inclination. The flag 1 objects have no clear structure remaining in their residuals, flag 2 objects have minor structures, flag 3 objects may have major structures in the residuals that could indicate a poor estimate of position angle and inclination, and flag 4 objects are clearly a poor fit. Our tests in Section 4 are run using objects with flags 1-2 and 1-3, with similar results obtained in each case.

Figures C3 and C4 show all non-group galaxies with reasonable GALFIT results (at least flag 3) that exhibit O VI absorption within  $500 \text{ km s}^{-1}$ . These illustrate the tendency for O VI absorption to be detected along the minor axis, as discussed in Section 4.2.

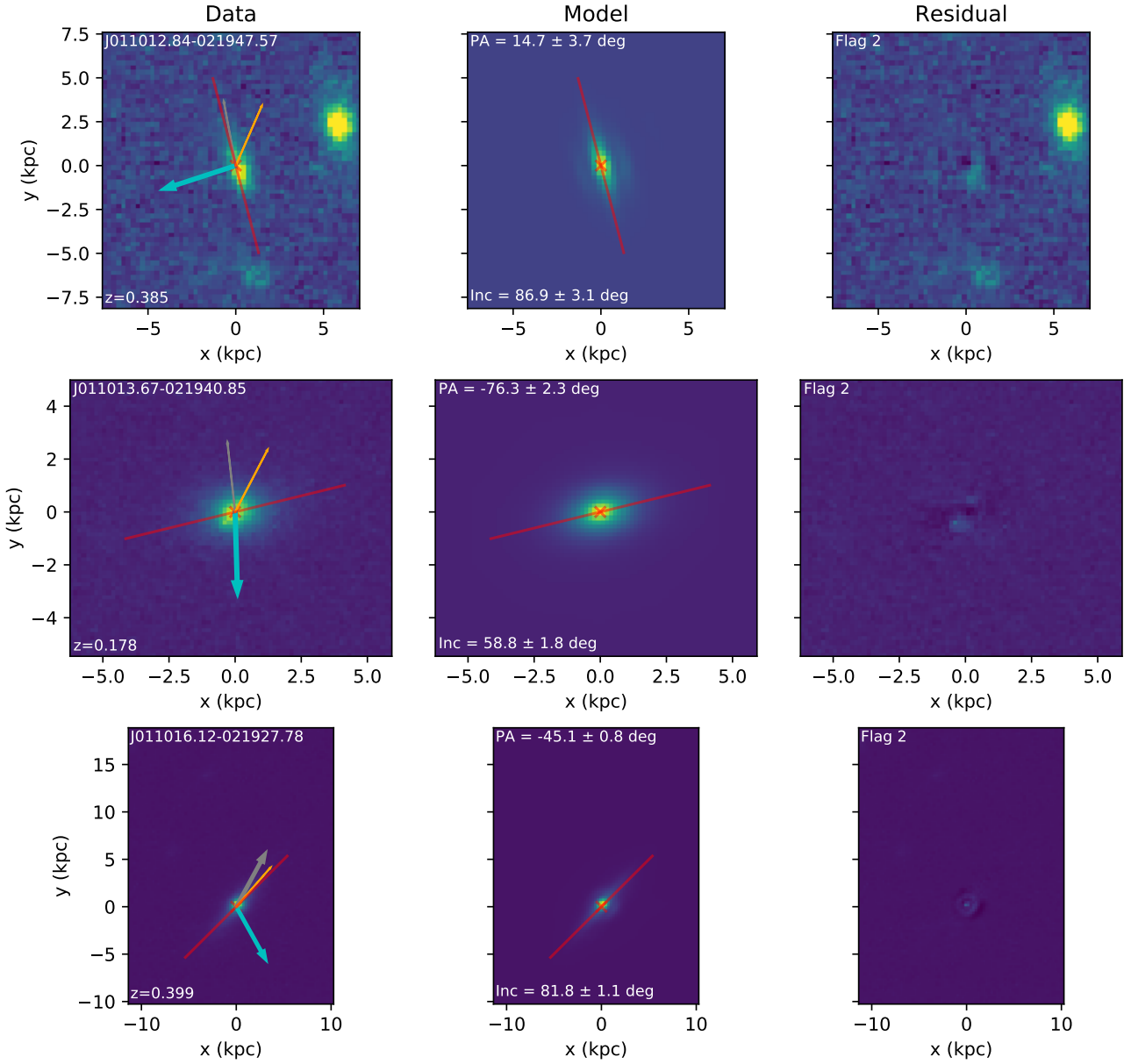
This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.



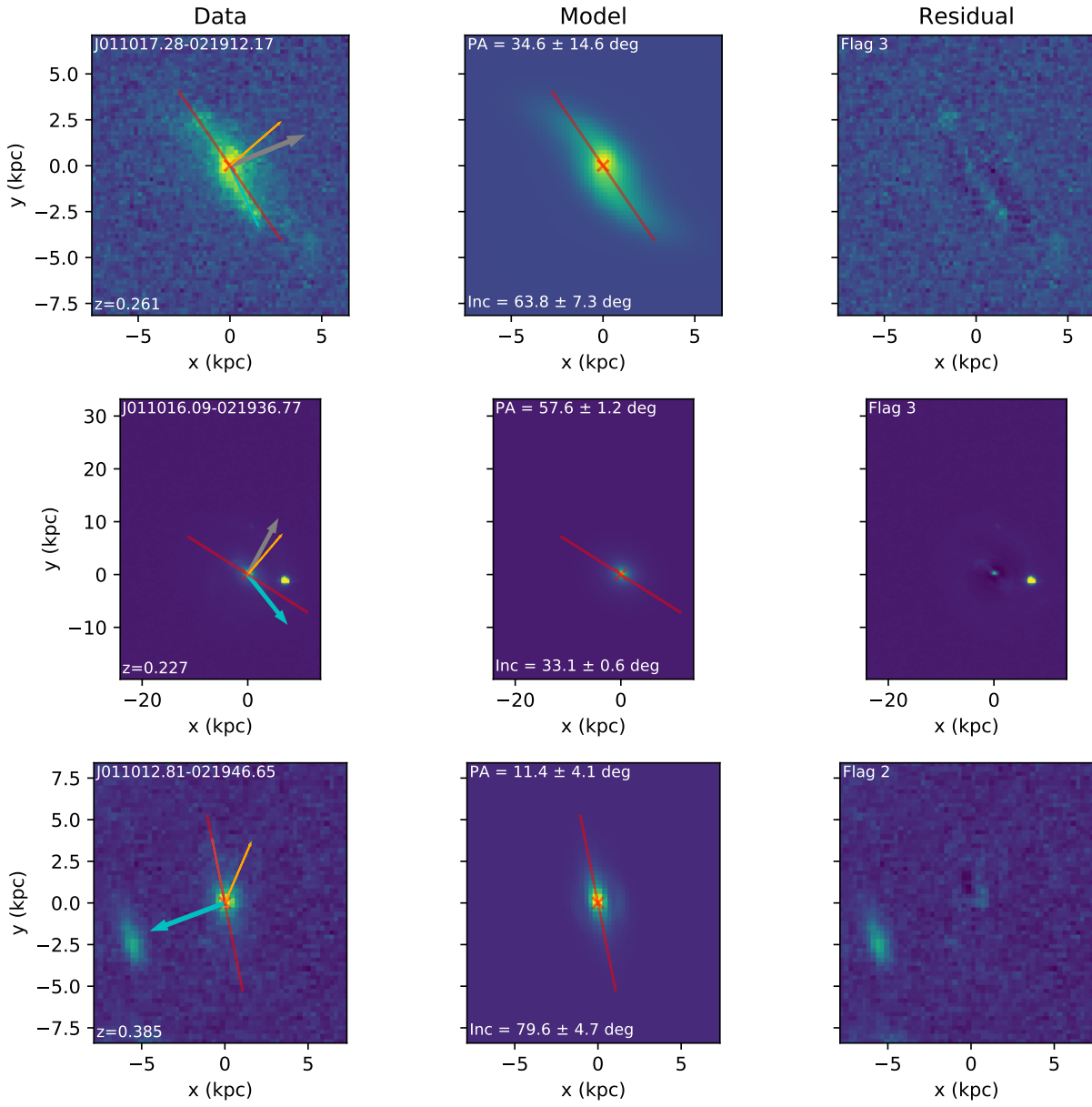
**Figure C1.** Several examples of galaxy cut-outs resulting from our GALFIT modelling. Each row shows a single galaxy, with all three panels on an identical flux scale. The left panel illustrates the data from HST, the middle panel the best fit GALFIT model, and the right the residual. The galaxy ID and redshift, as well as the resulting position angle (relative to the HST image, which is not aligned with north), inclination and quality flag, are shown in the panels. The projected major axis, as determined by GALFIT, is shown by the red line.



**Figure C2.** Further examples of galaxy cut-outs resulting from our GALFIT modelling, as in Figure C1



**Figure C3.** Galaxy cut-outs resulting from our GALFIT modelling, showing galaxies not in groups, but with observed O VI absorption (as discussed in Section 4.2). Each row shows a single galaxy, with all three panels on an identical flux scale. The left panel illustrates the data from HST, the middle panel the best fit GALFIT model, and the right the residual. The galaxy ID and redshift, as well as the resulting position angle (relative to the HST image, which is not aligned with north), inclination and quality flag, are shown in the panels. The projected major axis, as determined by GALFIT, is shown by the faded red line. The three coloured stubs in the left-hand panel point the direction of the three QSOs, with cyan towards A, grey towards B, and orange towards C. The thick stubs with arrowheads indicate that that sightline exhibits detected O VI absorption within  $500 \text{ km s}^{-1}$ , whilst the thin stubs without arrowheads indicate non-detections of O VI.



**Figure C4.** Further examples of galaxy cut-outs of objects with O VI absorption detected within  $500 \text{ km s}^{-1}$ , as in Figure C3.