

University of Milano-Bicocca

DEPARTMENT OF PHYSICS “GIUSEPPE OCCHIALINI”

Doctorate degree in Physics and Astrophysics



Galaxy evolution and the large-scale structure of the universe at high redshift

Candidate:
Marta Galbiati

Supervisor:
Prof. Sebastiano Cantalupo

Number:
800704

Contents

1	Introduction	5
1.1	Structures formation	5
1.2	Observing the cosmic gas	6
1.2.1	Absorption	7
1.2.2	Emission	10
1.3	Properties of the intergalactic gas	12
1.3.1	The Intergalactic Medium	12
1.3.2	The Cirumgalactic Medium	14
1.4	Linking gas and galaxies at cosmic noon	16
1.4.1	Evidences of a physical association	16
1.4.2	The cosmic gas observed in emission	20
1.4.3	Connecting the properties of gas and galaxies	21
1.5	The role of the large-scale environment	22
1.5.1	Impact on the gas	23
1.5.2	Impact on the galaxies	23
1.5.3	Protoclusters	24
1.6	Questions and goal of this thesis	27
2	MUSE Analysis of Gas around Galaxies (MAGG) - V: Linking ionized gas traced by C IV and Si IV absorbers to Lyα emitting galaxies at $z \approx 3.0 - 4.5$.	29
2.1	Introduction	29
2.2	Observations and Data Reduction	31
2.2.1	MUSE observations	31
2.2.2	High-resolution quasar spectroscopy	32
2.3	Data Analysis	32
2.3.1	Catalogue of C IV absorption lines	32
2.3.2	Catalogue of Si IV absorption lines	36
2.3.3	Identifying galaxies in MUSE data	37
2.3.4	Identifying galaxy groups	39
2.4	Results	39
2.4.1	Measuring correlations between C IV absorbers and LAEs	39
2.4.2	Correlation functions	43
2.4.3	Covering fraction of ionized gas	48
2.4.4	Detailed analysis of the C IV-LAE associations	50
2.4.5	Dependence on galaxy environment	52
2.4.6	Comparison with Si IV as a tracer	54
2.5	Discussion	55
2.5.1	Properties of C IV absorbers near galaxies as a function of mass and redshift	55
2.5.2	C IV and H I around $z > 3$ LAEs	58
2.6	Summary and conclusions	62

3	MUSE Analysis of Gas around Galaxies (MAGG) - VI: The cool and enriched gas environment of $z \gtrsim 3$ Lyα emitters	65
3.1	Introduction	65
3.2	Observations and data reduction	67
3.2.1	MUSE observations	67
3.2.2	Archival optical quasar spectroscopy	67
3.2.3	X-shooter near-infrared spectroscopy	67
3.3	Data analysis	69
3.3.1	Catalog of Mg II absorbers	69
3.3.2	Catalogs of LAEs, groups, and C IV absorbers	70
3.3.3	Overlap of Mg II with the H I and C IV samples in MAGG	71
3.4	Results	72
3.4.1	Searching for clustering of LAEs around Mg II absorbers	72
3.4.2	A lower LAE detection rate for Mg II gas	73
3.4.3	Characterizing the LAE overdensity around Mg II absorbers	74
3.4.4	The effect of the large-scale galaxy environment	76
3.5	Discussion	80
3.5.1	Evolution of the covering fraction with redshift	80
3.5.2	The effect of the large-scale galaxy environment across cosmic time	82
3.5.3	Radial distribution of different gas phases	84
3.6	Summary and conclusions	85
4	The MUSE Ultra Deep Field (MUDF) – VII: Probing high-redshift gas structures in the surroundings of ALMA-identified massive dusty galaxies	88
4.1	Introduction	88
4.2	ALMA observations and data reduction	89
4.2.1	Description of observations	89
4.2.2	Data calibration and imaging	91
4.3	Ancillary multi-wavelength observations	91
4.3.1	Catalog of HST-selected galaxies	92
4.3.2	Catalog of quasar absorbers	92
4.4	Extraction and characterization of ALMA sources	92
4.4.1	Continuum-selected sources in the 1.2-mm image	93
4.4.2	Line emitters in band 6 data cubes	94
4.4.3	Source detection in band 3 spectral scans	94
4.4.4	HST counterparts of ALMA continuum sources	95
4.4.5	Source fluxes and luminosities	97
4.5	Linking galaxies to absorbers in multiwavelength datasets	97
4.5.1	Gas-galaxy association rate	97
4.5.2	Properties of the ALMA galaxies around absorbers	99
4.5.3	The link between the large-scale galaxy environment and the absorbing gas	100
4.6	Summary & Conclusions	102
5	Connecting the growth of galaxies to the large-scale environment in a massive node of the Cosmic Web at $z \sim 3$	104
5.1	Introduction	104
5.2	Observations and data reduction	106
5.2.1	MUSE observations	106
5.2.2	FORS2 and HAWK-I mosaics	106
5.2.3	Multiwavelength photometry and ancillary data	108
5.3	Data analysis	109
5.3.1	Photometric selection of Lyman break galaxies	109
5.3.2	MUSE spectroscopic sample of galaxies	110
5.3.3	Control-field samples	111
5.3.4	Galaxies' physical properties and SED fitting	112
5.4	Results	112
5.4.1	A large overdensity of star-forming galaxies clustered around the quasar	112
5.4.2	Luminosity Function and amplitude of the overdensity	113
5.4.3	Clustering analysis	114
5.4.4	Galaxies' star formation activity compared to the field	116

5.4.5	Galaxies' stellar mass function	117
5.4.6	Connecting the local overdense environment to the galaxy properties	118
5.4.7	Probing the structure on larger scales	119
5.5	Discussion	120
5.5.1	MQN01 overdensity and clustering compared to other fields	120
5.5.2	The role of the local environment on galaxy formation and evolution	121
5.6	Summary and conclusions	121
6	Conclusions and future outlook	123
	Bibliography	124
A	Appendix to chapter 2	139
A.1	Theoretical comparison between C IV and Si IV absorbers as tracers of the ionized gas	139
A.2	Kinematics of C IV and Mg II absorbers	142
B	Appendix to chapter 3	144
B.1	Kinematics of C IV and Mg II absorbers	144
C	Appendix to chapter 4	146
C.1	SED of the ALMA-selected galaxies in the MUDF field	146
D	Appendix to chapter 5	148
D.1	Custom G-band and UGR selection	148
D.2	The UV luminosity function of LBGs	148
D.3	Alignment of the LBGs with the large-scale overdensity	149
D.4	MQN01 clustering compared to the MAGG survey	151
D.5	Results from SED fitting	152
E	List of publications	154

1.1 Structures formation

The majority of the inter-galactic matter, may it be baryonic or dark, is not uniformly distributed in the Universe, but rather spread in a network of elongated filaments, sheet-like walls and quasi-empty regions, called *Cosmic Web* (Bond et al., 1996). The anisotropic nature of this large-scale structure has formed as a consequence of the gravitational instabilities of the initial inhomogeneous Dark Matter (Zwicky, 1933) distribution in the Universe. Indeed, the Λ CDM cosmological model (Planck Collaboration et al., 2020a) predicts the presence of primordial quantum fluctuations in the Dark Matter energy density field which, in the overdense regions, collapsed under the effect of their own gravity and formed structures (Gamow & Teller, 1939; Lifshitz, 1946; Peebles, 1965; Silk, 1968; Peebles & Yu, 1970; Zel'dovich, 1970).

The density field of the Dark Matter, $\rho(\vec{r}, t)$, is a function of time, t , and position, $\vec{r}$. The average background density, $\bar{\rho}(t)$, as a function of time can be computed as the spatial average over a volume much larger than the typical size of the largest structure in the Universe:

$$\bar{\rho} = \frac{1}{V} \int_V \rho(\vec{r}, t) d^3r \quad (1.1)$$

The dimensionless density fluctuations can thus be written as the local density contrast against the expected average background value:

$$\delta(\vec{r}, t) = \frac{\rho(\vec{r}, t)}{\bar{\rho}(t)} - 1 \quad (1.2)$$

where $\delta < 0$ is for underdense regions, $\delta > 0$ indicate overdense regions and $\delta = 0$ corresponds to $\rho = \bar{\rho}$. The evolution and growth of such fluctuations can be modelled analytically only in the limit of small perturbations of the density field ($\delta \ll 1$), an approximation that is known as *linear perturbation theory* (see Bernardeau et al., 2002 for a complete review and mathematical treatment). However, this approximation is no longer valid on scales of tens of Mpc, where the perturbations grow in a non-linear regime and their evolution depends on the matter content of the perturbation itself (see e.g. Zel'dovich, 1970; Gunn & Gott, 1972; Press & Schechter, 1974). In case of small perturbations, $\delta \ll 1$, assuming that the Universe is dominated by non collision dark matter (that is, ignoring the pressure gradient term), the continuity, Euler and Poisson equations fully describe the motion of a fluid. Calling ϕ the gravitational potential and linearising the equations in the limit of small peculiar velocity $\vec{v}$ relative to the Hubble parameter $H(t)$, these read as:

$$\frac{\partial \delta(\vec{r}, t)}{\partial t} + \frac{1}{a} \Delta_{\vec{r}} \cdot \vec{v}(\vec{r}, t) = 0 \quad (1.3)$$

$$\frac{\partial \vec{v}(\vec{r}, t)}{\partial t} + \frac{\dot{a}}{a} \vec{v}(\vec{r}, t) = -\frac{1}{a} \nabla \phi(\vec{r}, t) - \frac{1}{a} \frac{\nabla P}{\bar{\rho}[1 + \delta(\vec{r}, t)]} \quad (1.4)$$

$$\Delta_{\vec{r}} \phi(\vec{r}, t) = 4\pi \bar{\rho} G \delta(\vec{r}, t) \quad (1.5)$$

where a is the scale factor accounting for the expansion of the Universe, $\vec{r}$ is the comoving coordinate and P the pressure. Assuming that the Universe is dominated by non-collisional Dark Matter, the

pressure term can be neglected. While Euler's equation link the changes in the fluid velocity to the gravitational acceleration, Poisson equation states the gravitational potential as a function of the overdensity. These equations lead to the following linear and homogeneous second order differential equation which fully captures the evolution with time of the density perturbations:

$$\frac{\partial^2 \delta(\vec{r}, t)}{\partial t^2} + 2H \frac{\partial \delta(\vec{r}, t)}{\partial t} - 4\pi \bar{\rho} G \delta(\vec{r}, t) = 0 \quad (1.6)$$

suggesting that the evolution of the perturbations is driven by the expansion of the Universe (second term) and gravity (third term). The solution of this equation can finally be written as a linear combination of two independent solutions, decoupling the spatial and time evolution of the perturbations: $\delta(\vec{r}, t) = D_+(t)\delta_+(\vec{r}, 0) + D_-(t)\delta_-(\vec{r}, 0)$. One solution leads to a decaying perturbation, while D_+ represent a perturbation that grows with time and possibly collapse.

However, the above linear regime is no longer a good approximation to model the evolution of the perturbations on scales of few tenths of Mpc. Starting from the 1970s (e.g. [Doroshkevich & Shandarin, 1978](#)), the evolution of the density field in the non-linear regime has been studied by using numerical simulations. They revealed that the initial density perturbations grew linearly up to reaching 1.68 times the critical matter density, then detached from the Hubble flow and collapsed under their gravity forming structures, called *dark matter halos* ([Gunn & Gott, 1972](#); [Press & Schechter, 1974](#)), that are gravitationally bound and in virial equilibrium. Within the Λ CDM cosmological model, density fluctuations of masses $\approx 10^5 - 10^6 M_\odot$ were the first to grow and collapse at $z \approx 30-20$ ([Barkana & Loeb, 2001](#), for a review). Larger halos formed hierarchically from the mergers of smaller structures or the collapse of larger density fluctuations ([White & Rees, 1978](#); [Kauffmann & White, 1993](#); [Lacey & Cole, 1993](#)).

The baryonic matter was initially fully ionized and coupled to the radiation through scatterings of photons and electrons. Such condition provided pressure support to the baryons against their own gravity and the additional gravitational pull exerted by the amplified density fluctuations of the Dark Matter. At $z_{\text{rec}} \approx 1100$, the average energy per photon gradually dropped below the ionization energy of Hydrogen (13.6 eV) due to the expansion of the Universe. Electrons and protons could thus recombine into a neutral state without being efficiently photoionized and soon after baryons decoupled from radiation. Free of any radiation pressure support, baryons started to fall in the potential wells created by the amplified fluctuations of the dark matter density field ([Bardeen et al., 1986](#); [Mo & White, 1996](#)) and here collapsed under its own self-gravity forming structures. Within this framework, the distribution of the inter-galactic matter in the Cosmic Web is the result of the combination of a Gaussian random initial density field ([Adler, 1981](#); [Bardeen et al., 1986](#)) and the evolution of the initial density fluctuations entirely due to gravity ([Peebles, 1980](#)). [Zel'dovich \(1970\)](#) considered the evolution of a pressureless and collisionless form of matter and introduced a Lagrangian approximation to solve Eq. 1.6 in non-linear regime in terms of how a particle moved from its initial position. He showed that the first structures to form from ellipsoid clouds of gas would be *pancake-like* due to the gravitational collapse occurring mostly along one of the axis. Subsequent perpendicular collapse forms filaments. From underdense regions called voids, matter flows along the pancake sheets into the higher-density filaments, and along the filaments toward the dark matter halos (see the review by [Primack, 2024](#)). [Bond et al. \(1996\)](#) later showed that the geometry of the structures of the Cosmic Web arises from small-amplitude quantum fluctuations in the primordial universe, which then expands during inflation from microscopic to cosmic scales. It was only in 1962 that [Eggen et al. \(1962\)](#) developed the first models of the formation of galaxies from the gravitational collapse of gas clouds. Since then, the modern theory of galaxy formation and evolution ([White & Rees, 1978](#)) states that galaxies form from baryons that accrete towards the center of the potential well, cool down and collapse under their own gravity to form stars (see e.g., [White & Frenk, 1991](#)). Once formed, galaxies grow by accreting from the filaments of the Cosmic Web gas that will be in part converted into stars and in part retained within the Interstellar medium as a reservoir to feed further star formation. Unveiling the fundamental link between galaxies and the gas locating within and outside their disks on different scales is thus essential to paint a complete picture of how galaxies form and evolve across cosmic time.

1.2 Observing the cosmic gas

The neutral and enriched gas surrounding galaxies on different scales plays an essential role in regulating their formation and evolution. If on the one hand, the cool and dense gas retained in

the Interstellar Medium (ISM) inside the disk of a galaxy is a direct reservoir of gas immediately available for the galaxy to form stars, accreting gas from the filaments of the Intergalactic Medium (IGM) is essential for a galaxy to grow and sustain its star formation activity (Kereš et al., 2005; Dekel et al., 2009). The gas at the interface between the galaxies’ ISM and the IGM is called Circum Galactic Medium (CGM). Although there is no clear definition of the physical boundaries of the CGM, it is broadly defined as the gas located between the disk of a galaxy, or the boundary of its ISM, and the virial radius of the dark matter halo. The intergalactic gas is typically diffuse and too low density to form stars, which makes it challenging to detect directly. The basics of the observing technique mostly adopted in the literature are described below in Sections 1.2.1 and 1.2.2, while the main properties (density and temperature) of the IGM and CGM estimated from such observations are reviewed in Sections 1.3.1 and 1.3.2, respectively.

Historically, the efforts for detecting the gas between galaxies begun with Field (1959), who attempted to detect the 21-cm absorption line¹ of neutral hydrogen along the line of sight to the extragalactic radio source Cygnus A. Although no detection was made at that time, the discovery of quasi-stellar objects (QSOs, Hazard, 1963; Schmidt, 1963) a few years later paved the way toward a significant improvement in the observations of the intergalactic neutral hydrogen (see, e.g., Gunn & Peterson, 1965; Bahcall & Salpeter, 1965) and in the study of the gas enriched by metals (that is, atoms heavier than He, see, e.g., Burbidge et al., 1966; Stockton & Lynds, 1966) by searching for the absorption signatures in the quasar spectra produced by clouds of gas in the foreground. Complementary to detect the cosmic gas in absorption, imaging with narrowband filters (e.g., Steidel et al., 2011; Cantalupo et al., 2014) and long-slit spectroscopy (e.g., Hennawi & Prochaska, 2013) have traditionally been employed as main techniques to also observe its emission lines. However, narrowband imaging allows observing a relatively large region of the universe, but is severely limited by the background, while spectroscopy is less affected by the background, but probe only a small area. Significant advancements have been made with the advent of sensitive integral field spectrographs (IFSs), such as the Multi-Unit Spectroscopy Explorer (MUSE, Bacon et al., 2010) at ESO Very Large Telescope (VLT) and the Keck Cosmic Web Imager (KCWI, Martin et al., 2010; Morrissey et al., 2018) at Keck II telescope. The products of IFSs are datacubes obtained by slicing a large field-of-view into small slitlets that are dispersed by a series of spectrographs and then re-assembled into three-dimensional data, which carries 2D spatial information and an additional dimension recording the spectrum of each pixel.

1.2.1 Absorption

As mentioned above, a powerful technique for probing the diffuse and low-density intergalactic gas is to search for resonant absorption features imprinted in the spectra of background bright objects, such as quasars, gamma ray (see, e.g., Wolfe et al., 2005; Bolmer et al., 2019; Heintz et al., 2019, 2023) or fast radio bursts (Prochaska et al., 2019), or galaxies themselves (see, e.g., Steidel et al., 2010; Cooke & O’Meara, 2015; Péroux et al., 2018). We review here the basic concepts of absorption and refer to Draine (2011) for a more complete and detailed description of the topic. As the light emitted by a distant bright source propagates across the Universe to an observer, it can eventually pass through clouds of intervening atoms, ions or molecules which can absorb a fraction of the incident radiation moving an electron from its lowest energy location, or *ground state*, to a higher energy state and then falls back to its original state re-emitting a photon in a random direction, scattered out of the observed line of sight. Given the intrinsic and unabsorbed flux density emitted by a background source, $F_0(\nu)$, at a given frequency, ν , this is related to the flux at the observer via the optical depth, τ_ν , of the intervening matter:

$$F(\nu) = F_0(\nu)e^{-\tau(\nu)} \quad (1.7)$$

The optical depth is a measure of the opacity of a medium: it increases with the path that light crosses in the cloud, ds , and depends on the attenuation coefficient, $k(\nu)$, according to the equation: $d\tau(\nu) = k(\nu)ds$. The optical depth can also be related to the column density, $N = \int n ds$, of the absorbing material, which is defined as the density of the particles, n , integrated along the

¹The hydrogen 21-cm line, also called h I line, is a hyperfine line associated with the spin-flip transitions. When the spin of the electron flips from being parallel to antiparallel relative to that of the proton, a photon is emitted at $\nu \approx 1.42$ GHz, corresponding to $\lambda \approx 21$ cm in vacuum. Detecting this line is of particular interest as it is expected to display angular and frequency structures due to the fluctuations of the matter density field (see, e.g., a review by Barkana & Loeb, 2007).

traversed path through the material, s :

$$\tau(\nu) = N\sigma(\nu) = N \frac{\pi e^2}{m_e c} f_{\text{osc}} \phi(\nu) \quad (1.8)$$

where $\sigma(\nu)$ is the absorption cross-section, f_{osc} is the oscillator strength of the atomic transition (see a compilation of tabulated values for the most relevant transitions in [Cashman et al., 2017](#)), e and m_e are the charge and the mass of the electron, respectively, c is the speed of light and $\phi(\nu)$ is the normalized profile defined as $\int \phi(\nu) d\nu = 1$. The line profile is typically modelled as a Voigt function, that is a convolution of a Gaussian, capturing the internal motion of the gas, with Lorentzian wings, reflecting the collisional broadening of the light. The column density of the absorbing gas can be estimated using the so-called rest-frame *equivalent width*, W_r , that is directly observable and defined as:

$$W_r = \int \frac{F_0(\lambda) - F(\lambda)}{F_0(\lambda)} d\lambda = \int [1 - e^{-\tau(\lambda)}] d\lambda \quad (1.9)$$

where the flux densities from Eq. 1.7 are now written as a function of the rest-frame wavelength, λ . At the redshift of the absorber, z_{abs} , this is related to the observed wavelength through $\lambda_{\text{obs}} = \lambda(1 + z_{\text{abs}})$. The equivalent width is a measure of what fraction of the intrinsic flux is missing at the observer and thus absorbed by an intervening cloud of gas. The *curve-of-growth* technique (see, e.g., [Draine, 2011](#)) is the classical approach used to map how the equivalent width varies as a function of the column density of the absorbing material in different regimes, depending on the optical depth of the gas cloud. In the optically thin regime, $\tau(\nu) \ll 1$, Eq. 1.9 can be solved analytically. Assuming that the absorbers have a Gaussian velocity distribution with dispersion σ_v , the optical depth can be expressed as a function of the velocity shift from the line center that is due to the motion of the particles, $\Delta\nu = c(\nu_0 - \nu)/\nu$ as:

$$\tau(\nu) = \tau_0 e^{[-\Delta\nu^2/(\sqrt{2}\sigma_v)^2]} \quad (1.10)$$

where the optical depth at the center of the line, τ_0 , is defined as:

$$\tau_0 = \sqrt{\pi} \frac{e^2}{m_e c} \frac{N f_{\text{osc}} \lambda}{\sqrt{2}\sigma_v} \quad (1.11)$$

The analytical solution of Eq. 1.9 in the optically thin regime can thus be obtained from the Taylor expansion of $(1 - e^{-\tau}) = \tau - \tau^2/2 + o(\tau^2)$ and reads:

$$W_r \approx \sqrt{\pi} \frac{\sqrt{2}\sigma_v}{c} \tau_0 \lambda = \left(\pi \frac{e^2}{m_e c^2} \right) f_{\text{osc}} \lambda^2 N \quad (1.12)$$

where we explicitly highlighted the linear dependence on the column density by using Eq. 1.11. The quantities in the brackets are all constants, while f_{osc} and λ varies depending on the transition. On the contrary, in the optically thick regime, $\tau(\nu) > 1$, only a lower limit on the column density can be estimated by inverting Eq. 1.12. The main advantage of deriving the column density by directly measuring the equivalent width of the line is that the latter does not depend on the resolution of the instrument, and can thus provide robust constraints also for under-resolved absorption lines.

This principle can be applied to both hydrogen and metal absorbers. In the first case, the atomic hydrogen is mostly found in the electronic ground state $1s$. The allowed transitions for this energy state are referred to as *Lyman series*: Lyman α ($1s \rightarrow 2p$, $\lambda_{\text{RF}} = 1215.67 \text{ \AA}$), Lyman β ($1s \rightarrow 3p$, $\lambda_{\text{RF}} = 1025.73 \text{ \AA}$), Lyman γ ($1s \rightarrow 4p$, $\lambda_{\text{RF}} = 972.54 \text{ \AA}$), and so on. The excited $2p$ state is actually a doublet, but the fine-structure splitting of these two levels is of the order of $\Delta\nu \approx 1.3 \text{ km s}^{-1}$ and the two lines thus appear to be always blended as their width is dominated by the thermal broadening. The Ly α line is extremely important as it is a direct tracer of the column density of atomic hydrogen. In a typical quasar spectrum, all these absorption lines are observed at bluer wavelengths relative to the Ly α emission line of the quasar, in a region that is called *Lyman alpha forest* (see Fig. 1.1 and a detailed discussion in Section 1.3.1). The absorption lines of the Lyman series converge to a wavelength limit called *Lyman limit* ($\lambda_{\text{RF}} = 911.75 \text{ \AA}$), which corresponds to the energy required for an electron in the ground state to escape out of the atom (13.6 eV), leaving the hydrogen ionized. All the photons above this limit, also called *Lyman continuum photons*, have energies higher than 13.6 eV and can thus be absorbed by neutral hydrogen atoms, resulting in a severe decrement of the observed flux of the quasar (see Fig. 1.1).

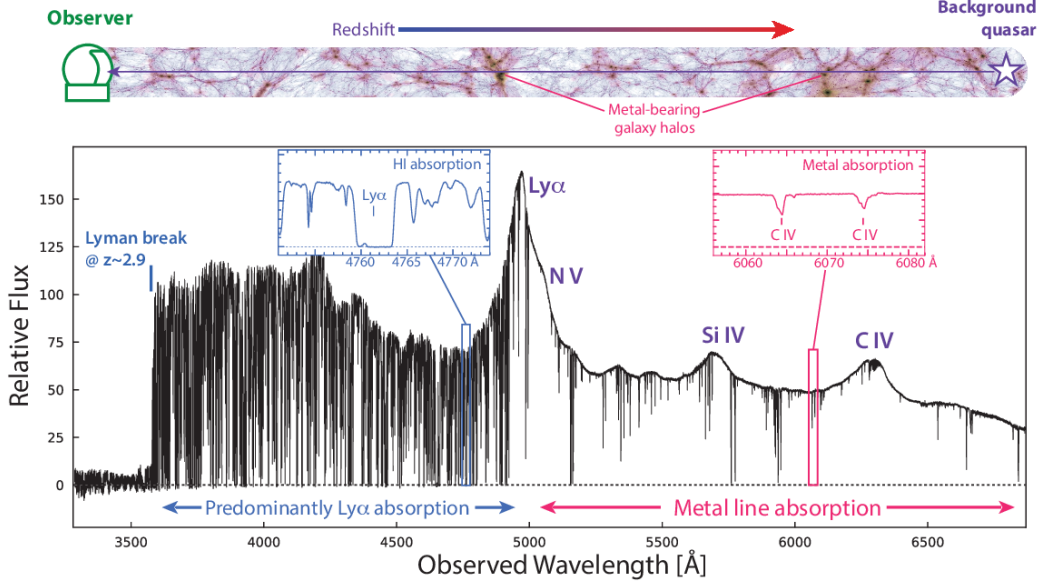


Figure 1.1: Example of using a background quasar to probe the foreground cosmic gas in absorption from Péroux & Howk (2020). Upper panel: illustration of a light ray traveling across the filaments of the cosmic web and crossing the halos of galaxies (darkest regions) to an observer. Bottom panel: VLT/UVES high-resolution spectrum of a $z \approx 3.1$ quasar from D’Odorico et al. (2016). The region dominated by Ly α absorption is shown blueward the Ly α emission of the quasar (left inset), while metal lines are detected redward the Ly α of the quasar (right inset).

Inside galaxies, dense clumps of gas are eventually able to cool down to temperatures that are low enough for the gas self-gravity to overcome the pressure support and condense into stars (see McKee & Ostriker, 2007, for a review). During their lifetime, stars synthesize metals in their cores starting from the fusion of hydrogen into Helium and proceed producing heavier and heavier elements until only a Carbon or Iron core, depending on their masses, is left surrounded by shells of lighter elements. Strong stellar winds and supernovae explosion can be powerful enough to transfer a large amount of energy to the gas enriched by metals that escape the disks and mixes with the surrounding intergalactic medium. Elements, such as C, Mg, and Si, must have formed in the first stars (Population III stars = metal-free) in the primordial Universe (see the review by Nomoto et al., 2013). Except for the inert gases He and Ne, all other neutral atoms have resonance lines with rest-frame wavelengths redder than the Lyman limit, which make them detectable in the same sightlines where neutral hydrogen is observed (see Fig. 1.1).

Moreover, spanning a broad range of densities and temperatures, this gas is not necessary neutral and gives rise to different ions. It is possible, by assuming a model for the ionization state of the gas (see, e.g., Fumagalli et al., 2016), to estimate the ionization fraction $X_{j,i}(T, n_H)$ for an element i in the ionization state j , as a function of the gas temperature and density. These two properties of the gas can be related to the column density of the parent element: $N_i = X_{j,i}^{-1}(T, n_H) N_{j,i}$. Since it can be used to retrieve the column density, and thus the abundance, of the original element from the observed ions, $X_{j,i}$ is often referred to as *ionization correction*. Given a set of ions, comparing the column densities of different elements, M , relative to that of hydrogen to a model is a powerful method to estimate not only the temperature and density, but also the metallicity of the gas in solar units: $[M/H] = \log(N_M/N_H) - \log(N_M/N_H)_\odot$. An accurate estimate of the ionization corrections requires solving the equations for the recombination and ionization rate due to both radiation and collisions, which depends on the radiation field, the geometry and composition of the material and on the thermal state of the gas. Numerical radiative transfer codes (e.g., CLOUDY Ferland et al., 1998, 2013; Gunasekera et al., 2023) are required to solve such a large set of coupled equations and to build grids of models (see, e.g., Bergeron & Stasińska, 1986; Prochaska et al., 2004; Werk et al., 2014) that can be compared to the observations and allow estimating several properties of the gas, such as density, temperature, metallicity and radiation field (see, e.g., Crighton et al., 2015; Fumagalli et al., 2016).

1.2.2 Emission

Although observing the cosmic gas in absorption is powerful and allows probing the gas down to relatively low density, it only provides one dimensional information which are limited to the line-of-sight direction. A complementary approach is thus to observe the gas directly in emission and preserve the spatial information in two dimensions. To this end, the Ly α transition is once again targeted as one of the major tracers of the cosmic gas as neutral hydrogen is the most abundant element in the Universe, and it is intrinsically the brightest hydrogen transition. The quantity that is used to solve the radiative transfer equations and estimate the Ly α luminosity is the emissivity, j , which is defined as the quantity of energy, dE , emitted by a source per unit volume, unit time and unit of solid angle ($d\Omega$):

$$j = \frac{dE}{dV dt d\Omega} \quad (1.13)$$

In the Universe, two mechanisms can provide electrons of enough energy to excite them up to the $2p$ state (and, in general, to higher energy state $n > 1$): collisions and recombination (see, e.g., [Cantalupo, 2017](#), for a review). In the first case, an electron can collide with a hydrogen atom, loose part of its kinetic energy and leave the atom in an excited state. In other words, this process converts the thermal energy of the gas into radiation. The efficiency at which this conversion happens depend on the relative velocity of the free electron and the hydrogen atom. The production rate of Ly α photons via collisional excitation is given by:

$$R_{\text{coll}} = n_{\text{r}} n_{\text{H}} q_{1s2p} \quad (1.14)$$

where n_{e} and n_{H} are the number densities of the two particles and q_{1s2p} is the rate coefficient (see, e.g., [Aggarwal, 1983](#); [Scholz et al., 1990](#); [Osterbrock, 1962](#)), which includes the dependence on the velocity of the electrons. Assuming that the velocity distribution of the electrons is given Maxwellian distribution, then it is uniquely determined by the temperature of the gas. As the electron transfers its thermal energy to the hydrogen atom, it cools down at a rate, per unit volume, that is given by:

$$\frac{dE_{\text{th}}}{dV dt} = n_{\text{e}} n_{\text{H}} C(T) \quad (1.15)$$

where the total cooling rate is defined as $C(T) = \sum_u q_{1s \rightarrow u} \Delta E_{1s \rightarrow u}$ summing over all the excited states, u . Due to the strong temperature dependence of the number density of electrons moving at velocity that are high enough to excite hydrogen atoms, $C(T)$ increases by ≈ 6 orders of magnitude from $T \approx 10^4$ K to $T \approx 10^6$ K (see the analytic formula by [Black, 1981](#)). At temperatures $T \lesssim 10^5$ K, $\approx 45 - 60\%$ of the cooling emerges as Ly α photons (see, e.g., fig. 7 in the lectures by [Dijkstra, 2019](#)). The cooling rate also depends on the products of the densities, that is, on the ionization state of the gas. If the gas in collisional ionization equilibrium, its ionization state, and thus the cooling rate per unit volume, is fully determined by the temperature and the overall gas density squared.

Recombination occurs due to the radiative decay of an electron from an excited quantum state (n, l) to the ground state ($n = 1, l = 0$) via intermediate states (n', l') (see, Fig. 1.2). The probability that this radiative cascade results in the emission of a Ly α photon is given by the product of the probability that a cascade from the (n', l') results in the emission of a Ly α photons times the probability that a higher (n, l) ionization state decays into the lower (n', l') one:

$$P_{\text{Ly}\alpha}^{(n,l)} = \sum_{n',l'} P_{(n',l')}^{(n,l)} P_{\text{Ly}\alpha}^{(n',l')} \quad (1.16)$$

where $P_{(n',l')}^{(n,l)}$ is the *branching ratio*, that is, the fraction of particles that decay from (n, l) to (n', l') over the total number of particles that decay into all the permitted ($|l - l'| = 1$) states (see, [Dijkstra, 2014](#)). Finally, the probability that an arbitrary recombination event results in the emission of a Ly α photon, assuming that it is known the atom is in the state (n, l) after a recombination event, is:

$$P(\text{Ly}\alpha) = \sum_{n_{\text{min}}}^{\infty} \sum_{l=0}^{n-1} \frac{\alpha_{nl}(T)}{\alpha_{\text{tot}}(T)} P_{\text{Ly}\alpha}^{(n,l)} \quad (1.17)$$

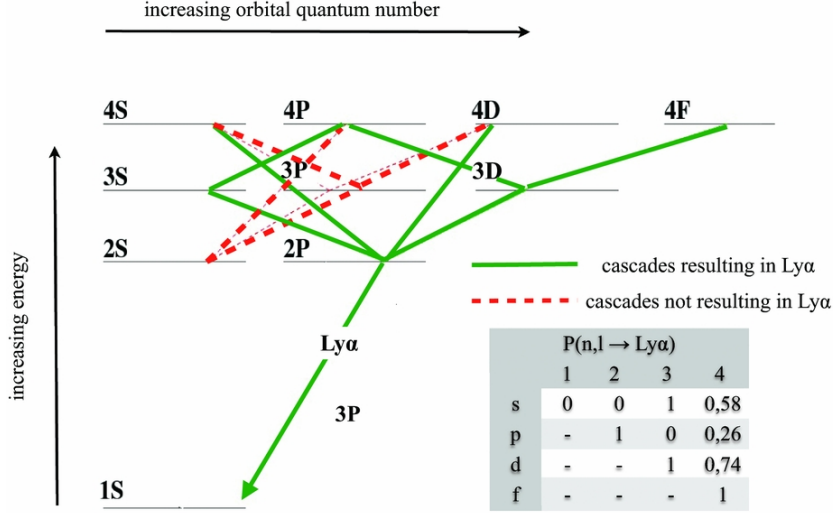


Figure 1.2: Energy levels of a hydrogen atom from Dijkstra (2014). The levels are ordered by increasing energy (to bottom) and orbital quantum number (left to right). Each state is described by two quantum numbers as (n, l) , where n and l are the principal and the orbital quantum numbers, respectively. The inset table records the probability $P_{Ly\alpha}^{(n,l)}$ (see, Eq. 1.16) that a cascade from the (n, l) level results in the emission of a Ly α photon.

where $\alpha_{nl}(T)$ is the recombination coefficient, defined as the number of recombination events into the state (n, l) per unit volume per unit time, and $\alpha_{\text{tot}}(T) = \sum_{n_{\text{min}}}^{\infty} \sum_{l=0}^{n-1} \alpha_{nl}(T)$ is the total recombination coefficient, both depending on the temperature (see, e.g., Rubiño-Martín et al., 2006). Here, n_{min} depend on the physical conditions of the medium, and we can distinguish between two different scenarios:

- *Case-A* recombination: the medium in which recombination takes place is optically thin to photons at any frequencies and direct recombination to the ground state is allowed and results in the emission of a photon with energy ≥ 13.6 eV. Therefore, in this case $n_{\text{min}} = 1$;
- *Case-B* recombination: the medium is optically thick to all the photons of the Lyman series ($n = 1$) and to all the ionizing photons emitted following Case-A recombination to the ground state, which are immediately absorbed by a nearby neutral hydrogen atom (*on the spot* approximation). In this case, $n_{\text{min}} = 2$.

Following the formalism from Cantalupo et al. (2008), we introduce the effective Ly α emission coefficient (in the *on the spot* approximation), which is given by the product of the total recombination coefficient of hydrogen, excluding those to the ground level, $\alpha^{\text{B}}(T)$ times the fraction of those recombinations that result in the emission of a Ly α photon, $\epsilon_{Ly\alpha}^{\text{B}}(T)$ (defined in eq. 1.17 and adapted for case B recombination), as: $\alpha_{Ly\alpha}^{\text{eff}} = \epsilon_{Ly\alpha}^{\text{B}}(T) \alpha^{\text{B}}(T)$. Cantalupo et al. (2008) found that $\epsilon_{Ly\alpha}^{\text{B}}(T)$ is only a weak function of the temperature, and at typical IGM temperatures $\epsilon_{Ly\alpha}^{\text{B}}(T = 10^4 \text{ K}) = 0.68$. The Ly α emissivity produced by case B recombination is given by:

$$j_{Ly\alpha}^{\text{rec}} = n_e n_p \alpha_{Ly\alpha}^{\text{eff}} \frac{E_{\alpha}}{4\pi} \quad (1.18)$$

where n_e and n_p are the number densities of the free electrons and protons, respectively. Eventually, before recombining, free electrons can interact and collide with neutral hydrogen atoms in the low density inter-galactic gas transferring a fraction of their energy that, if high enough, is able to excite these atoms. The efficiency of such collisional excitation is a strong function of the temperature. In this case, the Ly α emissivity produced by collisional excitation is given by the equation:

$$j_{Ly\alpha}^{\text{coll}} = n_e n_{\text{HI}} q_{Ly\alpha}^{\text{eff}} \frac{E_{\alpha}}{4\pi} \quad (1.19)$$

where n_e and n_{HI} are the number densities of free electrons and neutral hydrogen atoms, respectively, and $E_{\alpha} = 10.2$ eV is the energy difference from a $n = 2$ level to $n = 1$ and thus the

energy of a Ly α photon. The effective collisional coefficient is defined as $q_{\text{Ly}\alpha}^{\text{eff}} = q_{1s2p} + q_{1s3s} + q_{1s3d}$ including the contribution from all the excited states that can decay, emitting a Ly α photon in the *on the spot* approximation for case B recombination, up to $n = 3$. The total Ly α emissivity arising from recombinations and collisional excitation is given by $j_{\text{Ly}\alpha} = j_{\text{Ly}\alpha}^{\text{rec}} + j_{\text{Ly}\alpha}^{\text{coll}}$. The larger the fraction of ionized-over-neutral gas, the larger the contribution from the radiative recombinations, while the higher the temperature, the larger the contribution from collisional excitations, especially at temperatures $T \approx (2 - 5) \times 10^4$ K (see, fig. 1 in Cantalupo et al., 2008).

Searching for Ly α photons emitted from the inter-galactic gas would provide a direct detection of this diffuse medium in which galaxies form and evolve. In addition, Ly α emission from recombination typically arises from extended regions and can be powered by external sources of radiation, such as bright quasars or an ionizing background radiation produced by a large number of galaxies and AGN. In other words, a cloud of neutral hydrogen can be ionized by the incident radiation produced by an external source and be excited to higher energy levels, which then eventually decay to the ground state emitting a Ly α photon. This process of emission of recombination radiation following photoionization is called *fluorescence* (Hogan & Weymann, 1987; Gould & Weinberg, 1996) and is expected to be able to efficiently ionize the cosmic gas and boost the Ly α emission arising from radiative recombination, especially in presence of bright and powerful sources, up to detectable levels (see, e.g., Cantalupo et al., 2005, 2007; Kollmeier et al., 2010). In the most extreme case in which the incident radiation is bright enough to completely photo-ionize the cosmic gas (that is, $n_e = n_p = n$), the recombination Ly α emissivity is proportional to the gas density squared through Eq. 1.18: $4\pi j_{\text{Ly}\alpha}^{\text{rec}} = n^2 \alpha_{\text{Ly}\alpha}^{\text{eff}} E_\alpha$. What we observe is the Ly α surface brightness, $SB_{\text{Ly}\alpha}^{\text{rec}}$, measured in units of $\text{erg s}^{-1} \text{cm}^{-2} \text{arcsec}^{-2}$ and defined as the Ly α emissivity integrated along the line-of-sight direction over a depth l : $SB_{\text{Ly}\alpha}^{\text{rec}} = \int_l j_{\text{Ly}\alpha}^{\text{rec}} dz = \alpha_{\text{Ly}\alpha}^{\text{eff}} E_\alpha \int_l n^2 dz$. This observable depends on the distribution and inhomogeneities of the cold ($T \lesssim 10^5$ K) gas along the line of sight, (see, Pawlik et al., 2009; Cantalupo, 2017; Cantalupo et al., 2019).

Recently, Liu et al. (2024) investigated the detectability of the neutral hydrogen retained in filaments of the cosmic web in emission across the redshift range $z \approx 2 - 5$ by using the hydro-dynamical simulations ILLUSTRISTNG50. They estimated over 70% of the Ly α emission to have a surface brightness of the order of $\approx 10^{-21} \text{erg s}^{-1} \text{cm}^{-2} \text{arcsec}^{-2}$ assuming it originates from recombination and collisional excitation processes. By simulating realistic observations with MUSE, they concluded that this instrument is able to only detect Ly α emission arising from dense hydrogen in the galactic centers. Similarly, Tsai et al. (2024) calculated the expected surface brightness of the Ly α emission emerging from galaxies, CGM, and IGM across cosmic time by post-processes the cosmological simulations from ILLUSTRISTNG100 and THESAN. The Ly α emission from the diffuse gas in CGM and IGM is found to peak at $z \approx 4$ and decreases at lower redshifts. They also found that collisional excitation is the dominant powering mechanism for the Ly α emission at gas densities of $n_{\text{H}} \approx 10^{-5} - 10^{-2} \text{cm}^{-3}$ and temperatures of $T \approx 10^4 - 10^5$ K. Recombination, instead, mostly occurs in more diluted gas with $n_{\text{H}} \approx 10^{-8} - 10^{-5}$ and $T \approx 10^4$ K and yields a fainter Ly α emission. With future generation of instruments with improved detection sensitivity, the filaments of the cosmic web are expected to become more and more visible, especially at lower redshift.

1.3 Properties of the intergalactic gas

A deep understanding of the intergalactic gas as well as of the processes that link it to the formation and evolution of galaxies requires employing the techniques described in Section 1.2 to measure properties such as the density, temperature, kinematics, composition and ionization state of the gas located outside the galactic disks. Briefly reviewed in this section are the key properties of the Intergalactic (see, Section 1.3.1) and Circumgalactic (see, Section 1.3.2) gas, which pave the way to study this gas in connection with galaxies across cosmic time.

1.3.1 The Intergalactic Medium

Most of the cosmic matter in the Universe is located between galaxies in the IGM and spread in a complex network of voids, sheets, filaments and nodes (see Sec. 1.1). The filaments of the Cosmic Web connects, on small scales, the dark matter halos that host galaxies and, on larger scales, where multiple filaments intersect in the nodes, the halos that host more massive structures, such as groups and clusters (see, e.g., Colberg et al., 2005; Aragón-Calvo et al., 2010). By comparing different methods typically employed in the literature to reconstruct the Cosmic Web structure

from the galaxy overdensity distribution, Libeskind et al. (2018) found that, at $z = 0$, most of the volume fraction, $\approx 70\%$, is occupied by voids, which are only sparsely populated with halos with masses (M_h) $M_h \lesssim 10^{13.5} h^{-1} M_\odot$. The most massive halos, $M_h \gtrsim 10^{14} h^{-1} M_\odot$ reside in nodes. However, direct observations of the intergalactic gas are challenging, as this gas has a low surface brightness and too low density to form stars.

Statistically, the properties of the IGM have been derived from the observations of the so-called Ly α forest, that is a series of hydrogen Ly α absorption lines in the spectra of high-redshift quasars (see, e.g., Bahcall & Salpeter, 1965; Gunn & Peterson, 1965; Scheuer, 1965; Hu et al., 1995; Kim et al., 1997; Kirkman & Tytler, 1997) arising in the fluctuating IGM and tracing the dark matter distribution on large scales. Today, the Ly α forest has been targeted by high-resolution observations in hundreds, and even a hundred thousand, quasar sightlines (see, e.g., Lee et al., 2013; O’Meara et al., 2015) to study the IGM at $z \gtrsim 2$. As derived by Schaye (2001a), the observed column density of the neutral absorbing H I gas, N_{HI} , is a proxy for the total volume density of the gas in the Universe, n_{H} :

$$N_{\text{HI}} \simeq 2.3 \times 10^{13} \text{ cm}^{-2} \left(\frac{T}{10^4 \text{ K}} \right)^{-0.26} \left(\frac{\Gamma}{10^{-12} \text{ s}^{-1}} \right)^{-1} \left(\frac{n_{\text{H}}}{10^{-5} \text{ cm}^{-2}} \right)^{3/2} \left(\frac{f_g}{0.16} \right)^{1/2} \quad (1.20)$$

where T is the temperature of the gas, Γ is the photoionization rate for $\text{H}^0 + h\nu \rightarrow \text{H}^+ + e^-$ (see, e.g., Scott et al., 2000 at $2 \lesssim z \lesssim 4$ and Davé & Tripp, 2001 at $z \lesssim 0.5$), and f_g is the gas mass fraction. With $\bar{n}_{\text{H}}(z) \propto (1+z)^3$ the average density of the Universe, defined as in eq. 9 in Schaye (2001a), Equation 1.20 can be re-written as a function of the density contrast $\delta = (n_{\text{H}} - \bar{n}_{\text{H}})/\bar{n}_{\text{H}}$ (see eq. 10 in Schaye, 2001a):

$$N_{\text{HI}} \propto (1+\delta)^{3/2} T^{-0.26} \Gamma^{-1} (1+z)^{9/2} f_g^{1/2} \quad (1.21)$$

The low-column density absorbers in the Ly α forest trace gas of low overdensity. On the other hand, overdensities of $\delta \gtrsim 100$ give rise to absorption lines with high column density (e.g., of the order of $N_{\text{HI}} \gtrsim 10^{16} \text{ cm}^{-2}$ at $z \approx 3$) and trace virialized minihalos. The presence of dark matter has a minor effect on the scaling relation in Eq. 1.21 via the gas mass fraction, that is, if $f_g = 1$, the value of δ corresponding to a given N_{HI} would be smaller by only a factor of two. From the large dataset collected by modern 8-10m class telescopes (such as the Keck and the Very Large Telescope), it is possible to measure the column density distribution function of the neutral hydrogen, $f(X, N_{\text{HI}})$, defined as the number of H I absorbers per unit column density per unit path length, which is a powerful diagnostic of the distribution of gas density (see, e.g., Fumagalli et al., 2013; Rudie et al., 2013; Worseck et al., 2014). As reviewed by McQuinn (2016), $f(X, N_{\text{HI}})$ does not significantly evolve across $2 < z < 5$ (see also Prochaska et al., 2010; McQuinn et al., 2011). The shape of the column density distribution function reflects that of the mass distribution of the gas as a function of its density (Schaye, 2001a). In particular, at column densities $N_{\text{HI}} \gtrsim 10^{14} \text{ cm}^{-2}$ the distribution function assumes the functional form of a power-law (Kim et al., 1997), tracing the typical density distribution of collapsed structures. A flattening at $N_{\text{HI}} \gtrsim 10^{16} \text{ cm}^{-2}$ is also observed and interpreted as due to the virialization of the collapsed matter at overdensities $\delta \gtrsim 100$ (Petitjean et al., 1993). A break is visible at $N_{\text{HI}} \approx 1.6 \times 10^{17} \text{ cm}^{-2}$, due to the systems becoming self-shielded at higher densities (see, e.g., Zheng & Miralda-Escudé, 2002). Finally, a steep drop-off is observed at the largest densities, where molecules form and the diffuse gas is depleted onto stars (Schaye, 2001b). Intergalactic neutral hydrogen in the forest has typically column densities $N_{\text{HI}} \lesssim 10^{17.2} \text{ cm}^{-2}$. At higher densities (e.g. Lyman Limit Systems, LLS, at $10^{17.2} \lesssim N_{\text{HI}}/\text{cm}^{-2} \lesssim 10^{20.3}$ or Damped Ly α Absorbers, DLAs, at $N_{\text{HI}} \gtrsim 10^{20.3} \text{ cm}^{-2}$) are more likely found closer to galaxies (see, e.g., Møller et al., 2002, 2004; Fynbo et al., 2010; Srianand et al., 2016; Fumagalli et al., 2017; Christenson & Jorgenson, 2019; Mackenzie et al., 2019; Bielby et al., 2020; Lofthouse et al., 2023) and have thus been studied to derive the properties of the gas in their immediate surroundings (such as metal enrichment, dust and gas content, degree of ionization and kinematics, see, e.g., Péroux et al., 2003; Wolfe et al., 2005; Prochaska & Wolfe, 2009; Rafelski et al., 2012; Fumagalli et al., 2013, 2016).

The Ly α absorption features also encode information on the temperature of the low-density IGM: the higher is the temperature of the gas, the broader the line due to thermal broadening and pressure effects. Measuring the gas temperature from the width of the absorption lines can constrain different mechanisms that inject energy into the IGM. The line width is typically parametrized by the Doppler parameter, $b = (b_{\text{th}}^2 + b_{\text{kin}}^2)^{1/2}$, defined as a combination of a thermal contribution, b_{th} , and a kinematic component (e.g., turbulence), b_{kin} . This has been observed to

range within $10 < b/(\text{km s}^{-1}) < 100$ (see the review by Meiksin, 2009), with the majority of the Ly α absorption lines showing $b \approx 15 - 60 \text{ km s}^{-1}$ (see, e.g., Hu et al., 1995; Kim et al., 1997; Kirkman & Tytler, 1997). The thermal component of the Doppler parameter can directly relate the width of the line to the temperature of the gas via the relation $b = (2k_{\text{B}}T/m_{\text{H}} + b_{\text{kin}}^2)^{1/2}$, where k_{B} is the Boltzmann constant, and m_{H} is the Hydrogen atomic mass. Several studies (see, e.g., Becker et al., 2011; Garzilli et al., 2012; Boera et al., 2014; Bolton et al., 2014) have found that the IGM is highly ionized and the temperature of the gas is of the order of $T = (1 - 2) \times 10^4 \text{ K}$ at $2 < z < 4$.

While these studies succeeded in measuring the properties of the IGM, it still unclear how galaxies get their gas and what is the role played by the cosmic web in shaping the properties and the evolution of the embedded galaxy populations.

1.3.2 The Circumgalactic Medium

On smaller scales, the evolution of galaxies and their star formation activity is regulated by the local flows of gas between the ISM and the CGM. Accretion of gas from the IGM provides new material and replenishes the galaxies' gas reservoir in their ISM. In turn, stellar winds as well as powerful supernovae explosions and Active Galactic Nuclei (AGNs) are able to expel enriched material back to the CGM though feedback on kpc scales (see the review by Veilleux et al., 2005). In the most extreme cases, feedback processes can remove the gas reservoirs from the ISM and potentially prevent further gas accretion, quenching the star formation activity inside the galaxy. These events are thus believed to be responsible for a large fraction of the metals that we observe today outside the galaxy disks (see, e.g., Aguirre et al., 2001; Oppenheimer & Davé, 2006; McNamara & Nulsen, 2012; Shen et al., 2013; Hafen et al., 2019). Finally, the ejected metal-enriched gas can eventually fall back onto the galaxy and being recycled (see, e.g., Shapiro & Field, 1976; Bregman, 1980; Fraternali, 2017). This complex set of processes, which goes under the name of *baryon cycle*, occur during the entire lifetime of a galaxy and leaves a footprint in the properties of the CGM, making this gas multiphase in temperature, density, ionization structure and complex in its dynamics and kinematics (Tumlinson et al., 2017). Since the CGM is enriched by metals produced by nucleosynthesis in stars and ejected outside the galaxies by strong winds and feedback mechanisms, metal ions spanning with a difference of an order of magnitude in their ionization potential are powerful observational targets that can be used to probe the physical state of the gas. Indeed, neutral species (such as, H I, C I or Mg I) trace dense and cold gas with density $n_{\text{H}} \lesssim 10^{-2} \text{ cm}^{-3}$ and at temperature $T \lesssim 10^{4.5} \text{ K}$ (see e.g., Neeleman et al., 2015), while high ions (such as Ne VIII, O VII or Mg X) can be found in a much underdense and warmer gas phase, typically with $n_{\text{H}} \lesssim 10^{-4} \text{ cm}^{-3}$ and at $T \gtrsim 10^{5.5} \text{ K}$ (see e.g., Oppenheimer et al., 2012). Direct evidence for the presence of an even hotter gas phase with $T \gtrsim 10^6 \text{ K}$ comes from the detection of both galactic and extragalactic diffuse X-Ray emission (see e.g., Anderson & Bregman, 2010; Gupta et al., 2012; Anderson et al., 2013).

At low redshift, $z \lesssim 0.5$, a combination of space-based observations (for example, with the Cosmic Origin Spectrograph, COS, on board the Hubble Space Telescope, HST) and optical spectroscopy from the ground is instrumental to probe both neutral hydrogen and metals hosted in the galaxies' CGM. The COS-Halo survey (Werk et al., 2012, 2013; Tumlinson et al., 2013) is one of the first programs designed to study the CGM of several active and passive galaxies with $\log(M_{\star}/M_{\odot}) = 9 - 11.5$. By detecting H I and metals in different ionization states in absorption, Werk et al. (2013, 2014) constrained the metal budget of 44 star-forming galaxies up to distances of 150 kpc, which corresponds to $\approx 35\% - 110\%$ of the virial radius in this stellar mass range (Moster et al., 2010). They found that a cool ($T \approx 10^4 \text{ K}$) and warm-hot ($T \approx 10^5 - 10^6 \text{ K}$), photoionized and metal enriched gas phase in the CGM accounts for $\approx 40\%$ of the baryons in halos of $M_{\text{h}} \approx 10^{12} M_{\odot}$, representative of the typical galaxies in their sample, at $z \approx 0.2$ (see Fig. 1.3, left panel). Finally, the presence of a hot gas phase has been revealed by X-Ray observations (see, e.g., Mathur et al., 2023; Nicastro et al., 2023) and found to harbor for a significant fraction of baryons in the halos. At intermediate redshifts, $0.5 \lesssim z \lesssim 2.0$, Mg II absorbers are the most studied tracers of the cool CGM gas. Several observations have shown that the strength of the absorption decreases with increasing distance from the galaxies (see e.g., Lanzetta & Bowen, 1990; Bergeron & Boissé, 1991; Weiner et al., 2009; Chen et al., 2010a; Bordoloi et al., 2011; Churchill et al., 2013b; Nielsen et al., 2013; Lan et al., 2014; Rubin et al., 2014; Dutta et al., 2020). In addition, this cool gas appears to be concentrated close to the disks, while the gas is ionized up to higher states at

increasing distance from the host galaxy. As a result, warmer phases, such as the one traced by CIV absorbers, are found to be more extended than the Mg II gas in both observations (see e.g., Liang & Chen, 2014; Werk et al., 2014; Dutta et al., 2021) and simulations (Ford et al., 2013; Ho et al., 2020, 2021; Weng et al., 2024).

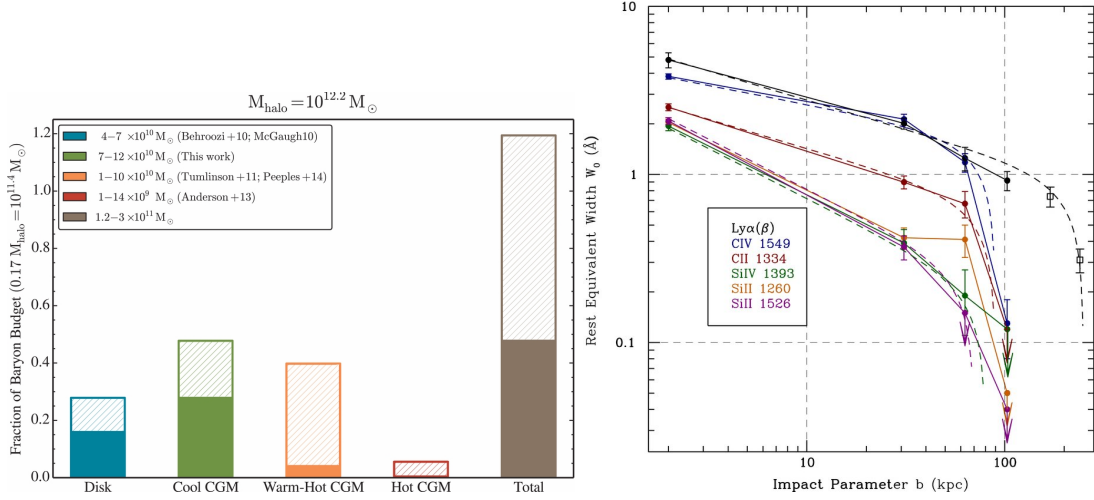


Figure 1.3: Composition and distribution of the CGM gas of galaxies. *Left panel:* fraction of baryons contributing to different gas phases in the CGM of halo mass $M_h = 10^{12.2} M_\odot$ at $z \approx 0.2$, typical of galaxies targeted by the COS-Halo survey. Lower limits are shown by filled thick bars, while the hatched region account for the possible additional contribution allowed by the data. Taken from Werk et al. (2014). *Right panel:* radial profile of the strength, probed by the equivalent width, of neutral hydrogen and metal absorbers around LBGs as a function of the projected distance between the galaxy-galaxy pairs at $z \approx 2.3$. Taken from Steidel et al. (2010).

At cosmic noon, $2 \lesssim z \lesssim 3$, the CGM gas has been mostly studied around Lyman break galaxies (LBGs), a population of massive (stellar mass $M_\star \approx 10^{10} - 10^{11} M_\odot$) and star forming (star formation rate $SFR \approx 10 - 100 M_\odot/\text{Yr}^{-1}$) sources (see e.g., Erb et al., 2006), with the Keck Baryonic Structure Survey (KBSS; Steidel et al., 2010) and the VLT LBG Redshift Surveys (VLRs; Bielby et al., 2011). Strong Hydrogen absorbers have been found to trace the CGM of LBGs on scales comparable with the typical virial radius of these galaxies (Bielby et al., 2017), while their column density is observed to decrease more than three order of magnitudes with increasing distance (Rudie et al., 2012), likely transitioning from tracing the CGM to the IGM. A similar declining profile is also shown by metal ions (see Fig. 1.3, right panel), suggesting that the gas in the halo of these galaxies is enriched and in multiphase in its ionization state, as found by Steidel et al. (2010). However, they studied the radial profile of the metals' distribution only in the transverse direction by using galaxy-galaxy pairs, which does not preserve the information of the relative kinematics of gas and galaxies. Turner et al. (2014) adopted an alternative approach to overcome this limitation and produced two-dimensional maps of the pixel optical depth both in the transverse and in the line-of-sight direction. As a result, they found that both neutral hydrogen and metals have similar extension in the transverse direction (≈ 180 kpc), while higher ions (such as, CIV, SiIV, OVI) are more extended compared to H I along the line-of-sight. In this direction, however, the excess of optical depth is more elongated than in the transverse direction due to the effect of peculiar motions (that is, inflows, outflows and virial motion) that add to the pure Hubble flow. Finally, by comparing these maps to predictions from the EAGLE simulation, they found that a fraction of the enriched gas found around LBGs is accreting, or re-accreting, onto the galaxies and not only outflowing away at larger distances.

A new frontier has recently opened with the advent of sensitive IR detectors, that is studying the CGM at $z \gtrsim 6$. At these times, the universe is closer to and within the epoch of reionization, corresponding to when the galaxies begun to form massive stars whose radiation was able to ionize the neutral hydrogen in the universe (see e.g., Gnedin & Madau, 2022, for a review). D'Odorico et al. (2018) detected a galaxy at $z \approx 6$ associated to a DLA by using the Atacama Large Millimeter/submillimeter Array (ALMA) and studied, for the first time, the chemistry of the CGM gas at the reionization epoch. Combined with other studies which employed MOS and IFS data (see e.g., Meyer et al., 2020; Díaz et al., 2021; Bielby et al., 2020), the emerging picture is that young and low-mass galaxies produced bubbles of ionized gas that grew up to filling the

universe and outflows that shaped the properties of the surrounding gas. An unprecedented view of the gas at these early epochs is expected to be accessible with the James Webb Space Telescope (JWST; see e.g., the early results from [Wu et al., 2023](#); [Bordoloi et al., 2024](#)).

These studies reveal the ubiquitous presence of the baryon cycle and unveil both the properties and the distribution of the halo gas across cosmic time. However, at cosmic noon, a little is known about the multiphase CGM of galaxies at the lower end of the mass function.

1.4 Linking gas and galaxies at cosmic noon

Studying the properties of the cosmic gas retained in the CGM of galaxies or spread in the cosmic web provides key information about the gas that potentially drives the formation and evolution of galaxies themselves and that regulate their growth and star formation activity. However, to ultimately trace the baryon cycle and its ability to affect the properties of the galaxies, an important step forwards consists on directly linking the gas and the galaxies on different scales (ideally from their ISM to the IGM) and across cosmic time. In particular, since the processes involved in the baryon cycle are strictly connected to the evolution of the galaxies, it is interesting to explore such connection at epochs close to the cosmic noon and the peak of the star formation rate density in the universe ([Madau & Dickinson, 2014](#)). Through this section, we present the main results obtained in the literature and introduce the questions that remain open.

1.4.1 Evidences of a physical association

Several techniques can be employed to explore how the cosmic gas and the galaxies trace and affect each other on different time and spatial scales. Here, we focus our review on those that we used in the thesis and allowed us to study i) the clustering of galaxies around the tracers of the cosmic gas, such as absorption line systems; ii) the radial distribution of different gas phases around these sources. The first case comes as a consequence of the tracers of the cosmic gas to arise from the CGM of galaxies or from their immediate surroundings, for instance from the filaments connecting the galaxies themselves. Therefore, gas and galaxies are not expected to be distributed randomly relative to each other, but to be located in similar regions of the universe.

Clustering along the line of sight

A first approach is *gas-centric*, that is, it consists on centering at the redshift and location of the gas and search for galaxies in its surroundings. This is particularly powerful when intervening absorption line systems identified in the sightlines of background bright objects, such as quasars or galaxies, are chosen as tracers of the cosmic gas. As described in Sec. 1.2.1, modelling these lines provides a measure of their redshifts, making it possible to connect these systems with galaxies along the line-of-sight. This is typically done by measuring their velocity separation: $\Delta v = c(\Delta z^2 - 1)/(\Delta z + 1)$ with $\Delta z = (1 + z_{\text{gal}})/(1 + z_{\text{abs}})$ and z_{gal} , z_{abs} the redshift of the galaxies and that of the absorbing gas, respectively. The clustering of galaxies around the absorbers can be further investigated on a more statistical ground as a function of the projected separation of the galaxies from the gas, and as a function of their properties, such as luminosity or stellar mass.

Correlation functions

A powerful tool to study the clustering of galaxies around the gas detected in absorption as a function of the projected separation is by computing the cross-correlation function. In three dimensions, this is defined as the conditional probability of finding a galaxy within a volume dV at a distance r from the absorber, which is given by the equation $PdV = P_0[1 + \xi(r)]dV$ where P_0dV is the probability of finding a galaxy in an average place in the universe. A similar definition holds for the galaxy-galaxy auto-correlation function, which is instead the conditional probability of finding a galaxy within dV and r from another galaxy. Following the formalism from [Trainor & Steidel \(2012\)](#), both the correlation functions can be modelled as a power law:

$$\xi(r) = (r/r_0)^{-\gamma} \quad (1.22)$$

where γ is the slope and r_0 is the co-moving distance at which the number density of galaxies around either the absorbers or other galaxies is twice that expected for a random region of the universe. However, since the line-of-sight velocities are affected by the proper motions of the gas and are thus not a reliable proxy of the radial distance, the three-dimensional correlation function $\xi(r)$ cannot be measured directly. It is thus convenient to integrate along the redshift direction and compute reduced angular correlation functions. With this method, the conditional probability is given by: $P(R)d\Omega = P'_0 d\Omega [1 + \omega_p(R)] = d\Omega \int_{\Delta z} P(r) dz$, where $R = D_A(z)\theta(1+z)$ is the projected co-moving distance and $D_A(z)$ is the angular diameter distance. If this is computed over a truncated redshift range, with half-width of the redshift window z_0 , it takes the form of a Gaussian hypergeometric function, ${}_2F_1$, instead of a simple power-law (as in the limits $\Delta z \rightarrow \infty$):

$$\omega_p(R) = \left(\frac{r_0}{R}\right)^\gamma {}_2F_1\left(\frac{1}{2}, \frac{\gamma}{2}; \frac{3}{2}; -\frac{z_0^2}{R^2}\right) \quad (1.23)$$

which is still a function of both γ and r_0 as the original three-dimensional form. This model reproduces what in the data can be estimated in circular annuli of increasing projected separation as $\omega_p(R) = \Sigma_{\text{obs}}(R)/\Sigma_{\text{pred}}(R) - 1$ (see e.g., [Davis & Peebles, 1983](#)), where Σ_{obs} and Σ_{pred} are the surface density of galaxies observed around the absorbers, or other galaxies, and the expected surface density in a random region of the universe, respectively. The correlation functions provide key information about the physical association between gas and galaxies. To this end, the three-dimensional galaxy-galaxy and absorber-absorber auto-correlation function, ξ_{GG} and ξ_{AA} , respectively, as well as the galaxy-absorber cross-correlation function, ξ_{GA} , can all be compared to reveal if the gas is preferentially located in the CGM of the galaxies or in IGM filaments. Following [Adelberger et al. \(2003, 2005\)](#); [Bielby et al. \(2011\)](#), galaxies and absorbers trace the same underlying matter density distribution if it holds the Cauchy-Schwarz relation: $\xi_{\text{GA}}^2 \leq \xi_{\text{GG}}\xi_{\text{AA}}$.

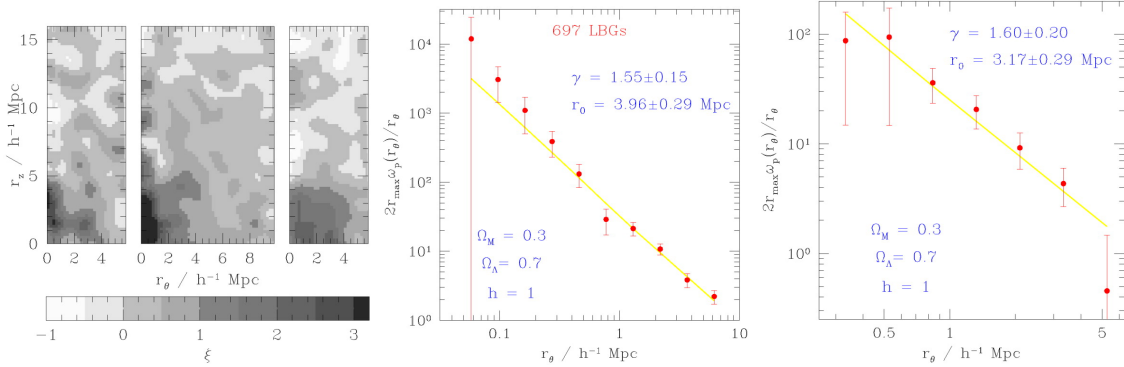


Figure 1.4: LBGs-LBGs auto-correlation function and LBGs-C IV cross-correlation function. *First panel:* two-dimensional LBGs autocorrelation function, LBGs-C IV cross-correlation function and LBGs-Ly α transmittivity cross-correlation function (multiplied by a factor -10). *Second panel:* projected LBGs-LBGs auto-correlation function. *Third panel:* projected LBGs-C IV cross-correlation function. The best-values estimated for the parameters γ and r_0 are annotated in the figure. Taken from [Adelberger et al. \(2003\)](#).

With these methods, [Chen & Mulchaey \(2009\)](#); [Tejos et al. \(2014\)](#); [Finn et al. \(2016\)](#); [Prochaska & Wolfe \(2009\)](#) have shown that galaxies are clustered to H I, O VI, and Mg II absorbers at $z \lesssim 1$. At cosmic noon, [Adelberger et al. \(2003, 2005\)](#); [Bouché & Lowenthal \(2004\)](#); [Cooke et al. \(2006\)](#) found that strong H I absorbers are significantly correlated with the location of LBGs at $z \approx 3$. In addition, [Adelberger et al. \(2003, 2005\)](#); [Crighton et al. \(2011\)](#) reported that the cross-correlation function of the C IV absorbers is similar to the auto-correlation function of the LBGs, suggesting that they inhabit in similar regions of the universe (see Fig. 1.4). An approach that is complementary to deriving the three-dimensional correlation function has been adopted by [Turner et al. \(2014\)](#). They produced two-dimensional maps of the pixel optical depth as a function of the projected and line-of-sight separation, which is related to the transmitted flux (see Eq. 1.7) and thus the presence of an excess of absorption along the sightline. As a result, they found that out to ≈ 180 kpc along the transverse direction and ≈ 1 Mpc along the line-of-sight away from the LBGs, the median optical depth is enhanced by one order of magnitude compared to random regions of the universe for H I absorbers as well as for O VI, N V, C IV, Si IV ions. Overall, the picture these works put forward is that both neutral hydrogen and the enriched and ionized gas co-exist and preferentially arise from regions in the universe close to where galaxies are found.

Luminosity functions

Complementary to studying the radial extent of the galaxy clustering around the absorbers, it is possible to explore such overdensity as a function of the properties of the galaxies themselves, such as their luminosity or stellar mass. One way to do it is by computing the luminosity functions of galaxies associated to the absorbing gas and compare it to that of galaxies located in random regions of the universe. The differential luminosity function, $\phi(L, z)$, is a measurement of the number of galaxies that emit a given luminosity at a given redshift, per unit volume. This is typically parametrized by a parametric [Schechter \(1976\)](#) function as follows:

$$\phi(L)dL = \phi^* \left(\frac{L}{L^*} \right)^\alpha \exp \left(-\frac{L}{L^*} \right) \frac{dL}{L^*} \quad (1.24)$$

which depends on three free parameters: the normalization L^* , with the units of a number density, the power-law slope α and the characteristic luminosity ϕ^* at which the transition occurs from a power-law form to an exponential cut-off. Following [Herenz et al. \(2019\)](#) and [Fossati et al. \(2021\)](#), the differential luminosity function can be derived from observational datasets by means of a non-parametric estimator, $1/V_{\max}$, from [Schmidt \(1968\)](#) and [Felten \(1976\)](#), according to which the luminosity function can be approximated in bin of luminosity by the following equation:

$$\phi(\langle \Delta \log L \rangle) = \frac{1}{\Delta \log L} \sum_i \frac{1}{V_{\max, i}} \quad (1.25)$$

where $\langle \Delta \log L \rangle$ is the average luminosity in each bin with width $\Delta \log L$. The sum runs over all the i sources in the current bin of luminosity, while $V_{\max}$ is the volume of the survey weighted by the selection function, which accounts for the completeness of the sample related to the sensitivity and depth of the observations (see e.g., [Johnston, 2011](#)), and is defined as:

$$V_{\max, i} = \omega \int_{z_{\min}}^{z_{\max}} \delta_{\text{obs}}(z) f_c(L_i, z) \frac{dV}{dz} dz \quad (1.26)$$

Here, $f_c(L, z)$ is the selection function, which depends both on the luminosity of the source i and on its redshift, dV/dz is the differential co-moving volume element (see e.g., [Hogg, 1999](#)), ω is the angular area of the survey. $\delta_{\text{obs}}(z)$ is introduced to recover the effective co-moving area of the survey which has been searched for gas-galaxy connections. In other words, if galaxies are considered associated to the gas within a given line-of-sight separation, Δz_{th} , from the redshift of the absorbers, z_{abs} , this parameter will only take into account the contribution of the interval $z_{\text{abs}} \pm \Delta z_{\text{th}}$. The analytical expression of the differential luminosity function, given by Eq. 1.24, is typically fitted to the data, Eq. 1.25, by using a maximum likelihood estimator (see e.g., [Mehta et al., 2015](#)) to estimate the best values for the free parameters ϕ^* , α , and L^* . In particular, encoded here are two key information: i) the normalization ϕ^* is a tracer of the amplitude of the overdensity of galaxies around the absorbers with respect to the expectations for random regions in the universe; ii) the slope α and the overall shape of the luminosity functions is indicative of an excess or deficiency of bright or faint galaxies (depending on if its flatter or steeper at the faint or the bright end). It is thus a tracer of the properties of the galaxy population and can be used to qualitatively assess any difference with respect to galaxies in the field.

By using this method, [Lofthouse et al. \(2023\)](#) studied the clustering and the properties of a large sample of low-mass galaxies detected via their Ly α emission line at $z \approx 3 - 4.5$ in the Muse Analysis of Gas around Galaxies (MAGG; [Lofthouse et al., 2020a](#)) survey and found to be associated to strong H I absorbers. They observed that the normalization of the Ly α luminosity function of these galaxies is, compared to that of similar galaxies in the field, ≈ 5 times higher. The shape, however, does not present significant differences, suggesting that the population of galaxies close to H I absorbers is not fainter nor brighter than in the field. Since the Ly α luminosity is broadly a tracer of the recent star formation activity that is ongoing in the galaxies, this suggest that they are forming star at a rate that is similar to galaxies in random regions of the universe, despite being found close to large reservoirs of neutral hydrogen.

Covering fraction

Both the correlation and the luminosity functions are derived for galaxies that are found to be associated to absorption line systems within small line-of-sight and projected separations. To complete the view of the gas-galaxy associations, it is useful to adopt the opposite *gas-centric* and

search for absorbers around the galaxies, which would provide key insights on the spatial extent of the gaseous halos as well as on the distribution of the gas. This can be achieved by computing the covering fraction as a function of the projected distance from the galaxies. The covering fraction is defined as the number of galaxies associated to an absorber with rest-frame equivalent width $W_r \geq W_0$ (see Eq. 1.9) relative to the total number of galaxies in the absorbers' redshift path, and thus potentially candidate to be connected to the gas:

$$f_c(R) = \frac{N_{\text{obs}}(R; W_r \geq W_0)}{N_{\text{tot}}(R)} \quad (1.27)$$

Such counting of galaxies in the vicinity of the absorbing gas is performed as a function of the projected distance, R , and can be differential, derived in circular annuli at increasing distance from the sources, or cumulative, counting all the associations within a given separation. Both these approaches succeed in tracing up to which distances from galaxies the gas observed in absorption extends and is distributed.

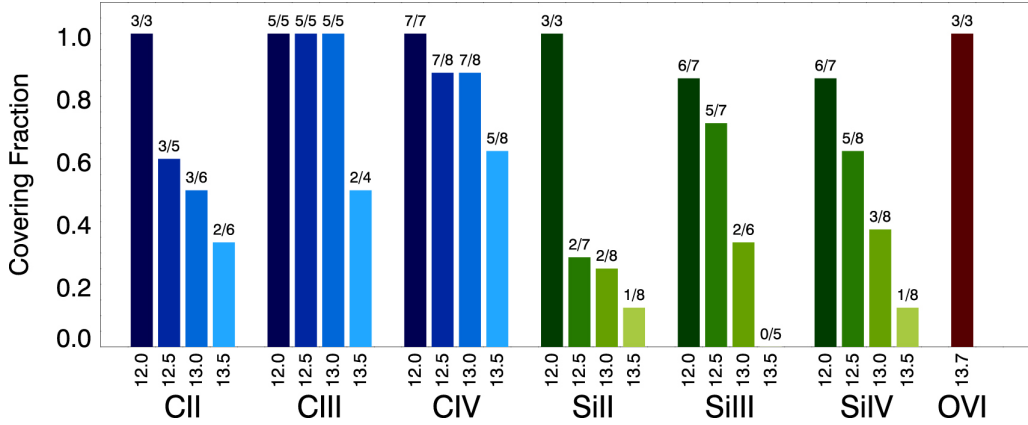


Figure 1.5: Covering fraction of metals in different ionization state found within 100 kpc and $\pm 100 \text{ km s}^{-1}$ from LBGs in the KBSS. Darker colors corresponds to higher column density thresholds for the absorbers. The covering fraction of the triply ionized species is $\gtrsim 87 - 100\%$, while that of CIV is $\gtrsim 50\%$ even for strong systems with $\log(N_{\text{CIV}}/\text{cm}^{-2}) \geq 13.5$. Taken from Rudie et al. (2019).

Several works in the literature have focused on the radial distribution of different gas phases around $z \lesssim 1$ galaxies, using H I, Mg II, and C IV absorbers as tracers of the cosmic gas. Bordoloi et al. (2014); Turner et al. (2014); Burchett et al. (2016); Rudie et al. (2019); Dutta et al. (2021) observed that the covering fraction decreases with increasing threshold on the strength of the absorption as well as with increasing distance, the gaseous halos extend up to a few time the galaxies' virial radii. In particular, at $z \lesssim 2$, Dutta et al. (2021) studied both the C IV- and the Mg II-bearing gas to trace different gas phases at the same time. They found that the covering fraction of C IV is ≈ 2 times larger compared to that of Mg II at a given projected separation, which may be a hint that the warm gas phase is more extended around galaxies than the cooler gas traced by Mg II absorbers. With the help of large surveys, like MAGG, it is also possible to explore if the distribution of different gas phases evolves with time. To this end, Schroetter et al. (2021) explored the connection between Mg II and C IV and $z \approx 1 - 1.5$ [O II] line emitters in the MusE GAs FLOW and Wind (MEGAFLOW; Schroetter et al., 2016) survey. Comparing the covering fraction of these two ions, they found that the distribution of the gas phases they are tracing evolves with time in different ways: the size of the Mg II halo decreases with time becoming smaller compared to that of the dark matter halo, while the size of the C IV halo increasing with time. They speculate that a scenario in which the C IV halo co-evolve with the dark-matter one would be possible if the warm gas is in hydrostatic equilibrium with the halo. Further hint of redshift evolution of the Mg II covering fraction was found by Lan (2020) for systems with equivalent width $W_r \geq 1 \text{ \AA}$ and appears to follow a trend given by $(1+z)^\alpha$, with $\alpha = 2$, which is consistent with the evolution of cosmic star formation rate density at $z \lesssim 2$ (see e.g., Prochter et al., 2006). At higher redshift, $z \approx 2$, Rudie et al. (2019) computed the covering fraction for multiply ionized ionic species around LBGs and estimated it to be of the order of $87 - 100\%$ for triply ionized metals within 100 kpc. For the C IV absorbers is found to be $\gtrsim 50\%$ for systems with column density $N_{\text{CIV}} \geq 10^{13.5} \text{ cm}^{-2}$, which does not deviate significantly from the results obtained at lower redshifts in the literature quoted above.

1.4.2 The cosmic gas observed in emission

Mapping the cosmic gas in emission provides a powerful complementary view of what we can learn by tracing it in absorption. Although it is challenging to observe such a low density and low surface brightness gas (see e.g., [Hennawi & Prochaska, 2013](#)), successful detections have been reported in the literature and multiplied with the advent of IFUs (see Sec 1.2.2). One possible approach to observe the gas in emission consists on directly imaging of specific emission lines with narrow band filters, the brightest of which is the Ly α (since $\approx 2/3$ of all the recombinations leads to the emission of Ly α photons, see Sec. 1.2.2). This technique led to the detection of Ly α emission arising from overdense regions (see e.g., [Steidel et al., 2000](#); [Hayashino et al., 2004](#); [Nilsson et al., 2006](#); [Yang et al., 2009](#)) and, for the first time, of the CGM around a $z \approx 2.3$ quasar by [Cantalupo et al. \(2014\)](#). It has also been employed to image the enriched and ionized gas detected via the fainter C IV and He II emission (see e.g., [Arrighi Battaia et al., 2015](#)). Despite the high transmission efficiency of modern filters, two limitations of the imaging technique are the following: i) that it performs poorly in background-limited observations; ii) it provides useful information on the morphology and extent of the gaseous structures, but cannot be used to trace the kinematics of the gas. Spectroscopy is a alternative approach, as it limits the contribution of the background in the high-resolution element (see e.g., [Hennawi et al., 2009](#); [North et al., 2012](#)). However, slit or fiber spectroscopy can only map regions not larger than a few arcseconds.

The development of IFUs, such as the Multi Unit Explorer (MUSE; [Bacon et al., 2010](#)) at the Very Large Telescope (VLT) and the Keck Cosmic Web Imager (KCWI; [Morrissey et al., 2018](#)) at Keck, revolutionized the observations of the cosmic gas in emission by combining the high signal-to-noise ratio that can be obtained with spectroscopy to the two-dimensional imaging of the gas. The basic concept of the Integral Field Spectroscopy is slicing a large field into small slitlets or fiber bundles and disperse each of them with a series of spectrographs. These spectra are then re-combined into three-dimensional datacubes that contain the spatial information on two axis, and provide the spectrum of each pixel along an additional axis for the wavelength dimension.

At $z \gtrsim 3$, MUSE reaches surface brightness levels of the order of $\approx 10^{-19} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-2}$ in both stacked and individual observations, and is the ideal instrument to target the Ly α emission arising from the CGM of galaxies. Extended Ly α halos has thus been observed in emission and found ubiquitously around LBGs and lower mass Ly α emitting galaxies (see e.g., [Patrício et al., 2016](#); [Kusakabe et al., 2022](#)). Such extended emission is typically modeled by an exponentially declining profile with a scale length of $\approx 4.5 \text{ kpc}$ ([Wisotzki et al., 2016](#); [Leclercq et al., 2017](#)) that is larger by one order of magnitude compared to that of the rest-frame UV stellar continuum and thus suggest that these halos extends beyond half the virial radius of the galaxies. [Kusakabe et al. \(2022\)](#) and [Wisotzki et al. \(2016\)](#) computed the fraction of the sky that is covered by Ly α emission. They found that, projected in the redshift range $3 \lesssim z \lesssim 6$, Ly α emission with a surface brightness $\lesssim 10^{-20} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-2}$ covers $\gtrsim 80\%$ of the sky. It is now possible to compare it to the covering fraction of a similar gas phase observed in the complementary form of strong hydrogen absorbers, such as DLAs and LLSs. As a result, Ly α halos are found to be responsible for the observed DLAs and LLSs within projected separations from the galaxies of $\approx 30 \text{ kpc}$ and $\gtrsim 60 \text{ kpc}$, respectively.

The observations of Ly α halos allow to direct imaging the CGM, but they only provide a view that is limited to the immediate surroundings of the galaxies. Searching for extended Ly α emission around quasars, which are able to ionize the gas (see Sec. 1.2.2), is expected to be a powerful strategy for both studying the CGM in halos that are typically more massive than those hosting LBGs and detecting the cosmic gas in emission on scales of several kpc. In the absence of local sources of radiation, such as galaxies or quasars, the UV Background (UVB; [Haardt & Madau, 2012](#)) is the only ionizing radiation that is able to illuminate the the cosmic gas. With this assumption, [Umehata et al. \(2019\)](#) shows that the surface brightness of the optically thin H I as scales linearly with its column density. On the contrary, in case of optically thick gas at $\log(N_{\text{H I}}/\text{cm}^{-2}) \geq 18$, $\approx 60\%$ of the ionizing photons are converted in Ly α radiation by collisions without any significant dependence on the column of the gas. However, due to the cosmological dimming $\propto (1+z)^4$, typical surface brightness levels expected for the Ly α emission arising from the cosmic web are of the order of $\approx 10^{-20} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-2}$. Additional sources of radiation, like quasars, are thus necessary to illuminate this gas at observable levels. Indeed, [Cantalupo et al. \(2014\)](#) and [Hennawi et al. \(2015\)](#) imaged the CGM of $z \gtrsim 2$ radio-quiet quasar using narrow-bands and detected Ly α emission extended up to $\approx 200 - 500 \text{ kpc}$, well beyond the typical virial radius of quasar host halos ($R_{\text{vir}} \approx 150 - 200 \text{ kpc}$ at $z \approx 2.5$ for halo masses of $M_{\text{h}} \approx 10^{12-12.5} M_{\odot}$). Furthermore, the nebula observed by [Cantalupo et al. \(2014\)](#) shows a filamentary morphology that

is extended towards a nearby quasar at the same redshift of the system. Large reservoirs of gas around $z \gtrsim 3$ quasars have been found to be ubiquitous, with 100% of detection rate in IFS surveys (see e.g., [Borisova et al., 2016](#); [Arrighi Battaia et al., 2019a](#); [Farina et al., 2019](#); [Fossati et al., 2021](#)). In the MUSE Ultra Deep Field (MUDF; [Fossati et al., 2019](#)) survey, with 142 hours on source, [Tornotti et al. \(2024\)](#) detected not only extended Ly α nebulae around a pair of quasars at $z \approx 3.22$, but also a Ly α -emitting filament of the cosmic web connecting the two quasars (see also, [Arrighi Battaia et al., 2019c](#)) and extending over $30 \times 90 \text{ arcsec}^2$. Such filamentary structure has a mean surface brightness of $\approx 8 \times 10^{-20} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-2}$ and extends $\approx 140 \text{ kpc}$ in the transverse direction at a surface brightness limit of $4.5 \times 10^{-20} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-2}$, which is the first measurement of the thickness of a filament. Similarly, for the first time, they also found that the transition from the CGM to the IGM roughly occurs at the virial radius of typical quasar host halos. At similar depths, [Bacon et al. \(2010\)](#) reported the detection of diffuse and extended Ly α emission in 5 out of 22 overdense regions and with average surface brightness of $\approx 5 \times 10^{-20} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-2}$. They observed that regions located at large distances from the galaxies, and likely tracing the IGM rather than the CGM, give a contribution of 70% of the total Ly α luminosity of the filaments.

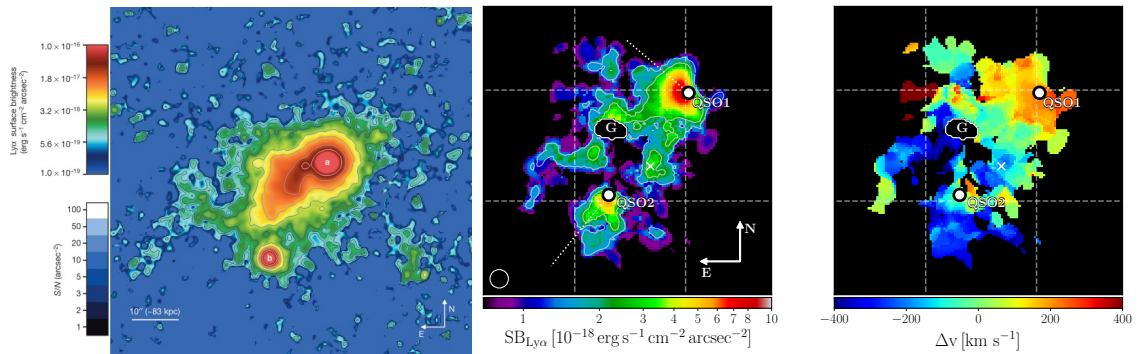


Figure 1.6: Extended Ly α emission arising from the cosmic gas around $z \gtrsim 2$ quasar. *Left panel:* Ly α nebula detected by [Cantalupo et al. \(2014\)](#) around the quasar UM 287 (labelled “a”) and extended across $\approx 460 \text{ kpc}$. The nebula shows a filamentary morphology extending towards a second quasar (labelled “b”) that is found at the same redshift of the systems. *Right panel:* Ly α emission around and connecting a $z \approx 3$ pair of quasars separated by 89 kpc (first panel). A clear velocity gradient of $\approx 400 \text{ km s}^{-1}$ emerges between the region of the nebula close to the quasar “QSO1” and that southern than “QSO2” (second panel). Taken from [Arrighi Battaia et al. \(2019c\)](#).

If on the one hand the Ly α ($n = 2 \rightarrow 1$) is the brightest recombination line, on the other hand it is also resonant and can easily be absorbed by a large fraction of electrons in the ground state before escaping the cloud. Due to resonant scattering, these photons are thus diffuse both in frequency and in space. Therefore, non-resonant lines are especially useful to overcome this limitations. This is case, for instance, of the H α ($n = 3 \rightarrow 2$) or He II ($\lambda_{\text{RF}} = 1640 \text{ \AA}$) lines. Comparing the Ly α with these non-resonant emission lines it is possible to: i) measure the physical properties of the emitting gas, such as kinematics (see e.g., [Steidel et al., 2010](#); [Gronke, 2017](#)), density, and clumpiness ([Cantalupo et al., 2019](#)); ii) get information about the mechanism that is powering the emission (see e.g. [Mas-Ribas et al., 2017](#); [Langen et al., 2023](#)).

1.4.3 Connecting the properties of gas and galaxies

Once established that there exist a physical association between the cosmic gas and the galaxies, it is interesting to investigate if and how they do affect each other’s properties. At redshift $z \lesssim 2$, [Chen et al. \(2010a\)](#); [Bordoloi et al. \(2011\)](#); [Nielsen et al. \(2013\)](#); [Lan et al. \(2014\)](#); [Rubin et al. \(2018\)](#); [Dutta et al. \(2020\)](#) have shown that the size of Mg II halos and the strength of the absorption depend on the galaxies’ stellar mass, star formation rate, and specific star formation rate. Similar findings have been reported for the more ionized gas traced by C IV absorbers. Indeed, [Borthakur et al. \(2013\)](#); [Bordoloi et al. \(2014\)](#); [Dutta et al. \(2021\)](#) found that the equivalent width of these absorbers is positively correlated with the galaxies’ stellar mass and star formation rate, although the strength of the correlation is weaker compared to the Mg II gas. At higher redshift, $z \gtrsim 2$, similar relationships have mostly been explored only for LBGs. [Adelberger et al. \(2005\)](#) found a hint that galaxies residing close to H I absorbers may have higher star formation rates. However, they did not found any significant difference in the colors, ages, dust reddening, stellar mass,

kinematics or environments. On the other hand, galaxies close to C IV absorbers appear to have bluer colors and younger ages. Instead, Prochaska et al. (2014) argued that the properties of a cool phase of the CGM ($T \approx 10^4$ K), traced by H I and C II absorbers, is shaped by the halo mass and mostly tracks the cumulative growth and enrichment history of the halo. The effect of processes such as AGN or supernova feedback, and accretion from the IGM is weaker. While these studies focus on the relatively massive and star-forming population of LBGs, Muzahid et al. (2021) extended these studies to a sample of Ly α -emitting galaxies, which have typically lower masses. They stacked the H I and C IV absorption at the redshift of the galaxies and found a hint of correlation between the strength of the absorption and the galaxies' star formation rate for the H I-gas, but not for the C IV. The existence of a link between the properties of the galaxies and the distribution of the H I gas at $z \approx 2$ is also predicted by cosmological simulations. For instance, Momose et al. (2021) found that the H I gas in the IGM is spatially correlated to the location of galaxies, with the most massive ones being clustered in higher-density regions, and its distribution strongly depend on the halo mass (see also, Turner et al., 2017; Martizzi et al., 2020).

Despite these promising results, it is still challenging to paint a complete picture on the link between gas and galaxies and track it across cosmic time. In particular, approaching the cosmic noon, studies are wither limited to LBGs or relies on stacks, which encodes information only averaged across the entire population. Moreover, the CGM is expected to be patchy, a property that cannot be fully captured by individual quasar sightlines. Therefore, it is essential to extend these studies across larger dynamic ranges in both star formation rate and stellar mass, and thus include sources that populate the low-end of the galaxies' mass function. In addition, statistical and larger samples are needed to study the gas-galaxy connection focusing on individual systems, rather than stacks.

1.5 The role of the large-scale environment

If on the one hand the formation and evolution of galaxies is regulated and connected to the properties and flows of gas, it is clear from the first spectroscopic surveys (see e.g. Kirshner et al., 1981; de Lapparent et al., 1986) that their distribution is inhomogeneous and a large fraction of them do not evolve in isolation, but rather in regions of high galaxy density. Such regions form in the densest part of the Cosmic Web and then grow by merging into the most extreme structures that we observe in the universe today, such as massive groups and clusters (see e.g., Geller & Huchra, 1989; Alpaslan et al., 2014; Libeskind et al., 2018). Indeed, in the local $z \approx 0$ universe, $\approx 40 - 50\%$ of all galaxies are found in groups (Eke et al., 2004). At low redshift, $z \lesssim 1$, significant effects of the galaxy density have been observed in their properties: an excess of massive, and passive galaxies is typically found in high density environments, while those inhabiting less dense regions are more gas rich, and star forming (see e.g., Dressler, 1980; Baldry et al., 2006; Raichoor et al., 2011; Lemaux et al., 2019; Chartab et al., 2020). Based on both observations and simulations, several processes are expected to affect the evolution and the properties of galaxies in rich environments. For instance, the observed quenching of the star formation activity inside galaxies in overdensities can take place even before they enter the clusters (see e.g., Cortese et al., 2006; Cluver et al., 2020; Estrada et al., 2023). This phenomenon, called *pre-processing* (Fujita, 2004), happen in lower density environments, such as in the outskirts, filaments or groups (see e.g., Haines et al., 2015; Bianconi et al., 2018; Lopes et al., 2024). Some processes, such as strangulation of cold gas (see e.g., Larson et al., 1980; Peng et al., 2015) or AGN feedback (see e.g., Beckmann et al., 2017; Zhang et al., 2021), are internal within a galaxy and goes under the name of *mass quenching*, as they are more common in massive galaxies and are able to efficiently suppress the star formation activity (see e.g., Peng et al., 2010). All the external processes that are related to the environment that surrounds galaxies are included in the *environmental quenching* and depends more on the density of the environment, rather than on the galaxies' stellar mass. This is the case of gravitational interactions, such as tidal stripping (see e.g., Gallagher & Ostriker, 1972; Richstone, 1976; Merritt, 1984; Coziol & Plauchu-Frayn, 2007; Lin et al., 2023), high-velocity encounters (see e.g., Moore et al., 1996; Bialas et al., 2015; Hwang et al., 2018), and gas removal due to the presence of a nearby galaxy (see e.g., Park & Hwang, 2009; Moon et al., 2019; de Vos et al., 2024). Efficient quenching and displacement of fuel for the star formation from the disks can happen as a consequence of hydrodynamical interactions. These occurs between different gas phases inside and outside a galaxy that is moving at high speed in the hot ($T \approx 10^7 - 10^8$ K; see e.g., De Grandi & Molendi, 2002; Vikhlinin et al., 2006; Ghirardini et al., 2019) intra-group or intra-cluster medium (ram-pressure stripping; see e.g., Book & Benson, 2010; Steinhauser et al., 2016; Boselli et al.,

2021, 2022).

At higher redshifts the picture is more unclear. Overdense structures are the progenitors of the most extreme environments, such as clusters in the local universe, and are thus expected to harbor key information about the onset of the differences in the galaxy populations between these regions and the field. However, such structures are not yet virialized. This makes the environmental effects described above less efficient in quenching the star formation activity and removing gas from the galaxies. We give in the next sections an overview of what is known, at present, on the effect of the large-scale environment on the properties and evolution of galaxies in overdensities approaching the cosmic noon.

1.5.1 Impact on the gas

Stripping and displacement of gas from galaxy disks into their surroundings is expected to affect the composition and the properties of the CGM and IGM gas. Several studies have investigated this scenario and searched for signature of environmental effects in the gas around groups and overdensities, especially at $z \lesssim 1$ (see e.g. Bordoloi et al., 2011; Nielsen et al., 2018a; Chen et al., 2019; Leclercq et al., 2022). For this purpose, dense spectroscopic galaxy surveys are required and ideal targets for IFU observations or slitless spectroscopy. In particular, at $z \lesssim 2$, Dutta et al. (2020, 2021) measured a factor of $\approx 2 - 3$ higher probability of detecting Mg II absorption around galaxies in groups compared to the isolated ones. Instead, the more ionized gas, traced by C IV absorbers, does not show any significant excess depending on the galaxy environment. These results have been tested for being independent on the galaxies' star formation rate and stellar mass. Moreover, the strength of both the Mg II and C IV absorption increases with increasing galaxies' number density. At $z \approx 3 - 4.5$, Lofthouse et al. (2023) found that the covering fraction of strong H I absorbers around galaxies in groups is ≈ 3 times higher compared to that of isolated ones (left panel in Fig. 1.7). However, the large-scale galaxy environment does not appear to significantly affect the column density nor the metallicity of the gas. Exploring a similar dataset, but stacking the H I and C IV absorption in the quasar spectra at the redshift of the galaxies, Muzahid et al. (2021) found that both the H I and the C IV absorption is stronger around groups (see also, Banerjee et al., 2023). While it is evident that the galaxy environment enhances the cross-section of the gas, the specific processes driving this effect remain unconstrained by observations. Nielsen et al. (2018a) and Dutta et al. (2020) showed that modeling the absorption as a superposition of halos of individual isolated-like galaxies (see e.g., Bordoloi et al., 2011), does not necessary succeed in reproducing the observations. This suggests that the role of environmental processes, such as gravitational or hydrodynamical interactions, is not negligible. However, at $z \gtrsim 3$ environmental interactions are expected not be efficient in these primordial overdensities. Therefore, Muzahid et al. (2021) interpreted the differences in the properties of the absorbers around groups as due to the large-scale structure in which the massive halos of groups are hosted.

1.5.2 Impact on the galaxies

Galaxies in overdensities may experience processes that are able to affect not only the surrounding gas, but also their properties. Several works in the literature studied how galaxies' star formation rate and stellar mass depend on the local overdensity. Grützbauch et al. (2011) assembled a sample of stellar mass-selected galaxies in the range $1.5 \leq z \leq 3.0$ (down to a completeness limit of $\log(M_\star/M_\odot) = 9.5$ at $z \approx 3$) and found that the influence of the galaxy overdensity on the star formation rate is weaker at $z > 1.5$ than it is in the local universe (see e.g., Lewis et al., 2002; Balogh et al., 2004; Patel et al., 2009). The star formation rate does not decreases monotonically with increasing density (see also, Scoville et al., 2013; Darvish et al., 2016), but drops by a factor of ≈ 3 only in the densest regions. Instead, there exist a strong correlation between the star formation rate and the stellar mass, which is insensitive to the galaxy overdensity. In particular, Darvish et al. (2016) argued that at $z \gtrsim 1$ the star formation in galaxies is independent on the overdense environment and mostly regulated by their stellar mass. Given the large fraction of quiescent galaxies observed at $z \lesssim 1$ and the strong dependence on the overdensity, however, it is possible that the quenching of star forming galaxies occurs in a short timescale ($\lesssim 1$ Gyr). On the other hand, Lemaux et al. (2022) found that the galaxies' star formation rate is positively correlated with the local overdensity of galaxies at $2 \leq z \leq 5$. This would imply that at $z \gtrsim 2$ the negative correlation between star formation rate and overdensity observed in the local universe reverses and denser regions begin to have higher star formation rates (see also, Elbaz et al., 2007; Cooper et al., 2008; Taamoli et al., 2024; Shi et al., 2024). They speculate that massive and highly

star forming galaxies in the densest environments are the main drivers of the reversal of the SFR–density relation. Indeed, if in the $z \approx 0$ clusters these massive galaxies are mostly quiescent, at earlier times they appear to be still actively forming stars at high rates. This evolution may be a consequence of the numerous processes that occur in overdense environments. It also interesting to investigate the link between the galaxy overdensity and the stellar mass, as it is, with the large-scale environment, the major driver and regulator of galaxy evolution. To this end, [Shi et al. \(2024\)](#) found that the galaxies’ stellar mass, at all redshifts, increases with the environment overdensity (see also [Hatch et al., 2011](#); [Koyama et al., 2013a](#); [Cooke et al., 2014](#); [Shimakawa et al., 2018a](#); [Ito et al., 2020](#); [Lemaux et al., 2022](#); [Taamoli et al., 2024](#); [Toshikawa et al., 2024](#)) leading to the conclusion that mass itself likely drives the SFR–density relation over cosmic time. Similar results have also been found by [Forrest et al. \(2024\)](#). They adopted a more statistical approach and computed the galaxies’ stellar mass function, that is the number of galaxies per unit stellar mass, as a function of the overdensity. The densest regions have ≈ 6 times more galaxies with stellar masses of $\log(M_*/M_\odot) \approx 10$ compared to the field, and only ≈ 3 time higher number density of galaxies with $\log(M_*/M_\odot) \approx 9.5$ (right panel in Fig. 1.7). Therefore, the massive end of the stellar mass function appears to be flatter in overdensities compared to the field. This may be a hint that “preprocessing” is occurring in these dense regions and is able to accelerate the growth of galaxies via mergers or efficient accretions of gas.

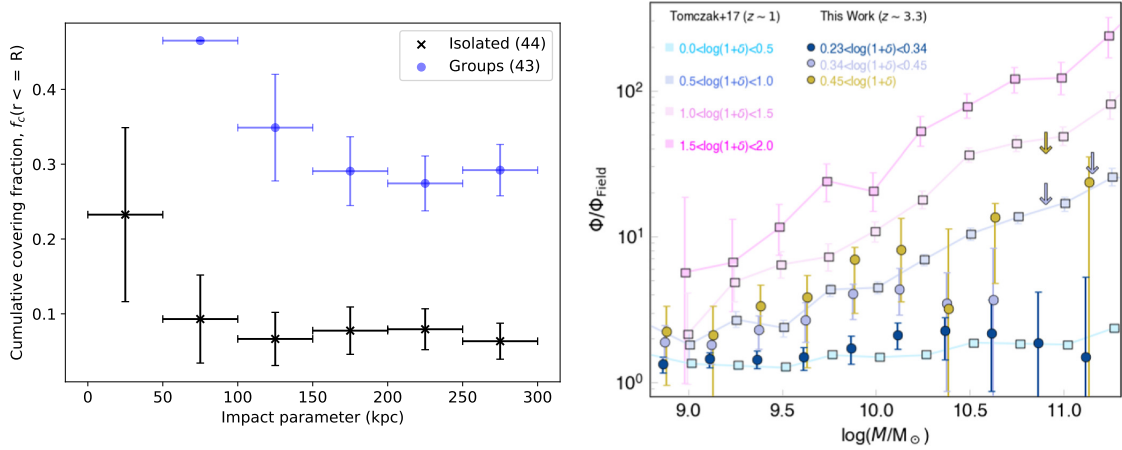


Figure 1.7: Effect of galaxy environments on the cosmic gas and on the galaxies’ properties *Left panel:* the cumulative covering fraction of strong H I absorbers is found to be ≈ 3 times larger around Ly α -emitting galaxies in groups (blue) compared to isolated ones (black). Taken from [Lofthouse et al. \(2023\)](#). *Right panel:* ratio between the stellar mass functions of galaxies in different interval of overdensities with respect to that of galaxies in the field. A clear excess of massive galaxies is found in the densest regions at $z \approx 3$. Such a trend is similar to the one observed at $z \approx 1$, although less pronounced. The larger number density of massive galaxies compared to low-mass ones suggests that sources in overdensities grow, or have grown, more rapidly than in the field. Taken from [Forrest et al. \(2024\)](#).

1.5.3 Protoclusters

Among all the galaxy overdensities at high redshifts, of particular interest are those that are eventually expected to evolve into virialized clusters in the present day universe. These are called *protoclusters* (see e.g., [Steidel et al., 1998](#); [Kurk et al., 2000](#); [Tanaka et al., 2011](#); [Toshikawa et al., 2020](#)), as they are candidate progenitors of $\log(M_h/M_\odot) \geq 14$ structures, typically identified as clusters, by $z \approx 0$ (see e.g., [Overzier, 2016](#); [Alberts & Noble, 2022](#), for reviews). At redshifts $z \gtrsim 2$, these structures consists of a large number of small halos that are clustered together, but filamentary and spread on scales of several Mpc (see e.g., [Steidel et al., 1998](#); [Pentericci et al., 2000](#); [Chiang et al., 2013](#); [Cucciati et al., 2018](#)) and not yet virialized. As time passes, a strong concentration of mass begins to develop at the center of the overdense regions. At $z \approx 1$, the halos start to collapse and merge forming the final clusters by $z \approx 0$. A similar scenario was proposed by [Shimakawa et al. \(2018b\)](#) and based on the properties of two protoclusters found at $z \approx 2$ (Fig. 1.8). According to this, at high redshifts galaxies are rapidly building up their masses and are clustered in fragmented groups. In a maturing phase, groups begin to merge and grow hierarchically. A high AGN fraction is expected to develop at this stage, with AGN feedback being one of the most efficient mechanisms able to quench the star formation activity in galaxies. At

$z \lesssim 2$ the structure becomes more compact and in ideal conditions for environmental quenching to settle. As important as the evolution of galaxies is that of the halo gas, which provides mass and fuel to sustain their star formation activity (Overzier, 2016). At $z \gtrsim 3$ protoclusters are connected to the cosmic web and able to efficiently accrete gas from the filaments, resulting in both a rapid growth and vigorous star formation activity. At $z \lesssim 2$ protoclusters detach from the cosmic web, the hot intra-cluster medium develops and heats up the gas suppressing further accretion. In on the one hand these scenarios paint a clear picture about the faith of protoclusters, identifying them in both observations and simulations and predict if they will effectively evolve into a cluster by $z \approx 0$ is still challenging. Indeed, Muldrew et al. (2015) combined the Millennium simulation with semi-analytical models to show that at fixed cluster mass at $z = 0$, there exist a $\gtrsim 0.5$ dex scatter in the mass of the most massive halo in a protocluster at $z = 2$. Regardless these uncertainties, as these structures are identified as candidate clusters progenitors, they are ideal laboratories where to study at which stage of their evolution the properties and the evolution of the member galaxies begin to differ from those of similar galaxies in the field.

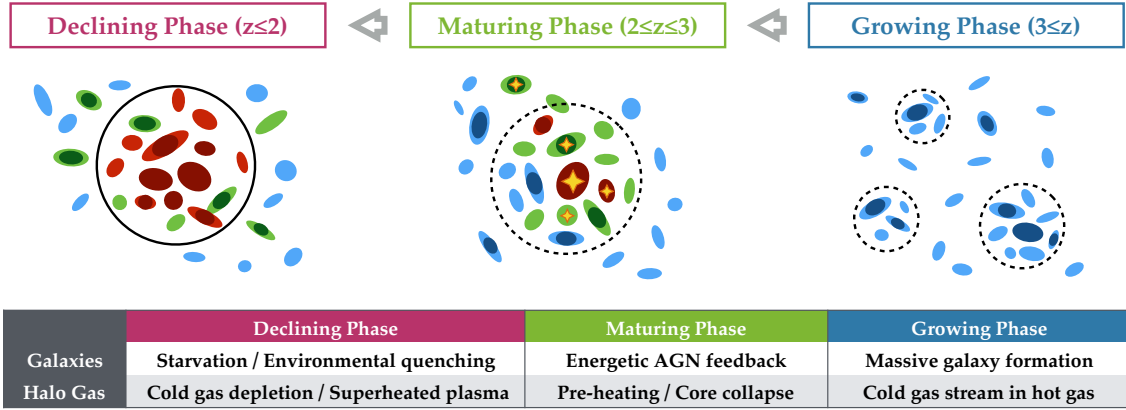


Figure 1.8: Simple schemes of the evolution of protoclusters proposed by Shimakawa et al. (2018b).

The search for protoclusters is based on looking for a concentration of galaxies clustered in redshift space and within a small angular separation. This can be done by targeting different populations, such as star forming, passive or dusty submillimeter galaxies, either blindly or around specific so-called *biased tracers*, which are expected to stand as signposts of galaxy overdensities. This approach typically consists on performing photometric selections and a spectroscopic follow-up. For instance, Chiang et al. (2014) used photometric redshifts to identify overdensities in three-dimensional space across 1.6 deg^2 in the COSMOS/UltraVISTA field. They found 36 systems that are candidate Coma-like protoclusters (that is, predicted to evolve into $\log(M_h/M_\odot) \gtrsim 15$ structures by $z \approx 0$) at $z \approx 2 - 3$, several of which have later been confirmed by follow-up spectroscopic observations. On the other hand, numerous protoclusters have been identified in the VIMOS Ultra Deep Survey (VUDS; Le Fèvre et al., 2015), one of the largest available spectroscopic surveys (see e.g., Cucciati et al., 2014; Lemaux et al., 2014, 2018), and in its follow up Charting Cluster Construction with VUDS and ORELSE (C3VO; Hung et al., 2024) survey. This is also the case of the so-called Hyperion proto-supercluster discovered by Cucciati et al. (2018) at $z \approx 2.45$, an extreme overdensity of galaxies spread in a volume of $\approx 60 \times 60 \times 150 \text{ cMpc}^3$ with a total mass of $\approx 4.8 \times 10^{15} M_\odot$. With similar methods, Steidel et al. (1998) discovered an exceptional overdensity of star forming Lyman break galaxies at $z \approx 3.08$, the SSA22 protocluster (see also, Steidel et al., 2003; Yamada et al., 2012; Topping et al., 2016; Zeng et al., 2024). Complementary to searching for overdensities of star-forming galaxies is the approach of targeting submillimeter dusty, massive and star forming sources (see e.g., Greenslade et al., 2018; Lammers et al., 2022; Polletta et al., 2022). For instance, the Spitzer Planck Herschel Infrared Cluster (SPHerIC) survey combines the data from *Planck* (Planck Collaboration et al., 2016), *Herschel* (Pilbratt et al., 2010) and *Spitzer* (Werner et al., 2004) to identify clusters and protoclusters at $1 \lesssim z \lesssim 3$ (see e.g., Martinache et al., 2018). While these results allow to identify a large number of protoclusters at high redshifts and get information about the properties of different galaxy populations therein, looking around biased tracers can further optimize the effort. A consistent number of protoclusters has been identified around radio galaxies (see e.g., Roettgering et al., 1994; Hayashi et al., 2012; Hatch et al., 2014; Rigby et al., 2014; Orsi et al., 2016), mostly by using narrowbands to select emission-line galaxies, such as $H\alpha$ or $\text{Ly}\alpha$ emitters, followed up by spectroscopy to confirm the

redshifts. Among these, one of the most extreme and known overdensity is the protocluster found in the close proximity to the Spiderweb galaxy at $z \approx 2.3$ (Pentericci et al., 2000; Miley et al., 2006; Dannerbauer et al., 2014; Tozzi et al., 2022b,a; Pérez-Martínez et al., 2023; Zhang et al., 2024). Other than in the surroundings of radio galaxies, large overdensities have been found close to Hot Dust-Obscured Galaxies (Hot DOGs; Jones et al., 2014; Assef et al., 2015; Penney et al., 2019; Luo et al., 2022; Ginolfi et al., 2022; Zewdie et al., 2023) and Sub-Millimeter Galaxies (SMGs; Alberts & Noble, 2022; Calvi et al., 2023; Cornish et al., 2024). Also luminous high redshift quasars (see e.g., Kashikawa et al., 2007; Overzier et al., 2008; Bañados et al., 2013; Adams et al., 2015; Mignoli et al., 2020; Kashino et al., 2023) have been suggested as tracers of galaxy overdensities and protoclusters. Finally, the cosmic gas itself can be used as a biased tracer. For instance, Lee et al. (2018) and Newman et al. (2020, 2022) mapped the H I gas in absorption in three dimensions using a technique, called “Ly α tomography”, to search for overdense structures in the IGM (left panel in Fig. 1.9). Alternatively, extended Ly α blobs detected in emission are found to be preferentially located in galaxy overdensities and expected to trace large reservoirs of cool gas, that is dense regions of the cosmic web, as well as the presence of bright sources that are able to ionize the gas, such as Active Galactic Nuclei or starburst galaxies (see e.g., Prescott et al., 2008, 2012; Ramakrishnan et al., 2023).

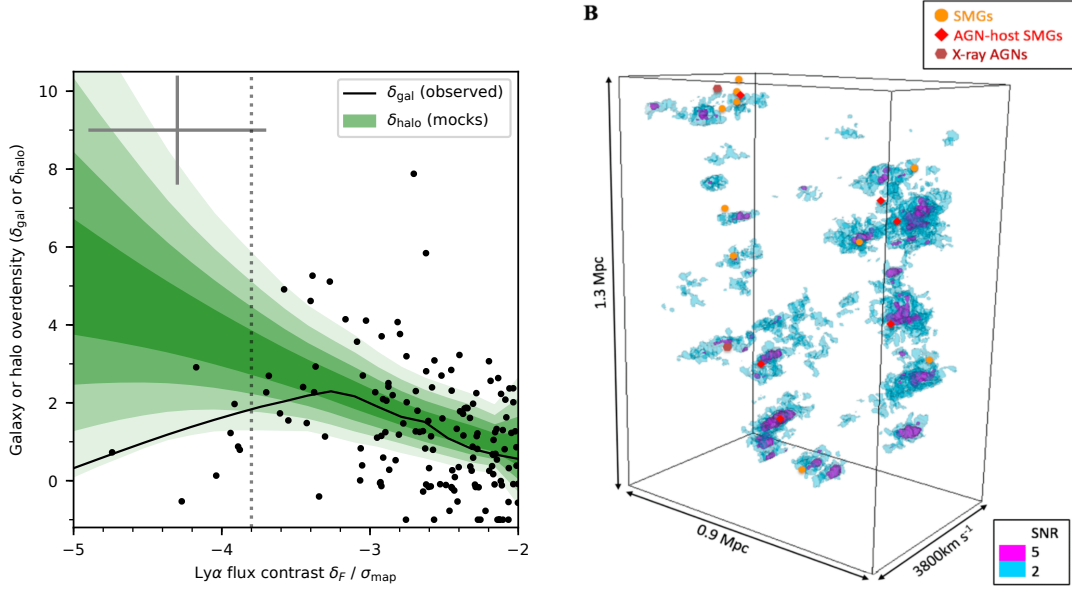


Figure 1.9: Protoclusters are located in the densest regions of the cosmic web, where multiple filaments intersect and provide them of a gas-rich reservoir. Tracing the overdense cosmic gas in these structures is thus a powerful method to both identify protoclusters and unveil what is its role in the evolution of the embedded galaxies. *Left panel:* correlation between the overdensity of galaxies and the location of Ly α flux absorption peaks (more negative values on the x-axis indicate stronger absorption), which identify overdensities of H I gas in the IGM, as obtained from Ly α tomography by Newman et al. (2022). Black points are the individual observed peaks, while the green lines are the expectations from simulated mock observations. White circles highlight the location of known galaxies. *Right panel:* Extended Ly α emission in SSA22 protocluster at $z \approx 3.09$ detected by Umehata et al. (2019).

Protoclusters are thus expected to be located in dense regions of the cosmic web, where multiple filaments intersect. A number of detections of the diffuse cosmic gas via Ly α emission exists in the literature (see e.g., Cai et al., 2016, 2017; Apostolovski et al., 2024). In particular, Umehata et al. (2019) observed extended and filamentary Ly α emission (right panel Fig. 1.9), with a surface brightness of $\approx 3 \times 10^{-19} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-2}$, connecting galaxies in the SSA22 protocluster (see also, Martin et al., 2014). In such a gas-rich environment, galaxy formation is accelerated and occurs at earlier times than in the field (Thomas et al., 2005; Alberts & Noble, 2022). Therefore, it is expected that galaxies in protoclusters develop different properties and follow a different evolutionary path compared to those in less dense regions. As already discussed in Sect. 1.5.2, the availability of a large gas reservoir, the efficient accretion of cold gas onto the galaxies and the higher merger rate contribute in favoring the galaxies to be, on average, more massive and with an older stellar population than those in the field (see e.g., Steidel et al., 2005; Hatch et al., 2011; Muldrew et al., 2018; Harshan et al., 2021). However, it is still unclear whether the availability of

large gas reservoirs also results in enhanced star formation activity (see e.g., [Alberts et al., 2014](#); [Shimakawa et al., 2018a](#); [Lemaux et al., 2022](#); [Pérez-Martínez et al., 2024](#)) or if it is comparable to field galaxies (see e.g., [Toshikawa et al., 2014](#); [Cucciati et al., 2014](#); [Koyama et al., 2021](#); [Shi et al., 2021](#); [Polletta et al., 2021](#); [Pérez-Martínez et al., 2023](#)). A unified picture of how the large-scale protocluster environment affects the evolution of galaxies therein is thus still missing. In addition to affecting the properties of the galaxies, overdense environments have an impact on the galaxy populations themselves. Indeed, overdensities of AGNs (see e.g., [Lehmer et al., 2009](#); [Macuga et al., 2019](#); [Tozzi et al., 2022b](#); [Gatica et al., 2024](#)) and SMGs (see e.g., [Wang et al., 2016, 2018](#); [Lee et al., 2017](#); [Castignani et al., 2019](#); [Tadaki et al., 2019](#); [Umehata et al., 2019](#); [Polletta et al., 2022](#); [Zeng et al., 2024](#)) are often found in protoclusters, suggesting that different galaxy populations co-exist in these structures. For instance, in the SSA22 protocluster, [Umehata et al. \(2019\)](#) found that AGNs and SMGs are aligned with the Ly α -emitting structure both in space and in velocity, possibly efficiently supplied of gas from the filaments. Similarly, [Steidel et al. \(2011\)](#) found an extremely large overdensity of LBGs at $z \approx 2.84$ around a hyperluminous quasar, together with overdensities of Ly α -emitting galaxies ([Kikuta et al., 2019](#)) and SMGs ([Lacaille et al., 2019](#); [Wang et al., 2024a](#)) spread across over 100 cMpc. Also in the Spiderweb protocluster at $z \approx 2.3$ the large overdensity of H α -emitting galaxies ([Shimakawa et al., 2018b](#)) has recently been found to be co-spatial with that of submillimeter sources detected in their dust continuum ([Zhang et al., 2024](#)), CO(1–0) emitters ([Jin et al., 2021](#)), and SMGs ([Dannerbauer et al., 2014](#)) (Fig. 1.10).

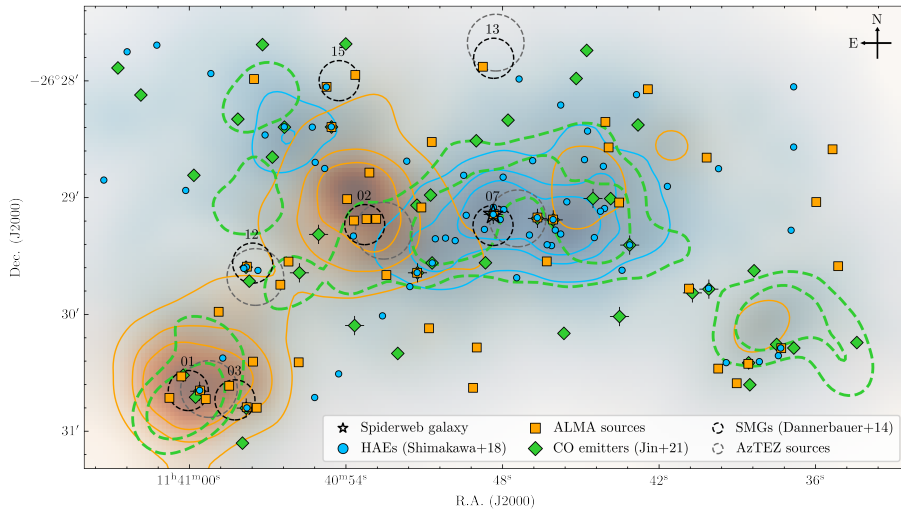


Figure 1.10: Spatial distribution of the H α -emitters (blue markers, contours, and density map), CO-emitters (green markers and contours), SMGs (dashed circles), and ALMA-selected sources (orange markers, contours, and density map) in the Spiderweb protocluster at $z \approx 2.3$ from [Zhang et al. \(2024\)](#).

1.6 Questions and goal of this thesis

What emerges from the discussion above is that a complex interplay between galaxies and the surrounding circum- and inter-galactic gas is able to shape their properties and regulate their evolution across cosmic time. However, it is challenging to trace all the processes that are involved as well as the instantaneous and cumulative effects on both the galaxies and the gas. On the other hand, a large fraction of galaxies is observed not to evolve in isolation, but as a component of more complex structures and overdensities. The main goal of this thesis is thus to unveil the link between gas and galaxies and explore the role of the large-scale environment in their evolution focusing on epochs close to cosmic noon, at the peak of the star formation rate density in the universe. We will address the following open questions:

1. what are the properties of multiphase gas around low-mass galaxies?
2. what characterizes the gas surrounding submillimeter dusty galaxies, often undetected in optical surveys?
3. how does the large-scale galaxy environment shape the properties of the gas?

4. how is the growth of galaxies regulated by the large-scale environment in extreme, overdense regions where multiple filaments intersect?

Chapters 2 and 3, expand the study on different gas phases of the CGM, typically focused on massive LBGs, to a population of more than 1000 Ly α emitting galaxies identified in the MUSE Analysis of Galaxies (MAGG) survey. The Ly α -emitters identified in MAGG have typical stellar masses of $\log(M_*/M_\odot) \approx 9$ and thus sample the low-mass of the galaxies' mass function at $z \approx 3$. We detect the cosmic gas in absorption in the sightlines of the quasars by assembling samples of more than 200 C IV, 100 Si IV, and 50 Mg II absorbers. We therefore investigate the connection between low mass galaxies and different gas phases up to ≈ 250 kpc. Within the dense and complete sample of LAEs assembled in the MAGG survey we also distinguish between galaxies that are in groups from the isolated ones. This allow us to explore how the galaxy environment impact the distribution and the properties of the different gas phases observed simultaneously around the galaxies. These results are published in Galbiati et al. (2023) and Galbiati et al. (2024b).

Chapter 4, presents the results of recent ALMA observations targeting the MUSE Ultra Deep Field (MUDF). With this data, we search for dusty, massive and star forming galaxies at $1 \lesssim z \lesssim 3$ detected in the 1.2-mm dust continuum or as CO(3-2) line emitters to investigate the properties of the gas around a population of galaxies that is typically not targeted by optical surveys. We thus studied the connection between these galaxies and both the large-scale environment the gas detected in absorption in the spectra of a $z \approx$ quasar pair. The results have been submitted in Pensabene et al. (2024b)².

Chapter 5, investigate the role of the large-scale galaxy and gaseous environment in the extreme case of a protocluster found around a $z \approx 3$ quasar. This overdensity is embedded in Ly α -emitting filaments converging to a massive node of the cosmic web. We followed-up this field with a mosaic of four MUSE pointings, each one deep 10 hours, VLT/FORS2, VLT/HAWK-I, ALMA (Pensabene et al., 2024a), *Chandra* (Travascio et al., 2024), HST, JWST/NIRCam and NIR/Spec (Wang et al., 2024b) observations. In a region of 4×4 cMpc², we measure the amplitude of the overdensity and investigate the properties of the population of star forming galaxies. By comparing them to similar galaxies found in the field, we explore if this special environment plays any role in galaxy evolution. Finally, we search for LBGs on larger scales of 24×24 cMpc² and found overdensities aligned with that in the central region and with the Ly α emitting gas. The results are presented in the recently submitted paper Galbiati et al. (2024a).

²In this paper I am listed as a second author, but with equal contribution.

MUSE Analysis of Gas around Galaxies (MAGG) - V: Linking ionized gas traced by C iv and Si iv absorbers to Ly α emitting galaxies at $z \approx 3.0 - 4.5$.

2.1 Introduction

Based on our current understanding of the onset of the first episodes of star formation in halos hosting population III stars, heavy elements are produced in a primordial Universe. Due to the massive nature of these first stars, the newly-produced metals are not always retained in compact objects, but are ejected into the interstellar medium (ISM) and, owing to energetic events, also in the circumgalactic medium (CGM) and the intergalactic medium (IGM), contributing to the early enrichment of the Universe (e.g. Scannapieco et al., 2002; Schneider et al., 2002; Maio et al., 2011; Wise et al., 2012). Likewise, stellar winds and supernova explosions associated with the formation of the subsequent generations of stars in galaxies are believed to be the source of a substantial fraction of the heavy elements we observe today near and outside galaxies (e.g. Aguirre et al., 2001; Oppenheimer & Davé, 2006; Shen et al., 2013), with mechanisms like winds from active galactic nuclei, gravitational interactions or ram-pressure stripping accounting for an additional fraction of the metals found outside galaxies (e.g. Fossati et al., 2016; Hafen et al., 2019). Due to the tight correspondence between the production of metals by stars and the subsequent ejection into the CGM and IGM, the study of the chemical enrichment of the more diffuse gas outside and around galaxies as a function of time is a powerful tool to complement our view of the star formation history in galaxies (e.g. Bouché et al., 2006; Rafelski et al., 2014; Madau & Dickinson, 2014; Fumagalli et al., 2016), and a fundamental step for developing complete chemical evolution models (e.g. Tremonti et al., 2004; Finlator & Davé, 2008; Welsh et al., 2019).

The most effective way to map the chemical enrichment of the low-density IGM and CGM is, at present, the study of hydrogen and metal absorption lines imprinted on the spectra of background sources, such as quasars. Compilations of high-resolution spectra, as well as moderate-to-low resolution spectroscopy in larger surveys, paint a view of a widespread metal enrichment of the Universe at least for gas around or above the mean density (e.g. Schaye et al., 2003; Simcoe et al., 2004), with only a few rare exceptions of unpolluted regions observed below $z \lesssim 4.5$ (e.g. Fumagalli et al., 2011; Robert et al., 2019). The analysis of the C iv doublet with its two strong transitions at 1548.195Å and 1550.770Å and a characteristic ratio of the optical depth for unsaturated lines equal to two (given by the ratio of the corresponding oscillator strengths) has been particularly effective for this type of study, enabling the identification of this ion also in moderate signal-to-noise and low resolution data. In particular, recent searches of C iv doublets in Sloan Digital Sky Survey quasar spectra by Cooksey et al. (2013) and in quasar spectroscopy from the Keck and Very Large Telescope (VLT) archives by Hasan et al. (2020) consistently show a steady increase of the number of C iv absorbers per unit redshift with time. As the observed incidence of C iv is proportional to the comoving number density of C iv bearing systems times their cross section, one can learn about the link between C iv and galaxies in a model-dependent fashion. For instance, assuming that C iv absorbers with equivalent width (EW) > 0.06 Å arise in the CGM of Lyman break galaxies (LBGs), Cooksey et al. (2013) estimated that the observed incidence can be modelled by filling the inner ≈ 50 kpc of LBGs halos with ionized carbon with time. More recently, Hasan et al. (2021) expanded this argument both in mass and redshift, combining C iv absorption statistics with

dark matter halos to constrain the link between gas and galaxies. Leveraging the deployment of echellette spectrographs in the near infrared, [Simcoe et al. \(2011\)](#) have expanded previous efforts to characterise the evolution of C IV in quasar surveys at even higher redshift (see also [Pettini et al., 2003](#); [Songaila, 2005](#)), reaching the epoch of reionization at $z \approx 6$ and confirming the presence of a downturn in the mass density of C IV beyond $z \approx 5$. These studies in large samples are complemented by the orthogonal approach followed by ? who have obtained an extremely-high S/N spectrum of a single quasar, measuring the distribution of some of the weakest C IV lines currently detected around column density $N_{\text{CIV}} = 10^{11.4} \text{ cm}^{-2}$ (see also [Ellison et al., 2000](#)). Their analysis reveals a factor ≈ 2 higher incidence of C IV lines associated with H I than what is expected in LBG halos, implying the presence of metals in IGM filaments or in the CGM of galaxies not selected as LBGs (e.g. at lower mass; see also [Pieri et al., 2006](#); [Hasan et al., 2021](#)).

While absorption line studies currently remain the most powerful approach to trace the entire distribution of the selected ions regardless of the astrophysical environment in which they reside (e.g. IGM, CGM at different halo masses), more direct links between the production sites of heavy elements and their current location requires an explicit correlation between the ions detected in absorption and the surrounding galaxy population traced in emission. Focusing on C IV as a tracer, [Dutta et al. \(2021\)](#) have completed the largest survey to date of galaxies in quasar fields between $0.5 \lesssim z \lesssim 2$, finding that C IV is relatively more extended than Mg II around galaxies (see also [Liang & Chen, 2014](#); [Schroetter et al., 2021](#)). Further, they reported a higher incidence of C IV around massive and star-forming galaxies (see also [Chen et al., 2001](#); [Bordoloi et al., 2014](#); [Burchett et al., 2016](#)), although this ion appears comparably less sensitive than Mg II with respect to the properties of the galaxies and of the environment. However, in the local Universe, [Burchett et al. \(2016\)](#) have uncovered an environmental dependence, with an higher detection rate in lower-density regions. Moreover, [Schroetter et al. \(2021\)](#) further explored the connection between C IV absorbers and galaxies, finding a non-negligible fraction of instances where ionized gas can be associated to the IGM or undetected low-mass galaxies. At redshift $2 \lesssim z \lesssim 3$, [Adelberger et al. \(2005\)](#) have provided the first clear evidence of an association between strong C IV absorbers and massive (with halo mass $M_{\text{H}} \approx 10^{12} M_{\odot}$) star-forming galaxies (with star formation rate, SFR, $\approx 10 - 50 M_{\odot} \text{ yr}^{-1}$) selected as LBGs (see also [Crighton et al., 2011](#)), with these types of halos accounting for $\approx 1/3$ of all C IV absorbers with $\text{EW} > 0.4 \text{ \AA}$. Their study has also suggested that galaxies in denser environments are more likely to show C IV in absorption compared to isolated galaxies. The ubiquitous presence of C IV near LBGs is also confirmed by later studies building on the same sample presented by [Adelberger et al. \(2005\)](#), with [Steidel et al. \(2010\)](#) using galaxy pairs to infer that LBGs account for almost 50% of strong ($\text{EW} > 0.15 \text{ \AA}$) C IV systems at $z \approx 2$ (see also [Turner et al., 2014](#)).

These studies rely on spectroscopic follow-up of relatively-bright continuum-selected galaxies and, especially at $z > 2$, focus on the massive end of the galaxy population. For a complete view of the correlation between galaxies and metals, this type of analysis needs to be extended at the low-mass end, and to encompass also passive or highly-obscured galaxies. The serendipitous discovery of a faint Ly α emitting galaxy (LAE) associated with a metal-rich absorption system in an LBG survey by [Crighton et al. \(2015\)](#) provides a clear example of the need to extend this analysis to lower-mass galaxies without relying on continuum selection (see also [Díaz et al., 2015](#)), an approach that is now possible (see e.g. [Fumagalli et al., 2016, 2017](#)) thanks to large integral field spectrographs and in particular the Multi Unit Spectroscopic Explorer (MUSE; [Bacon et al., 2010](#)) at VLT. To this end, the recent $z \approx 3 - 4$ MUSE observations by [Muzahid et al. \(2021\)](#) have identified 96 LAEs in 8 quasar fields. By correlating these galaxies with C IV absorption, these authors found an elevated C IV optical depth near LAEs similarly to the case of LBGs. They also report evidence of an excess of absorption near LAEs in groups compared to isolated ones. Examples of similar studies extending to higher redshift can also be found in the literature ([Bielby et al., 2020](#); [Díaz et al., 2021](#)).

In this paper, we exploit the larger MUSE Analysis of Gas around Galaxies (MAGG, [Lofthouse et al., 2020a](#)) survey to study in detail the correlation between 292 LAEs and 220 C IV absorbers. To verify whether the link between LEAs and ionized gas traced by C IV is specific to the selected transition or general for a ionized gas phase, we also expand our study to the associations between LAEs and Si IV, which is an additional doublet that arises from moderately-ionized gas and it is conveniently accessible in a comparable wavelength range to C IV. MAGG is built on a MUSE Large Programme (ID 197.A-0384; PI Fumagalli) that was explicitly designed to study the link between gas and galaxies at $z \approx 3 - 4$ by targeting 28 fields with quasars at $z \approx 3.2 - 4.5$ for which archival high-resolution ($R \gtrsim 30,000$) spectroscopy is available ([Lofthouse et al. 2023](#), hereafter

MAGG IV; see also Dutta et al., 2020; Fossati et al., 2021). The MAGG selection results in a sample of quasars with magnitudes $m_r \leq 19$ mag and at least one strong hydrogen absorption line system ($N_{\text{HI}} \geq 10^{17} \text{ cm}^{-2}$) at redshift $z \geq 3.05$.

The structure of this paper is as follows: in Sect. 2.2-2.3 we present an overview of the observations and data analysis, including a description of how the LAEs, C IV and Si IV samples have been assembled. Readers not interested in technical details on data can skip these two sections. In Sect. 2.4 we turn to the main question of studying in detail the correlation between LAEs and C IV absorbers as a function of galaxy properties and environment, also comparing with the results obtained by relying on Si IV rather than C IV as a tracer of ionized gas. We discuss our main findings and conclusions in Sect. 2.5-2.6. Throughout, we assume a flat Λ CDM cosmology with $H_0 = 67.7 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $\Omega_m = 0.307$ (Planck Collaboration et al., 2016), magnitudes are expressed in the AB system, distances are in physical units unless explicitly stated (for instance when computing the correlation functions), and errors are at 1σ confidence level.

2.2 Observations and Data Reduction

2.2.1 MUSE observations

Each quasar field included in the MAGG survey has been observed with MUSE between period 97 and period 103, for a total on-source observing time of ≈ 4 hours per field using the wide field mode with extended wavelength coverage (465 – 930 nm, $R \approx 2000 - 3500$). As part of this survey, we also include archival data from the GTO observations (PI Schaye; Muzahid et al., 2021) that matched the selection criteria of the original MAGG sample (i.e., the presence of $z > 3$ H I optically-thick absorption systems along the line of sight; see details in table 1 of Lofthouse et al. 2020a). All the observations have been executed on clear nights at airmass ≤ 1.6 so that the image quality is better than 0.8 arcsec full width at half maximum (FWHM). Image quality for individual fields can be found in table 1 of Lofthouse et al. 2020a. In wide field mode, the MUSE field of view (FoV) is $\approx 1 \times 1 \text{ arcmin}^2$. In our observations, the quasar typically lies at the center of the FoV, enabling deep spectroscopic surveys of regions of $\approx 500 \times 500 \text{ kpc}^2$ at redshift $z \approx 3$. As documented in the MUSE instrument manual, sensitivity variations across the FoV may arise from small differences in the performance of the MUSE spectrographs, and these are mitigated by including small dithers and instrument rotations of 90 degrees.

The detailed procedure of MUSE data reduction is described by Lofthouse et al. (2020a), and only the key steps are summarized here. The raw data are first processed with the ESO MUSE pipeline (Weilbacher et al., 2014, version 2 or greater) which applies bias and flat calibrations, reduces the sky flats, and reconstructs the cube of individual exposures following wavelength calibration. Each individual exposure is then post-processed with the CUBEXTRACTOR package (CUBEX, Cantalupo et al., 2019), which is one of the available tools to improve the quality of the datacubes and mitigate the known imperfections arising after the basic ESO pipeline reduction (see figure 1 in Lofthouse et al., 2020a). In particular, the residual differences in the relative illumination of the 24 MUSE IFUs are corrected with the CUBEFIX tool in CUBEX which scans the cube as a function of wavelength to re-align the relative illumination of the IFUs. Next, the CUBESHARP tool (also part of CUBEX) is used to perform local sky subtraction removing the residuals left from the ESO pipeline reduction, by taking into account spatial variations in the line spread function of the instrument. A second iteration using both these tools is then performed masking continuum-detected sources to enable an improved determination of the background. At the end of the reduction process, in order to allow a more accurate identification of possible contaminants, all the single exposures are combined in four final products: an average cube of all the exposures (mean cube), a median cube and two cubes, each containing only one half of all the exposures. These last two products are useful to confirm the detection of sources in two independent datasets.

As fully described by Lofthouse et al. (2020a), the uncertainties of each pixel are computed by propagating the detector noise through the whole reduction process. Since the individual pixels are transformed across several steps, including the non-linear interpolation leading to the final data cube, the resulting pixel variance does not accurately reproduce the effective standard deviation of each volumetric pixel. Following the procedure detailed by Fossati et al. (2019), we derive a series of data cubes with accurate estimates of the standard deviation by bootstrapping the pixels in individual exposures in order to compute the noise in each final product.

2.2.2 High-resolution quasar spectroscopy

All the quasars included in the MAGG survey have high-resolution ($R \gtrsim 30,000$) and moderate or high signal-to-noise ($S/N \gtrsim 10$ per pixel, for details see table 2 of Lofthouse et al. 2020a) spectra obtained with the Ultraviolet and Visual Echelle Spectrograph (UVES, Dekker et al., 2000) at the VLT, the High Resolution Echelle Spectrometer (HIRES, Vogt et al., 1994) at Keck, or the Magellan Inamori Kyocera Echelle (MIKE; Bernstein et al., 2003) at the Magellan telescopes. Moderate dispersion spectra from X-SHOOTER (Vernet et al., 2011) at the VLT and Echelle Spectrograph and Imager (ESI, Sheinis et al., 2002) at Keck are also used to extend the observed wavelength range. Details on the reduction procedure for each instrument and a summary of the data available for each quasar are described by Lofthouse et al. (2020a) in section 3.1 and listed in their table 2. The main observational information that are summarized therein include: the wavelength range covered by the data, the spectral resolution, and the final signal-to-noise ratio at selected wavelengths.

2.3 Data Analysis

2.3.1 Catalogue of C IV absorption lines

We visually inspect the high-resolution quasar spectra searching for C IV $\lambda\lambda$ 1548, 1550 resonant doublets, identified through the characteristic rest wavelength separation of $2.575 \text{ \AA}$, corresponding to 489 km s^{-1} , and the equivalent width ratio 2:1 for $W_r^{1548} : W_r^{1550}$ in the unsaturated regime. For the sightlines where multiple observations from different instruments are available, we choose to inspect the spectrum with the highest signal-to-noise ratio for each sightline in a specific wavelength range. The inspection is limited to in the wavelength range lying redward the quasar Ly α forest, so as to exclude wavelengths in which metals are only the minority of the observed transitions and thus hard to detect due to the severe blending. We also exclude small windows that overlap with known strong telluric bands at $6870\text{--}6935 \text{ \AA}$ and $7595\text{--}7700 \text{ \AA}$ since these contaminant lines of the Earth’s atmosphere could be more easily misinterpreted as false positives. To obtain the highest possible purity of the sample, we also reject all the candidate C IV doublets that show nonphysical line ratios (i.e., not matching the ratio set by the oscillator strengths) in the range $7180\text{--}7295 \text{ \AA}$, a further window that is contaminated by the presence of additional weak telluric features. Lastly, we mask a velocity window $|\Delta v| < 3000 \text{ km s}^{-1}$ from the redshift of each quasar to avoid proximity effects (see, e.g., Hennawi & Prochaska, 2007). The resulting sample includes 467 candidate absorbers in the redshift range $z \approx 2 - 4$. To reduce the subjectivity of human classification performed at this stage, candidates absorbers are vetted by two authors (MG and RD) independently. Since the galaxies targeted in the optical band observed by MUSE lie at $z > 3$, we further restricted our sample to the systems above this redshift, for a total of 332 candidate C IV absorbers.

The availability of high resolution spectra, with intrinsic widths of the lines often resolved, allows us to model the absorption lines as Voigt profiles, see Figure 2.1 for a few examples. Once the spectra are continuum-normalized (see Lofthouse et al., 2020a for the detailed procedure), we fit each candidate doublet with Voigt profiles by running the MC-ALF code (Fossati et al., 2019). Using a Bayesian formalism in which the likelihood is sampled with a nested sampling algorithm and the best model is chosen via the Akaike Information Criterion, this code models the absorption profiles assuming an initial number of components within an interval based on visual inspection of each line (e.g, typically ranging between $\approx 5 - 10 \text{ \AA}$ for most profiles) and then estimates the minimum number of Voigt components required to reproduce the observed flux. Each component is defined by the redshift, the Doppler parameter and the column density. We set priors restricting the line parameters between $1 \leq b/(\text{km s}^{-1}) \leq 30$ for the Doppler parameter and $12 \leq \log(N/\text{cm}^{-2}) \leq 16$ for the column density. To account for possible uncertainties in continuum-fitting, we multiply the normalized continuum by a constant (allowed to vary in the range $0.98 - 1.02$) which is included as a free parameter. To improve the quality of the fit, we allow the spectral resolution to vary slightly around the nominal value. Specifically, the range extends over $\pm 1 \text{ km s}^{-1}$ and $\pm 5 \text{ km s}^{-1}$ for high-resolution (HIRES, UVES, MIKE) and moderate-resolution (X-SHOOTER, ESI) spectra, respectively. The model is then convolved with a Gaussian kernel to match the line spread function of the observed data. To reproduce the shape of blended absorbers, we introduce filler Voigt components that account for absorption lines physically uncorrelated with C IV doublets.

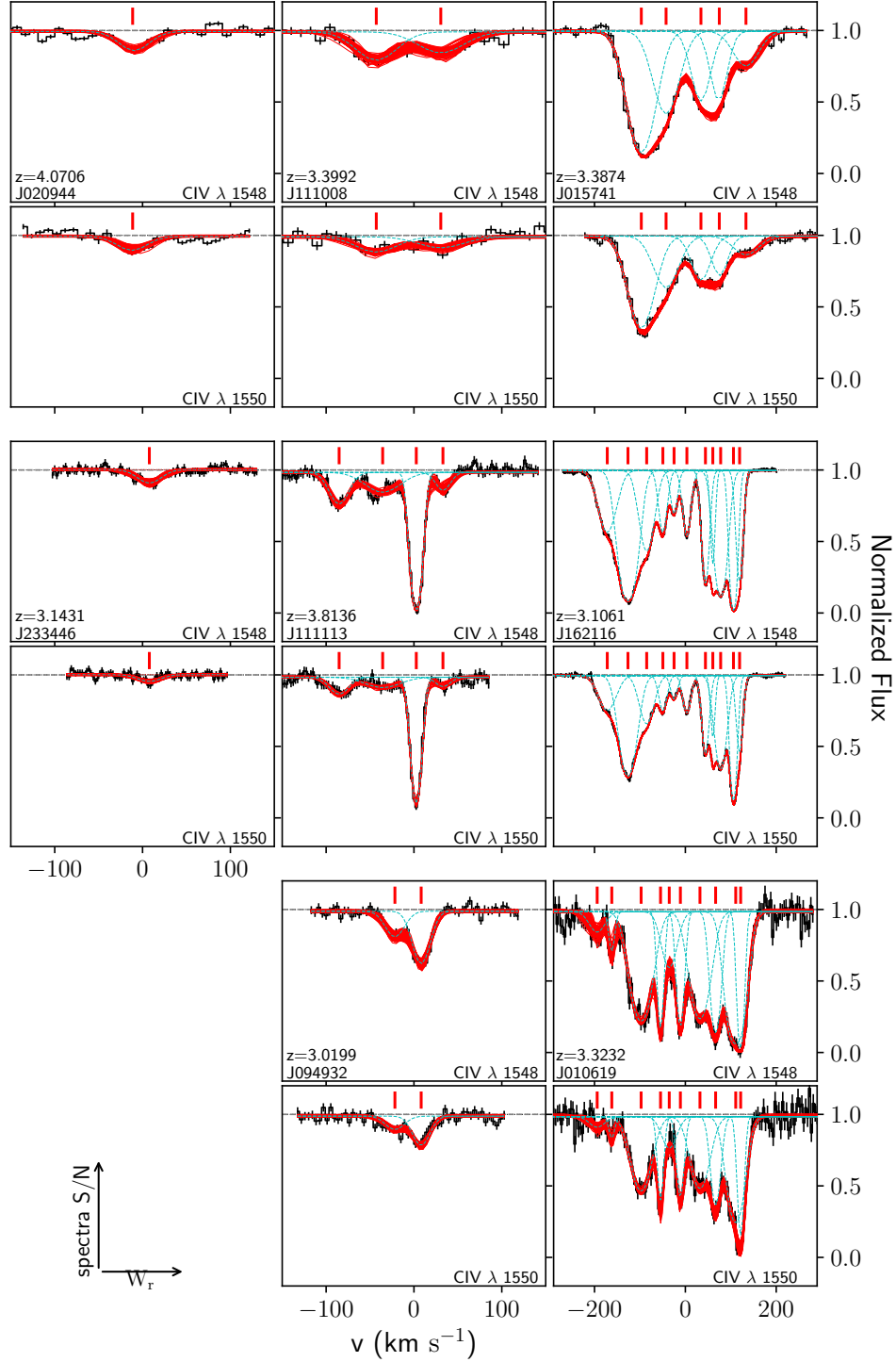


Figure 2.1: Examples of C IV doublet absorption line systems detected in MAGG quasar spectra at different S/N ($S/N \leq 20$, $20 < S/N < 40$ and $S/N \geq 40$) and randomly extracted from three bins of C IV equivalent width ($W_r/\text{\AA} \leq 0.03$, $0.03 < W_r/\text{\AA} < 0.20$ and $W_r/\text{\AA} \geq 0.20$). The normalized spectrum (step black lines with 1σ error bars) is reproduced by Voigt profiles sampling the posterior distribution of the parameters obtained by MC-ALF (solid red lines). Each individual C IV component is marked by a vertical red tick and shown as cyan dotted line.

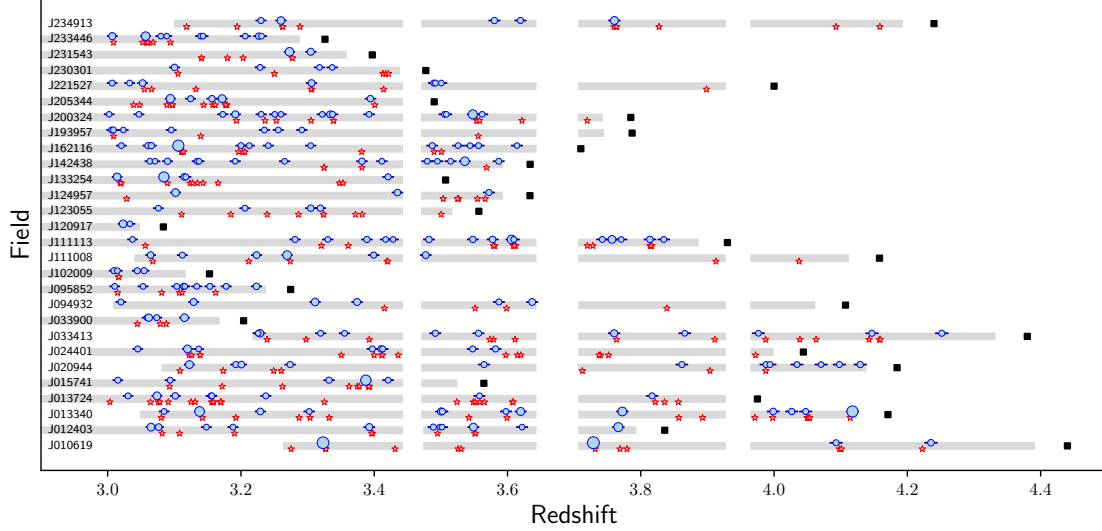


Figure 2.2: Summary of the final MAGG-C IV sample. The grey bands mark the searchable C IV redshift path of each sightline. The gaps correspond to known telluric bands we masked to avoid false positives. The doublets at $z \geq 3$ are drawn as blue dots and sized proportionally to the equivalent width of each line. The linking window corresponding to $|\Delta v| < 500 \text{ km s}^{-1}$ is marked as a blue line around each CIV absorber. LAEs detected in the MUSE cubes $z \geq 3$ and lying in the C IV redshift path are shown as red stars. The central quasar of each field is shown as a black square. A velocity window $|\Delta v| < 3000 \text{ km s}^{-1}$ from the redshift of each quasar is also masked to avoid proximity effects.

Once the Voigt model is computed for all the candidate doublets, we identify the significant detections by computing the rest-frame equivalent width for all the posterior samples returned by the chains of each MC-ALF fit. Filler components are excluded. Based on this estimate, we inspect again all the candidate doublets and consider significant detections those systems with an equivalent width of the weakest line, W_r^{1550} that, once compared to its error, is above 3σ . This step leaves us with 264 systems at $z > 3$. All the selected doublets are visually inspected one last time to manually exclude heavy blended lines for which we cannot reliably reconstruct a unique model. We also exclude the weakest features that passed the automatic selection but appear to be arising from imperfect continuum normalization and for which the doublet line ratio deviates from the modelled one set by the ratio of the oscillator strengths.

Sightline	Instrument	z	$W_{1548} (\text{\AA})$	$\Delta v_{90} (\text{km s}^{-1})$
J010619+004823	MIKE	3.3232	1.108 ± 0.008	247.707 ± 4.198
J010619+004823	MIKE	3.7289	1.146 ± 0.013	285.509 ± 14.692
J010619+004823	MIKE	4.0931	0.093 ± 0.007	37.800 ± 8.400
J010619+004823	ESI	4.2354	0.188 ± 0.006	229.808 ± 10.793
J012403+004432	UVES	3.0650	0.393 ± 0.003	155.025 ± 2.501

Table 2.1: List of the properties (z , W_{1548} , Δv_{90}) measured for the C IV absorption line systems detected at 3σ and included in the sample studied in this work. We also provide information about the sightline and the spectrum in which each absorber is detected. Only the first five lines are shown here, the full table is available as online supplement.

The final sample includes 220 C IV absorption-line systems with redshift $3.0 \lesssim z \lesssim 4.3$ (median $z \approx 3.31$; see Figure 2.3). In assembling this sample, all the components found within a velocity separation of $\pm 500 \text{ km s}^{-1}$ from the highest column density one are considered as an individual absorption system. We show a summary of the final sample in Figure 2.2, where the doublets are plotted in the C IV redshift path of each sightline and sized as a function of the rest frame equivalent width. For each C IV system in the catalogue we provide a measure of the rest-frame equivalent width (median $W_r^{1548} \approx 0.089 \text{ \AA}$) and a measure of the kinematics traced by the line width in velocity space (Prochaska & Wolfe, 1997), defined by the interval enclosing the 90% of the optical depth (median $\Delta v_{90} \approx 74.1 \text{ km s}^{-1}$). The C IV absorption systems included in this sample and their properties (redshift, W_r and Δv_{90}) are listed in Table 2.1 and available as online

material.

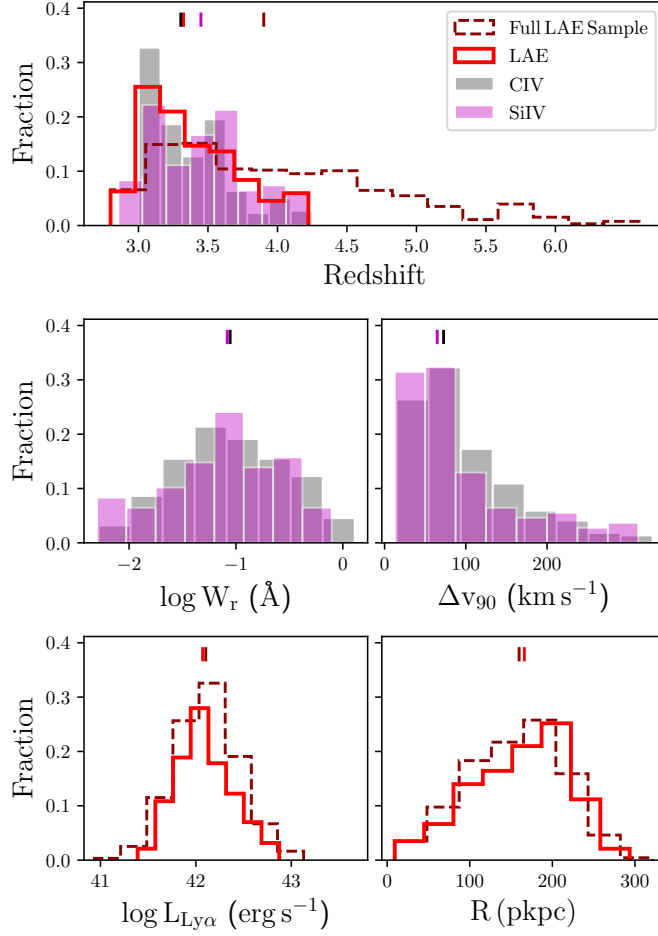


Figure 2.3: Summary of the properties measured for the C IV (black) and Si IV (magenta) doublets (redshift, rest-frame equivalent width and velocity width) and those of the emitters (redshift, $L_{\text{Ly}\alpha}$ luminosity and impact parameter) identified in the MAGG fields at $\text{SNR} > 7$ (dashed red line) and in the absorbers redshift path (solid red line). Vertical ticks mark the median of each distribution.

To assess the completeness of this sample we compare the final observed equivalent width distribution from this search to known completeness-corrected functions from the literature. Specifically, we compute the equivalent width frequency distribution, $f(W_r, X)$, which counts the number of C IV systems per unit equivalent width per unit survey co-moving path, X . The uncertainties on the estimate of $f(W_r, X)$ are derived as the 10th and 90th percentiles of the distribution obtained by bootstrapping over the sightlines with 10^3 repetitions. To gauge the completeness of our search, in Figure 2.4 we compare the observed equivalent width frequency distribution obtained for the C IV line at 1548 Å in MAGG, as a function of the rest-frame equivalent width, with the completeness corrected function derived by Hasan et al. (2020) and Cooksey et al. (2013), which are selected as significant examples in the literature in terms of statistical modelling of $f(W_r, X)$. In detail, Hasan et al. (2020) inspected 369 quasar sightlines at $1.1 \leq z_{\text{qso}} \leq 5.3$ observed with Keck/HIRES and VLT/UVES, compiling a sample of 1318 absorbers in the redshift range $1.0 \leq z \leq 4.75$ that is 50% complete at $W_r \geq 0.05$ Å. Cooksey et al. (2013) searched ≈ 26000 sightlines from the SDSS DR7 quasar catalogue at $z_{\text{qso}} \geq 1.7$. The resulting sample is a collection of 16459 absorbers with $1.46 \leq z \leq 4.55$ that is 50% complete at $W_r \geq 0.6$ Å.

According to the comparison shown in Figure 2.4, our $f(W_r, X)$ is in good agreement with the observations reported in the literature. Overall, this good agreement validates the procedures adopted to extract our C IV sample. We also note that the slope of MAGG data starts decreasing around $W_r \approx 0.08$ Å and flattens significantly at lower column densities, a feature we attribute to the progressively-increasing incompleteness of our sample. Moreover, the MAGG quasar spectroscopy and the data from Hasan et al. (2020) are comparable in quality, both being derived from echelle data at $S/N \gtrsim 10$. Therefore, if we take as reference the sensitivity function shown in figure

4 from Hasan et al. (2020) where the 50% completeness limits are highlighted for different redshift ranges, we can assume that the MAGG C IV is approximately 50% complete around an equivalent width of $W_r^{1548} \approx 0.05 \text{ \AA}$, and generally complete for $W_r^{1548} \gtrsim 0.08 \text{ \AA}$.

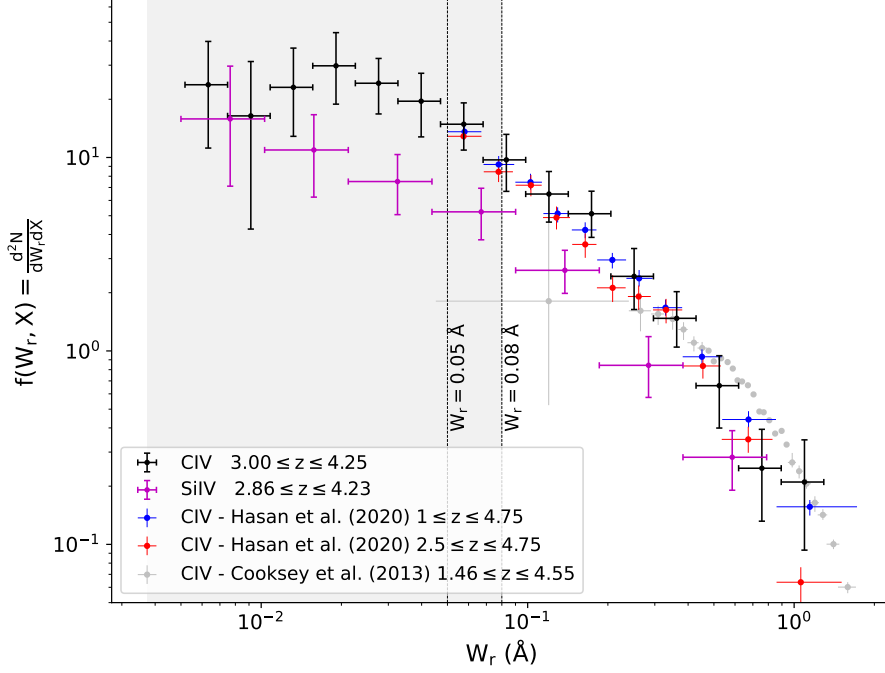


Figure 2.4: Equivalent width frequency distribution function $f(W_{r,1548}, X)$ for the MAGG C IV sample (black dots). Vertical error bars are the 10th and 90th percentiles from bootstrap re-sampling, while horizontal error bars reproduce the width of each redshift bin. The MAGG sample is compared with the completeness-corrected results from Hasan et al. (2020) (red and blue dots) and Cooksey et al. (2013) (grey dots). The shaded region mark the equivalent width interval over which our search suffers from incompleteness. Also shown (magenta dots), the equivalent width distribution for the Si IV line at 1393 Å.

2.3.2 Catalogue of Si iv absorption lines

With the goal of identifying an additional tracer of ionized gas comparable to C IV in the wavelength range covered by our observations, the same strategy developed to assemble the catalogue of C IV doublets is implemented to search and analyze Si IV $\lambda\lambda$ 1393, 1402 absorption-line systems. A first sample of 145 Si IV absorbers with redshift $2.9 \lesssim z \lesssim 4.4$ is built by visually inspecting the quasar spectra. Running MC-ALF code with the same priors and setting defined for C IV doublets, we fit each of these lines and estimate the minimum number of Voigt components required to reproduce the observed flux. We exclude all the candidates that do not show a characteristic rest wavelength separation of $\approx 9 \text{ \AA}$ and optical depth ratio $\tau_{1393}(\nu)/\tau_{1402}(\nu) \approx 2$ (as set by the ratio of the oscillator strengths) between the two transitions of the doublet in the unsaturated regime. The final sample includes 108 Si IV systems with a rest-frame equivalent width of the weakest part of the doublet, W_r^{1402} , detected above 3σ . For each absorber we provide a measure of the redshift (median $z \approx 3.45$), rest-frame equivalent width (median $W_r^{1393} \approx 0.077 \text{ \AA}$) and the width in velocity space (median $\Delta v_{90} \approx 64.7 \text{ km s}^{-1}$). The properties of these systems are listed in Table 2.2 and available as online material, while a summary of the distribution of these properties is shown in Figure 2.3, with median values listed in Table 2.3. To estimate the completeness of the sample, we derive the equivalent width frequency distribution, $f(W_r, X)$, which we compare with the C IV one in Figure 2.4. We note that the shape of the Si IV is similar to what we obtained for the C IV sample, and we conclude that the completeness estimates above can be also applied to the Si IV catalogue, which we take to be $\approx 50\%$ complete for $W_r^{1393} \gtrsim 0.05 \text{ \AA}$, and complete for $W_r^{1393} \gtrsim 0.08 \text{ \AA}$.

Sightline	Instrument	z	W_{1548} (Å)	Δv_{90} (km s ⁻¹)
J010619+004823	MIKE	4.2348	0.065 ± 0.009	55.649 ± 10.501
J012403+004432	UVES	3.3922	0.106 ± 0.002	79.997 ± 2.500
J012403+004432	UVES	3.5488	0.321 ± 0.002	197.532 ± 2.501
J012403+004432	UVES	3.6755	0.388 ± 0.002	115.727 ± 1.874
J012403+004432	UVES	3.7661	0.057 ± 0.001	105.014 ± 1.250

Table 2.2: Same as Table 2.1 for the Si IV absorbers detected at 3σ . The first 5 lines are shown, while the full table is available as online material.

Property	z	W_r (Å)	Δv_{90} (km s ⁻¹)
C IV	3.31	0.089	74.1
Si IV	3.45	0.077	64.7

Table 2.3: Summary of the median properties of C IV and Si IV absorbers, where W_r refers to the strongest part of each doublet.

2.3.3 Identifying galaxies in MUSE data

To link the ionized gas detected in absorption via C IV with the surrounding galaxy population, we proceed to compile catalogues of galaxies detected in the MUSE data. The first step is to identify continuum-bright galaxies. Lofthouse et al. (2020a) and MAGG IV provide extensive details on how continuum sources are handled. Briefly, we first run SEXTRACTOR (Bertin & Arnouts, 1996) on the reconstructed white-light image obtained by collapsing the cube along the wavelength axis. We then extract 1D spectra from the cubes following the 2D segmentation maps generated by SEXTRACTOR and use the M. Fossati branch¹ of the MARZ tool (Hinton et al., 2016) to measure spectroscopic redshifts for the extracted sources. At the end of this procedure, we identify over 1200 sources with redshift, corresponding to 90% completeness down to ≈ 24.85 mag.

With the purpose of extending the study of the CGM at high redshift to lower-mass star-forming galaxies, we next search the MUSE field of view for LAEs, here defined as any galaxy emitting bright Ly α emission lines on a typically faint or non-detected continuum. Following the procedure described by Lofthouse et al. (2020a) and MAGG IV, we first run the three-dimensional automatic extraction performed by CUBEX to identify connected groups of voxels (three-dimensional pixels) that lie above a defined $S/N = 3$ on the cubes with subtracted continuum sources, including both galaxies and the quasar point spread function (PSF). For the subtraction of the PSF, we follow the non-parametric method detailed in section 3.1 of Arrighi Battaia et al. (2019b).

The presence of residuals from the continuum of stars and galaxies is mitigated by masking the spatial position of all the continuum-detected sources with known redshift (see Borisova et al., 2016). This process produces a set of cubes free of any continuum contamination which can be used to build a sample of compact LAEs.

More specifically, to be identified by the CUBEX algorithm, each line emitter candidate requires: i) a minimum volume, imposed by a number of voxels ≥ 27 above the threshold S/N ; ii) a lower limit of the width in the wavelength direction of at least at least 3 wavelength channels; iii) an upper limit in the wavelength direction of 20 channels to exclude any possible contamination from residuals arising in continuum sources. Since these selection criteria do not remove all the contaminants, such as residual cosmic rays, visual inspection of each extracted LAE is required to weed out the remaining contaminants. Compact LAEs are also distinguished from residuals of sky lines, significant noise fluctuations at the edge of the field or lower redshift emission lines (typically, C IV $\lambda\lambda$ 1548, 1550; C III] $\lambda\lambda$ 1907, 1909; [O II] $\lambda\lambda$ 3727, 3729; [O III] λ 5008; H β λ 4862). In order to assure the best quality sample, a first classification is performed by two authors (MG and EL) and confirmed by other two authors (MF and MfO), independently.

The visual inspection of the 28 MAGG fields results in a sample of 994 LAEs at redshift $2.81 \lesssim z \lesssim 6.60$ and detected at integrated $S/N \geq 7$. All the identified emitters are divided into three confidence levels, which updates the original classification by Lofthouse et al., 2020a based on the integrated signal-to-noise ratio (ISN) and confidence of the classification: i) *Confidence 1* class contains emitters with $ISN \geq 7$ which are confidently detected and unambiguously recognized as LAE due to e.g. asymmetric line profiles, or presence of additional absorption lines other than Ly α ; the *confidence 1.5* sub-class contains instead confidently-detected sources at $ISN \geq 7$ that

¹matteofx.github.io/Marz

are recognized as LAEs but for which there is less confidence in ruling out lower-redshift sources (e.g. with double peaked profiles possibly mimicking [OII] emitters but not recognized as such due to the lack of other oxygen lines or with atypical line separations/ratios); ii) *Confidence 2* emitters have $ISN < 7$ but are deemed to be LAEs, with a sub-class, marked as *confidence 2.5*, containing any system that is noisy and shows anomalies due to noise in its segmentation map and half-exposure cubes; iii) *Confidence 3* emitters are recognized as LAEs but lie at the edge of the FoV in noisy regions or only partially overlap with the detector, making any measurement of their properties (e.g. integrated flux, centroid) unreliable. Figure 2.5 provides an example of sources classified under class 1 and 1.5, which are used in this work. Based on this classification, we assembly a sample of LAEs detected at higher integrated S/N ($ISN \geq 7$) compared to the MAGG IV sample ($ISN \geq 5$) by excluding all the emitters with confidence ≥ 2 from the following analysis. The final sample includes 921 high-confidence LAEs.

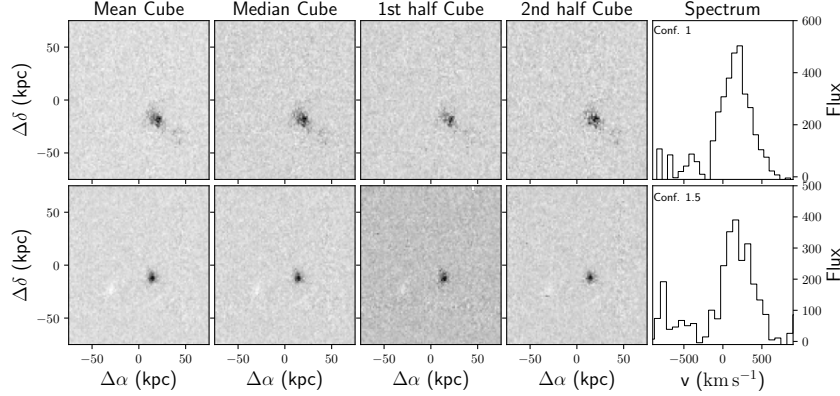


Figure 2.5: Examples of high-confidence LAEs identified in the fields. Columns from left to right show the images extracted by CUBEX in the mean, median and the half-exposure cubes. The last column shows the 1D spectra limited to the wavelength channels centered on the redshift of the Ly α . The flux is measured in units of $10^{-20} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$.

Field	RA (deg)	DEC (deg)	z	$\log[L_{\text{Ly}\alpha}/(\text{erg s}^{-1})]$	R (pkpc)	SNR	Confidence
J010619+004823	16.5788	0.8118	2.9263	41.620 ± 0.083	157	8.3	1.0
J010619+004823	16.5834	0.8042	3.1082	42.289 ± 0.023	112	20.4	1.0
J010619+004823	16.5860	0.8001	3.1213	41.627 ± 0.099	242	8.5	1.0
J010619+004823	16.5834	0.8075	3.2081	42.033 ± 0.036	95	13.9	1.0
J010619+004823	16.5782	0.8011	3.2112	42.169 ± 0.044	159	15.0	1.0

Table 2.4: List of the LAEs detected in MAGG at $SNR \geq 7$. Information of the sightline, coordinates, SNR and confidence from our classification are included as well as the properties we measured (redshift z , luminosity $\log[L_{\text{Ly}\alpha}/(\text{erg s}^{-1})]$, projected distance to the quasar R). Typical errors for redshifts are of the order of $(6 - 7) \times 10^{-4}$. Only the first five lines are shown here: the full table is provided as online material.

For each LAE included in the final MAGG sample, we derive a measure of the redshift, Ly α luminosity and projected distance from the line-of-sight (i.e., the impact parameter). For this purpose, from the 3D segmentation cubes produced by CUBEX, the 1D spectra is extracted along the full wavelength range and used to derive an estimate of the redshifts. Ly α photons are known to be subject to radiative transfer effects which affect both their spatial and frequency distribution, resulting in typical asymmetric or double-peaked emission lines. Therefore, we apply the following convention (see also Verhamme et al., 2018; Muzahid et al., 2020, MAGG IV): in case of a noticeable double-peaked line ($\approx 10\%$ of the sample), we take the redshift of the red peak; if a single peak is observed instead, we assume that this is the red peak since the blue one is more easily absorbed.

To derive the Ly α luminosity we adopt instead a curve of growth (CoG) analysis. This method better adapts to the typical Ly α emission profile that is spatially extended and characterized by two distinct components: a bright core and a faint diffuse halo (Wisotzki et al., 2016; Leclercq et al., 2017). The latter is easily excluded in the flux measured from the segmentation cube at moderate S/N. Here, we follow the method described by Fossati et al. (2019) (see also Marino et al., 2018). For each emitter we build a pseudo narrow-band image by summing up the spectral channels within

Property	z	$\log[L_{\text{Ly}\alpha}/(\text{erg s}^{-1})]$	R (kpc)
Full Sample	3.91	42.10	160
In C IV z-path	3.24	42.07	166

Table 2.5: Summary of the median properties of LAEs reported for the full sample and for the emitters detected in the C IV redshift path.

$\pm 14 \text{ \AA}$ from the source redshift and masking the neighbours. A series of circular apertures with increasing radii, over which we produce the CoG flux, is centered on the CUBEX coordinates of each source. As a result, we take as reliable CoG flux the estimate obtained at the last radius where the total flux increases by more than 2.5% than the previous step. All the diagnostic plots are visually inspected to check that no contamination other than the LAEs has been included in the fluxes. Fluxes are then corrected for Milky Way extinction using the re-calibrated extinction map (Schlafly & Finkbeiner, 2011) and assuming the Milky Way extinction law by Fitzpatrick (1999). Finally, we compute the luminosity distance at the redshift of each source and convert the total flux into a Ly α luminosity. A complete list of the Ly α emitters identified in the MAGG survey at $SNR > 7$, including a measure of the redshift, Ly α luminosity and impact parameter, is provided in Table 2.4 and available as online material.

In the whole MAGG sample, we identify 292 LAEs lying in the C IV redshift path. This subset also includes Ly α emission from 23 galaxies detected in continuum and forms the candidates for associations with the absorbers. Although we provide a complete survey of UV-selected galaxies, including continuum-detected LBGs and LAEs, our sample significantly favors line emitters. Thus, additional $z > 3$ sources, e.g. LBGs with absorption features but not Ly α emission (1 detected in C IV redshift path), passive or heavily obscured galaxies, are not well represented in this work. In Figure 2.3 and Table 2.5 we summarize the properties measured for LAEs (redshift and Ly α luminosity) in the full MAGG sample and in the sub-set lying in the C IV redshift path. The latter are observed in a redshift range $2.80 \lesssim z \lesssim 4.23$, have a median $z \approx 3.24$ ($z \approx 3.91$ for the full MAGG sample) and emit a luminosity $41.39 \lesssim \log(L_{\text{Ly}\alpha}/\text{erg s}^{-1}) \lesssim 42.87$, with a median $\log(L_{\text{Ly}\alpha}/\text{erg s}^{-1}) \approx 42.07$ (median $\log(L_{\text{Ly}\alpha}/\text{erg s}^{-1}) \approx 42.10$ for the full MAGG sample). The emitters are detected at a projected distance from the line of sight $20 \lesssim R/\text{kpc} \lesssim 295$, with a median $R \approx 166 \text{ kpc}$ ($R \approx 160 \text{ kpc}$ for the full MAGG sample).

2.3.4 Identifying galaxy groups

According to the hierarchical scenario of structure formation, galaxies assemble in groups or clusters, with only a fraction evolving in isolated environments. In dense galaxy environments, interactions between galaxies and with the surrounding medium are observed to affect the properties of different gas phases of the CGM, both at low and high redshift (see e.g., Nielsen et al., 2018a; Fossati et al., 2019; Dutta et al., 2020; Muzahid et al., 2021). Therefore, a better understanding of the gas-galaxy connection and co-evolution requires galaxy groups to be identified in the MAGG survey. For this purpose, without including any constraint on the halo mass, we recognize a galaxy to be part of a group if it is not isolated within a line-of-sight separation $|\Delta v| \leq 500 \text{ km s}^{-1}$ in the MUSE FoV. By linking galaxies with this criterion, we identify 53 groups in the C IV absorbers redshift path shown in Figure 2.2 (152 in the full MAGG sample) so that $\approx 47\%$ of the galaxies are part of a group ($\approx 41\%$ for the full MAGG sample). The majority of these identified groups ($\approx 57\%$) are composed of two galaxies ($\approx 61\%$ for the full MAGG sample), but groups of three to six members (seven for the full MAGG sample) are also found.

2.4 Results

2.4.1 Measuring correlations between C IV absorbers and LAEs

Connecting C IV absorbers with galaxies

To establish the extent with which the detected C IV trace ionized gas in proximity of LAEs, we cross-match the catalogue of C IV absorbers with the one of the LAEs in redshift space. For this purpose, we center on each absorption system and we search for the galaxies within $\pm 2500 \text{ km s}^{-1}$ along the line of sight, reserving the definition of a more stringent criterion (i.e., a smaller velocity separation) for the following analysis. In Figure 2.6 we show the gas-galaxy connections as a

function of the line-of-sight separation (horizontal axis) and as a function of the transverse distance (vertical axis). We distinguish between the LAEs detected with high confidence (red dots) and those observed at the edge of the field of view (blue dots). We also highlight the galaxies associated to C IV absorbers that are connected to Lyman limit systems (LLSs) from the MAGG IV paper (gold dots). The MAGG IV sample includes 127 LAEs detected at $S/N \geq 5$, with 79 of these galaxies being detected at $S/N \geq 7$ (the threshold adopted in our study) and being associated to 61 LLSs. As $\approx 70\%$ of the $S/N \geq 7$ LAEs are in the vicinity of a C IV absorber, this fraction is also independently recovered by our search.

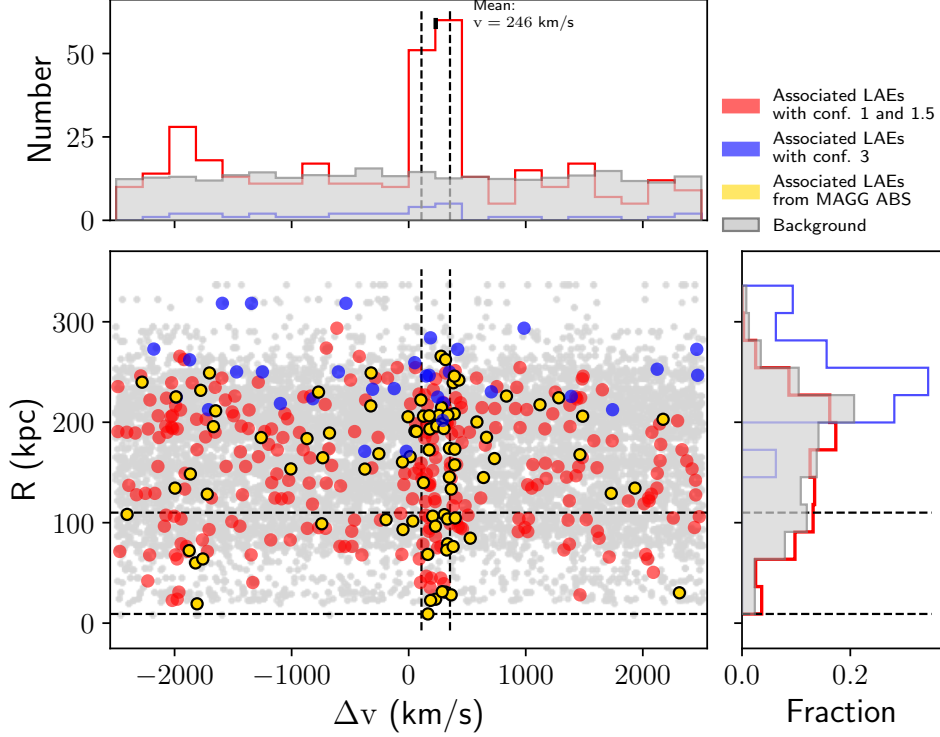


Figure 2.6: Distributions of LAEs near C IV absorbers as a function of line-of-sight velocity separation, within $\Delta v = \pm 2500 \text{ km s}^{-1}$, and projected distance. The high-confidence LAEs are marked in red, while the galaxies at the edge of the field are in blue. In gold, we mark the galaxies connected to C IV absorbers that are also associated to high-column hydrogen absorbers from the MAGG IV paper. A random population of LAEs, which is representative of background and foreground population at larger velocity separations from C IV absorbers, is shown for reference in grey. The dashed lines highlight the regions where a clear excess signal of LAEs around C IV is detected over the uncorrelated background.

In order to characterize the significance of the gas-galaxy connection, we build a control sample (in grey within Figure 2.6) which aims at reproducing the typical distribution of the galaxies around random regions in the Universe not selected in absorption, where we do not necessarily expect C IV systems. To do so, we bootstrap over the sightlines, randomly selecting 28 fields from the full MAGG sample for a total of 10^3 realizations. At each iteration and in each selected field, we measure the distance of LAEs to the redshift of a real C IV absorber, which is however randomly extracted from all the systems detected in the remaining sightlines. Using this technique, we are effectively building a control sample using real data (both for LAEs and C IV absorbers), optimally controlling for systematics like the presence of sky line residuals.

Examining the distribution of velocities along the line of sight in the upper histogram of Figure 2.6, we observe a clear excess of LAEs near C IV absorbers compared to the random sample distribution at offset velocities of $\Delta v = 0 - 500 \text{ km s}^{-1}$. Assuming that the galaxies are randomly distributed relative to the line of sight, one would expect to find a relative velocity separation between LAEs and C IV that is symmetric around the C IV redshift. However, being based on the Ly α emission line, our estimate of the galaxies redshift suffers from the resonant scattering that affects the frequency and spatial distribution of the Ly α photons emitted by the galaxies. To measure this offset, we fit the velocity separation distribution with a Gaussian model plus a background as a function of the velocity separation, obtaining a mean value of $\Delta v_{\text{Ly}\alpha} \approx 246 \text{ km s}^{-1}$. Similar trends are indeed observed in the literature. Rakic et al. (2011) measured an offset $\Delta v_{\text{Ly}\alpha} \approx (295 \pm 30) \text{ km s}^{-1}$

between $z \sim 2.3$ LBGs and stacked H I absorption lines. More recently, [Muzahid et al. \(2020, 2021\)](#) derived a velocity offset of the order of $\Delta v_{\text{Ly}\alpha} \approx (171 \pm 8) \text{ km s}^{-1}$ between LAEs and stacked H I absorption at redshift $z \approx 3$. They also obtained a similar velocity offset, although with larger uncertainties, between LAEs and stacked C IV profile. Lastly, the velocity separation between LAEs and H I absorbers observed in MAGG IV peaks at $\Delta v_{\text{Ly}\alpha} \approx 250 \text{ km s}^{-1}$. In light of this result, in the following we establish associations by correcting the redshift of the LAEs for the mean offset $z'_{\text{LAE}} = z_{\text{Ly}\alpha} - \Delta v_{\text{Ly}\alpha} (1 + z_{\text{Ly}\alpha})^{-1} c^{-1}$.

Since the majority of the LAEs is observed within $0 - 500 \text{ km s}^{-1}$ from the absorber, we can restrict the velocity window in which a galaxy is considered associated with an absorber to build a sample of connected systems. In the following, as done in previous papers of this series, we perform our analysis on galaxies that are $\pm 500 \text{ km s}^{-1}$ from a C IV absorber, an interval that becomes $-254 < \Delta v / (\text{km s}^{-1}) < 746$ once corrected for the Ly α velocity offset. We note that key results of this paper are often interpreted as excess signals over background values computed in comparable volumes and thus are not affected by the choice of the velocity window within which we connect the absorbers to the galaxies, as long as the excess of LAEs around C IV systems is fully included. We detect at least one LAE within this velocity separation for 79 C IV absorbers, corresponding to a detection rate of 36 ± 5 per cent (79/220), which is statistically significant compared to the rate of 17 ± 3 per cent we measured for the control sample by connecting C IV absorbers to LAEs taken from different fields.

The right histogram of Figure 2.6 shows instead the transverse separation distribution between the C IV absorbers and the LAEs. The overall trend is consistent with an increasing number of detections with increasing area of the annulus ($\propto R^2$), and a steep decrease due to edge effects at $R > 200 \text{ kpc}$. Although with a significant scatter, only a very small excess of LAEs is detected at transverse separations $R < 100 \text{ kpc}$ over the control distribution, suggesting that some, but not many, C IV systems preferentially arise from the inner CGM of the associated galaxy. Overall, however, the impact parameter distribution looks similar to the one of the control population.

In summary, the connection in redshift space between the C IV absorbers and the LAEs reveals that a factor of ≈ 2.6 more galaxies are observed within $\pm 500 \text{ km s}^{-1}$ from an absorption line system relative to the number of galaxies detected in random regions of the Universe. This result implies that the C IV absorbers do not trace a random region of the Universe, but regions with a preference for hosting LAEs.

Number density of LAEs associated to C IV absorbers

The detection of a local clustering signal along the line-of-sight is indicative of the existence of a physical connection between the C IV absorbers and the LAE population. In this section, we measure statistically the LAE number density by deriving the luminosity function (LF), $\phi(L)dL$. A comparison of LFs from different environments (e.g., in the field or near C IV absorbers) further allows us to study the extent to which the LAE number density depends on the proximity to the absorption-line systems. The procedure we follow in the computation of the LF is described in detail by [Fossati et al. \(2021\)](#). We assume any redshift dependence to be negligible in the probed range $3.0 < z < 4.5$ as, based on the results from [Herenz et al. \(2019\)](#), no significant redshift evolution is observed in the wide range $2.9 < z < 6.0$. In the analysis, we compare a non-parametric estimator of the LF, $1/V_{\text{max}}$ from [Schmidt \(1968\)](#) and [Felten \(1976\)](#), with the results obtained from fitting the full sample with no binning using a parametric Schechter function.

The V_{max} non-parametric estimator relies on the survey selection function $f_c(L, z)$, defined as the probability to find an LAE with a given luminosity at a given redshift. The selection function is derived performing a simulation based on the injection of 500 mock LAEs and measuring the fraction of sources we recover running CUBEX and applying our search algorithm. The flux, redshift and spatial positions of the mock sources are tabulated. This procedure is repeated for 1000 iterations over all the MAGG fields, with the only exception of J142438+225600 that is lensed ([Patnaik et al., 1992](#)).

In order to compute the effective co-moving area of the survey, we re-scale the MUSE field area by a factor $\delta_{\text{obs}}(z)$ that accounts for the number of fields where we searched for LAEs at a given redshift. Therefore, we center on the redshift of each C IV absorber and consider only the contributions from the redshift range enclosed within $\pm 500 \text{ km s}^{-1}$ from the absorber. The factor $\delta_{\text{obs}}(z)$ is equal to zero in the ranges excluded from the searched path. The luminosity bins are 0.2 dex wide. We limit the uncertainties on the weight given by V_{max} by excluding from the analysis all the LAEs identified below $f_c(L, z) < 0.1$. The results of the luminosity function using this non-parametric estimate for the LAEs connected to the C IV absorbers are shown as black

Sample	S/N limit	$\log[L^*/(\text{erg s}^{-1})]$	α	$\log[\phi^*/(\text{Mpc}^{-3} \text{ dex}^{-1})]$	$\log(I/\text{Mpc}^{-3})$
QSO	5	42.573 ± 0.221	-1.164 ± 0.319	-1.429 ± 0.276	-1.239
ABS	5	42.558 ± 0.242	-1.350 ± 0.309	-1.788 ± 0.343	-1.540
C IV	7	42.575 ± 0.212	-1.418 ± 0.323	-2.175 ± 0.163	-1.887
Field	7	42.466 ± 0.081	-1.339 ± 0.158	-2.422 ± 0.121	-2.260

Table 2.6: Parameters of the Schechter function describing the LAE luminosity function around the different tracers studied in MAGG. The last column is the integral of the Schechter function computed at luminosities $\log[L_{\text{Ly}\alpha}/(\text{erg s}^{-1})] \geq 41.8$.

dots in Figure 2.7 with respective uncertainties.

We also derive a parametric estimate of the LAE luminosity function by fitting a Schechter function (Schechter, 1976) to the non-binned sample:

$$\phi(L) = \ln(10) \phi^* 10^{(\log L - \log L^*)(1+\alpha)} \exp[-10(\log L - \log L^*)]. \quad (2.1)$$

For the fit, we proceed as described in Fossati et al. (2021). The result from the parametric modelling of the luminosity function is shown as a solid black line in Figure 2.7 for the LAEs associated to the C IV absorbers, with the shaded areas marking the 1σ confidence intervals. In order to study how being in the vicinity of the C IV absorbers affects the LAEs number density, we apply the procedure described above to a sample of LAEs from a random sample (referred to as “MAGG field”) that is not selected by its proximity to a specific tracer. This field sample is built by including 617 LAEs detected in the range $3.0 < z < 4.5$, corresponding to the redshift interval in which we observe C IV absorption in MAGG. The non-parametric estimate and the parametric Schechter fit are shown in red in Figure 2.7.

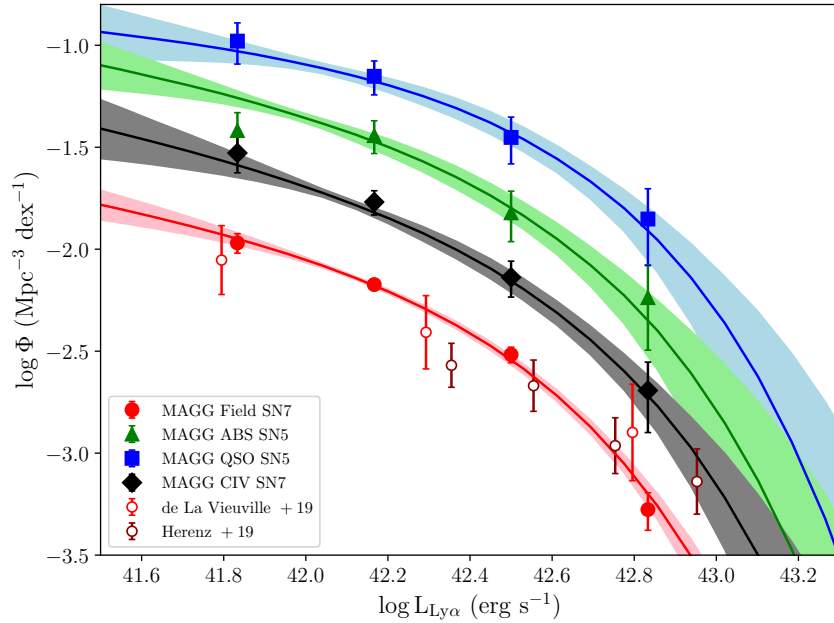


Figure 2.7: Luminosity function derived for the LAEs associated to the C IV absorbers (black diamonds) in a luminosity range $41.5 < \log(L_{\text{Ly}\alpha}/\text{erg s}^{-1}) < 43.0$, using non-parametric (data points) and parametric (lines) estimates. The results from this work are compared with the results from a field control sample (red circles), the environment of high-column density H I systems (green triangles) from MAGG IV, and the environment around quasars from Fossati et al. (2021) (blue squares). For comparison, we also show the field luminosity function from Herenz et al. (2019) and de La Vieuville et al. (2019).

To test the consistency of the field luminosity function with previous determinations, the results are compared to the literature. The estimate from Herenz et al. (2019) is based on a sample of 237 LAEs spread in a redshift range $2.9 < z < 6.0$ and detected in ~ 1 hour of MUSE observations. Here we report their results for a sub-sample restricted to the range $3.0 < z < 4.0$, chosen to be consistent with the C IV absorbers detected in MAGG. To extend the comparison at lower luminosities, we consider the results from de La Vieuville et al. (2019). This sample contains 156 LAEs at $2.9 < z < 6.7$ with luminosities $39 \lesssim \log(L_{\text{Ly}\alpha}/\text{erg s}^{-1}) \lesssim 43$. As done before, we consider

the luminosity function derived for a sub-sample of LAEs detected in the range $2.9 < z < 4.0$. As we adopt $H_0 = 67.7 \text{ km s}^{-1} \text{ Mpc}^{-1}$ while the literature studies chose $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, the luminosity functions derived from these two samples are scaled to be consistent with our cosmological parameters (a $\lesssim 10\%$ correction). As shown in Figure 2.7, the luminosity function derived for the MAGG field sample is consistent within the uncertainties with the results from the literature. Furthermore, the large size of the MAGG field sample provides a statistically significant modelling of the luminosity function at $3.0 < z < 4.5$ for a population of lower-mass star forming galaxies traced by LAEs with halo mass $M_H \approx 10^{11} \text{ M}_\odot$ compared to, e.g., continuum detected populations like LBGs with $M_H \approx 10^{12} \text{ M}_\odot$ (see Section 2.5.1 for additional details).

Having computed the field luminosity function, we derive an estimate of the LAE number density in the vicinity of the C IV absorbers extending the comparison with other populations in MAGG. In MAGG IV, we studied the luminosity function of 97 LAEs detected within $\pm 500 \text{ km s}^{-1}$ from 45 strong hydrogen absorbers selected as tracers of the dense CGM (the “MAGG ABS” sample). We note the MAGG ABS and the MAGG C IV samples are partially overlapping (not all C IV absorbers are associated to LLSs, see Section 2.5.2), thus the comparison is not independent. In addition, the “MAGG QSO” sample from Fossati et al. (2019) includes 86 LAEs found within $\pm 500 \text{ km s}^{-1}$ from the central quasar of each field. For consistency, these two samples include LAEs detected in a redshift range $3.0 \lesssim z \lesssim 4.0$ that is comparable to the MAGG C IV sample. The use of different S/N thresholds for the various samples ($S/N \geq 5$ for MAGG ABS and QSO samples, $S/N \geq 7$ for MAGG C IV and field samples) is justified by noticing that a higher S/N limit does not change the overall shape either of the non-parametric or the parametric estimate (as expected from the completeness correction based on the survey selection function), but only broadens the 1σ confidence regions. The best Schechter parameters estimated from the parametric fit are listed in Table 2.6 for each MAGG sample.

The parameters $\log L^*$ and α of the different samples are all consistent within 1σ . This suggests that the overall shape of the LAE luminosity function does not depend on which tracer is used to select the environment within which these galaxies reside. Despite the normalization constant, ϕ^* , having the units of a LAEs number density, it is not computed at fixed luminosity and thus cannot be directly compared among the different tracers. Therefore, we derive a measure of the relative LAEs overdensity by computing the integral I of the Schechter function over the luminosities $\log[L_{\text{Ly}\alpha}/(\text{erg s}^{-1})] \geq 41.8$. The results are shown in Table 2.6, where we find that the cumulative number of LAEs at close separation from the C IV absorbers is a factor ≈ 2.4 higher than the field. The MAGG ABS sample is associated to an intermediate LAE overdensity that is a factor ≈ 2.2 higher than the C IV sample. Finally, the normalization for the MAGG QSO sample is a factor ≈ 4.5 higher relative to the MAGG C IV sample and ≈ 2.0 compared to the MAGG ABS sample, pinpointing a richer overdensity of LAEs.

In conclusion, the analysis of the distribution of LAEs in velocity space and of the luminosity function in different environments underscores a systematic preference for LAEs to cluster around C IV absorbers with equivalent width $\gtrsim 0.05 \text{ \AA}$, implying in turn that C IV systems do not always trace average regions of the Universe but, often, regions near LAEs. Indeed, as we will show in detail below, $\approx 36\%$ of the total number of the absorbers is in fact associated to at least an LAE, with $\approx 48\%$ of these C IV systems being associated to multiple LAEs.

2.4.2 Correlation functions

Having established the presence of an overdensity of LAEs near C IV absorbers, we now quantify the connection between the galaxies and the ionized gas statistically, by deriving the LAEs auto-correlation and the LAEs-absorber cross-correlation functions. The main goal of this analysis is to establish whether the C IV absorbers are physically connected to the galaxies, and actually populate halos, or whether they are preferentially observed in IGM filaments. To do so, following the interpretation of Adelberger et al. (2003, 2005), we exploit the Cauchy-Schwarz relation that compares the galaxy-galaxy and the absorber-absorber auto-correlation functions, $\xi_{\text{GG}}(r)$ and $\xi_{\text{AA}}(r)$ respectively, with the galaxy-absorber cross-correlation function, $\xi_{\text{GA}}(r)$:

$$\xi_{\text{GA}}^2 \leq \xi_{\text{GG}} \xi_{\text{AA}}. \quad (2.2)$$

This relation can be used to compare the galaxies and absorbers underlying matter density distribution since the equality holds if the C IV absorbers and the galaxies originate from density fluctuations that are linearly dependent. A similar approach has been widely adopted in the literature to measure the clustering of the LBGs at redshift $z \sim 3$ (Adelberger et al., 2003, 2005;

Bielby et al., 2011) and establish the correlation of galaxies with various ions, such as H I and O VI at redshift $z < 1$ (Chen & Mulchaey, 2009; Tejos et al., 2014; Finn et al., 2016; Prochaska et al., 2019) or C IV at $z \sim 3$ (Adelberger et al., 2003, 2005). Furthermore, by relying on the k -estimator (Adelberger et al., 2005) that counts the number of pairs without requiring the assembly of a random sample as an alternative to the typical two-point correlation function, Diener et al. (2017) and Herrero Alonso et al. (2021) detected, for the LAEs auto-correlation, a clustering signal extended up to ≈ 3.0 pMpc in a large sample of emitters at $3.0 \lesssim z \lesssim 6.0$.

Formalism and random samples

The correlation functions are derived as a function of the pairs projected distance R and their separation along the line-of-sight R_{los} . The latter is computed as velocity separation and converted into a distance assuming pure Hubble flow. We use the Davis & Peebles (1983) estimator to compute both the auto-correlation and the cross-correlation functions, which compares the number of data-data pairs and random-random pairs at a given R and R_{los} .

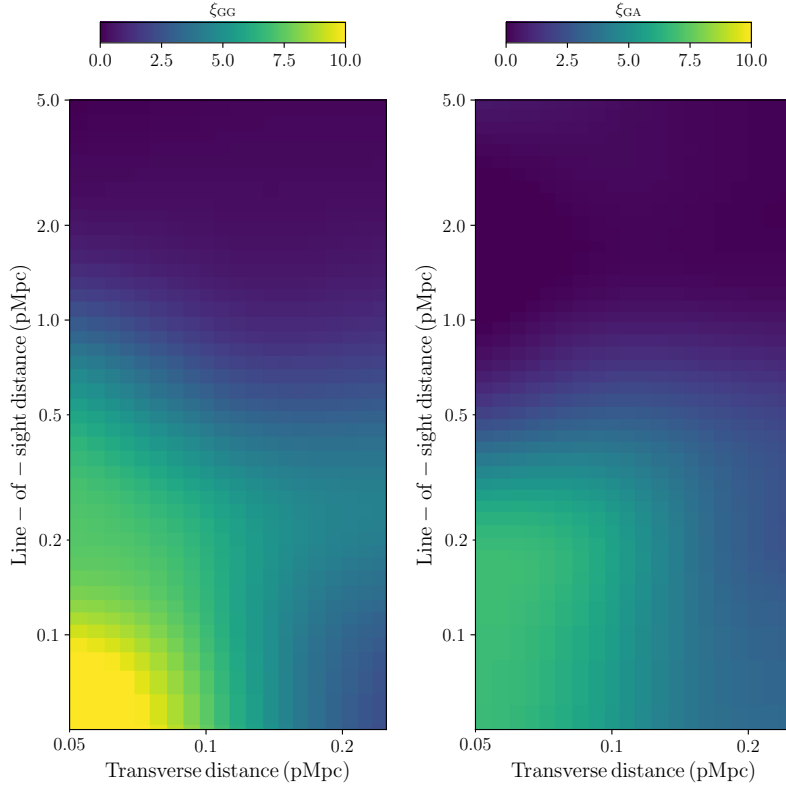


Figure 2.8: Two dimensional auto-correlation function for LAEs (left panel) and LAE-CIV cross-correlation function (right panel).

For the computation of the random-random pairs, we proceed as follows. In order to preserve the original geometry of the survey, we perform a random extraction of the galaxy coordinates weighted by the sum of the exposure maps of the 28 MAGG fields. The redshift of each mock galaxy is randomly extracted from a flat distribution limited to the range $3.0 \leq z \leq 4.5$. We require the size of the random sample to be a multiple of the number of real galaxies: $N_{\text{R,g}} = \beta_{\text{g}} \times N_{\text{D,g}}$. The factor $\beta_{\text{g}} = 3$ is chosen as it is sufficient to populate the coordinates space and optimize the Poissonian fluctuations of the random sample. We assemble the random sample of absorbers in a similar fashion. Lying on the line-of-sight, the three-dimensional position of each absorber is defined by the quasar coordinates and the redshift measured from the Voigt fit of the line. To preserve the geometry of the survey for the random sample of absorbers, we avoid the telluric bands and generate the redshift of the mock absorbers by randomly extracting values in the C IV redshift path. Given a number $N_{\text{D,a}}$ of observed C IV systems, we build a sample of $N_{\text{R,a}} = \beta_{\text{a}} \times N_{\text{D,a}}$ random absorbers, where $\beta_{\text{a}} = 5$ to compensate for the smaller sample size.

Galaxy auto-correlation

The galaxy auto-correlation function is derived by counting the number of real galaxies and random pairs in each bin in two dimension (transverse distance and line-of-sight separation). The intervals of transverse distances are logarithmically spaced between $0.05 - 0.50$ pMpc, where the lower bound accounts for the typical resolved size of the emitters and the upper bound is set by the field dimension. Along the line of sight, we consider pairs with a separation $5 \leq \Delta v / \text{km s}^{-1} \leq 2500$, corresponding to $0.02 \lesssim D_{\text{los}} / \text{Mpc} \lesssim 8.2$ at $z = 3$. To optimally extract the signal, the binned correlation function is oversampled and smoothed with a Gaussian filter of standard deviation $\sigma = 5$, corresponding to the oversampling factor. The result is shown in the left panel of Figure 2.8, up to a transverse distance $R = 0.25$ pMpc to make it comparable with the galaxy-absorber cross-correlation function. A clear clustering signal is detected at a transverse distance $R \lesssim 0.1$ pMpc and line-of-sight separation $R_{\text{los}} \lesssim 0.1$ pMpc as expected for galaxy populations that are not randomly distributed in the Universe. A weaker clustering signal is elongated up to $R_{\text{los}} \approx 0.5$ pMpc, likely due to the proper motions of the galaxies.

Comparison with the galaxy-absorber cross-correlation

We use the same procedure employed for the galaxy auto-correlation function to derive the LAE-C IV correlation function in two dimensions, as a function of the transverse distance and the line-of-sight separation. Since the C IV absorbers are detected along the sightline to the central quasar in each field, we choose the interval $0.05 - 0.25$ pMpc for the transverse separations, where the lower limit accounts for the quasar PSF and the upper limit corresponds to half the size of a field at $z \approx 3.0$. The line-of-sight separation between the LAEs and the C IV systems are derived by correcting the redshift of the galaxies for the observed offset from the absorbers, $\Delta v \approx 246 \text{ km s}^{-1}$. The cross-correlation function is shown in the right panel of Figure 2.8. At transverse distances $R \lesssim 0.1$ pMpc and line-of-sight separations $R_{\text{los}} \lesssim 0.5$ pMpc, that is at small separations from a C IV absorber, the probability to observe an LAE is enhanced with respect to a random point in the field, reinforcing the evidence that the LAEs are clustered to C IV absorbers on this scale. Both the galaxy auto-correlation and the galaxy-absorber cross-correlation functions are elongated in redshift space, possibly due to the galaxies proper motions and the gas proper motions relative to the galaxies, as it is typically observed in the literature (Adelberger et al., 2003, 2005; Turner et al., 2014; Herrero Alonso et al., 2021). In detail, Turner et al. (2014) performed a similar analysis measuring the optical depth as a tracer of metals in different ionization stages (e.g., H I, C IV O VI) around galaxies and found evidence of absorption enhancement extending up to ≈ 180 kpc in the transverse direction and on scales a factor ≈ 5 larger along the LOS.

A complete analysis involving the Cauchy-Schwarz relation would be possible by deriving the C IV absorbers auto-correlation function. In our survey, the absorption systems are observed along the line-of-sight of the central quasar of the MAGG fields which are separated by large transverse distances that exclude any possible physical correlation between the absorbers. Therefore, the C IV absorber auto-correlation function can only be computed in one-dimension as a function of the line-of-sight separation. The amplitude of the C IV absorbers auto-correlation function is almost flat with small fluctuations around $\xi_{\text{AA}} \approx 0.0 - 0.2$ at any line-of-sight separation. This suggests that the clustering of C IV absorbers may be too weak to be observed with our sample since it collects a high number of absorption systems, but spread across a large redshift range and along 28 different sightlines. For this reason, we can only derive a 3σ upper limit for the absorbers auto-correlation function that results in $\xi_{\text{AA}} \lesssim 2.0$ at any line-of-sight separation in the range $\Delta v = 200 - 2500 \text{ km s}^{-1}$ (the lower limit accounts for the rest frame separation of the C IV doublets and the median width of the lines), corresponding to $R_{\text{los}} \approx 0.6 - 7.0$ pMpc.

The comparison between the LAEs auto-correlation with the galaxy-absorber cross-correlation function can be used to test whether galaxies and C IV systems trace the same underlying matter distribution. Previously, Adelberger et al. (2003, 2005) detected metals around galaxies up to a transverse distance of ≈ 300 pkpc and found evidence of a clustering signal of C IV absorption systems near LBGs from the analysis of the galaxy-absorber cross-correlation function at $z \sim 3$. They used the analogies in the galaxy auto-correlation and the galaxy-absorber cross-correlation function as a hint to conclude that the galaxies and the absorption systems show a tight correlation and likely arise in the same regions of the Universe. However, a comparison between the LAEs auto-correlation and the galaxy-absorber cross-correlation suggests that the two functions do not show identical shapes nor the same amplitude in MAGG. Based on this result, we do not have direct evidence in support of the fact that LAEs and the C IV absorbers selected by MAGG are

exclusively tracing the same regions of the Universe. It is of course possible that the C IV auto-correlation function at small line-of-sight separations combines with the galaxy auto-correlation function to yield the Cauchy-Schwarz equality, but we are fundamentally limited on measuring the C IV auto-correlation below few hundreds of km s^{-1} . Therefore, from this analysis, we can conclude that a correlation between the two populations is evident from what is shown so far, but there is no evidence that would exclude that at least a fraction of the C IV systems arise beyond the halos of LAEs, from other regions such as the IGM filaments.

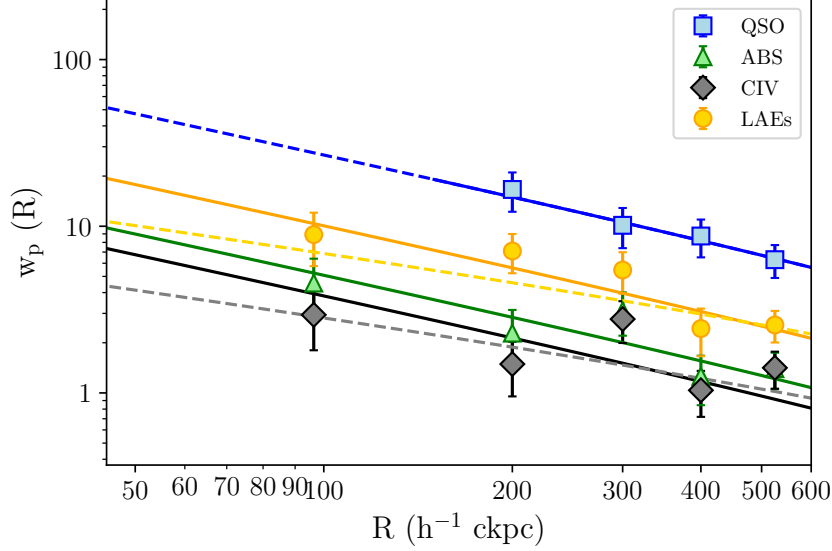


Figure 2.9: Reduced angular cross-correlation function for the LAEs associated to the C IV absorbers (black diamonds) with $\gamma = 1.8$ (solid black line) and $\gamma = 1.5$ (dotted black line) at $S/N > 7$. We compare this result with the ABS sample from MAGG IV (green triangles) and the QSO sample from Fossati et al. (2021) (blue squares), both of which are computed for $S/N > 5$ and assuming $\gamma = 1.8$. The LAEs auto-correlation function (gold circles) is obtained by integrating the 2-dimensional function along the line-of-sight direction.

Radial extent of the LAE overdensity near C iv

As a complement to the analysis of the luminosity function, we now study the LAE-C IV clustering signal in terms of the reduced angular cross-correlation function, also comparing with different tracers of the LAE environment (C IV and H I absorbers, and quasars). As found before, the two-dimensional cross-correlation function is elongated in the line-of-sight direction (as seen in the right panel of Figure 2.8) due to the uncertainties in the redshift measurement via $\text{Ly}\alpha$ combined with the underlying peculiar velocities. Therefore, the line-of-sight velocities are not an ideal measure of the radial distance for the galaxy-absorber cross-correlation. We thus focus next on the angular cross-correlation function, $\omega_p(R)$, as a function of the transverse distance R by integrating the original three-dimensional cross-correlation $\xi(r) = (r/r_0)^\gamma$ over a small redshift window. To do so, we follow the method described in Trainor & Steidel (2012), also adding the results for the correlation functions derived in MAGG using the central quasars (Fossati et al., 2021) and the LLSs (MAGG IV) as tracers. The full description of this analysis is detailed in Fossati et al. (2021) and only briefly summarized below.

Our estimate of the angular cross-correlation function in the data is, consistent with what is adopted in other papers of the MAGG series, defined as the number of observed LAEs, N_{obs} , in excess with respect to the predicted number, N_{pred} :

$$\omega_p(R) = \frac{N_{\text{obs}}}{N_{\text{pred}}} - 1 \quad (2.3)$$

The estimator in Equation 2.3 requires a measure of the comoving mean density of LAEs in the fields. To derive this value, we first compute the number of LAEs within $\pm 500 \text{ km s}^{-1}$ from a random redshift obtained from shuffling the real redshifts of the C IV absorbers observed in other fields. We then divide this number by the comoving area of the survey, assuming each field to

Sample	r_0 (h^{-1} cMpc)	γ
QSO	$3.58^{+0.34}_{-0.39}$	1.8
LAE	$2.07^{+0.23}_{-0.24}$	1.8
	$2.47^{+0.38}_{-0.40}$	1.5
ABS	$1.42^{+0.23}_{-0.27}$	1.8
C IV	$1.23^{+0.25}_{-0.27}$	1.8
	$1.39^{+0.32}_{-0.36}$	1.5

Table 2.7: Parameters of the best power-law function that reproduces the reduced angular cross-correlation function for the tracers studied in MAGG.

be modeled as a square with 0.98 arcmin on a side. The measure and the relative uncertainty of $n_{\text{pred}} = (35.16 \pm 6.94) h^{-2} \text{ cMpc}^{-2}$ is derived by bootstrapping the C IV redshift with random extractions over 10^3 repetitions and by taking the 16th and 84th percentiles. Finally, the expected number of galaxies in each interval of distance is derived by multiplying the mean density of galaxies in the fields, n_{pred} by the comoving area of the circular annulus corresponding to the i -th radial bin: $N_{\text{pred}} = n_{\text{pred}} \pi (R_{i+1}^2 - R_i^2)$.

The results are shown as black dots in Figure 2.9 with the respective uncertainties and summarized in Table 2.7. Due to the central quasar PSF contamination, we mask the region at distances $R < 2 \text{ arcsec}$ corresponding to $R \lesssim 43 h^{-1} \text{ ckpc}$ at redshift $z \approx 3.0$. The solid and dashed black lines are a fit to the projected correlation function with free parameter r_0 . Due to the degeneracy of the two parameters r_0 and γ , we fix the slope of the power law, and study the implication of this assumption by choosing two different values $\gamma = 1.8$ (Hennawi et al., 2006; Bielby et al., 2011; Diener et al., 2017; García-Vergara et al., 2017, 2019) and $\gamma = 1.5$ (Trainor & Steidel, 2012). Consistently with what Trainor & Steidel (2012) pointed out, we observe that the correlation length decreases with increasing slope of the power law from $r_0 = 1.39^{+0.32}_{-0.36} h^{-1} \text{ cMpc}$ for $\gamma = 1.5$ to $r_0 = 1.23^{+0.25}_{-0.27} h^{-1} \text{ cMpc}$ for $\gamma = 1.8$. Despite the model suggesting that a positive correlation extends towards the edge of the MUSE field of view, at distances $R \gtrsim 400 h^{-1} \text{ ckpc}$ the larger uncertainties make the number of LAEs at larger separation from the C IV absorbers consistent with expected field number.

The observed LAEs overdensity at small distances from the C IV absorbers is compared in Figure 2.9 with the excess of LAEs observed in the surroundings of the central quasars of each field (Fossati et al., 2021) and the LLSs from MAGG IV. Differently from the connection criteria adopted by Fossati et al. (2021), in order to perform a consistent comparison, we restrict the QSO sample to the LAEs at $S/N > 5$ observed at line-of-sight separations within $\pm 500 \text{ km s}^{-1}$ from the quasars (which is half the window in the original analysis). To do this, we correct the LAE mean density from Fossati et al. (2021) as $n_{\text{LAE, pres}}^{\text{QSO}} = 0.5 (17.15 \pm 2.28) = (8.58 \pm 1.14) h^{-2} \text{ Mpc}^{-2}$. For the C IV and H I MAGG samples, the region at distances $R < 50 h^{-1} \text{ ckpc}$ is masked due to contamination of the central quasars PSF. The only exception is represented by the MAGG QSO sample, for which the masked region is extended to $R < 150 h^{-1} \text{ ckpc}$. In this case, the behaviour of the correlation function on smaller scales is an extrapolation from the best-fit power law (dashed blue line in Figure 2.9). In the entire MUSE FoV, the correlation length with respect to quasars is $r_0 = 3.58^{+0.34}_{-0.39} h^{-1} \text{ cMpc}$, which is higher than it is around the LLSs, $r_0 = 1.42^{+0.23}_{-0.27} h^{-1} \text{ cMpc}$, and the C IV absorbers, $r_0 = 1.23^{+0.25}_{-0.27} h^{-1} \text{ cMpc}$.

A similar analysis has been performed to study the radial profile of C IV absorption around LBGs. Adelberger et al. (2003, 2005) selected C IV absorbers with column densities $N_{\text{C IV}} \gtrsim 10^{12.5} \text{ cm}^{-2}$ (corresponding to a rest-frame equivalent width $W_r \approx 0.01 \text{ \AA}$, of the same order of magnitude as the 50% completeness limit of this work sample) and derived a C IV-LBGs cross-correlation length that is a factor ≈ 2.2 larger than what we measured for the LAEs in our sample, assuming a power-law slope $\gamma = 1.8$. Adopting a different approach, Turner et al. (2014) produced 2D optical depth map showing that metal absorption is enhanced at small transverse separations from $z \approx 2 - 4$ galaxies. A significant excess of C IV absorption is also observed near LAEs at higher redshift $z \gtrsim 3$. Muzahid et al. (2021) recovered C IV optical depth profiles by stacking lines in the spectra of 8 bright quasars and found a significant enhancement of C IV absorption within $\approx 500 \text{ km s}^{-1}$ and $\approx 250 \text{ kpc}$ from 96 LAEs at $z \approx 3.3$.

To complete the analysis, we derive the projected LAEs auto-correlation function by integrating the two-dimensional function shown in Figure 2.8 along the line-of-sight direction up to a separation of 500 km s^{-1} . Uncertainties are derived by applying a bootstrap procedure over the assembly

of the random galaxy sample. Following the same strategy we applied for the other MAGG tracers, we measure the correlation length from a power-law modelling of the projected function for fixed slopes. The resulting LAEs correlation length $r_0 = 2.07^{+0.23}_{-0.24} h^{-1} \text{ cMpc}$ for $\gamma = 1.8$ (and $r_0 = 2.47^{+0.38}_{-0.40} h^{-1} \text{ cMpc}$ for $\gamma = 1.5$) is a factor ≈ 1.5 and ≈ 1.7 larger compared to that obtained for the H I and C IV sampled, respectively and ≈ 1.7 times lower compared to that obtained for LAEs clustering around quasars. Our estimate of the LAEs correlation length is found to be consistent, within the uncertainties, with the values measured in the literature for LAEs at redshift $z \sim 3$ (corrected for the cosmology adopted in this paper): $r_0 = 2.5^{+0.6}_{-0.7} h^{-1} \text{ cMpc}$ (Gawiser et al., 2007), $r_0 = 1.78^{+0.41}_{-0.48} h^{-1} \text{ cMpc}$ (Ouchi et al., 2010) and $r_0 = 2.99^{+0.35}_{-0.35} h^{-1} \text{ cMpc}$ (Bielby et al., 2016). We also observe that our result is consistent with the correlation length measured by Herrero Alonso et al. (2021) for a sample of ≈ 700 LAEs at redshift $3 < z < 6$. At first, they optimized the k -estimator to measure the LAE clustering for a survey with large redshift range, but limited angular coverage, and measured $r_0 = 3.76^{+3.24}_{-0.94} h^{-1} \text{ cMpc}$ and $\gamma = 1.3^{+0.36}_{-0.45}$. However, employing a more traditional method, they modelled the projected two-point correlation function with a power law and measured $r_0 = 2.34^{+0.26}_{-0.37} h^{-1} \text{ cMpc}$ with a slope $\gamma = 1.85^{+0.25}_{-0.25}$ that appears to fully agree with our findings.

Altogether, the analysis of the correlation functions, of the luminosity function and of the excess of LAEs in velocity space, provide firm evidence that LAEs cluster differently around the various tracers found across the entire MUSE FoV. The strongest signal is found near quasars and the amplitude of the clustering decreases progressively for LLSs and then for the C IV absorbers. While LLSs and C IV absorbers show a comparable level of clustering as expected for an overlapping population (43/220 of the C IV selected in this work are in fact associated with the LLSs from MAGG IV), there is a sufficient evidence for an excess of LAEs around LLSs compared to C IV, suggesting that high-column density H I absorbers are more prominently associated to the (outer) CGM of halos compared to the C IV absorbers with $\gtrsim 0.08 \text{ \AA}$, which are more likely to also trace gas at larger distances from LAEs.

2.4.3 Covering fraction of ionized gas

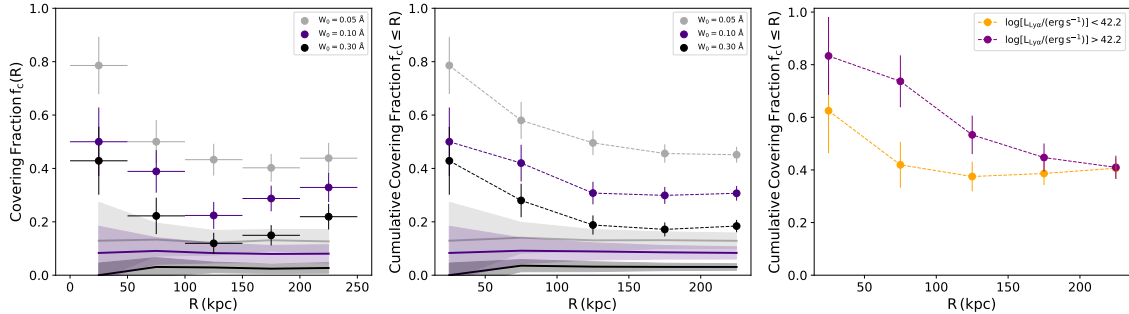


Figure 2.10: Differential (first panel) and cumulative (second panel) covering fraction, $f_c(R)$, as a function of the projected distance from the line-of-sight assuming three different equivalent width limits (see legend). The vertical error bars reproduce the 1σ Wilson-score confidence interval, while horizontal error bars in the left panel mark the width of each radial bin. The covering fraction is $\approx \times 4$ higher compared to what we obtained for LAEs around C IV absorbers taken from a different field (solid lines mark the median bootstrap value, while the shaded area show the 1σ interval derived from the 16th and 84th percentiles of the bootstrap distribution). The third panel shows the cumulative covering fraction dividing the LAEs in two bins of Ly α luminosity, assuming an equivalent width threshold $W_0 = 0.05 \text{ \AA}$.

In the previous sections, we explored the existence of a physical association between the C IV absorbers and the LAEs using an *absorber-centered* approach and searching for galaxies in their surroundings. Carrying on the analysis of the connection between C IV absorbers and LAEs with a statistical method, we now turn to a *galaxy-centered* point of view. To avoid proximity effects, we masked all the LAEs observed within a LOS separation $|\Delta v| \leq 3000 \text{ km s}^{-1}$ from the redshift of the central quasar of each field. We thus derive the C IV covering fraction as a function of the transverse distance in order to measure the observed occurrence of C IV absorbers around the LAEs. The covering fraction is defined as the number of LAEs connected to a C IV absorber (according to the above definition, i.e. within a 500 km s^{-1} window centred on the peak seen in Figure 2.6) with rest-frame equivalent width above a given threshold $W_r^{1548} \geq W_0$, relative to the total number of

LAEs lying in the C IV redshift path:

$$f_c(R) = \frac{N_{\text{obs}}(R; W_r^{1548} \geq W_0)}{N_{\text{tot}}(R)}. \quad (2.4)$$

In Figure 2.10 we show the differential covering fraction derived in radial annuli (first panel), and thus including LAEs with transverse distances in the i -th interval $R_i < R < R_{i+1}$, and the cumulative covering fraction, at radial separations $R < R_{i+1}$ (second panel). The vertical error bars account for the $1\sigma = 68\%$ Wilson-score confidence intervals, while in the left panel the horizontal bars reproduce the width of each radial interval. The covering fraction is computed for three rest-frame equivalent width thresholds (0.05, 0.10, 0.30) Å. As expected from the frequency distribution function, the C IV covering fraction decreases with increasing threshold from $f_c \approx 80\%$ with $W_0 = 0.05$ Å, to $f_c \approx 50\%$ with $W_0 = 0.10$ Å and $f_c \approx 35\%$ with $W_0 = 0.30$ Å at $R \lesssim 50$ kpc. For each C IV equivalent width threshold, the covering fraction decreases with increasing transverse separation between LAEs and the connected absorbers from $f_c \approx 60\%$ at $R < 100$ kpc with $W_0 = 0.05$ Å to $f_c \approx 50\%$ at $R > 100$ kpc. For the strongest C IV absorbers in the sample, assuming a threshold $W_0 = 0.30$ Å, the covering fraction is smaller at any transverse separation and decreases from $f_c \approx 30\%$ at $R < 100$ kpc to $f_c \approx 20\%$ at $R > 100$ kpc. We also observe that the differential covering fraction (first panel in Figure 2.10) increases in outer radial interval ($R > 200$ kpc). Although it is not statistically significant, Dutta et al. (2020) observed a similar trend studying Mg II absorbers at lower redshift and explained it as possibly due to the superposition of individual galaxy halos in case of multiple galaxies associated with the same C IV system. In the end, the observed trend suggests that the probability to observe a C IV within $\pm 500 \text{ km s}^{-1}$ of an LAE is enhanced at small transverse separations and decreases at higher distances roughly up to $R \approx 100$ kpc before flattening. Hence, C IV is present around LAEs on scales that extend well beyond the virial radius, which typically measures $R_{\text{vir}} \approx 25 - 53$ kpc for halo mass of the order of $M_h \approx 10^{10} - 10^{11} M_\odot$ (Herrero Alonso et al., 2021) respectively. To further support these results, we derive the covering fraction of a control sample assembled as described in 2.4.1, i.e. connecting LAEs to C IV absorbers selected from a different field (see Figure 2.10). We found a flat covering fraction at any transverse separation which is, on average, a factor of ≈ 4 lower compared to the one of LAEs around real C IV systems for any equivalent width threshold.

Similar trends in the covering fraction as a function of the transverse separation and the absorption strength are observed in the literature (Bordoloi et al., 2014; Turner et al., 2014; Burchett et al., 2016; Rudie et al., 2019; Dutta et al., 2021). In particular, Bordoloi et al. (2014) derived the covering fraction at $z < 0.1$ for C IV absorbers detected up to ≈ 100 kpc. They measure, at transverse separations of $R < 50$ kpc, $f_c \approx 73\%$ for C IV absorber with $W_r > 0.1$ Å and $f_c \approx 36\%$ for $W_r > 0.3$ Å. Both these results are consistent within 1σ with our findings at $R < 50$ kpc for the same equivalent width thresholds. At higher redshift, $z \approx 2.0$, Rudie et al. (2019) derived the covering fraction by computing the fraction of LBGs with a C IV system within line-of-sight separation up to 1000 km s^{-1} and within a transverse separation 100 kpc. The covering fraction decreases from $f_c \approx 50\%$ to $f_c \approx 33\%$ for increasing column density threshold from $\log(N_0/\text{cm}^{-2}) = 13.0^2$. This result is consistent at 1σ with the fraction of LAEs for which we detect C IV absorption above the threshold $W_0 = 0.05$ Å within a similar transverse separation $R < 100$ kpc. Using a large sample of galaxies up to redshift $z \approx 1.5 - 1.6$, Dutta et al. (2021) measured a C IV covering factor beyond 200 kpc of around 40% for $W_0 = 0.03$ Å which is just 25% below our determination for $W_0 = 0.05$ Å at comparable distances. Thus, comparing with our MAGG survey, the analysis of the C IV covering fraction over ≈ 10 Gyr of cosmic time could potentially imply small evolution with redshift, which appears to be driven by the lowest equivalent width systems. In order to study whether the incidence of C IV absorbers depends on the properties of the galaxies, we derive the cumulative covering fraction for a threshold of $W_0 = 0.05$ Å in two bins of Ly α luminosity above and below $\log(L_{\text{Ly}\alpha,0}/\text{erg s}^{-1}) = 42.2$ (see the third panel in Figure 2.10). We find that the fraction of C IV within $\pm 500 \text{ km s}^{-1}$ of LAEs is enhanced up to $f_c \approx 80\%$ for luminous galaxies relative to $f_c \approx 63\%$ for the fainter LAEs within transverse separation $R < 50$ kpc. The covering fraction decreases to $f_c \approx 75\%$ and $f_c \approx 41\%$ at higher separations $R < 100$ kpc for the two galaxies population, respectively, before flattening at $f_c \lesssim 50\%$ and $f_c \lesssim 40\%$ for distances $R > 100$ kpc. A larger covering factor around LAEs with higher SFR (assuming Ly α as a proxy of star formation activity) is in line with what is observed for Mg II at $z \approx 1$ by Dutta et al. (2020) in MAGG and by Dutta et al. (2021) at $z < 2$ in QSAGE survey. As for lower redshift, however, it is difficult to discern whether this difference arises because of the different star formation or whether

² $W_0 = 0.05$ Å corresponds to approximately $\log(N_0/\text{cm}^{-2}) = 13.5$.

– assuming LAEs lie on the star formation main sequence – it is a mere reflection of the different size of the halos probed as a function of Ly α luminosity. Additional information, and in particular independent mass estimates from IR observations, are required to investigate this trend further.

2.4.4 Detailed analysis of the C IV-LAE associations

Since the link between the C IV absorption-line systems and LAEs encodes useful information about the enrichment of the CGM, we now explore the possible correlations between the properties of the ionized gas and those of the galaxies. The upper panels in Figure 2.11 show the absorbers’ rest-frame equivalent width and velocity width as a function of the Ly α luminosity of the associated galaxies. To visualize any difference, we divide the LAEs in two sub-samples based on a threshold Ly α luminosity ($L_{\text{Ly}\alpha} = 10^{42.2} \text{ erg s}^{-1}$) and we compare the resulting distributions. For those C IV absorbers that are not connected to any LAEs within $\pm 500 \text{ km s}^{-1}$, we take an upper-limit corresponding to the Ly α luminosity at which our search is 10% complete. The brightest LAE of each group is considered in the analysis.

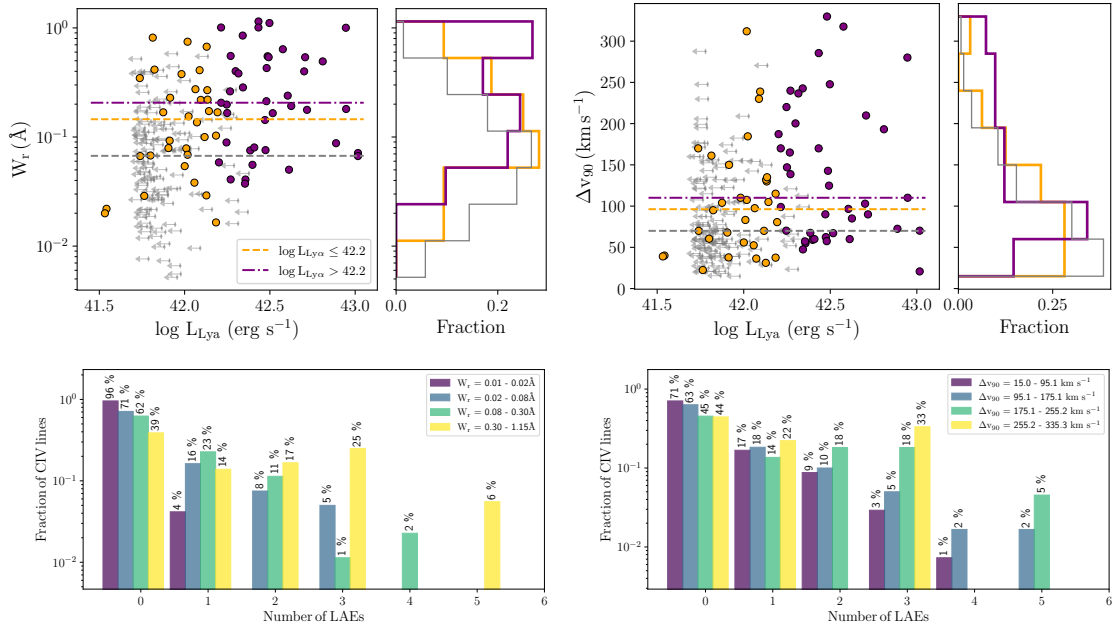


Figure 2.11: Upper panels: C IV absorbers’ equivalent width (left) and velocity width (right) as a function of the Ly α luminosity of the galaxies found within $\pm 500 \text{ km s}^{-1}$. For easier comparison between a bright and faint subset, we divide the sample below (orange) and above (purple) a threshold of $L_{\text{Ly}\alpha} = 10^{42.2} \text{ erg s}^{-1}$. Dashed lines mark the median of each distribution, shown as histograms along the right hand-side axis. The grey arrows mark instead the Ly α luminosity upper-limit for the C IV systems without any associated LAE. Lower panels: distribution of the equivalent width (left) and velocity width (right) as a function of the number of galaxies connected to each absorber.

In the upper-left panel in Figure 2.11 we observe that the distribution of the equivalent width of the C IV absorbers not connected to a galaxy reveals that a higher number of systems is weak (with $W_r \lesssim 0.05 \text{ Å}$) compared to those with LAEs detected, and drops off at larger equivalent widths. The equivalent width distribution of the absorbers connected to bright LAEs is more skewed toward stronger systems with $W_r \gtrsim 0.1 \text{ Å}$, suggesting that the strongest absorbers show a preference to be connected to luminous galaxies. Weak C IV systems with $W_r \lesssim 0.02 \text{ Å}$ are completely absent in this more luminous sub-sample. To quantify the statistical correlation between C IV equivalent width and Ly α luminosity of LAEs, we compute the Spearman’s correlation coefficient and its associated p -value, finding a probability of 0.042 that uncorrelated variables produce a trend similar to the one observed in figure. We thus conclude that there is at least tentative evidence of a preference for brighter LAEs to be associated with higher equivalent width absorbers. A similar analysis, focusing instead on the C IV velocity width, is presented in the second panel of Figure 2.11. Here, comparing the two sub-samples, we find that the LAE non-detections are mostly limited to the narrowest C IV systems with $\Delta v_{90} \lesssim 100 \text{ km s}^{-1}$ with a sharp drop-off at $\Delta v_{90} \gtrsim 200 \text{ km s}^{-1}$. The distribution of the velocity width of the C IV absorbers connected to bright LAEs shows a tail extending at $\Delta v_{90} \gtrsim 150 \text{ km s}^{-1}$ relative to those associated to less luminous galaxies. The

p -value associated with the Spearman's correlation coefficient is 0.168 in this case, revealing a more marginal difference compared to the equivalent width case.

Finally, in the lower panels of Figure 2.11 we investigate the C IV absorbers equivalent width (left figure) and velocity width (right figure) as a function of the number of associated LAEs within $\pm 500 \text{ km s}^{-1}$. To do so, we divide the absorbers into 4 intervals of W_r and Δv_{90} . We derive the fraction of C IV absorbers connected to a certain number of LAEs and with equivalent-width (velocity width) in a certain range, over the total number of systems with W_r (Δv_{90}) in that interval. The bottom-left panel of Figure 2.11 supports the existence of a marked correlation between the strength of the C IV absorption and the number of connected galaxies. We do not detect any LAE at close separation from the $\approx 96\%$ of the weakest absorbers with $W_r \lesssim 0.02 \text{ \AA}$, while the remaining $\approx 4\%$ is connected to isolated galaxies. Conversely, $\approx 48\%$ of the strongest absorbers ($W_r \gtrsim 0.3 \text{ \AA}$) are found to be associated with ≥ 2 LAEs, with the few rich groups hosting 4 and 5 galaxies being found exclusively near high equivalent width C IV absorbers. The same analysis is then repeated for the absorbers' velocity width in the bottom right panel of Figure 2.11. In this case, we do not find a correlation that is as significant as above. Indeed, we do not observe any LAE in the vicinity of $\approx 71\%$ of the narrow systems with $\Delta v_{90} \lesssim 95 \text{ km s}^{-1}$, but $\approx 13\%$ of them are connected to ≥ 2 galaxies. However, $\approx 44\%$ of the broadest systems with $\Delta v_{90} \gtrsim 255 \text{ km s}^{-1}$ are not associated with any LAE, $\approx 22\%$ is found in proximity to 1 galaxy. Only $\approx 33\%$ of this sample is connected to 3 LAEs. A significant fraction of both narrow and wide absorbers, $\approx 19\%$ of the systems with $\Delta v_{90} \sim 95 - 175 \text{ km s}^{-1}$ and $\approx 41\%$ with $\Delta v_{90} \sim 175 - 255 \text{ km s}^{-1}$, is associated to ≥ 2 galaxies.

In conclusion, the C IV absorbers found within $\pm 500 \text{ km s}^{-1}$ from a bright galaxy appear on average stronger (and, to a lesser extent, more kinematically complex) compared to the systems connected to faint galaxies or without any LAE detected in the field-of-view. A high fraction ($\approx 44\%$) of the strongest absorbers with $W_r \gtrsim 0.30 \text{ \AA}$ is found to be connected to ≥ 2 galaxies and the number of associated LAEs decreases with decreasing equivalent width. The same analysis applied to the absorbers velocity width revealed no such clear correlation, with only the $\approx 33\%$ of broad systems with $\Delta v_{90} \gtrsim 255 \text{ km s}^{-1}$ connected to multiple galaxies.

Radial profile of the C IV absorption strength

To complete the picture of the distribution of the ionized gas around LAEs, we turn to investigate how the rest-frame equivalent width depends on the projected distance from the galaxy associated to each C IV absorber. The results for C IV are shown in the left panel of Figure 2.12.

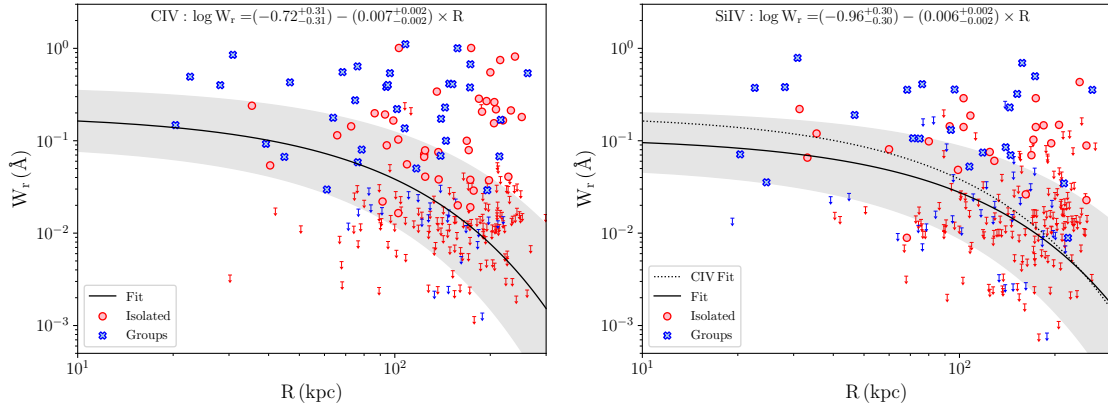


Figure 2.12: Rest-frame equivalent width as a function of the transverse separation from the associated galaxies for the C IV (left panel) and the Si IV (right panel) absorbers. The isolated LAEs are shown in red, while the closest galaxy of a group is shown in blue. For the galaxies without any associated absorber we derive a 3σ upper limit (arrows with the same color-coding). The observed relationship is fitted with a log-linear function taking upper limits into account. The best fit result is drawn as a solid black line and the 1σ uncertainties are marked by shaded regions. In the right panel, the best fit obtained for the Si IV absorbers (solid line) is compared to the best fit derived for the C IV (dotted line).

Many studies in the literature support the evidence of an anti-correlation between the rest-frame equivalent width of the C IV absorbers and the transverse separation from the associated galaxy. Most of these are focused on galaxies at low redshift $z \lesssim 1$ (Chen et al., 2001; Bordoloi et al., 2014; Burchett et al., 2016) and $z \sim 1 - 1.5$ (Dutta et al., 2021). Similar results are observed for LBGs at higher redshift $z \sim 2 - 3$ with a statistical approach based on spectral stacking and

optical depth analysis (Steidel et al., 2010; Turner et al., 2014). Motivated by these findings, we search for an anti-correlation following the procedure described by Dutta et al. (2020) and briefly summarized here. The rest-frame equivalent width is expected to decrease with increasing distance from the associated galaxies following a log-linear relation that models an expected steep decrease of the strength of the absorption regulated by a scale factor that is linked to the galaxy virial radius:

$$\log(W_r/\text{\AA}) = a + b \cdot (R/\text{kpc}). \quad (2.5)$$

For those LAEs that lie in C IV redshift path, but that are not connected to any C IV absorber within $\pm 500 \text{ km s}^{-1}$, we derive a 3σ upper limit (arrows in Figure 2.12). We fit the full sample with the function in Eq. 2.5 applying a Bayesian method based on a likelihood that takes both measurements and upper limits into account (see for more details and applications Chen et al., 2010a; Rubin et al., 2018; Dutta et al., 2020, 2021). The result is shown in Figure 2.12 as solid black line, where the shaded grey region marks the 1σ confidence interval. We notice that the modelling is strongly driven by the upper limits at large distances where a significant fraction of the data is scattered upward the relation and the equivalent width profile appears to be flatter.

Dutta et al. (2021) performed a similar analysis on 123 C IV absorption-line systems at lower redshift, $z \sim 0.1 - 2.4$. The best log-linear model resulting from their fit is shifted significantly upward compared to the one derived in this work and parametrized by $a = -0.42^{+0.25}_{-0.28}$ and $b = -0.002^{+0.001}_{-0.001}$. The QSAGE (Bielby et al., 2019; Stott et al., 2020) dataset explored by Dutta et al. (2021) results in a 3σ sensitivity limit of $W_r^{\text{C IV}} \approx 0.1 \text{\AA}$ for the detection of C IV absorbers, dominated by the measurements based on UV spectra with a lower resolution and S/N compared to the optical spectra available in MAGG. In this work, the upper limits at transverse distances $R \lesssim 100 \text{ kpc}$, that are absent in the sample from Dutta et al. (2021), follows from the higher resolution of the MAGG spectra and pushes down the fit towards lower equivalent width values. Moreover, the size of the field-of-view at low redshift is about a factor of ≈ 2 larger than it is at high redshift, making it possible to detect galaxies up to transverse separations $R \approx 750 \text{ kpc}$, far beyond the typical virial radius of a dark matter halo with $M_H \approx 10^{11} M_\odot$. Extending the analysis to larger distances lead to a fit that is flatter compared to the one at high redshift. Thus, the differences in the two datasets make it difficult to directly compare our findings with low redshift results and explain the discrepancies as due to a physical evolution with redshift. We note, however, that detections occupy a comparable region of parameter space, all clustering between $0.1 - 1 \text{\AA}$.

2.4.5 Dependence on galaxy environment

The ability of MUSE observations to obtain a complete survey inside the FoV reveals that a large fraction of the observed galaxies are part of a group (see Sect. 2.3.4 for the definition and identification of groups), allowing us to study the multiphase circumgalactic gas as a function of galaxy environment (see e.g. Bielby et al., 2017; Fossati et al., 2019; Muzahid et al., 2021; Dutta et al., 2021). In Figure 2.12, we distinguish between isolated galaxies (red) and groups (blue), considering only the closest galaxy of each group, and we place an equivalent width upper-limit at 3σ for the LAEs that are detected in the C IV redshift path, but are not associated to any absorber within $\pm 500 \text{ km s}^{-1}$ (arrows). An excess of galaxies in groups, relative to the number of isolated LAEs, is observed at small transverse separations for $R \lesssim 60 \text{ kpc}$. The observed C IV absorbers connected to groups show, on average, a rest-frame equivalent width that is ≈ 2.5 times larger than that associated to isolated galaxies at any separation, consistent with the trend shown in third panel in Figure 2.11. This difference is suppressed approaching the edges of the field-of-view at $R \gtrsim 200 \text{ kpc}$. Galaxies without C IV absorption detected within $\pm 500 \text{ km s}^{-1}$ are mostly found at large distances ($R \gtrsim 100 \text{ kpc}$) from the sightlines. We recall that our classification of isolated galaxies and groups is limited by the size of the FoV and thus it is not possible to know if an isolated galaxy, especially those connected to strong absorbers, does not have any companion outside the observed footprint.

Next, we derive the differential and cumulative covering fraction for the isolated LAEs (to the completeness limit of our survey) and the galaxies that are part of a group (i.e. with at least a companion within $\pm 500 \text{ km s}^{-1}$ and within the MUSE FoV). The results are shown Figure 2.13 for the full sample of galaxies in groups (first column) and for sub-samples assembled including only the closest or the brightest component of each group (second and third columns respectively). The cumulative covering fraction of isolated and group galaxies decreases with increasing transverse distance and flattens above $R > 100 \text{ kpc}$. We observe, however, that the fraction of galaxies in

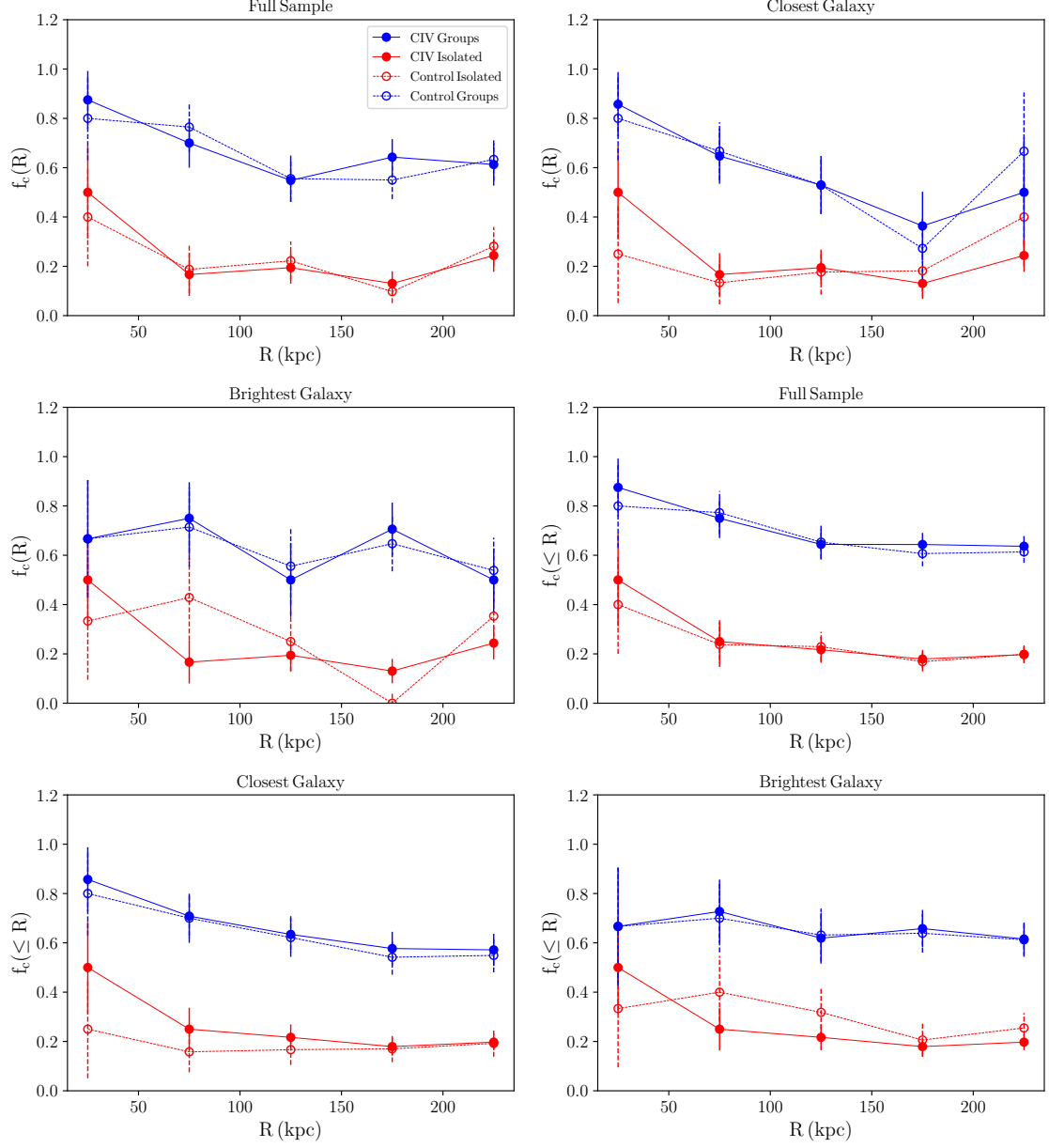


Figure 2.13: Differential (top row) and cumulative (bottom row) covering fraction of CIV around isolated galaxies and galaxies in groups that are connected to CIV absorbers within $\pm 500 \text{ km s}^{-1}$, as a function of the transverse separation. All the galaxies in groups are considered in the *Full Sample*. The analysis is reproduced by taking only the closest (second column) and the brightest (third column) component of each group. These measurements (solid lines) are compared to the covering fraction estimated for control samples (dashed lines) assembled by requiring the properties of isolated galaxies and groups to match.

Property	Isolated	Groups	Groups (R)	Groups (L)
LAEs in C IV redshift path				
R (kpc)	163.54	169.42	128.45	162.88
$\log[L_{\text{Ly}\alpha}/(\text{erg s}^{-1})]$	42.12	42.08	42.07	42.20
LAEs with C IV within $\pm 500 \text{ km s}^{-1}$				
R (kpc)	168.62	168.41	105.53	168.10
$\log[L_{\text{Ly}\alpha}/(\text{erg s}^{-1})]$	42.20	42.03	41.97	42.13

Table 2.8: Median properties of LAEs in the C IV redshift path (rows 1 – 2) and of LAEs with C IV absorbers detected within $\pm 500 \text{ km s}^{-1}$ (rows 3 – 4). Columns 4 – 5 are referred to the samples containing only the closest (R) and the brightest galaxy (L) of each group respectively.

groups with detected C IV absorption within $\pm 500 \text{ km s}^{-1}$ is on average a factor ≈ 3 higher relative to the fraction of isolated galaxies for any transverse separation up to $R \approx 250 \text{ kpc}$.

Our result is, at small separations, consistent with the findings from Dutta et al. (2021) at $z < 2$ on a sample of ≈ 750 galaxies detected in MAGG and QSAGE. Indeed, their analysis revealed that the covering fraction of C IV absorbers around group galaxies is a factor ≈ 2 higher relative to that of the isolated galaxies up to $R < 100 \text{ kpc}$. In their analysis, the difference is less pronounced than in our sample for larger impact parameters, where we observe a persistent difference up to the edge of the FoV. Exploiting the large spectral coverage of MAGG and QSAGE, they also compared the covering fraction of the intermediate ionized gas phase, traced by the C IV absorbers, with that of a low ionized phase traced by the Mg II absorption, both for groups and isolated galaxies, finding that the covering fraction is enhanced by a factor $\approx 2 - 3$ for the Mg II gas in group galaxies. A substantial difference is, however, observed once control samples are built for the isolated galaxies by requiring them to match the impact parameter, the stellar mass and the redshift of the galaxies in groups. For these control samples, the enhancement of the covering fraction in group galaxies relative to that of the isolated galaxies is preserved only for the Mg II absorption, but it is suppressed for the C IV gas.

In light of this result, we define a control sample to study in detail whether the excess in covering factor of group galaxies is an intrinsic property or depends on the properties of the sample. We build control-matched samples from groups by selecting the members of the groups with properties resembling those of the isolated galaxies. Specifically, we select group galaxies within $\pm 20 \text{ kpc}$ from the impact parameter and $\pm 0.15 \text{ dex}$ from the log of the Ly α luminosity of the isolated LAEs, without allowing for repetitions. The same strategy is then applied to assemble a control sample of isolated galaxies. The result is shown in Figure 2.13 (blue and red dashed line for groups and isolated galaxies, respectively). In order to test the quality of the control samples, we perform a KS test comparing the properties of the control samples to the ones of groups and isolated galaxies and measure a p -value > 0.95 for all the cases shown in Figure 2.13. In building these control samples, we notice (see Table 2.8) that isolated galaxies and groups do not differ significantly in their properties, but stronger differences are observed only if considering the closest or the brightest component of each group. In these cases, we assemble control samples of groups and isolated galaxies with the same strategy described above to reproduce the properties of the isolated sample and of the closest or the brightest component of each group respectively. The results are shown in the middle and right panels in Figure 2.13. The comparison to the covering fraction of group galaxies (blue-dashed line) and that of the isolated galaxies (red-dashed line) in matched samples show that the enhancement in groups is preserved for the control samples.

A similar study on the effect of environment was also conducted by Muzahid et al. (2021) who performed a statistical analysis on the strength of C IV absorption along the line of sight based on spectral stacking. In line with our findings, they report a stronger C IV absorption connected to galaxies in groups, relative to those observed in the vicinity of isolated galaxies, up to velocities of a few hundreds km/s . These authors interpret the observed excess as a consequence of galaxies being embedded in a large-scale structure which give rise to the C IV absorption, and not simply due to the properties of the two samples which appear comparable (see their table 4).

2.4.6 Comparison with Si IV as a tracer

The analysis presented in the previous sections has focused on C IV as tracer of moderately ionized and possibly warm gas. An alternative tracer of gas in the same ionization stage that is accessible for observation with MAGG spectroscopy is the Si IV doublet, which has a slightly lower but overall

comparable ionization potential (≈ 45.1 eV for Si IV and ≈ 64.5 eV for C IV).

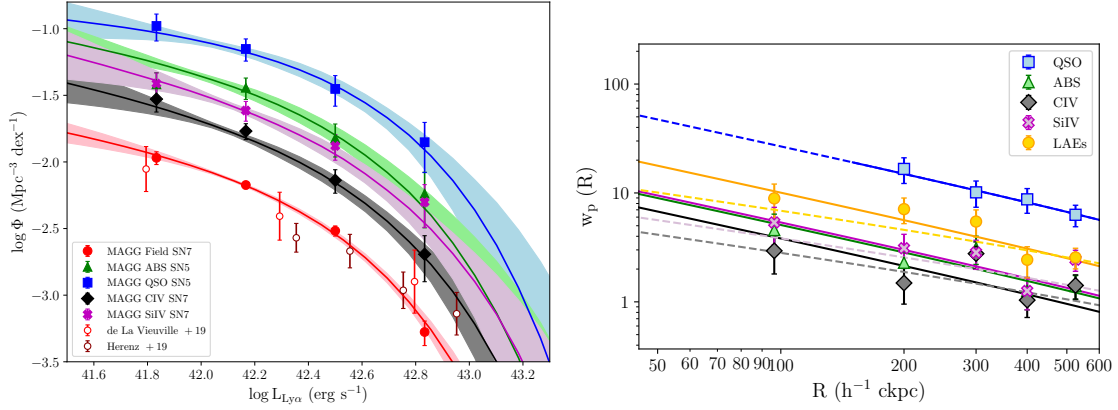


Figure 2.14: Measure of the amplitude and the radial profile of the LAEs overdensity observed within $\pm 500 \text{ km s}^{-1}$ from Si IV absorbers using the luminosity function (left panel, purple markers) and the projected cross-correlation (right panel, purple markers). Markers and color-coding are the same as Figures 2.7 and 2.10, respectively, for the other tracers studied in MAGG.

Making use of the detection of both C IV and Si IV absorption-line systems in MAGG, we now test explicitly whether differences in the correlation between Si IV and LAEs exist in comparison to C IV (arising, for instance, from differences in the nucleosynthesis or the shape of the ionizing spectrum), or whether the results presented above largely apply for Si IV as well. The latter scenario would support the hypothesis that C IV and Si IV absorbers are both tracers of the same gas phase in the CGM or IGM. To this end, we derive the galaxies luminosity and cross-correlation functions of LAEs and Si IV systems in comparison with what is found with C IV. We perform a blind search of LAEs within $\pm 500 \text{ km s}^{-1}$ from a Si IV absorber with the same procedure adopted for the C IV analysis. We find that 49 Si IV absorbers are connected to 86 LAEs, corresponding to a detection rate of 46 ± 8 per cent (49/108) that is a factor ≈ 1.4 higher compared to the C IV detection rate but still consistent within 1σ uncertainty. Once we identified the LAEs associated to Si IV absorbers, we compute the LAE luminosity function and the projected cross-correlation to measure the amplitude and the radial profile of the galaxy overdensity. The results are shown in Figure 2.14. We find that the luminosity function of LAEs around Si IV absorbers is modelled by a Schechter function whose best-fit parameters are the characteristic luminosity $\log[L^*/(\text{erg s}^{-1})] = 42.645 \pm 0.273$, slope $\alpha = -1.466 \pm 0.370$ and normalization $\log[\phi^*/(\text{Mpc}^{-3} \text{dex}^{-1})] = -2.058 \pm 0.431$. These estimates are all consistent with the C IV results within 1σ uncertainties. Furthermore, the projected cross-correlation function is modelled by a power law with correlation length $r_0 = 1.48^{+0.29}_{-0.32} h^{-1} \text{cMpc}$ assuming a slope $\gamma = 1.8$ and increases to $r_0 = 1.71^{+0.40}_{-0.42} h^{-1} \text{cMpc}$ for $\gamma = 1.5$. All these parameters are consistent within 1σ with the results derived for the C IV-LAE cross-correlation.

Complementary to this empirical analysis, we further support these findings by adopting a more theoretical approach, detailed in Appendix A.1, to show that C IV and Si IV absorbers likely arise from the same gas. In conclusion, this analysis provides evidence that the C IV and Si IV systems are comparable tracers of ionized gas, and that the results obtained in studying the galaxy overdensity around these absorption-line systems are not strongly sensitive to the tracer adopted

2.5 Discussion

2.5.1 Properties of C IV absorbers near galaxies as a function of mass and redshift

Comparison of the MAGG and MUSEQuBES samples

Following the pioneering survey by Adelberger et al. (2003, 2005), several studies have reported evidence of metal enrichment in the form of C IV in the CGM of star-forming galaxies at the mass scale traced by $z \approx 2 - 3$ LBGs ($M_H \approx 10^{12} M_\odot$; Steidel et al., 2010; Turner et al., 2014; Rudie et al., 2019). Only recently, this analysis has been extended to the lower-mass galaxy population traced by LAEs at $z \approx 3$, with $M_H \approx 10^{11} M_\odot$. Besides MAGG, the other major survey that studies the distribution of C IV near LAEs is MUSEQuBES (Muzahid et al., 2021). By observing

samples of LAEs at comparable redshift and sensitivity to our study, these authors followed a complementary approach to ours by stacking quasar spectra at the wavelength of H I and C IV near 96 LAEs at redshift $2.9 < z < 3.8$. Through this technique, they observe an excess of H I and C IV absorption within a line-of-sight separation $\Delta v \leq \pm 500 \text{ km s}^{-1}$, i.e., on comparable scales to the velocity clustering seen in MAGG (Figure 2.6). Consistently with the covering fraction shown in Figure 2.10, the excess of C IV absorption is extended at large projected distances ($R \geq 250 \text{ kpc}$), corresponding to $R \approx 7 R_{\text{vir}}$ for halo masses $M_{\text{H}} \approx 10^{11} M_{\odot}$ with virial radius $R_{\text{vir}} \approx 35 \text{ kpc}$. However, through stacking, Muzahid et al. (2021) report no evident difference in the line-of-sight optical depth at different impact parameters, while our detailed analysis of the equivalent width as a function of impact parameter suggests a mild decline in the absorption strength as one moves away from galaxies. Compared to the results by Muzahid et al. (2021), MAGG data reveal a higher fraction of stronger absorbers near brighter LAEs. This observation hints at a correlation between SFR and C IV equivalent width, that is only observed in MUSEQuBES in H I. We note, however, that Muzahid et al. (2021) rely on rest-frame UV continuum emission for measuring SFRs, while we use Ly α emission.

These authors also report that the C IV absorbing gas shows a strong dependence as a function of the different galaxy environments connected to the absorbers. In their sample, $\approx 33\%$ of the LAEs are in groups (i.e., not isolated within $\pm 500 \text{ km s}^{-1}$) and this fraction is even higher ($\approx 44\%$) in the MAGG sample. They found that both H I and C IV absorption are enhanced around galaxies in groups with respect to the isolated LAEs. According to our estimate of the covering fractions (Figure 2.13), the C IV absorbers are preferentially detected around groups rather than isolated galaxies, at least up to the edges of the field of view. Furthermore, they observe that the H I absorption is stronger and broader in groups, but the weakness of the C IV absorption prevents from deriving a significant difference between isolated galaxies and groups. The larger size of the MAGG sample allows us to detect, on average, ≈ 2.5 times stronger and ≈ 1.5 times broader C IV absorbers around groups compared to those around isolated galaxies, suggesting that the galaxy environment does also affect the strength of the absorption and, to a lesser extent, the kinematics of the gas.

In summary, the study of both the MAGG and MUSEQuBES samples lead to mostly similar conclusions with independent data and different analysis techniques: there is clear clustering of LAEs and C IV absorbers on scales up to $\pm 500 \text{ km s}^{-1}$ along the line of sight and $\approx 200 \text{ kpc}$ in the transfer direction, and the environment appears to modulate this signal. Owing to the study of individual associations, MAGG adds further insight into this analysis, unveiling stronger/wider C IV absorption near brighter LAEs, and a correspondence between the strongest absorbers and the presence of groups. Finally, by studying the LAE population near absorbers (and not the strength of absorption near LAEs), we confirm that only a small fraction of C IV is in fact associated to galaxies ($\approx 36 \pm 5 \text{ per cent}$) and this fraction increases with the strength of the absorption.

Comparison of MAGG galaxies and $z \approx 2 - 3$ LBGs

Many studies in literature have shown there is a physical connection between continuum detected galaxies with the surrounding ionized gas detected in absorption in background quasars or galaxy pairs (Turner et al., 2014; Bordoloi et al., 2014; Burchett et al., 2016; D’Odorico et al., 2016; Rudie et al., 2019; Schroetter et al., 2021). Compared to LBGs, our study extends this investigation to LAEs at $z \approx 3 - 4$ in lower-mass halos ($M_{\text{H}} \approx 10^{11} M_{\odot}$). We now investigate analogies or differences between these two galaxy populations and the properties of the ionized and enriched gas around them. Considering first galaxies alone, Bielby et al. (2016) offer a direct comparison between LAEs and LBGs using data from the VLT Lyman break galaxies (LBG) redshift survey. From their measurement of the LAEs correlation length, $r_0 = (2.99 \pm 0.35) h^{-1} \text{ cMpc}$, they derive a halo mass $M_{\text{H}} = 10^{11 \pm 0.3} M_{\odot}$, which is lower compared to the typical values measured for LBGs, $r_0 \approx 4.18 h^{-1} \text{ cMpc}$ and $M_{\text{H}} \approx 10^{12} M_{\odot}$ (Adelberger et al., 2005; Bielby et al., 2011). However, the correlation length of the two population are found to be consistent at the fainter magnitudes, suggesting that LAEs inhabit on average low-mass haloes but that the two populations are overlapping (see, e.g., Steidel et al., 2011). Thus, statistics of C IV near LAEs and LBGs are not independent.

Turning to the association between C IV and LBGs, Adelberger et al. (2003, 2005) studied the spatial distribution of the ionized gas by means of the galaxy auto-correlation, ξ_{gg} , and the galaxy-absorbers cross-correlation, ξ_{ga} , functions. Modeling the two-point correlation functions as a power law with a slope $\gamma_{\text{ga}} \approx 1.6$, their measurement of the cross-correlation length $r_{\text{ga}} \approx 3.34 h^{-1} \text{ cMpc}$ is a factor ≈ 2.4 higher compared to the value $r_0 \approx 1.39 h^{-1} \text{ cMpc}$ estimated for LAEs, assuming

a fixed slope $\gamma = 1.5$, and a factor ≈ 2.7 higher compared to the estimate $r_0 \approx 1.23 \text{ h}^{-1} \text{ cMpc}$ with a fixed slope $\gamma = 1.8$ (Figure 2.9). They also found that the shape of the galaxy-absorber cross-correlation is consistent with the galaxy auto-correlation down to C IV systems with column density $N_{\text{C IV}} \gtrsim 10^{12.5} \text{ cm}^{-2}$ suggesting that LBGs and C IV absorbers inhabit similar regions of the Universe. However, this is not completely the case for the sample of LAEs and C IV systems studied in this work. As mentioned above, our results (Figure 2.8) show that the galaxy auto-correlation and the galaxy-absorbers two dimensional cross-correlation functions have different shapes and amplitudes, suggesting that the LAEs and the ionized C IV gas are unlikely to trace exclusively the same underlying matter distribution and thus, do not always inhabit the same regions. The enhancement of C IV absorption around LBGs is also observed by Turner et al. (2014), who adopted the optical depth method to measure the absorption strength and produce 2D maps of the average distribution of metal absorption around these galaxies. These maps reveal a strong enhancement of metal absorption at small transverse separations from the galaxies at $R \lesssim 180 \text{ kpc}$, which is elongated in the LOS direction up to $\Delta v \lesssim \pm 240 \text{ km s}^{-1}$ ($\approx 1 \text{ Mpc}$) due to the gas peculiar motions.

The analysis performed on larger scales suggests that the C IV absorbing gas is typically extended far beyond the virial radius of LBGs, generalizing the finding of Rudie et al. (2019) who measured the covering fraction of strong ($N \geq 10^{13.5} \text{ cm}^{-2}$) C IV systems to reach a value $f_c \approx 50\%$ up to $R \approx 100 \text{ kpc}$, that is $\approx R_{\text{vir}}$. Indeed, Turner et al. (2014) detected C IV absorption up to the survey limit ($\approx 2 \text{ Mpc}$) in the transverse direction, corresponding to $\approx 20 R_{\text{vir}}$ for an average halo mass $M_{\text{H}} \approx 10^{12} M_{\odot}$. The C IV absorbing gas shows a similar tendency to be extended far beyond the typical virial radius ($R_{\text{vir}} \approx 25 - 35 \text{ kpc}$ for halo mass $M_{\text{H}} \approx 10^{10} - 10^{11} M_{\odot}$ respectively) also around LAEs. The covering fraction in Figure 2.10 shows that $\approx 50\%$ of the LAEs are connected to C IV absorbers within $\pm 500 \text{ km s}^{-1}$ at $R > 100 \text{ kpc}$, reaching as far as the survey limit of $R \approx 250 \text{ kpc}$ which corresponds to $\approx 10 R_{\text{vir}}$. As the separation between the ionized and enriched gas and the associated LBGs approaches the virial radius, the strength of the absorption is commonly observed to decrease monotonically with increasing distance. The spectral stack performed by Steidel et al. (2010) shows a steep decrease of the rest-frame equivalent width of strong absorbers (with $W_r \geq 0.1 \text{ \AA}$) up to the galaxy virial radius, which can be extrapolated leading to a sharp drop-off at larger distances. Considering weaker absorption, Turner et al. (2014) showed that the equivalent width decreases more slowly with the transverse separation relative to the predictions from Steidel et al. (2010) and flattens around $W_r \approx 0.1 \text{ \AA}$ at $R > 100 \text{ kpc}$, up to the edges of their survey.

Turning to the LAEs population at $z \sim 3$, Muzahid et al. (2021) observed that the absorption strength as a function of the transverse distance, normalized by the virial radius, is at the level found near LBGs by Turner et al. (2014), but without any significant monotonic decrease. In Figure 2.12, the log-linear modeling of the equivalent width as a function of the distance shows that the absorption strength decreases at $R \gtrsim 30 \text{ kpc}$, corresponding to the typical virial radius in which LAEs reside. However, this trend is strongly driven by the non detections, with measured values remaining somewhat flat between $0.1 - 1 \text{ \AA}$ also beyond $\approx 100 \text{ kpc}$. In this context, Hasan et al. (2021) studied the statistical distribution of C IV gas relative to different galaxy populations and its evolution across the redshift range $0 \leq z \leq 5$ by means of an absorption model. Their model predicts that all the C IV absorbers are confined within the virial radius of low mass halos, while weak systems with $W_r \geq 0.05 \text{ \AA}$ may live beyond the virial radius of the most massive galaxies at redshift $z > 2.5$. These predictions are consistent with the observed extension of C IV gas around massive galaxies, with strong $W_r \geq 0.3 \text{ \AA}$ absorbers detected in the inner regions of the CGM and the weakest systems spread at distances larger than the virial radius. However, the model does not reproduce the large extension of weak and strong C IV absorbing gas around low mass haloes that host LAEs. A possible explanation may reside in the assumptions the model is based on, i.e. each individual C IV absorption line system is co-spatial with a single halo that hosts a single galaxy. We found that $\approx 20\%$ of the C IV absorbers in our sample are connected to a multiple galaxies within $\pm 500 \text{ km s}^{-1}$. Given the limited size of the FoV, it is also possible that isolated galaxies have companions at larger separations. Thus, the galaxy environment might play a role in shaping the differences between the observed and predicted C IV extension around low mass galaxies.

Altogether, this comparison highlights how both LAEs and LBGs coexist with a significant distribution of C IV gas which is clearly associated to the galaxy themselves on small scales, but that appears more weakly correlated as one moves to larger distances and lower equivalent widths.

Comparison with lower redshift results

The MAGG survey has been also employed to investigate the properties of the cool gas, probed by Mg II absorption, around galaxies at $z \sim 0.8 - 1.5$ (Dutta et al., 2020) for which the authors infer solar masses of $M_\star \approx 10^9 M_\odot$. Moreover, they combined results with the QSAGE survey, to trace both the more neutral and ionized gas phases around galaxies through Mg II and C IV absorption at $z < 2$ (Dutta et al., 2021). Their findings allow us to compare the properties of the cooler Mg II gas with the more ionized C IV gas both at low and high redshift.

At $z < 2$, the strength of the absorption is found to decrease with increasing distance. The anti-correlation is steeper for the Mg II absorbers compared to the C IV gas. However, the radial profile of both the Mg II and C IV absorption strength shows a significant scatter and flattening once galaxies in groups are included in the analysis (see also Bordoloi et al., 2011; Fossati et al., 2019), with strong absorption detected even at large transverse separations around groups. This is consistent with our findings shown in Figure 2.12, where the anti-correlation is sensitive to the upper limits, but strong systems are observed at any separations. A direct comparison as a function of redshift is made difficult by the different size of the FoV and the different sensitivities of these surveys. However, a general trend appears: a significant contribution to the flattening in the $W_r - R$ relation of C IV absorbers may come at all redshift from the larger extension of the highly-ionized gas which, as argued at lower redshift (Dutta et al., 2021), is possibly embedding the cooler phase. Similar results are indeed found around LBGs at $z \sim 2$ by Rudie et al. (2019), with singly-ionized ions commonly detected at smaller distances from the galaxies relative to higher ionization species.

Finally, the galaxy environment is observed to play an important role in shaping the properties of the absorbers both at low and high redshift. At $z \lesssim 2$, the Mg II systems connected to multiple galaxies are ≈ 5 times stronger and show a ≈ 3 times higher covering fraction up to $R \approx 200$ kpc ($\sim 2 R_{\text{vir}}$) compared to those associated to isolated galaxies. Fossati et al. (2019) found stronger Mg II absorption around 4 groups at redshift $z \sim 0.5 - 1.5$ and invoked gravitational interactions between the components of a group as a possible mechanism responsible for the enhanced Mg II absorption. Moreover, Dutta et al. (2021) found that the C IV covering fraction, both connected to groups and isolated galaxies, is ≈ 2 times higher relative to Mg II systems, and enhanced in groups up to $R \approx 750$ kpc ($\approx 5 R_{\text{vir}}$). However, once a control sample of isolated galaxies is assembled to reproduce the properties of the groups, the excess is still significant for Mg II absorbers, but is suppressed for the C IV systems. In contrast, at $z \gtrsim 3$ (Figure 2.13), we notice that the C IV covering fraction of control samples of galaxies in groups is confidently detected to be higher compared to isolated galaxies. Our findings are also supported by similar results from Muzahid et al. (2021). Therefore, there seems to be a hint of a possible redshift evolution in how more neutral and ionized gas phases respond to different galaxy environments. However, the C IV sample at $z < 2$ suffers from more limited statistics compared to $z \approx 3$. Likewise, there is currently no systematic study of the Mg II distribution near $z \approx 3$ LAEs. This evolution scenario should therefore be tested further with better spectroscopic coverage for C IV at $z < 2$ and for Mg II at $z \gtrsim 3$. The latter effort is the objective of an ongoing X-Shooter campaign (PID 0109.A-0559; PI M. Galbiati).

2.5.2 C iv and H i around $z > 3$ LAEs

As a complement to the analysis presented here for ionized gas traced by C IV absorbers around LAEs, in the MAGG IV paper we have studied the link between these galaxies and the neutral gas traced by optically-thick absorption line systems. The MAGG IV sample is a collection of 61 strong H I absorbers ($N_{\text{HI}} \gtrsim 10^{16.5} \text{ cm}^{-2}$), ≈ 84 per cent of which are associated to 127 LAEs within $\pm 1000 \text{ km s}^{-1}$ at redshift $z \approx 2.9 - 4.2$. Reducing the velocity separation to the limit we chose to associate absorbing gas and galaxies to the one adopted in this work, 45 H I absorbers are connected to 97 LAEs within $\pm 500 \text{ km s}^{-1}$ centered on the peak of the velocity separation distribution (found at $\Delta v \approx 250 \text{ km s}^{-1}$ in MAGG IV). These numbers correspond to an LAE detection rate of $\approx 74 \pm 15$ per cent around H I gas, a factor ≈ 2 higher relative to the fraction of C IV systems connected to at least one galaxy within the same velocity window.

Searching for the presence of metal absorption-lines as a probe of the metallicity of H I systems, in MAGG IV we identify a sample of 58 C IV absorbers that have been analysed independently from this work. Of these C IV absorbers, excluding upper limits for non detections and systems falling within telluric regions (5/58), we found that all these absorbers have a counterpart in the sample assembled in this work within a redshift separation $\Delta z = 0.0045$ ($\Delta v \approx 300 \text{ km s}^{-1}$ at $z = 3.5$) with ≈ 81 per cent (43/53) detected at $W_r^{1550} > 3\sigma$. We observe that ≈ 51 per cent

Samples	W_r (Å)	Δv_{90} (km s $^{-1}$)	$\log[L_{Ly\alpha}/(\text{erg s}^{-1})]$	R (kpc)
With H I - Without H I	8.69×10^{-6}	1.12×10^{-5}	0.52	0.90
With H I - Full Sample	1.52×10^{-4}	6.59×10^{-3}	0.66	0.87
Without H I - Full Sample	0.46	0.57	~ 0.99	~ 0.99

Table 2.9: Results of KS test (p -values) that measures the probability that the distribution of C IV and LAEs properties (equivalent width, velocity width, LAE luminosity, impact parameter) for systems that are (With H I) and are not (Without H I), connected to strong H I absorbers are drawn from the same parent population.

(22/43) of these systems are connected to at least 1 galaxy, with ≈ 33 per cent (14/43) connected to galaxies that are part of a group. Based on this different detection rate, we conclude that if LLSs trace both the CGM and the filaments connecting several LAEs as argued in MAGG IV, C IV absorbers only partially trace these same structures. This conclusion is reinforced by the statistical estimates of the LAE luminosity function (Figure 2.7) and the galaxy-absorber cross-correlation function (Figure 2.9). Our analysis shows that the LAE number density is higher around strong H I absorbers relative to C IV systems both as a function of the galaxy luminosity and the transverse separation, indicating a stronger connection between strong H I and LAEs than between C IV and LAEs.

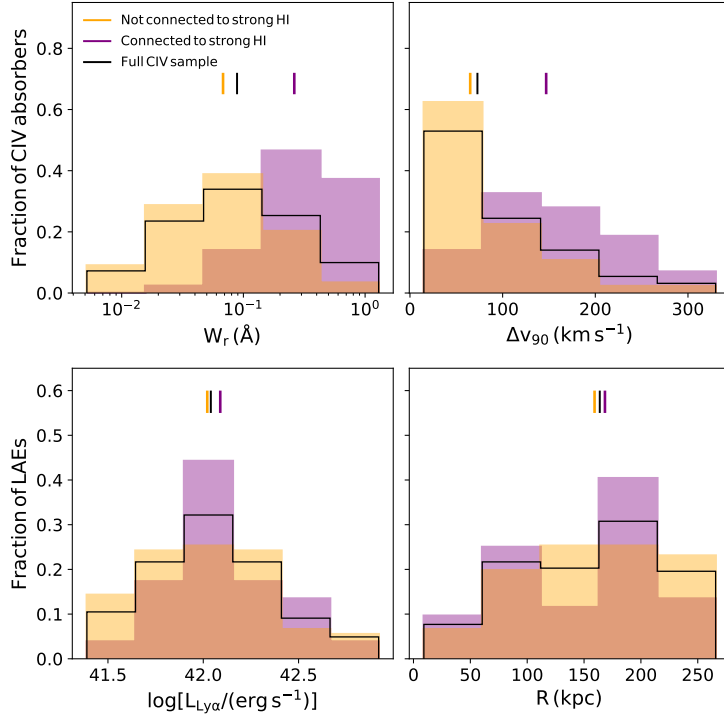


Figure 2.15: Upper panels: comparison between the distribution of the rest-frame equivalent width (left panel) and velocity width (right panel) of the full C IV sample (black) with the sub-sample overlapping MAGG IV catalogue (purple) and the remaining systems (orange). Lower panels: same as above, but comparing the Ly α luminosity (left panel) and the transverse separation (right panel) of the LAEs associated to the different C IV samples. Median values are shown as ticks.

The role of cosmic filaments

We investigate whether differences exist between the full sample of C IV absorbers relative to the sub-samples of C IV systems that are connected to strong H I systems or not. The results are shown in the upper panels of Figure 2.15, while the p -values resulting from a KS test are listed in Table 2.9. We observe a lack of weak ($W_r \lesssim 0.01$ Å) C IV systems from MAGG IV and a small excess of strong absorbers $W_r \gtrsim 0.1$ Å, with $\approx 40\%$ of the $W_r \gtrsim 0.1$ Å C IV systems included in this work connected to strong H I absorbers within ± 500 km s $^{-1}$. This suggests that the C IV systems connected to H I absorbers are, on average, a factor ≈ 3.7 stronger compared to the C IV that is

not associated to H I absorbing gas. The KS test suggests that the systems with H I detection, compared to those without H I absorption and the full C IV sample, show an equivalent width that is unlikely to be drawn from the same parent distribution, with p -value $< 10^{-6}$ and p -value $< 10^{-4}$ respectively. On the other hand, the strength of the C IV absorbers without H I detection is more similar to that of the full C IV sample (p -value > 0.46). The proximity to H I absorbers seems to affect also the kinematics of the gas, since the C IV systems have a velocity width that is a factor ≈ 2.3 broader if connected to H I absorbing gas. As found before, the KS test suggests that the Δv_{90} of the C IV absorbers with H I detection, compared to C IV without H I absorption and the full C IV sample, are unlikely to be drawn from the same parent population (p -value $< 10^{-5}$ and p -value $< 10^{-3}$ respectively). C IV absorbers without H I detection do not significantly deviates (p -value > 0.57) from the full C IV sample.

A similar analysis is performed to compare the transverse separation and the Ly α luminosity of the LAEs connected to the full C IV sample and to the systems that are or are not associated to H I absorbers. From the results shown in the lower panels of Figure 2.15, we notice that the C IV systems are associated to the population of LAEs that does not show any difference (p -value > 0.50) in the Ly α luminosity they emit and in their transverse separation from the absorbing gas. Thus, this comparison suggests that C IV systems matched to strong H I absorbers have higher equivalent widths and are more broad kinematically, but this change in absorption properties does not influence significantly the emission properties of the associated LAEs. As mentioned above, MAGG IV results support the picture in which LAEs are clustered within large scale structures and strong H I absorbers are tracers of both the galaxies CGM and of the dense optically thick gas contained in the filaments connecting LAEs.

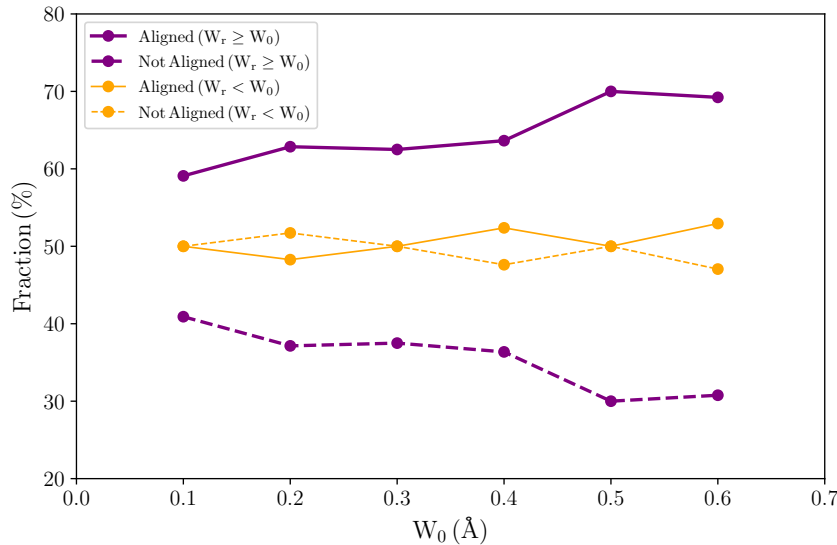


Figure 2.16: Fraction of LAEs aligned ($0^\circ - 45^\circ$ or $135^\circ - 180^\circ$, solid line) and not aligned ($45^\circ - 135^\circ$, dashed lines) with the axis defined by the absorber and the closest LAE, as a function of the absorbers equivalent width. We show results for LAEs connected to $W_{\text{C IV}} \geq W_0$ and $W_{\text{C IV}} < W_0$ as purple and orange lines, respectively.

We have shown so far that strong C IV systems are connected to H I absorbers, suggesting that these systems might be tracers of the same large-scale structure as high-density neutral hydrogen. In support of this hypothesis, we would expect the galaxies to be aligned with respect to the positions of the C IV absorbers, similarly to what found when analysing strong H I absorbers (see MAGG IV). We thus follow the procedure adopted in MAGG IV and measure the offset between the galaxy positions and the axis connecting the C IV absorbers to the closest LAE. Only the 38 C IV systems connected to 2 or more LAEs are included in this analysis. Results are shown in Figure 2.16 as a function of C IV equivalent width threshold, W_0 . We found an excess of aligned ($< 45^\circ$ and $> 135^\circ$) LAEs connected to strong ($W_r \geq W_0$) C IV absorbers. The fraction of aligned LAEs increases with increasing absorption strength, while LAEs connected to weaker C IV ($W_r < w_0$) does not show any preference to be aligned with the absorbers.

While strong C IV absorbers appear to overlap at least in part with optically-thick H I gas, the above analysis suggest that weaker C IV systems are tracer of a more homogeneous medium, found also at larger distances from galaxies where the H I column density is expected to decline (Rudie

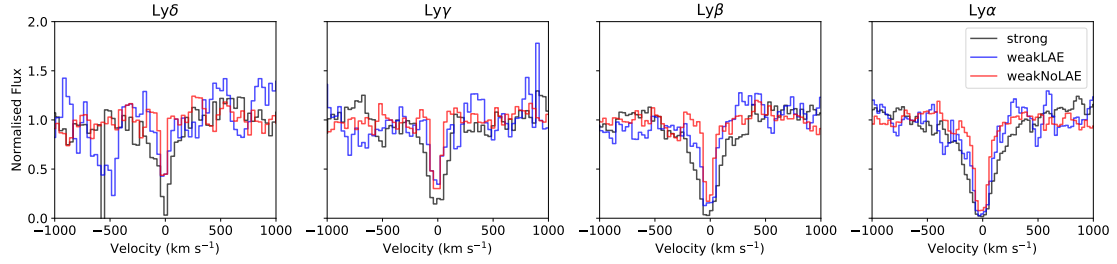


Figure 2.17: Fraction of LAEs aligned (Stacks of $\text{Ly}\alpha$, $\text{Ly}\beta$, $\text{Ly}\gamma$ and $\text{Ly}\delta$ absorption at the redshift of each C IV system for strong absorbers ($W_r \geq 0.1 \text{ \AA}$, black), weak absorbers ($W_r < 0.1 \text{ \AA}$) connected to at least one LAE and weak absorbers that are not connected to any galaxies within $\pm 500 \text{ km s}^{-1}$ (red).

et al., 2012). We thus investigate if there is any difference in the average H I profiles associated with different ‘types’ of C IV absorbers. To do so, we stacked the $\text{Ly}\alpha$, $\text{Ly}\beta$, $\text{Ly}\gamma$ and $\text{Ly}\delta$ absorption at the redshift of each C IV system (where data were available) and compare the results we obtained for strong C IV absorbers ($W_r \geq 0.1 \text{ \AA}$), weak C IV absorbers ($W_r < 0.1 \text{ \AA}$) connected to at least one LAE and weak C IV absorbers that are not connected to any galaxies within $\pm 500 \text{ km s}^{-1}$. The result is shown in Figure 2.17 and suggests that strong C IV absorbers arise from higher column density regions, likely close to the spine of the cosmic filaments where both galaxies and optically-thick absorbers reside (see MAGG IV), while weaker absorbers arise from lower column density regions away from the central parts of the filaments, where the presence of galaxies is more stochastic. Figure 2.17 further hints at a progressive less significant hydrogen absorption in weak C IV systems far from galaxies compared to a similar population of metal absorbers near LAEs.

In force of this analysis, we propose a new picture of the distribution of metals around LAEs in which strong C IV absorbers, as well as H I absorbers, trace dense regions of the filaments connecting galaxies. On the contrary, weaker C IV absorbers seem to lie progressively farther from the galaxies and hence farther from the densest regions of the filaments, pinpointing to the existence of an enriched, but more diffuse and homogeneous medium that degrades towards the IGM.

The role of outflows

A variety of models has been studied in the literature in order to understand the physical processes that lead to the C IV enrichment of the CGM and IGM, spanning from Population III stars at high redshift ($z \sim 25 - 30$), to more recent ejections of galactic winds and feedback powered by star formation activity in galaxies at $z \sim 3 - 10$ (e.g. Simcoe et al., 2004). The C IV absorbers we observe at redshift $z \sim 3 - 4$ may be the result of cumulative feedback processes that occurred at higher redshift and ongoing feedback ejections at the epoch near the peak of the star formation rate density towards $z \approx 2$. We attempt to disentangle these two effects by exploring if it is plausible that the bulk of outflows powered by recent star formation in galaxies has instantaneously enriched all the observed absorption systems near LAEs. To this end, in Figure 2.18 we compare the transverse separation between the LAEs and the C IV absorbers found within $\pm 500 \text{ km s}^{-1}$ with the distance traveled by outflows, assuming they are launched at redshifts between $z = 4$ and $z = 6.5$ with steps $\Delta z = 0.5$ and travel at average velocities $\langle v \rangle = 100, 250 \text{ km s}^{-1}$ (solid and dashed lines respectively). We found that to reach C IV absorbers at separations $R \gtrsim 150 \text{ kpc}$, outflows must be launched at redshift $z \gtrsim 6.5$ if travelling at $\langle v \rangle = 100 \text{ km s}^{-1}$, and are thus linked to episodes of star formation that happened in the galaxies $\Delta t \gtrsim 1 \text{ Gyr}$ before the epoch of observations. Outflows launched $\Delta t \lesssim 750 \text{ Myr}$ before the observations and linked to more recent star formation activity in the galaxies must instead travel at average velocities $\langle v \rangle \gtrsim 250 \text{ km s}^{-1}$ to reach the C IV absorbers up to the edge of the FoV ($R \approx 250 \text{ kpc}$).

In a similar analysis, Díaz et al. (2021) interpreted the connection between 11 C IV absorbers and LAEs at $z > 4.7$ as potentially due to the relatively recent ($\Delta t \lesssim 700 \text{ Myr}$ at $z \approx 5$) galaxy activity, modelled as outflows started at $z = 10$ and travelling at average velocities $\langle v \rangle \lesssim 400 \text{ km s}^{-1}$ for 89% (8/9) of the LAE sample. This scenario suggests that only outflows traveling at average high velocities ($\langle v \rangle \gtrsim 250 \text{ km s}^{-1}$) are able to enrich the gas around a fraction of the LAEs if powered by recent ($\Delta t \lesssim 500 \text{ Myr}$) star formation activity in galaxies.

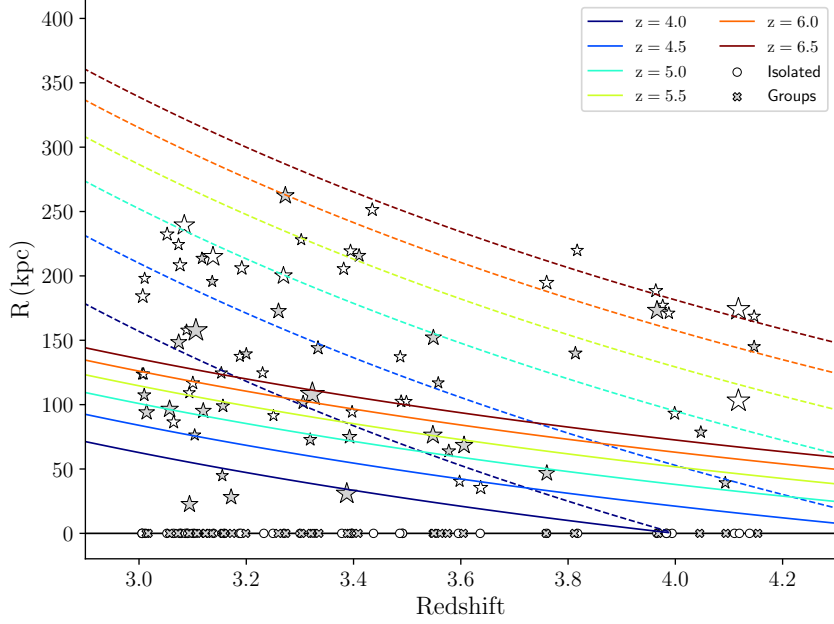


Figure 2.18: Comparison between the distance traveled by outflows ejected during episodes of recent star formation activity (e.g. launched at redshift between $z = 4$ and $z = 6.5$ with steps $\Delta z = 0.5$) in the galaxy and the transverse separation between the LAEs (shown at $R = 0$ kpc) and the connected C IV absorbers (stars), assuming the outflows travel with average velocities $\langle v \rangle = 100, 250 \text{ km s}^{-1}$ (solid and dashed lines respectively). We distinguish between C IV associated to isolated galaxies (white-filled markers) and groups (gray-filled markers). The size of the marker is weighted by the rest-frame equivalent width.

2.6 Summary and conclusions

In this work, we explored the connection between an ionized gas phase, traced by absorption-line systems in quasar spectra, and a population of Ly α emitting galaxies identified in the MUSE Analysis of Gas Around Galaxies (MAGG) survey, which combines medium-deep VLT/MUSE observations of 28 quasar fields with high S/N and high resolution spectroscopy of the central sources. We traced the ionized gas by assembling a sample of 220 C IV and 108 Si IV absorption-line systems at redshift $3.0 \lesssim z \lesssim 4.5$. To extend the analysis of gas-galaxies connection to the lower end of the galaxy mass function, we identified > 1000 LAEs in the MUSE cubes at $3.0 \lesssim z \lesssim 6.0$, 292 of which lie in the C IV redshift path and are candidates for an association with the absorbing gas. We found 143 LAEs within a velocity separation $|\Delta v| \leq 500 \text{ km s}^{-1}$ from C IV absorbers. Given the large size of this sample, we were able to study, for the first time with significant statistics, both the global correlation between LAEs and C IV/Si IV and the properties of the individual associations. The main findings of this work are summarized next.

- We measured, for the first time, the luminosity function of LAEs connected to C IV absorbers. Compared to an updated version of the field luminosity function derived from MAGG, we found a normalization ≈ 2.4 higher for LAEs associated with the C IV absorbing gas. This analysis revealed that the LAE number density is sensitive to the proximity to different tracers studied in MAGG and progressively increases from the field, to the surrounding of C IV gas, strong H I absorbers and central quasars.
- We measured the two-point LAE auto-correlation and LAE-C IV cross-correlation functions and the respective reduced angular auto- and cross- correlation functions. We found a clustering length $r_0 = 2.07^{+0.23}_{-0.24} \text{ h}^{-1} \text{ cMpc}$ for LAEs and $r_0 = 1.23^{+0.25}_{-0.27} \text{ h}^{-1} \text{ cMpc}$ for LAE-C IV absorbers, assuming a fixed slope $\gamma = 1.8$. Employing the Cauchy-Schwarz relation, we found that a fraction of the C IV absorbers is likely to arise not only in the halo of the LAEs they are connected to, but also from other regions, such as locations in the IGM not connected to galaxies. The LAE-C IV cross-correlation function lies below the cross-correlation function measured between LAEs and optically-thick H I absorbers, indicating a larger contribution from more underdense regions in the C IV case.
- Moving to a galaxy-centered approach (i.e., centering on the galaxies and exploring the sur-

rounding gas distribution) we measure the LAE covering fraction as a tracer of the probability to observe C IV absorbers within $\pm 500 \text{ km s}^{-1}$ from a galaxy. The covering fraction is observed to decrease with increasing C IV absorption strength and increasing impact parameter up to $\approx 100 - 150 \text{ kpc}$ from galaxies, at which point it is observed to flatten with elevated values far beyond the virial radius and up to the edge of the FoV. We also found a higher covering fraction for galaxies with luminosity $\log[L_{\text{Ly}\alpha}/(\text{erg s}^{-1})] > 42.2$ compared to fainter LAEs.

- Among the galaxies connected to the C IV absorbers, ≈ 30 per cent are isolated within $\pm 500 \text{ km s}^{-1}$, while the remaining ≈ 70 per cent are part of a group. The absorbers connected to 1 or at least 2 LAEs show, on average, a rest-frame equivalent width that is a factor ≈ 1.5 and ≈ 4.6 higher than the median of the full sample, respectively. However, the effect on the gas kinematics is less significant, with only the absorbers connected to more than 1 LAE exhibiting a Δv_{90} that is a factor ≈ 1.7 above the median of the full C IV sample. C IV absorbers with larger equivalent width and Δv_{90} tend to be associated with brighter LAEs. The probability to observe a C IV absorber near an LAE is a factor ≈ 3 higher around galaxies in groups compared to the isolated ones at any transverse separation. As opposed to what is observed at $z \lesssim 2$, the excess of C IV absorption around galaxies in groups is present even if controlling for the properties of isolated galaxies and groups.
- Comparing the statistics of C IV and Si IV absorbers, we found that ≈ 29 per cent of the C IV systems show Si IV absorption with the same number of components aligned with the C IV ones. The fraction of Si IV systems' matched to C IV decreases with decreasing equivalent width, and the ratio between the equivalent width of Si IV and C IV is found to be less than unity, as expected from ionization models for densities $n_{\text{H}} \lesssim 10^{-2} \text{ cm}^{-3}$. Thus, we conclude that the two ions, with a similar ionization potential and in the same ionization stage, mostly arise from the same gas phase when detected. This is further supported by the fact that an independent analysis of gas-LAE connection leads to similar results if using either C IV or the Si IV ions as a tracer.
- Despite the large projected distance from LAEs at which C IV absorption is detected (up to the edge of the FoV, corresponding to $R \approx 250 \text{ kpc} \gtrsim 7 R_{\text{vir}}$ for a dark matter halo $M_{\text{H}} \approx 10^{11} M_{\odot}$), we have found that at least a fraction of the absorbers can be instantaneously enriched by outflows traveling at $< v > \gtrsim 250 \text{ km s}^{-1}$ and launched $\Delta t \lesssim 750 \text{ Myr}$ before the observations. This makes the CGM and IGM, and in particular an intermediate ionized gas-phase, at least partially linked to the recent star formation activity in galaxies.
- Compared to similar studies about the connection of metals and LBGs, LAEs and (strong) C IV absorbers are less clustered, with a correlation length that is a factor of ≈ 2 lower than LBGs at similar redshifts. For both galaxy populations, the ionized gas is extended far beyond the typical virial radius of the halo with a mild (if any) decline in the equivalent width beyond a few hundred kiloparsecs from the galaxy position.
- We found that 43 out of 220 C IV systems are also connected to strong H I absorbers within $\pm 500 \text{ km s}^{-1}$ in MAGG. We found that these systems show, on average, higher rest-frame equivalent width and more complex kinematics compared to those that are not associated to H I absorption. We did not observe any difference between the properties of the LAEs connected to C IV absorber with or without H I detection, but we found that it is more likely ($\approx 2\times$) for strong H I absorbers to be associated with LAEs than C IV systems.

The combination of this study on gas at intermediate ionization near LAEs and the findings presented in our MAGG IV paper on the link between LAEs and optically-thick H I absorbers, together with literature results, allows to paint a clearer picture for how hydrogen and metals are distributed around $z \approx 3 - 4$ star-forming galaxies, for the first time extending to lower mass scales ($M_{\text{H}} \approx 10^{11} M_{\odot}$) with a complete and coherent survey. As argued in MAGG IV, data support a model in which star-forming galaxies lie within large-scale structures (i.e., cosmic filaments) and strong H I absorbers are a tracer of the neutral phase both within the CGM of the embedded star-forming galaxies and along the filaments connecting halos. This paper confirms a similar picture for the metals traced by C IV, but only when considering relatively strong absorbers, with $W \gtrsim 0.1 \text{ \AA}$, which are more clearly correlated to galaxies and the LLS themselves. For weaker absorbers, our MUSE observations support the idea of a more widespread and homogeneous medium that extends

far beyond the CGM of galaxies, also reaching regions away from the densest parts of the filaments where optically-thick H I absorbers are found.

We thus propose a ‘three-component model’ to fit all the principal features uncovered by MAGG near LAEs, which agree with the findings and conclusions by Muzahid et al. (2021). The first component is represented by the CGM of individual systems that gives rise to the strongest H I and C IV absorbers observed at close separation from galaxies and which accounts for the elevated covering factors of hydrogen and metals near LAEs. On these scales, it is more likely that galaxies have a direct impact on the properties of the CGM, as seen for instance in some correlations between emission and absorption properties (e.g., C IV velocity width and absorption strength with Ly α luminosity), although as argued in MAGG IV the patchy nature of this gas leads to high scatter in the observed relations. The second component describes the overdense gas along filaments, which accounts for a large fraction of LLSs and for the remaining fraction of high equivalent-width absorbers. This gas, together with the outer CGM components, is what drives the excess of galaxies located at distances greater than a few times the virial radius and sets the normalization and shape of the luminosity and cross-correlation function for strong H I and C IV systems. The third component is represented by the lower equivalent-width but enriched medium, farther from the denser CGM and the central parts of the filaments connecting galaxies. This medium, which is not easily selected by LLSs, i) accounts for the substantial fraction of C IV for which no LAE counterparts are found; ii) lowers the normalization of the luminosity function and the correlation length of the cross-correlation function; iii) leads to a high covering factor extending to large distances from galaxies; iv) yields a weak dependence of equivalent width with impact parameter. This more diffuse gas is harder to detect with less sensitive tracers like Si IV, which explains the stronger clustering/higher normalization of LAEs near Si IV than C IV (in other words, Si IV is a better tracer of the first two components of this model).

Finally, environment shapes the distribution of gas, particularly the stronger absorbers that seem to be more affected by the presence of groups. Currently, however, it is not particularly clear how environment affects different gas phases (e.g. ionized and neutral components) differently as a function of redshift, and which physical mechanisms are primarily responsible for the observed trends. Refining this picture is the subject of ongoing observational efforts. Observations of other galaxy populations that reside in the same structures are expected to yield comparable results. For instance, LBGs, which are not completely independent from LAEs and relatively more clustered, are too embedded within these structures and carry their own CGM. Hence, most of the trait discussed for LAEs apply to LBGs (as observed), with the exception of a higher clustering of this population.

Biases against the detection of particular systems (e.g., dusty or passive galaxies) are surely present and will inevitably lead to changes in this picture. As argued in MAGG IV, however, the number density of LAEs is sufficiently high that this population becomes a very good tracer of the large-scale structures surrounding galaxies, within which additional undetected populations will reside. We therefore speculate that it is implausible that populations at much lower number density will drastically alter the model put forward here. Future multiwavelength surveys, particularly at infrared/millimetre wavelengths, are required to settle this point. Sustained effort should also be invested in theoretical work to develop a more quantitative model, e.g. using numerical simulations, that can be compared more quantitatively with our observations to test this ‘three-component model’ rigorously.

MUSE Analysis of Gas around Galaxies (MAGG) - VI: The cool and enriched gas environment of $z \gtrsim 3$ Ly α emitters

3.1 Introduction

The gas reservoir retained within the intergalactic medium (IGM) and circumgalactic medium (CGM) stands as a critical regulator in the formation, growth, and evolution of galaxies. While metal-poor gas flows through filaments connecting the galaxies and provides fuel for sustaining their star formation (Kereš et al., 2005; Dekel et al., 2009), the balance between accretion and feedback processes enriches and shapes the properties of the CGM gas, rendering it multiphase in temperature, density, kinematics, and composition (Oppenheimer & Davé, 2008; Lilly et al., 2013; Tumlinson et al., 2017). A complete view of the connection and the coevolution of galaxies with the surrounding cosmic gas is thus crucial to paint a comprehensive picture of how galaxies form and evolve through time. In addition to the galaxies' internal processes, the large-scale galaxy environment is expected to affect the activity of the galaxies themselves and the properties of their surrounding gas via hydrodynamic and gravitational interactions (see e.g., Gunn & Gott, 1972; Merritt, 1983; Boselli & Gavazzi, 2006).

Targeting the diffuse cosmic gas in absorption in the spectra of background bright sources, such as quasars, has been widely adopted as a powerful technique to connect the multiphase CGM and (often massive star-forming) galaxies across cosmic time (see Chen et al., 2017 for a review at $z < 0.5$ and Steidel et al., 2010; Rudie et al., 2012; Turner et al., 2014; Rudie et al., 2019 at $z \approx 2 - 3$). Doublet transitions of metals, such as Mg II $\lambda\lambda 2796, 2803$ and C IV $\lambda\lambda 1548, 1550$, are ideal probes of the different gas phases ranging from the cool ($T \approx 10^4$ K) and enriched gas to a more ionized and, sometimes, warmer phase ($T \approx 10^5$ K). At the same time, large spectroscopic surveys, especially with modern integral field spectrographs (IFSs) that allow the assembly of dense and complete samples of galaxies, have been instrumental in investigating the gas-galaxy connection down to the lower-mass galaxy population.

Bergeron (1988) and Bergeron & Boissé (1991) pioneered the observations of Mg II absorption systems in quasar sightlines and reported the detection of a large number of galaxies at the redshift of the absorbers. Since then, the Mg II has been widely used as an effective tracer of cool gas around galaxies across cosmic time (see Barnes et al. 2014 for a review), even up to the reionization epoch with recent *James Webb* Space Telescope observations (Bordoloi et al., 2023). Mg II-bearing gas is indeed found to be an effective tracer of inflows and outflows (Tremonti et al., 2007; Weiner et al., 2009; Martin & Bouché, 2009; Rubin et al., 2010; Nestor et al., 2011; Ribaud et al., 2011; Thom et al., 2011; Rubin et al., 2012, 2014; Xu et al., 2023), and to exhibit tight correlations with the galaxies' stellar mass and ongoing star formation activity (Gauthier et al., 2009; Chen et al., 2010a; Bordoloi et al., 2011; Ménard et al., 2011; Churchill et al., 2013a; Nielsen et al., 2013; Bordoloi et al., 2014; Liang & Chen, 2014; Lan et al., 2014; Lan, 2020; Dutta et al., 2021).

A more thorough perspective emerges when different gas phases are observed simultaneously. In the COS-Halos survey Werk et al. (2013, 2014) reported that low and medium ionized metals tracing $T \approx 10^4 - 10^5$ K gas are common features in the CGM of $L \sim L^*$ galaxies at $z \approx 0.2$. They also found a difference in the radial distribution of different gas phases around galaxies: the gas's ionization state is higher at larger distances from the host, possibly due to the gas being more diffuse and less shielded. Several results from both simulations (Ford et al., 2013; Ho et al., 2020, 2021; Weng et al., 2024) and observations (Schroetter et al., 2021; Dutta et al., 2020; Qu et al.,

2023) do indeed show that the presence of low-ionization ions (e.g., Mg II, C II, and O II) is confined at closer distances from the galaxies, while higher ions (e.g., C IV, Si IV, O IV, and Ne VIII) extend up to several times the galaxies' virial radius. Furthermore, a tomographic analysis of absorption lines performed by Dutta et al. (2024b) revealed that the C IV gas shows a higher degree of coherence compared to Mg II. Schroetter et al. (2021) investigated the connection between C IV and Mg II absorbers and a population of [O II] emitters in the MuSE GAs FLOW and Wind (MEGAFLOW) survey (Schroetter et al., 2016). They found that if, on the one hand, the size of the Mg II halo decreases with time and becomes smaller compared to the dark-matter halo, on the other hand, the C IV gas progressively increases in size, possibly indicating a coevolution with the dark-matter halo. At $z < 2$, Dutta et al. (2021) showed that different gas phases do not only vary in their distribution, but also in how they respond to the properties of the galaxies. They did indeed find that the correlation between the absorption strength and the galaxies' stellar mass and star formation rate is stronger for Mg II than for C IV.

Besides individual gas-galaxy connections, an additional dimension must be taken into account to fully trace the coevolution of gas and galaxies as a function of cosmic time, especially in the context of hierarchical structure formation and interactions in dense galaxy environments. At $z \lesssim 1$ several detections of strong H I absorbers connected to multiple galaxies have been reported as tracers of an extended intragroup medium (see e.g., Péroux et al., 2017; Fumagalli et al., 2017; Chen et al., 2019; Hamanowicz et al., 2020). The cool and enriched phase traced by ions such as Mg II has also been detected around group galaxies. The presence of multiple galaxies, at $z \lesssim 2$, raises the detection rate of the Mg II absorption-line systems, compared to what is observed around isolated galaxies (Nielsen et al., 2018b; Dutta et al., 2020, 2021; Qu et al., 2023). These systems are also stronger and show more complex kinematics (Chen et al., 2010a; Nielsen et al., 2018b; Fossati et al., 2019; Nelson et al., 2021). Both the higher incidence rate of the absorbers and the effects on the gas properties are not fully explained in superposition models, requiring a contribution from environmental interactions that boost the gas cross section instead (see e.g., Nielsen et al., 2018b; Dutta et al., 2020). At $z \gtrsim 3$, when massive structures are not yet collapsed and environmental interactions are not expected to affect the properties of the surrounding gas as in the local Universe, the picture is less clear. Given the accessibility of H I and C IV transitions to optical ground-based spectrographs, the cosmic gas traced by these two ions has been studied as a function of the large scale-galaxy environment around massive Lyman break galaxies (LBGs; see, e.g., Adelberger et al., 2005) and, recently, Ly α emitters (LAEs). By using the Multi Unit Spectroscopic Explorer (MUSE; Bacon et al., 2010) at VLT, Muzahid et al. (2021) targeted a population of LAEs at $z \gtrsim 3$ and reported that both the H I and C IV absorption are stronger in the stacked spectra of groups, relative to isolated LAEs. Similarly, Lofthouse et al. (2023) (MAGG IV hereafter) and Galbiati et al. (2023) reported elevated covering fractions of H I and C IV around group galaxies. To better understand the origin of the excess absorption in overdensities at $z \approx 3 - 4$ and the evolution with cosmic time, we expanded the core dataset of the MUSE Analysis of Galaxy (MAGG) survey Lofthouse et al., 2020a; Dutta et al., 2020; Fossati et al., 2021; Lofthouse et al., 2023; Galbiati et al., 2023 with a new VLT/X-shooter program (PID 0109.A-0559; PI M. Galbiati). The MAGG survey is based on a MUSE Large Programme (ID 197.A-0384; PI M. Fumagalli) targeting 28 fields centered on $z \approx 3.2 - 4.5$ quasars with magnitudes $m_r \leq 19$ mag and at least one $N_{\text{HI}} \geq 10^{17} \text{ cm}^{-2}$ hydrogen absorption line system at redshift $z \geq 3.05$. By complementing the MUSE observations with archival high-resolution ($R \gtrsim 30,000$) and moderate or high signal-to-noise ($S/N \gtrsim 10$ per pixel) spectroscopy of the central quasars, this survey is specifically designed for studying the connection between galaxies and gas across the cosmic time blindly. Leveraging new data, this study focuses on the cool gas traced by 47 Mg II absorbers in the near-infrared (NIR), complementing the view of more ionized C IV-bearing gas from Galbiati et al. (2023) (MAGG V hereafter).

We describe the dataset and the new NIR observations in Section 3.2. In Section 3.3 we report the details of the samples of Mg II absorbers, strong H I Lyman limit systems from Lofthouse et al. (2020a), C IV systems and LAEs from MAGG V. The correlation between LAEs and the Mg II-bearing gas is explored in Section 3.4.1, and in Section 3.4.4 we investigate the effect of the large-scale galaxy environment on the multiphase gas. Finally, in Section 3.5, we discuss how different gas phases evolve with time in connection to galaxies and the large-scale environment. Throughout, unless otherwise noted, we quote magnitudes in the AB system, distances in physical units, and adopt the *Planck* 2015 cosmology ($\Omega_m = 0.307$, $H_0 = 67.7 \text{ km s}^{-1} \text{ Mpc}^{-1}$; Planck Collaboration et al., 2016).

3.2 Observations and data reduction

3.2.1 MUSE observations

The 28 quasar fields included in the MAGG survey have been observed during ESO periods 97-103 with an integration time of ≈ 4 hours per field using MUSE in the wide field mode with extended wavelength coverage (465 – 930 nm, $R \approx 2000 - 3500$). All the observations have been carried out on clear nights at airmass ≤ 1.6 , resulting in better image quality than $\text{FWHM} = 0.8$ arcsec. Small dithers and relative rotations of 90 degrees for each exposure have been included to mitigate the effects of sensitivity variations across the $\approx 1 \times 1$ arcmin² field of view (FoV). Longer exposures, up to ≈ 10 hours, characterize fields with available archival exposures from GTO observations. All the details, including the exposure times and the image quality for each field, are listed in table 1 of [Lofthouse et al. \(2020a\)](#).

We summarize the MUSE data reduction next, and additional details can be found in [Lofthouse et al. \(2020a\)](#). We first applied bias subtraction, sky flats correction, and flux and wavelength calibrations to the raw data using the ESO MUSE pipeline ([Weilbacher et al., 2014](#), version 2 or greater). The datacubes were then post-processed with the CUBEXTRACTOR package (CUBEX, [Cantalupo et al., 2019](#)), one of the tools currently available to improve the imperfections arising from the sky subtraction and different response of MUSE spectrographs (see details in [Lofthouse et al., 2020a](#)). The single reduced exposures are combined into the final products using mean and median statistics. Additional datacubes, containing only half of the exposures each, are produced to facilitate the detection of sources and efficiently identify possible contaminants. We described our method to propagate the uncertainties and estimate the standard deviation in [Lofthouse et al. \(2020a\)](#) and [Fossati et al. \(2019\)](#).

3.2.2 Archival optical quasar spectroscopy

The MAGG survey includes high-resolution ($R \gtrsim 30,000$) and moderate or high signal-to-noise ($S/N \gtrsim 10$ per pixel) archival spectroscopy at optical wavelengths for the central quasar of each field. The available observations are obtained with the Ultraviolet and Visual Echelle Spectrograph (UVES, [Dekker et al., 2000](#)) at the VLT, the High Resolution Echelle Spectrometer (HIRES, [Vogt et al., 1994](#)) at Keck, the Magellan Inamori Kyocera Echelle (MIKE; [Bernstein et al., 2003](#)) at the Magellan telescopes. Complementary medium-resolution ($R \gtrsim 5,000$) spectra are also available from the Echellette Spectrograph and Imager (ESI, [Sheinis et al., 2002](#)) at Keck and X-shooter [Vernet et al., 2011](#) at the VLT. A detailed description of the data and the reduction steps can be found in section 3.1 in [Lofthouse et al. \(2020a\)](#) with the wavelength coverage, spectral resolution, and signal-to-noise ratio at the selected wavelengths listed in table 2.

3.2.3 X-shooter near-infrared spectroscopy

Extending the wavelength coverage up to the near-infrared is essential to tracing the cool gas via Mg II absorption at $z > 3$. As shown in Figure 3.1, the updated redshift coverage of the MAGG survey for H I (green), Mg II (orange for optical spectroscopy, dark-orange for the extended range probed by NIR spectroscopy), C IV (blue) and Si IV (purple) overlap with that of the low-mass galaxy population traced by LAEs (red), making it feasible to explore different gas phases around the LAEs simultaneously.

Archival X-shooter spectroscopy is available for a subset of MAGG quasars (13 of the 28 quasars) from the data release of the XQ-100 survey ([López et al., 2016](#)) and, in case of the sightlines J015741 – 010629 and J020944 + 051713, from the ESO archive. We complete NIR observations of the MAGG survey by observing the remaining 15 quasars with a 15.9h X-shooter program (PID 0109.A–0559; PI M. Galbiati). The observations were executed during ESO periods 109-110 with clear sky, $\leq 80\%$ of lunar illumination, and at airmass ≤ 1.5 , requiring the image quality to be better than 1.1 arcsec. The nominal resolving power is ≈ 5600 for the 0.9 arcsec slit.

The archival XQ-100 spectra were reduced as described by [López et al. \(2016\)](#) by using a custom IDL pipeline to perform accurate sky subtraction and optimal extraction of the science spectra on unrectified 2D arrays of slit position and wavelength. Each frame is dark subtracted and flat fielded. A b-spline fit ([Kelson, 2003](#)) is used to model the sky emission that is then subtracted. The counts in each 2D frame is flux calibrated using observations of spectro-photometric standard stars taken, whenever possible, close in time to the science observations. As a last step, a final 1D spectrum is extracted for each exposure, rebinned into a common grid, and finally co-added.

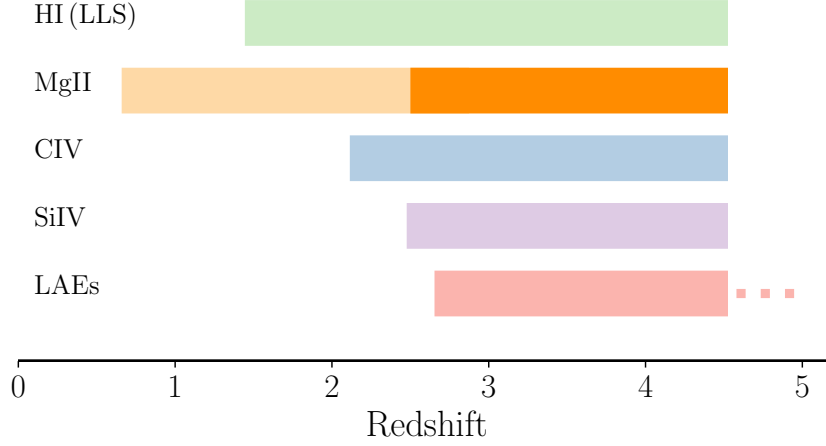


Figure 3.1: Redshift path of the H I (green), Mg II (orange), C IV (blue), and Si IV (purple) absorption line systems which trace the multiphase gas around LAEs (red) covered by optical-to-NIR spectroscopy in MAGG and shown up to the redshift of the most distant quasar. The coverage of the Mg II at $z \gtrsim 3$ (dark orange) is enabled by the new X-shooter observations (Section 3.2.3). The redshift range of the absorbers extends from redward the Ly α emission of the lowest redshift quasar up to 3000 km s^{-1} from the redshift of the highest quasar of the sample. Telluric regions are not masked in this figure.

We used the open-source python package PYPEIT (Prochaska et al., 2019, 2020) to reduce the new X-shooter observations included in MAGG. The key steps of the process are summarized as follows:

1. each raw frame is overscan and bias subtracted;
2. the master arc frame is used for wavelength calibration;
3. a 2D model of the master flat frame is applied to the science frames and used to trace the echelle orders and correct for the detector illumination;
4. cosmic rays are removed using the Laplacian Cosmic Ray Identification algorithm, L.A. COSMIC (van Dokkum, 2001);
5. the sky emission is subtracted from the 2D frames and modeled by both the AB-pair nodding patterns and iterative b-spline fit to remove residuals;
6. the quasar trace is detected automatically and optimally extracted (Horne, 1986) to generate the 1D and 2D science spectra;
7. the flux is calibrated using X-shooter standard stars.

If multiple exposures for each object are available, these are resampled into a common wavelength grid and then co-added. We used PYPEIT to correct for telluric absorption. As a final step, we modeled the quasar continuum using CONTFIT¹ which provides a python-based GUI to manually select points in the regions of the quasar spectra that are free of absorption, to be used as knots for a cubic spline.

We report in Table 3.1 the detailed list of the 15 MAGG quasar targeted by the new VLT/X-shooter spectra. The Table includes the redshift of each quasar as well as the median signal-to-noise ratios per pixel at the two rest-frame wavelengths $\lambda_{\text{RF}} = 3000 \text{ \AA}$ and $\lambda_{\text{RF}} = 3600 \text{ \AA}$ (computed within the wavelength ranges $\Delta\lambda = 2990 - 3010 \text{ \AA}$ and $\Delta\lambda = 3590 - 3610 \text{ \AA}$, respectively), to easily compare the new observations with the archived data of the remaining 13 quasars from López et al. (2016).

¹<https://github.com/trystynb/ContFit>

Sightline	Redshift	$S/N_{3000\text{\AA}}^{(a)}$	$S/N_{3600\text{\AA}}^{(b)}$	Wavelength ($\text{\AA}$) ^(c)
J010619+004823	4.4402	14	12	10195-24768
J033413-161205	4.3800	23	17	9817-24767
J033900-013318	3.2040	21	18	10194-24766
J094932+033531	4.1072	26	28	10195-24767
J095852+120245	3.2746	16	21	10195-24768
J102009+104002	3.1528	23	20	10195-24768
J120917+113830	3.0836	15	13	10195-24769
J123055-113909	3.5570	15	19	10195-24771
J142438+225600	3.6340	128	115	10194-24765
J193957-100241	3.7870	31	52	9816-24767
J200324-325144	3.7850	16	24	10195-24768
J205344-354652	3.4900	12	18	9817-24766
J230301-093930	3.4774	22	25	10195-24768
J231543+145606	3.3971	18	15	10194-24767
J233446-090812	3.3261	19	23	10195-24768

Table 3.1: MAGG sightlines targeted by the new VLT/X-shooter program PID 0109.A-0559. ^(a) Median signal-to-noise ratio per pixel at rest-frame $\lambda_{\text{RF}} = 3000 \text{ \AA}$ computed within the range $\Delta\lambda_{\text{RF}} = 2990 - 3010 \text{ \AA}$. ^(b) Median signal-to-noise at $\lambda_{\text{RF}} = 3600 \text{ \AA}$ within the range $\Delta\lambda_{\text{RF}} = 3590 - 3610 \text{ \AA}$. ^(c) NIR wavelength coverage.

3.3 Data analysis

3.3.1 Catalog of Mg II absorbers

To trace a cool and enriched gas phase surrounding LAEs, we assemble a sample of $z \gtrsim 3$ Mg II $\lambda\lambda$ 2796, 2803 absorption line systems identified by visually inspecting the near-infrared quasar spectra available in MAGG. We restrict the search of Mg II doublets between the wavelengths blueward the quasars' Ly α emission line, being the metals the minority of the transitions in the Ly α forest. We also mask the region at $\Delta v < 3000 \text{ km s}^{-1}$ from the redshift of the quasars (corresponding to observed wavelength range $13153 - 13285 \text{ \AA}$ at the mean redshift of the quasars included in MAGG, $z_{\text{QSO}}^{\text{mean}} = 3.75$), to avoid proximity effects (see, e.g., Hennawi & Prochaska, 2007). In addition, we mask the wavelength region $13500 - 14200 \text{ \AA}$ to preserve the sample from any possible contamination due to strong telluric features. Candidates with a strong-to-weak Mg II line optical depth ratio inconsistent with the expected value $\approx 2 : 1$ in the unsaturated regime are identified as interlopers and excluded from the following analysis. The final sample includes 47 Mg II absorbers² at $2.9 \lesssim z \lesssim 3.8$ (median $z \approx 3.31$) and rest-frame equivalent width $W_r \gtrsim 0.07 \text{ \AA}$.

Following the procedure described in MAGG V we model the absorption lines identified in the continuum-normalized quasar spectra with Voigt profiles using the Bayesian code MC-ALF (Longobardi et al., 2023). Each system is decomposed into the minimum number of individual Voigt components required to reproduce the observed flux and then convolved with a Gaussian kernel to match the line spread function of the observed data. We assign the following priors to the free parameters of each Voigt component: $1 \leq b/(\text{km s}^{-1}) \leq 30$ for the Doppler parameter and $12 \leq \log(N/\text{cm}^{-2}) \leq 16$ for the column density. To minimize the effect of imperfections in the continuum normalization, we include a multiplicative constant ranging between 0.98 and 1.02 as a free parameter. In addition, we have extensively tested (see MAGG IV and MAGG V) that allowing the spectral resolution to vary between $\pm 5 \text{ km s}^{-1}$ around the nominal value significantly improves the fit quality. In the case of line blending, we use additional *filler* Voigt components to model the absorption features arising from different ions at different redshifts that are uncorrelated with the targeted absorption-line system. For each Mg II system we provide a measure of the rest-frame equivalent width (median $W_r \approx 0.58 \text{ \AA}$), column density (median $\log(N/\text{cm}^{-2}) \approx 14.3$) and width of the line in velocity space, defined as the velocity interval enclosing the 90% of the absorption line optical depth (median $\Delta v_{90} \approx 110 \text{ km s}^{-1}$).

To assess the completeness of the sample, as done in MAGG V, we compute the equivalent width distribution function of the Mg II per unit redshift, $f(W_r, z) = d^2N/dW_r/dz$. To quantify the uncertainties, we bootstrap over the sightlines 10^3 times and derive the 10th and 90th percentiles

²A complete record of all the Mg II absorption systems included in this work as well as of the modeling of the lines is available at <https://doi.org/10.5281/zenodo.12794688>.

from the resulting distribution. We compare the results derived for the MAGG Mg II sample to the completeness corrected functions estimated for statistical surveys over a similar redshift range in Figure 3.2. In particular, we refer to the sample of Mg II systems found in $\approx 10^5$ quasar spectra from SDSS at $z \approx 2.05 - 2.30$ from [Zhu & Ménard \(2013\)](#) and 75 Mg II absorbers identified at $z \approx 3.15 - 3.81$ in the Magellan/FIRE spectra of 100 quasars from [Chen et al. \(2017\)](#). The distribution function of the MAGG sample lacks any completeness correction. However, since it does not deviate significantly compared to the expectation from the literature, we conclude that the MAGG Mg II sample is reasonably complete down to equivalent width $W_r \approx 0.10 \text{ \AA}$.

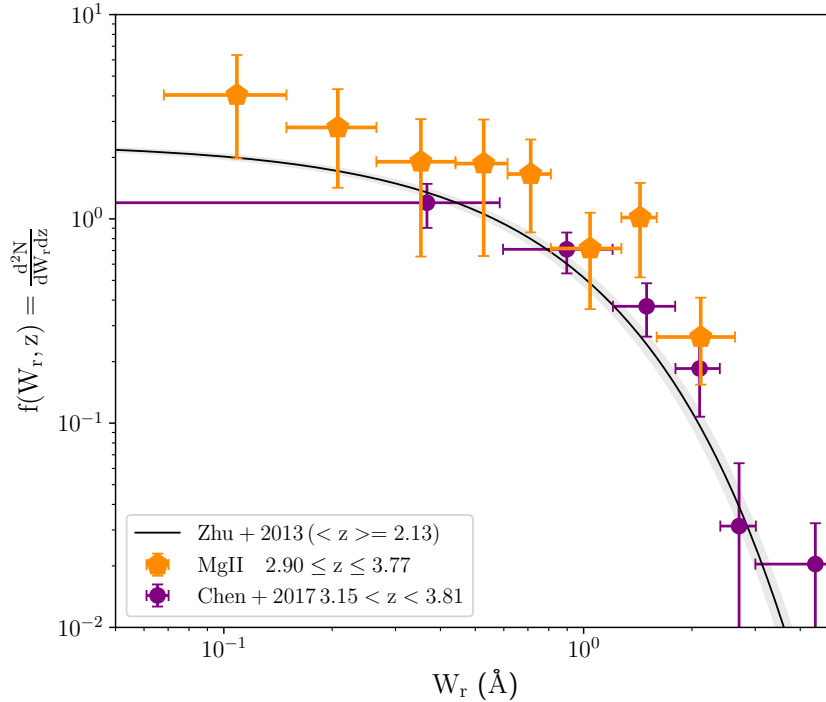


Figure 3.2: Equivalent width distribution function per unit redshift of the MAGG Mg II absorption line systems (orange pentagons). The 10th and 90th percentiles from the bootstrap resampling and the width of each bin of equivalent width are shown as vertical and horizontal error bars, respectively. The results for the MAGG sample are consistent with the completeness-corrected functions from [Zhu & Ménard \(2013\)](#) (black solid line) and [Chen et al. \(2017\)](#) (purple dots) down to $W_r \approx 0.1 \text{ \AA}$.

3.3.2 Catalogs of LAEs, groups, and C IV absorbers

In MAGG V we focused on studying how an intermediate ionized gas phase, traced by C IV absorption line systems, is linked to a population of low-mass galaxies identified as Ly α emitters in the MUSE cubes. We refer to MAGG V (section 3 and tables 1-5) for the detailed description of the identification and classification processes of the two samples, providing only the essential information here. Both catalogs are available as supplementary online material to MAGG V.

The galaxy sample (“MAGG LAEs” hereafter) consists of 921 high-confidence Ly α emitters observed at $2.81 \lesssim z \lesssim 6.60$ (median $z \approx 3.91$), with a median Ly α luminosity $\log[L_{\text{Ly}\alpha}/(\text{erg s}^{-1})] \approx 42.10$. This sample is designed to extend the study of the gas-galaxy connection down to typically low stellar and halo masses, on the order of $M_\star \approx 10^9 - 10^{10} M_\odot$ and $M_h \approx 10^{10} - 10^{11} M_\odot$, respectively (see e.g., [Ouchi et al., 2020](#)). The strategy we used to build this sample follows a two-step approach. For continuum-detected course, we identify rest-frame UV-bright sources by running SEXTRACTOR ([Bertin & Arnouts, 1996](#)) on the white-light images, and we used the output segmentation map to extract the spectra and derive a reliable measure of the redshift using the M. Fossati’s fork ³ of MARZ ([Hinton et al., 2016](#)) for 1200 sources. For line emitters, we first subtracted all the continuum sources in the MUSE cubes and searched for line emission running CUBEX to identify groups of ≥ 27 connected voxels with signal-to-noise ratio $S/N \geq 3$. We then visually inspected all the candidates to avoid contamination from cosmic rays and other emission lines at

³<https://matteofox.github.io/MarZ>

lower redshift, assigning high confidence to the sources detected at integrated signal-to-noise ratio $ISN \geq 7$ and lacking the presence of double peaks with separation consistent with [OII] emitters. The Ly α flux is measured using the curve of growth analysis in pseudo-narrow band images and then corrected for the Milky Way extinction using the map from [Schlafly & Finkbeiner \(2011\)](#) and the extinction law from [Fitzpatrick \(1999\)](#).

To characterize the galaxy environment, we identified 152 groups, broadly defined as ≥ 2 LAEs that are not isolated within a LOS separation of $|\Delta v| < 500 \text{ km s}^{-1}$, without any specific constraint on the halo mass. This evidence suggests that only a fraction of the galaxies in our sample (corresponding to ≈ 44 percent of the LAEs found in MAGG) appear to evolve in isolation. In contrast, for a large number of LAEs that reside in overdense regions, the gravitational and hydrodynamic interactions that take place in dense galaxy environments, as well as the large-scale structure in which groups reside, are found to affect the properties of different CGM gas phases in a large redshift interval (see e.g., [Fossati et al., 2019](#); [Dutta et al., 2020, 2021](#); [Muzahid et al., 2021](#); [Banerjee et al., 2023](#) and MAGG V). The sample of C IV systems identified in MAGG (“MAGG C IV” hereafter) includes 220 doublets ($\lambda\lambda 1548, 1550 \text{ \AA}$) at $3.0 \lesssim z \lesssim 4.3$ (median $z \approx 3.31$) identified in the high and medium resolution optical spectra available for each quasar of the survey. This sample is 50% complete at rest-frame equivalent widths $W_r^{\text{C IV}} \approx 0.08 \text{ \AA}$ for the strongest component at $\lambda_{\text{RF}} = 1548 \text{ \AA}$. The absorption strength of the systems identified in MAGG spans a broad range of equivalent width, from the weak systems at $W_r \lesssim 0.08 \text{ \AA}$ up to strongest ones with $W_r \gtrsim 0.30 \text{ \AA}$ and a median of $W_r \approx 0.09 \text{ \AA}$.

3.3.3 Overlap of Mg II with the H I and C IV samples in MAGG

The selection of Mg II, C IV and H I absorbers in MAGG has been conducted independently, but these samples partially overlap. Knowing the degree of overlap between various ions helps compare this analysis with the results of previous series papers. To quantify this effect, we matched the Mg II and C IV samples in redshift space within $|\Delta v| = 100, 150, 200, 250 \text{ km s}^{-1}$. The fraction of C IV absorbers matched to Mg II systems rises from ≈ 12 percent to ≈ 14 percent for increasing linking velocity, corresponding to $58 - 66$ percent of the Mg II lines sample ($28 - 31$ out of 47). In addition, independent of the linking velocity, $\gtrsim 90$ percent of the C IV systems matched to the Mg II absorbers show larger rest-frame equivalent width (black datapoints and histograms in Figure 3.3) and width in velocity space compared the median values of the full C IV sample. Consistently, [Steidel & Sargent \(1992\)](#) and [Schroetter et al. \(2021\)](#) found that the strength of C IV systems that are not matched to any Mg II absorber is peaked toward lower equivalent widths due to the ionized gas being more diffuse than a neutral and metal-enriched phase around galaxies.

We also caution that due to the difference in rest-frame wavelength, oscillator strength, and relative abundances, the Mg II absorption saturates at column densities ≈ 1 dex smaller compared to the C IV. At fixed rest-frame equivalent width, the two ions do not trace the same gas density. In addition, we used CLOUDY grids from [Fumagalli et al. \(2016\)](#), constructed by assuming photo-ionization equilibrium, to verify that at typical CGM and IGM densities ($\log[(n_{\text{H}}/\text{cm}^{-3})] \lesssim -2.5$) Mg II absorbers are stronger than the C IV systems, with a rest-frame equivalent width $W_{\text{Mg II}} \approx 2 \times W_{\text{C IV}}$ in the linear regime. These expectations are consistent with the equivalent width ratios we directly measure from the matched Mg II and C IV systems exhibiting a similar number of components, as detailed in Appendix B.1, resulting in $W_{\text{Mg II}} = (3.33 \pm 1.82) \times W_{\text{C IV}}$. In this appendix, motivated by the detailed comparison of multiple gas phases at $z \approx 2$ from [Rudie et al. \(2019\)](#), we also explore the differences in the kinematics between low and medium ions that are aligned in redshift space and likely arise from the same gas clouds.

Due to the severe contamination of the sky emission in the NIR and the lower spectral resolution compared to the optical quasar spectra in which we have identified the C IV doublets, the sample of the Mg II absorbers includes only stronger systems. To quantify the effect of the different selection functions between the two samples, we estimated the 3σ upper limits on the rest-frame equivalent width of the ions at the redshift of LAEs that are not connected to any absorber. Taking the median value and the spread given by the 16th and 84th percentiles of the upper limits distribution, we obtain $W_r^{3\sigma} = 0.02 \pm 0.01 \text{ \AA}$ for the C IV systems, ≈ 3 times lower and with a much smaller spread compared to the Mg II gas for which $W_r^{3\sigma} = 0.06 \pm 0.05 \text{ \AA}$.

Finally, to connect the cool enriched gas phase traced by the Mg II absorbers to the neutral and dense gas, we consider the sample of 61 strong ($N_{\text{H I}} \gtrsim 10^{16.5} \text{ cm}^{-2}$) hydrogen absorbers identified in MAGG by [Lofthouse et al. \(2023\)](#). As done for C IV, we investigate the degree of overlap between the samples of Mg II and H I absorbers. Depending on the choice of a linking velocity window ranging from 100 km s^{-1} to 250 km s^{-1} , we report that the $57 - 60$ percent (corresponding

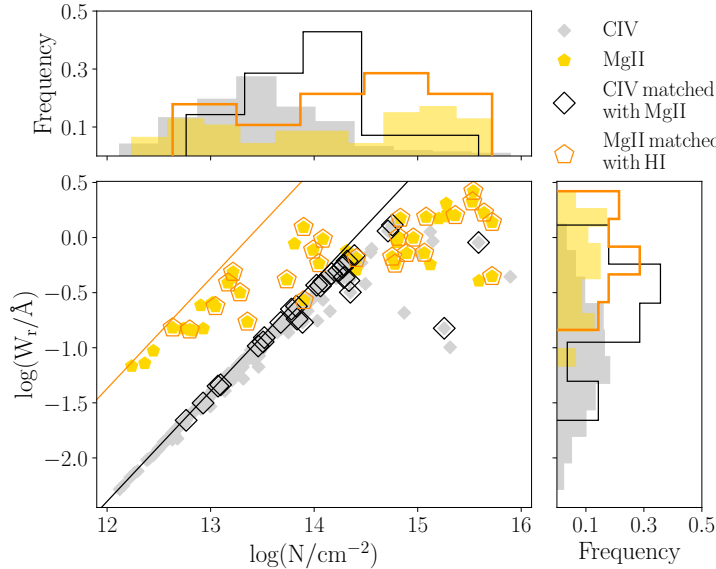


Figure 3.3: Column density as a function of the rest-frame equivalent width of the absorbers for the full sample of Mg II (gold) compared to those that overlap with the strong HI systems from MAGG IV (orange) and for the full sample of C IV lines from MAGG V (gray) compared to those matched to the Mg II sample (black). Solid lines show the equivalent width as a function of the column density in the linear regime of the curve of growth. The distributions for column density and the equivalent width are shown on the right and top panels. We matched the samples within a linking velocity window of $|\Delta v| = 150 \text{ km s}^{-1}$.

to 27/47 to 28/47) of the Mg II systems overlap with a strong HI. About 65 percent of these Mg II absorbers show a rest-frame equivalent width above the median of the entire sample (orange data points and histograms in Figure 3.3). At the same time, we do not observe any significant difference in the kinematics as traced by the profile width.

3.4 Results

3.4.1 Searching for clustering of LAEs around Mg II absorbers

We start our investigation into the physical connection between LAEs and the cool gas traced by the Mg II systems by cross-matching the two catalogs with velocity window $|\Delta v| \leq 1500 \text{ km s}^{-1}$. Figure 3.4 shows the galaxies' line-of-sight separation (left-hand panel) relative to the redshift of the Mg II absorbers (orange). A clear peak is visible in the velocity distribution, suggesting LAEs are locally clustered around the Mg II. To gauge the significance of this detection, we reproduce the expectations for random regions in the Universe that are not selected via Mg II absorption (light gray). As described in MAGG V, we achieve this by connecting LAEs observed in a random field to the Mg II absorbers found in a sightline randomly chosen between the remaining ones. We repeated this bootstrap procedure 10^3 times. We then model the peak observed for the LAEs-Mg II connections with a Gaussian function plus an additional term to account for the background. We list the best-fit parameters in Figure 3.4.

According to these results, the majority of the LAEs are found to be clustered to the Mg II absorbers around $\mu \approx 300 \text{ km s}^{-1}$ and within $2\sigma \approx \pm 256 \text{ km s}^{-1}$ (vertical dotted lines in Figure 3.4), while at larger velocities the distribution becomes consistent with the expectation for the random background. This also suggests that the velocity distribution is not symmetric around the redshift of the Mg II absorbers but significantly offset due to the systematic shift of the LAEs' redshift as a consequence of the resonant scattering of the Ly α photons. We thus correct the redshifts for the mean offset $z'_{\text{Ly}\alpha} = z_{\text{Ly}\alpha} - \mu(1 + z_{\text{Ly}\alpha})^{-1}c^{-1}$. Having corrected the LAE redshifts for the velocity offset, we consider LAEs within $\pm 500 \text{ km s}^{-1}$ relative to the Mg II absorbers as associated with the cool gas and included in the following analysis. With this criterion, we identify 43 LAEs connected to 18 Mg II absorbers, corresponding to a detection rate of $\approx 36 \pm 10$ percent (17/47) at 1σ confidence level. We also derive a first-order estimate of the amplitude of the clustering signal above the background by integrating the two distributions within the linking velocity window, finding a factor of ≈ 4 of LAEs near Mg II absorbers compared to random expectation. We reserve

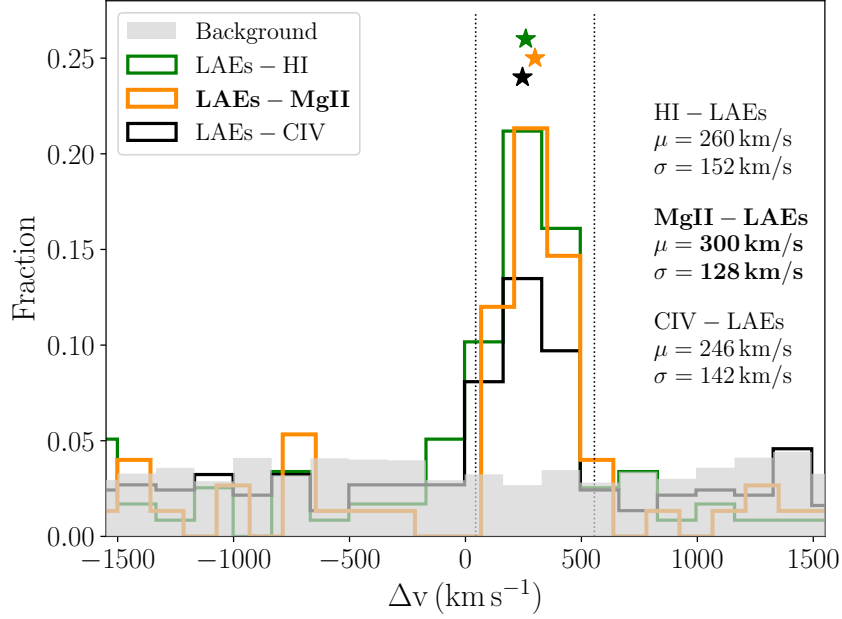


Figure 3.4: Distribution of LAEs within $|\Delta v| \leq 1500 \text{ km s}^{-1}$ of Mg II systems (orange) as a function of the line-of-sight separation. As a comparison, we also show the LAEs connected to C IV absorbers from MAGG V (black), those associated to strong H I absorbers from MAGG IV (green) and the distribution of a random sample representative of our expectations for the field (gray). The annotations show the best fit of the distributions considering a Gaussian function plus a background term, with the peaks shown as stars. The vertical dotted lines mark the 2σ interval around the peak, μ , of the Mg II-LAEs distribution, where the largest excess of galaxies is found.

a more statistical derivation of the overdensities of LAEs around the Mg II absorbers to the next section.

We compare these results with the redshift space connection between LAEs and C IV systems from MAGG V (black). In MAGG V we reported a detection rate of 36 ± 5 for the $z > 3$ C IV absorbers, consistent within 1σ with the result we obtained for the Mg II gas. The excess of LAEs within $\pm 500 \text{ km s}^{-1}$ from the Mg II absorbers is a factor $\approx \times 1.5$ larger compared to the C IV-LAE clustering amplitude, similar to the clustering of LAEs around strong H I absorbers from MAGG IV (green in Figure 3.4). Given the results from Section 3.3.3, we understand this finding because the Mg II sample significantly overlaps with optically thick absorbers that trace higher overdensities more correlated to galaxies. On the other hand, the overlap between Mg II and C IV absorbers is only partial and skewed to higher equivalent widths. The weak C IV absorbers that, based on MAGG V results, trace a diffuse enriched medium that is not always connected to galaxies dilute the clustering signal.

3.4.2 A lower LAE detection rate for Mg II gas

The rate at which we can find one or more galaxies close in velocity to the absorbers provides information about which type of galaxies are surrounded by different gas phases. We thus estimate the fraction of absorbers that are connected to at least one LAE within $\pm 500 \text{ km s}^{-1}$, as a function of their rest-frame equivalent width, $W_r \geq W_0$. As shown in Figure 3.5, the detection rate increases monotonically with the strength of the C IV absorbers (black). Instead, we do not find any significant trend for the Mg II gas (orange). To strengthen this analysis with a more statistical evidence, we perform the Spearman-rank correlation test. It returns the degree of monotonic correlation, expressed by the correlation coefficient ρ , and the probability that no monotonic correlation exists between the variables, p -value⁴. The Spearman-rank correlation test results in $\rho = 0.99$, p -value = 0.003 for the C IV and $\rho = -0.26$, p -value = 0.63 for the Mg II gas, supporting these conclusions. To further explore this difference, we focus on the strong $W_r^{\text{Mg II}} \geq 1 \text{ \AA}$ Mg II absorbers that have corresponding kinematically aligned C IV systems (see the Appendix B.1). In

⁴The closer the correlation coefficient ρ is to the values ± 1 , the stronger the monotonic correlation. $|\rho| < 0.20$ is considered a very weak correlation, while a correlation is strong for $|\rho| > 0.60$. On the other hand, the significance of the correlation depends on the p -value: if p -value < 0.05 , the correlation is significant at least at 2σ level.

this subsample, we observe that systems linked to at least one LAE have on average strong C IV equivalent width. Conversely, C IV absorbers with $W_r^{C\text{IV}} \lesssim 0.6 \text{ \AA}$ are less correlated to galaxies. In this case, the weakness of the C IV absorbers and the strength of the Mg II systems suggest either an excess of cool gas or a lower ionization; indeed, these systems largely overlap with the DLA population in MAGG.

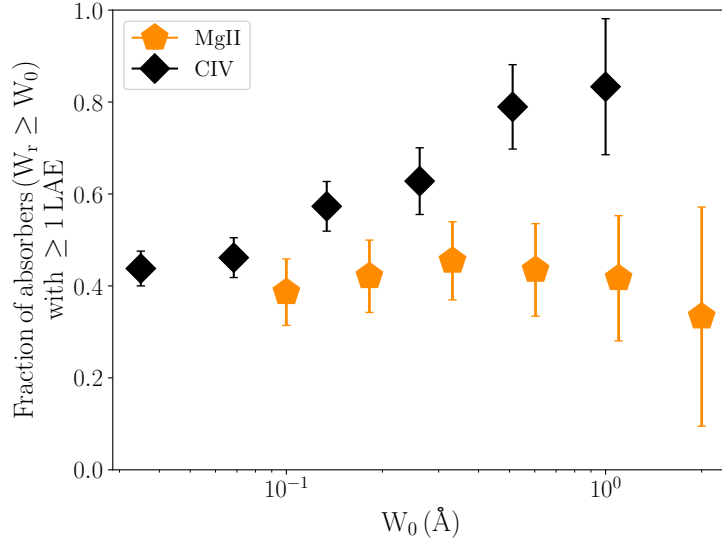


Figure 3.5: Fraction of Mg II (orange) and C IV (black) absorbers with equivalent width $W_r \geq W_0$ that are found to be connected to at least one LAE within $\pm 500 \text{ km s}^{-1}$ as a function of the threshold W_0 .

This lower detection rate in strong Mg II absorbers is unexpected, as this gas is largely believed to trace overdense and neutral regions in close proximity to galaxies. We thus explore the possibility that the galaxies connected to these strong absorbers lie at close separation from the quasar sightline (see, e.g., Fynbo et al., 2010, 2011; Nielsen et al., 2013; Schroetter et al., 2019; Guha & Srianand, 2023). We project 14 layers around the wavelength of the Ly α emission at the redshift of the absorbers (corresponding to about $\pm 1000 \text{ km s}^{-1}$, on average) to build pseudo-narrow band cut-outs around the central quasar of each field from the quasar point spread function (PSF) and continuum-subtracted datacubes. By visually inspecting these images, we did not find any Ly α emitting galaxies that were not identified by our selection due to their proximity or overlap with the quasar PSF up to impact parameters of 20 kpc.

Having excluded the presence of UV bright or Ly α emitting galaxies, a plausible explanation is the existence of a dusty galaxy population excluded from our UV selection. At present, there are no systematic investigations into a high-redshift obscured population near strong Mg II absorbers, an area of parameter space that should be explored with the Atacama Large Millimeter/submillimeter Array (ALMA) and the *James Webb* Space Telescope (JWST). At lower redshift, Ménard & Fukugita (2012) studied the dust content of strong $W_r \geq 0.8 \text{ \AA}$ Mg II absorbers. At $z \approx 0.5$, they found that it is similar to the dust content of galactic disks and showed that the Mg II absorbers carry $\approx 1/2 - 1/3$ of the total dust expelled by galaxies. Similarly, Budzynski & Hewett (2011) reported that in SDSS the ≈ 20 percent of $W_r \geq 1 \text{ \AA}$ Mg II systems are missed due to heavy dust obscuration. Once selected by high-column density hydrogen, Neeleman et al. (2017, 2018, 2019) reported a high detection rate of [C II] emitters around clumpy, neutral and enriched regions at $z \approx 4$, attributing the previous non-detection of galaxies to the significant dust obscuration. Systematic follow-up of a sample of strong Mg II absorbers with high resolution and at IR/submm frequencies will therefore be instrumental to test the presence of UV-dark galaxies, or galaxies lying closer than $\approx 0.8 \text{ arcsec}$ to the sightline.

3.4.3 Characterizing the LAE overdensity around Mg II absorbers

The peak in the line-of-sight separation between LAEs and Mg II absorbers suggests that the low-mass galaxies are preferentially clustered around regions harboring the cool gas rather than inhabiting random regions of the Universe. This aligns with results from earlier MAGG papers (IV and V), which consistently point to a preference for LAEs in tracing overdense cosmic structures selected by strong ions in quasar spectra. Following the methods of previous MAGG analyses, we

quantify this link with the luminosity function of LAEs connected to Mg II absorbers at $z \approx 3$. Referring to the formalism described by Fossati et al. (2021), we compare the nonparametric estimator $1/V_{\text{max}}$ (Schmidt, 1968; Felten, 1976) to a parametric Schechter (Schechter, 1976) fit of the unbinned data. Incompleteness is accounted for by performing injection-recovering simulations of 500 mock LAEs for 1000 iterations over all the MAGG fields. The lensed field J142438+225600 (Patnaik et al., 1992) is excluded from the analysis. As a final step, we compute the effective volume by considering only the velocity window we inspected to connect the absorbers to the galaxy. Once the LAE redshifts are corrected for the $\mu \approx 300 \text{ km s}^{-1}$ offset due to the radiative transfer effects, this corresponds to $\Delta v = \pm 500 \text{ km s}^{-1}$ around the redshift of each Mg II absorbers. We then bin all the galaxies with a step of 0.2 dex in luminosity, assuming that the luminosity function of LAEs does not depend on the redshift (Herenz et al., 2019). The nonparametric $1/V_{\text{max}}$ estimator and the parametric fit are both shown in Figure 3.6 (orange points and solid line, respectively). For the luminosity function of LAEs around the Mg II absorbers we obtain the following best-fit parameters for a Schechter (Schechter, 1976) function

$$\begin{aligned} \log[\phi^*/(\text{Mpc}^{-3}\text{dex}^{-1})] &= -2.307 \pm 0.760 \\ \log[L^*/(\text{erg s}^{-1})] &= 42.811 \pm 0.497 \\ \alpha &= -1.665 \pm 0.460. \end{aligned} \quad (3.1)$$

To quantify the amplitude of the clustering of LAEs within $\pm 500 \text{ km s}^{-1}$ around the Mg II absorbers, we integrate the luminosity function to compute the LAEs number density and compare it relative to the expectation for random regions of the Universe, represented by the ‘MAGG field’ sample that includes 617 galaxies at $3.0 < z < 4.5$, a range chosen to be consistent with the redshift distribution of the absorbers. In a similar way as done with the clustering in velocity space in Section 3.4.3, we thus estimate that the surroundings of the Mg II absorbers are a factor of ≈ 4 overdense in LAEs compared to random regions of the Universe.

Compared to different gas tracers (C IV and H I absorbers from MAGG IV, V) and the quasar population from Fossati et al. (2021), we observe no difference in the shape of the luminosity function, implying that the same galaxy population is identified around different tracers, with no excess in a given luminosity interval. Instead, we observe differences in normalization, as already noted in the previous MAGG papers. The LAEs near Mg II absorbers, in particular, have a $\approx \times 1.3$ lower normalization compared to H I systems. Therefore, the two populations largely overlap, which is unsurprising given that both ions trace a cool and partially neutral medium. LAEs linked to Mg II appear to have a marginally ($\approx \times 1.7$) higher normalization compared to around the C IV gas, which - as done in MAGG V - we attribute to the dilution in the signal induced by lower equivalent width C IV absorbers being less correlated to galaxies.

To complete the picture of the overdensity of LAEs around the Mg II gas, we follow the method described by Trainor & Steidel (2012) and computed the reduced angular cross-correlation function. Based on the random sample described in Section 3.4.1, we estimated the surface density of randomly distributed LAEs connected to Mg II absorbers within $|\Delta v| \leq 500 \text{ km s}^{-1}$ to be on the order of $\Sigma_{\text{rand}} = (8.160 \pm 2.896) h^2 \text{ cMpc}^{-2}$. We fit the data with the reduced angular cross-correlation function defined as:

$$\sigma_P(R) = (R/r_0)^\gamma \times {}_2F_1(1/2, \gamma/2; -3/2; -R_{\text{los},0}^2/R^2), \quad (3.2)$$

where R is the projected separation between the LAEs and the central quasar of each field, ${}_2F_1$ is the Gaussian hypergeometric function and $R_{\text{los},0}$ is half the amplitude of the redshift window we searched for gas-galaxies connection, in physical units (see Fossati et al., 2021). We fixed the power law index $\gamma = 1.8$, as done in similar studies (see e.g., Diener et al., 2017; García-Vergara et al., 2019; Fossati et al., 2021), to reduce the degeneracy with normalization. We derive a clustering length $r_0 = 1.53^{+0.43}_{-0.48} h^{-1} \text{ cMpc}$, ≈ 1.3 times higher than the C IV gas from MAGG V, but still consistent within 1σ uncertainties. The difference between the two gas phases is larger within $R \lesssim 250 h^{-1} \text{ ckpc}$ and decreases for increasing distance from the galaxies. A steeper power-law index, which we fix to $\gamma \approx 2.2$ in line with the best-fit model found by Dutta et al. (2021) for the Mg II absorbers at $z \lesssim 2$, yields a lower correlation length, $r_0 = 1.32^{+0.29}_{-0.33} h^{-1} \text{ cMpc}$, and better capture the cross-correlation for small angular separations. On the other hand, values of $\gamma > 1.8$ imprecisely model the flatter cross-correlation function of the C IV gas. As suggested by Dutta et al. (2021), this is consistent with the strong Mg II absorbers to be preferentially concentrated close to the associated galaxies compared to the more ionized gas phases. Indeed, we show in Figure 3.12 this seems to be the case also at higher redshift. In Figure 3.6, we also compare with

the other gas tracers introduced in previous MAGG papers, for which fitting parameters can be found in table 7 of MAGG V. Indeed, the clustering length of LAEs around the Mg II systems is consistent with those of LAEs clustered to LLSs and Si IV absorbers and $\approx \times 1.4$ smaller than the self-clustering of LAEs, although still consistent within 1σ level due to the large uncertainties and the limited statistics of the samples.

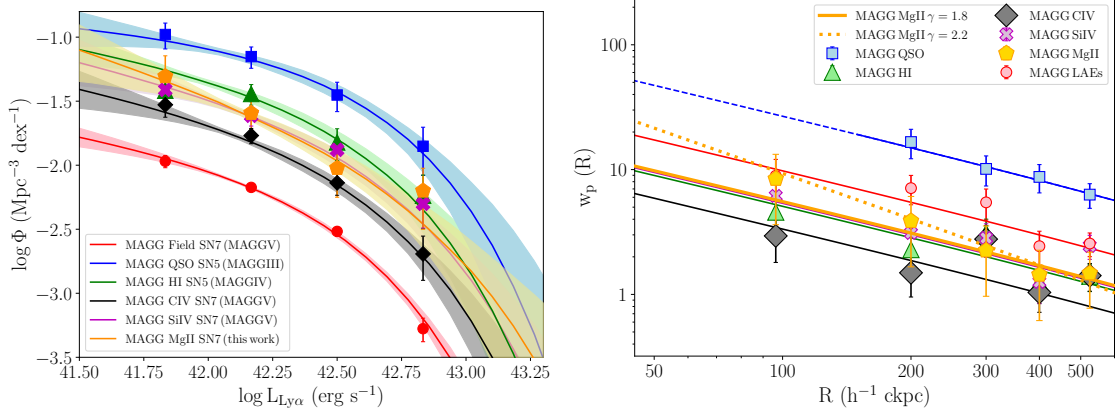


Figure 3.6: Clustering of the absorbers around LAEs. Left panel: LAE luminosity function within $|\Delta v| \leq 500 \text{ km s}^{-1}$ of the Mg II absorbers (orange) compared to the field control sample (red, MAGG V) and to the results for LAEs around other tracers studied in MAGG (high-column density hydrogen absorbers in green from MAGG IV, C IV, and Si IV systems in black and purple from MAGG V and quasars in blue from Fossati et al., 2021). We show the parametric (line) and nonparametric (data points) estimates for each sample. Right panel: Projected cross-correlation function for LAEs within $|\Delta v| \leq 500 \text{ km s}^{-1}$ of the Mg II absorbers (orange) assuming $\gamma = 1.8$ (solid line) and $\gamma = 2.2$ (dotted line). The other populations explored in MAGG are shown with the same colors adopted in the left panel, while the LAEs auto-correlation function derived in MAGG V is shown in red. Vertical error bars are derived by bootstrapping the redshift of the Mg II absorbers over 1000 iterations.

Combining all these findings from the MAGG survey, we find that the number density of LAEs surrounding the C IV absorbers is lower relative to the Mg II and the LLSs at any projected separation. This is expected considering that the H I and the Mg II absorbers trace a similar cool gas phase, although with different densities, and that the MAGG Mg II sample is selecting preferentially strong partially neutral systems. In MAGG V, we attributed the difference in the ionized phase to the weaker C IV absorbers, which appear less connected to galaxies and to the denser medium selected by lower-ionization ions. These results suggest that the LAE number density around different tracers depends on the absorbers’ strength and ionization, which vary with different gaseous components, such as the galaxies’ inner CGM, the overdense gas along the filaments, or an enriched and diffuse medium at larger distances. Moreover, different gas tracers are not fully independent: within this picture, LAEs - by virtue of their high number density - become themselves tracers of cosmic structures on a range of scales, containing multiple galaxy populations and gas phases.

3.4.4 The effect of the large-scale galaxy environment

So far, we have established and explored the physical connection between LAEs and different gas phases traced by Mg II and C IV absorbers. The IFU observations included in the MAGG survey identified more than 150 groups of LAEs across the 28 quasar fields (see Section 3.3.2). Throughout this work, we recognize as part of a group LAEs with at least one other galaxy within $\pm 500 \text{ km s}^{-1}$ in redshift space. This makes it possible to introduce an additional dimension in studying the gaseous surroundings of LAEs: the large-scale galaxy environment. In previous works, it has been shown that both the strong H I (Muzahid et al., 2021; Lofthouse et al., 2023) and the C IV absorption (Muzahid et al., 2021; Galbiati et al., 2023; Banerjee et al., 2023) is enhanced in groups compared to around isolated LAEs at $z \approx 3 - 4$. At lower redshift $z \approx 0.5 - 2.0$ a similar effect has been observed in MUSE surveys (e.g., Péroux et al., 2019; Fossati et al., 2019; Dutta et al., 2020, 2021).

With our new Mg II sample, we complete this picture by examining the role played by the large-scale galaxy environments on the cool gas and compare it to the effect on the more ionized phase traced by the C IV systems. Out of all the Mg II absorbers associated to LAEs within a line-of-sight

separation $|\Delta v| \leq 500 \text{ km s}^{-1}$, we find that the 65 ± 11 percent is connected to galaxies that are part of a group, while only the remaining fraction is associated to isolated ones. In this section, we explore the impact of the large-scale environment on different gaseous phases as a function of the galaxies' properties and the projected separation.

Covering fraction of different gas phases

We compute the cumulative covering fraction, $f_c(\leq R)$, which is defined as the probability of detecting absorbers with $W_r \geq W_0$ within a projected separation R and line-of-sight velocity in the range $|\Delta v| \leq 500 \text{ km s}^{-1}$ around a galaxy. We split the sample of galaxies connected to the Mg II absorbers into groups (34 LAEs that are part of 11 different groups) and isolated (6 LAEs). In computing the covering fraction, we only consider LAEs connected to Mg II absorbers with a rest-frame equivalent width above the threshold $W_0 = 0.10 \text{ \AA}$, that is, the completeness limit. The 1σ uncertainties account for the 68 percent Wilson-score confidence intervals. In MAGG IV, as sightlines were selected to host H I absorbers, we accounted for this selection bias by renormalizing the number of systems per unit redshift by the corresponding cosmological value, $\ell(X) = dN/dX$. For metal absorbers, due to the higher number of systems per sightline and the fact that we do not explicitly select these populations, we find that the incidence of Mg II and C IV absorbers in MAGG is comparable to the cosmic values within uncertainties, and hence we do not apply additional renormalizations.

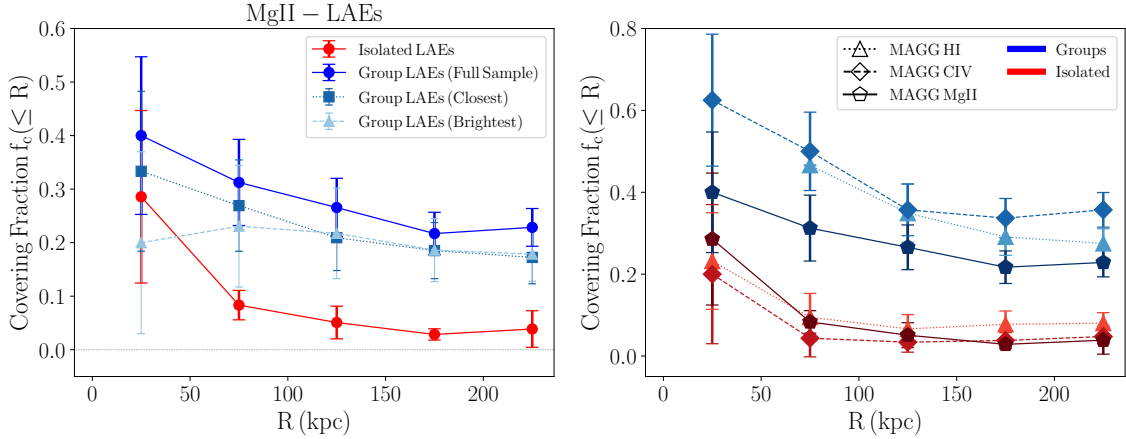


Figure 3.7: Covering fractions around isolated LAEs and groups. Left panel: Cumulative covering fraction of LAEs connected to $W_r \geq 0.10 \text{ \AA}$ Mg II absorbers as a function of the projected separation. We distinguish between isolated (red) and group (shades of blue) galaxies. The results obtained considering only the brightest (triangles) and the closest (squares) member of each group are consistent with each other and with the full sample within 1σ . Right panel: Comparison between the cumulative covering fraction for isolated (red) and group (shades of blue) LAE around the strong H I hydrogen absorbers (triangles), the C IV (diamonds) and Mg II (pentagons) systems, assuming $N_{\text{H I}} > 10^{17.5} \text{ cm}^{-2}$ and $W_r \geq 0.30 \text{ \AA}$.

Results in Figure 3.7 (left panel) show that the cool gas traced by the Mg II absorption is ≈ 4 more likely to be found around galaxies in groups than around isolated LAEs. At $R \leq 100 \text{ kpc}$ the covering fraction is $f_c = 0.27 \pm 0.09$ around group LAEs considering only the closest member, and $f_c = 0.08 \pm 0.03$ around isolated LAEs, decreasing to $f_c = 0.17 \pm 0.05$ and $f_c = 0.04 \pm 0.03$ at $R \leq 250 \text{ kpc}$ for the two subsamples, respectively. Taking either the full sample of groups or the brightest component leads to results that are all consistent within 1σ (see the left panel in 3.7). The observed excess of Mg II absorption in groups is consistent with the findings at $z \lesssim 2$ by Nielsen et al. (2018b) and Dutta et al. (2020, 2021). Section 3.5.2 discusses more details on the redshift evolution.

At $z \lesssim 2$, Dutta et al. (2021) has shown that the galaxy environment impacts the gas distribution differently according to the gas phase. We thus compare explicitly the cool and enriched gas phase traced by the Mg II-bearing gas, the neutral phase traced by strong H I systems, and the more ionized gas traced by C IV (right panel in Figure 3.7). For this analysis, we set the equivalent width threshold to $W_0 = 0.3 \text{ \AA}$ to limit the effects of the different selection functions of the Mg II and C IV samples. This corresponds to a difference in column density of $\log(N_{\text{C IV}}/\text{cm}^{-2}) - \log(N_{\text{Mg II}}/\text{cm}^{-2}) \approx 1$ dex in linear regime. The probability of observing any of these absorption systems is $\approx 3 - 4$ times higher in groups compared to isolated LAEs, regardless

of the choice of the subsample of groups. In detail, if, on the one hand, the C IV and H I gas show a similar covering fraction around group LAEs, on the other hand, that of the Mg II absorbers is consistent with the other tracers within 1σ confidence level, but ≈ 25 percent lower despite the large degree of overlap among the different samples in MAGG. This may depend on the equivalent-width threshold, which selects relatively strong C IV (corresponding to the 84th percentile of the C IV equivalent width distribution), and intermediate equivalent-width Mg II absorbers (25th percentile). However, at $z \lesssim 2$ Dutta et al. (2021) found that the Mg II covering fraction is ≈ 2 times lower and steeper compared to that of the C IV gas, which flattens at large impact parameters. They interpreted these differences as due to the warm and ionized C IV gas being more extended around galaxies than the Mg II cold phase. Altogether, these results support a picture in which the effect of the large-scale galaxy environments on the surrounding gas depends on the gas phase.

Motivated by the previous studies at lower redshift and by our findings in MAGG V for LAEs around the C IV absorbers, we compare the properties of groups to those of the isolated LAEs to verify if any intrinsic difference in the Ly α luminosity or projected separation of these two galaxy populations may be responsible for enhancing the metal absorption in LAEs overdensity. To this end, we performed a KS-test, finding a p -value $\gtrsim 0.40$ and p -value $\gtrsim 0.13$ (p -value $\gtrsim 0.54$ and p -value $\gtrsim 0.36$), depending on the different subsample of groups, when comparing the impact parameter and the Ly α luminosity of isolated galaxies and of those in groups that are connected to Mg II (C IV) absorbers, respectively. These values suggest no statistical difference between the distributions of luminosity and impact parameters of groups and isolated LAEs. This result clears the way for two possible scenarios to explain the excess of absorption in groups: the cooler Mg II gas may be either affected by the large-scale structure in which groups reside, as found for the ionized C IV phase at $z \gtrsim 3$ in MAGG V (see also Muzahid et al., 2021), or is more sensitive to the local galaxy environment and shaped by the gravitational and hydrodynamic interactions that occur in galaxy overdensities, as found for the Mg II-bearing gas at $z \lesssim 2$ (see, e.g., Dutta et al., 2021).

Environmental effects on the absorption properties

Having observed that the galaxy environment affects the covering factor near LAEs, we now investigate whether the number of LAEs connected to each individual Mg II and C IV system plays any role in shaping the strength of the absorption and the kinematics of different gas phases. In Figure 3.8, we show the rest-frame equivalent width (upper panel) and Δv_{90} (lower panel) as a function of the number of LAEs within $\pm 500 \text{ km s}^{-1}$ of the absorbers (gray points). The medians (data points) and standard deviations (error bars) are also shown. In MAGG V, we showed how C IV systems connected to ≥ 2 LAEs are, on average, stronger and mildly more kinematically complex than those associated with ≤ 1 galaxy. Taking into consideration that the sample of Mg II lines tends to select stronger systems, we observe a trend where both the strength and the kinematics of the Mg II absorbers rise with the number of connected LAEs, exceeding those of the C IV systems in the richest groups.

To support this evidence, we performed a nonparametric Spearman rank correlation test to quantify the degree of monotonic correlation between the properties of the absorbers and the number of associated LAEs, together with the probability of no correlation. For the Mg II absorbers, we measure correlation coefficients of $\rho \approx 0.30$ and p -value ≈ 0.045 and $\rho \approx 0.27$ and p -value ≈ 0.049 for the equivalent-width and Δv_{90} . For the C IV systems, we find $\rho \approx 0.39$ and p -value ≈ 0.001 and $\rho \approx 0.22$ and p -value ≈ 0.002 for the equivalent-width and Δv_{90} . The positive correlation of the strength and the kinematics of the Mg II and C IV lines with the number of LAEs is thus statistically significant for the two ions. At $z \approx 0.5$, Lan et al. (2014) similarly found a positive correlation between the strength of the Mg II absorbers and the mean number of star forming-galaxies at small impact parameters. Likewise, at $z \lesssim 1$, Dutta et al. (2021) reported that the strength and the incidence of the Mg II absorption increase with the galaxies number overdensity (see also Dutta et al., 2020) but found only a weaker dependence of the C IV covering fraction and equivalent width on the large-scale galaxy environment (see Section 3.5.2 for a detailed discussion).

Besides the overdensity of galaxies clustered to the absorbers, we also study whether the intrinsic property of Ly α luminosity shows trends with the absorption characteristics of different gas phases. We thus directly correlate, in Figure 3.9, the strength of the absorption (upper panels) and the kinematics (bottom panels) with the luminosity of isolated galaxies (red) and groups (blue). Reported in Table 3.2, as a statistical diagnostic, are the results of the nonparametric Spearman's rank correlation test measuring the degree of monotonic correlation, ρ , and the probability that the two variables are not correlated, the p -value. We tested the correlation considering for each

group the brightest galaxy and the one closest to the line-of-sight separately, as we noticed that they coincide only in 2/11 cases for LAEs around the Mg II and in 12/34 cases for those around the C IV absorbers).

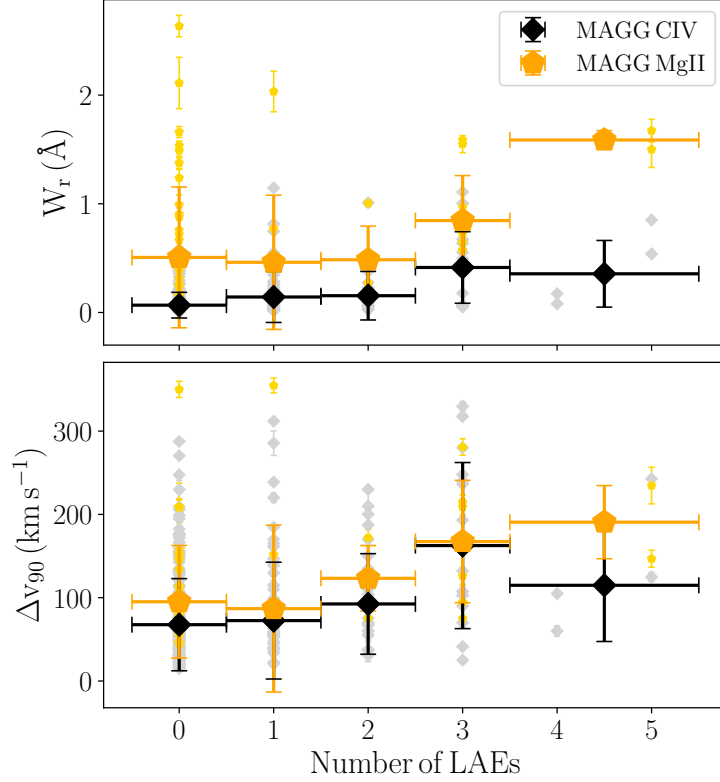


Figure 3.8: Effect of the large-scale galaxy environment on the absorption properties. We show the absorbers’ rest-frame equivalent width (upper panel) and the width in velocity space (lower panel) as a function of the number of LAEs connected to the absorbers. Individual C IV-LAE (gray) and Mg II-LAE (gold) connections are shown, as well as the binned median and standard deviation (black and orange, respectively).

Line	LAEs	W_r (Å)	Δv_{90} (km s $^{-1}$)
Mg II	Closest	$\rho = 0.43$ $p = 0.05$	$\rho = 0.34$ $p = 0.31$
	Brightest	$\rho = 0.54$ $p = 0.03$	$\rho = 0.04$ $p = 0.89$
	Full	$\rho = 0.06$ $p = 0.72$	$\rho = 0.47$ $p = 0.64$
	Isolated	$\rho = 0.89$ $p = 0.02$	$\rho = 0.31$ $p = 0.54$
C IV	Closest	$\rho = 0.43$ $p = 0.01$	$\rho = 0.42$ $p = 0.01$
	Brightest	$\rho = 0.33$ $p = 0.06$	$\rho = 0.33$ $p = 0.06$
	Full	$\rho = 0.26$ $p = 0.01$	$\rho = 0.29$ $p = 0.01$
	Isolated	$\rho = 0.16$ $p = 0.33$	$\rho = 0.01$ $p = 0.95$

Table 3.2: Correlations between absorbers’ properties and Ly α luminosity of isolated LAEs and groups. Results of the Spearman’s test on the correlation between the rest-frame equivalent width and the kinematics, traced by the width in velocity space, of the Mg II and C IV absorbers and the Ly α luminosity of isolated galaxies and groups. In the case of multiple galaxies connected to the same absorber, we consider the full sample, the brightest, or the closest member only.

Both the strength and the kinematics of the C IV absorbers show a tentative ($\approx 2\sigma$ level) positive correlation ($\rho \gtrsim 0.25$ and p -value $\lesssim 0.06$) with the luminosity of galaxies in groups, without any significance dependence on the chosen subsample. On the other hand, the strength of the Mg II absorption is marginally sensitive ($\rho \gtrsim 0.40$ and p -value $\lesssim 0.06$) to the luminosity of the closest and the brightest component of groups, while no strong dependence is found (p -value $\gtrsim 0.30$) for the kinematics of the cool gas. Finally, we observe that the Mg II equivalent width also positively correlates with the luminosity of isolated galaxies ($\rho = 0.89$ $p = 0.02$). To better understand the underlying physical motivation for these results, we recall that a KS test yields a p -value $\approx$

0.13, 0.40, 0.85 (p -value $\approx$ 0.92, 0.71, 0.36) when comparing the luminosity of isolated LAEs to that of the brightest, closest component and the full sample of groups, respectively, around the Mg II (C IV) absorbers. This is indicative of a high likelihood that the distributions of the two galaxy populations are similar and originate from the same parent one (see also [Nielsen et al., 2018b](#)), consistent with our findings for LAEs around the C IV gas and the high degree of overlap between the two samples.

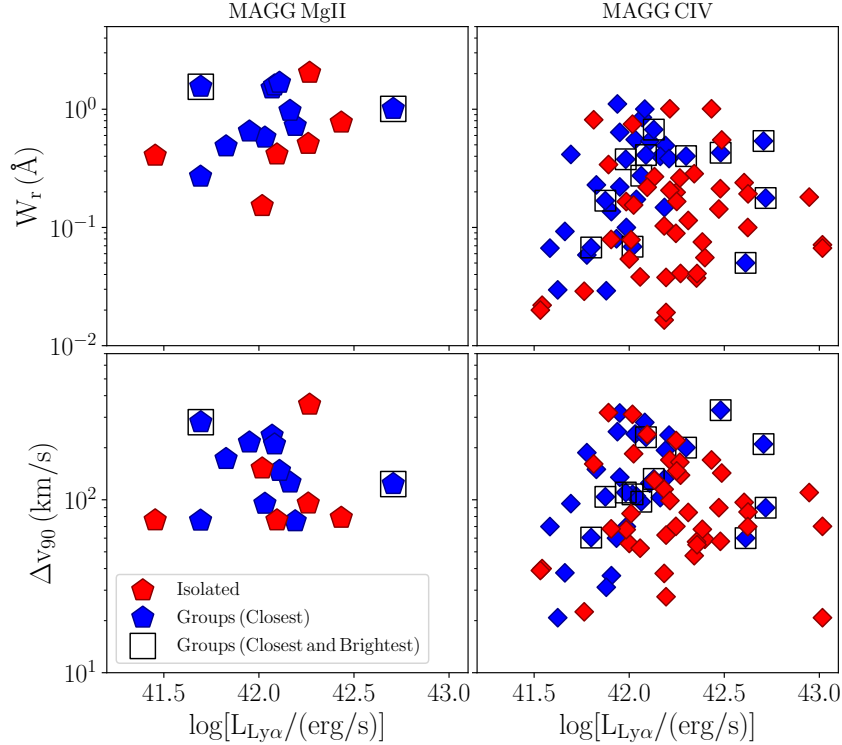


Figure 3.9: Mg II (left panels) and C IV (right panels) absorbers’ rest-frame equivalent width (top panels) and width in velocity space (bottom panels) as a function of the Ly α luminosity for isolated galaxies (red) and the closest member of each group (blue). Empty squares mark the case in which the closest galaxy of a group is also the brightest. The results from the Spearman rank test to assess the presence and the significance of possible correlations are listed in Table 3.2.

In light of this, since there is no statistical difference between the Ly α luminosity of isolated LAEs and groups, the possible dependence of the Mg II equivalent width on the luminosity of both the groups and the isolated galaxies can be interpreted as a hint that the absorption strength of the cool gas is sensitive to the galaxies’ activity and the local environment. On the contrary, the more ionized C IV gas, which does not correlate significantly with isolated LAEs, seems more affected by the large-scale environment.

3.5 Discussion

3.5.1 Evolution of the covering fraction with redshift

The distribution of the cool and metal-enriched gas traced by Mg II absorbers around star-forming galaxies has been extensively studied at $z \lesssim 2$, but less so at times closer to the peak of the cosmic star formation rate density and in comparison to other gas phases, such as the more ionized gas traced by the C IV systems. The statistical sample of $z \approx 3$ Mg II-LAEs systems surveyed in this work provides a key step forward in our understating of the evolution of the multiphase gas around galaxies at $z \gtrsim 3$. Many authors (e.g., [Turner et al., 2014](#); [Rudie et al., 2019](#); [Schroetter et al., 2021](#); [Dutta et al., 2021](#)) compared the incidence of different gas phases around star-forming galaxies at $z \lesssim 2$, with [Schroetter et al. \(2021\)](#) and [Dutta et al. \(2021\)](#) directly focusing on the comparison between Mg II and C IV gas. The picture emerging from these studies is that the cool gas is concentrated at close separations from the galaxies, while the gas phase traced by higher ions is extended up to larger distances. However, this scenario is less clear with redshift, mostly due

to the lack of large surveys of the Mg II-galaxies connection at $z \gtrsim 2$. A first step in this direction comes from [Lan \(2020\)](#), who proposed an analytical model of the evolution of the Mg II covering fraction as a function of redshift and stellar mass. By calibrating the model with a sample of 15,000 pairs at $z \approx 0.4 - 1.3$, they found that the covering fraction of strong ($W_r \geq 1 \text{ \AA}$) absorbers evolves as $(1+z)^\alpha$ with $\alpha \approx 2$, following a trend that is similar to the evolution of the cosmic star formation rate density at $z \lesssim 2$ (see, e.g., [Prochter et al., 2006](#); [Matejek & Simcoe, 2012](#); [Zhu & Ménard, 2013](#); [Chen et al., 2017](#)). The evolution of weak ($W_r = 0.4 - 1 \text{ \AA}$) systems is less significant. However, several studies ([Steidel et al., 2010](#); [Bordoloi et al., 2011](#); [Chen, 2012](#); [Schroetter et al., 2021](#); [Dutta et al., 2020](#)) did not find a significant redshift evolution when considering the contribution of weak and strong systems.

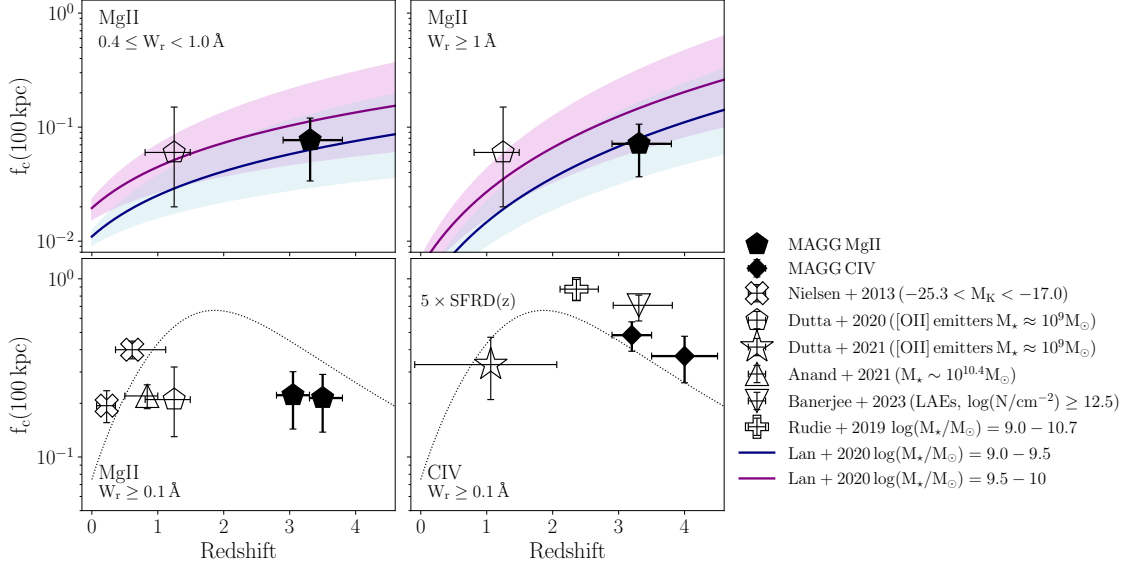


Figure 3.10: Mg II and C IV cumulative covering fraction within $R < 100 \text{ kpc}$. Upper panels: The results for strong ($W_r \geq 1 \text{ \AA}$, left panel) and weak ($0.4 \leq W_r < 1.0 \text{ \AA}$, right panel) Mg II absorbers observed around LAEs in MAGG (black) are consistent with the redshift evolution model calibrated at $z \lesssim 1.3$ by [Lan \(2020\)](#) for star-forming galaxies with stellar mass $\log(M_*/M_\odot) = 9.0 - 9.5$ and $\log(M_*/M_\odot) = 9.5 - 10$. Also shown is the covering fraction of Mg II gas observed around [O III] emitters by [Dutta et al. \(2020\)](#) at $z \approx 1$. Lower panels: Results for $W_r \geq 0.1 \text{ \AA}$ Mg II (right panel, black pentagon) and C IV (left panel, black diamond) samples, without any distinctions between weak and stronger absorbers. Also shown are the results from [Nielsen et al. \(2013\)](#) (crosses), [Dutta et al. \(2020\)](#) (pentagon) and [Anand et al. \(2021\)](#) (up triangle, $W_r \geq 0.4 \text{ \AA}$), [Dutta et al. \(2021\)](#) (star), [Rudie et al. \(2019\)](#) (cross, $\log(N/\text{cm}^{-2}) \geq 13.0$ corresponding to $W_r \gtrsim 0.04 \text{ \AA}$ in linear regime), and [Banerjee et al. \(2023\)](#) (down triangle, $\log(N/\text{cm}^{-2}) \geq 12.5$) corresponding to $W_r \gtrsim 0.01 \text{ \AA}$). Horizontal error bars mark the explored redshift range. The evolution of the C IV covering fraction is consistent with the star formation rate density as parametrized by [Madau & Dickinson \(2014\)](#) (dotted line, arbitrary normalization).

The sample of $z \approx 3 - 4$ Mg II lines presented in this work, along with the dense and complete set of LAEs identified in MAGG, offers an ideal dataset for testing the analytical predictions from [Lan \(2020\)](#), which we extrapolated from their results based on observations at $z \approx 0.4 - 1.3$ to higher redshifts. To this end, we estimate the covering fraction of both strong and weak Mg II absorbers around LAEs within a projected separation $R \lesssim 100 \text{ kpc}$: $f_c(100 \text{ kpc})$. In the following, we consider the entire sample of LAEs, and study in the next section 3.5.2 the evolution of the covering fraction of different gas phases around group and isolated galaxies. In the upper panels in Figure 3.10 we show that the $f_c(100 \text{ kpc})$ estimated for the Mg II gas is consistent, within the uncertainties, with the predictions from [Lan \(2020\)](#) for galaxies of stellar mass within the range $9.0 < \log(M_*/M_\odot) < 9.5$ and $9.5 < \log(M_*/M_\odot) < 10$, corresponding to typical stellar mass values expected for Ly α emitters ([Ouchi et al., 2020](#)). However, when incorporating the results from [Dutta et al. \(2020\)](#) at $z \approx 1$, we find that these are consistent with the covering fraction estimated at $z \approx 3$. While this could be in tension with the evolutionary trend suggested by [Lan \(2020\)](#) especially for $W_r \geq 1 \text{ \AA}$, favoring a scenario where the cross section of cool Mg II gas around galaxies remains relatively stable over cosmic time, it is important to note that large uncertainties in both observations and modeling prevent firm conclusions.

Complementary to the analysis above, we considered the covering fraction of Mg II absorbers

that are selected in equivalent width down to the completeness limit. In the lower-left panel of Figure 3.10, we show the results for $W_r \geq 0.1 \text{ \AA}$ Mg II lines, corresponding to the full sample of absorbers connected to LAEs in MAGG, which were computed in two redshift bins centered at $z = 3.05$ and $z = 3.55$. To explore the time evolution, we compared to lower redshift observations (Nielsen et al., 2013; Dutta et al., 2020; Anand et al., 2021) obtained with the same threshold on the absorbers' equivalent width. The incidence of the cool gas was derived around star-forming galaxies with stellar masses on the order of $M_\star \approx 10^9 M_\odot$ from Dutta et al. (2020) and $M_\star \approx 10^{10.5} M_\odot$ in all the other cases, which corresponds to $\approx 0.5 - 1.0$ dex above the typical mass of LAEs (except for the sample of Nielsen et al., 2013 only characterized by an absolute K-band magnitude, a proxy for galaxies' stellar mass, ranging between $-25.3 < M_K < -17.0$). This comparison shows no evolution across cosmic time, with $f_c(100 \text{ kpc}) \approx 0.2$ at all redshifts. The current results do not rule out an evolution between $z \approx 1.5 - 3$ due to lack of observations in this redshift range.

We subsequently investigate how the cool and enriched gas distribution evolves as a function of redshift compared to the more ionized gas phase traced by C IV absorbers (lower panels in Figure 3.10). Schroetter et al. (2021) modeled the redshift evolution of the C IV covering fraction based on the connection of 19 C IV systems to [O II] emitters at $z \approx 1 - 1.5$ and did not find any significant evolution across cosmic time. However, they observed that if, on the one hand, the size of the Mg II halo becomes progressively smaller with the cosmic time relative to that of the dark-matter halo, on the other hand, the evolution of C IV is more similar to that of the dark matter. At a higher redshift, Banerjee et al. (2023) found that the covering fraction of C IV absorbers around LAEs, considering either the full sample or the isolated ones, does not evolve within a small redshift range between $z \approx 2.9 - 3.8$ for C IV systems with $\log(N/\text{cm}^{-2}) \geq 12.5$ (corresponding to $W_r \gtrsim 0.01 \text{ \AA}$ in the linear regime). However, by combining the results from MAGG V in two bins of redshift (centered at $z \approx 3.2$ and $z \approx 4$), to those at $z \approx 1 - 2$ and $W_r \geq 0.1 \text{ \AA}$ from Dutta et al. (2021) and at $z \approx 2$ from Rudie et al. (2019) (assuming $\log(N/\text{cm}^{-2}) \geq 13.0$ which correspond to $W_r \gtrsim 0.04 \text{ \AA}$ in the linear regime), we observed that the C IV covering fraction increases with time up to $z \approx 2$ and then decrease at $z \lesssim 2$.

In support of this result, Dutta et al. (2021) found that, within the virial radius, the covering fraction of the C IV gas increases by a factor of ≈ 4 from $z \lesssim 1$ to $z \gtrsim 1$, possibly driven by the redshift evolution of the gas density and the ionizing background or due to the galaxies at higher redshift having higher star formation rates. Indeed, we observe that the C IV covering fraction increases from the local Universe up to $z \approx 2$, close to the peak of the cosmic star formation rate density (Madau & Dickinson, 2014) (Figure 3.10), and then decreases. Conversely, the covering factor for the Mg II systems remains constant across cosmic time. As mentioned earlier, however, the large uncertainties in the observations allow the results at different redshifts to remain consistent, not ruling out a scenario where no evolution occurs. Moreover, additional uncertainty arises from the lack of a uniform equivalent width threshold across these studies, which is known to influence the measurement of the covering fraction (MAGG V and Banerjee et al., 2023), and from the possibility that the samples could probe different galaxy populations at different redshifts.

Our observations suggest that the covering fraction of a cool and enriched gas phase around star-forming galaxies remains constant, within the uncertainties, throughout cosmic time, regardless of whether we consider strong or weak subsamples of Mg II absorbers. The covering fraction of C IV absorbers seems to resemble more the cosmic star formation rate density evolution. Even larger surveys are needed in the future to reduce uncertainties and test against the scenario of no evolution, which is not yet completely ruled out.

3.5.2 The effect of the large-scale galaxy environment across cosmic time

We have shown in Figure 3.7 that, regardless the gas phase they are tracing, both the Mg II and the C IV absorption is enhanced around groups, relative to isolated LAEs, at any projected separation up to $R \approx 250 \text{ kpc}$. How does the large-scale galaxy environment affect the covering fraction and the distribution of different gas phases across cosmic time is still an open question. To address this, we now compare the results of Section 3.4.4 with the findings from Dutta et al. (2021), who studied the environmental effects on different gas phases at $z \lesssim 2$ (see Figure 3.11). These authors found that $W_r \geq 0.1 \text{ \AA}$ Mg II absorption-line systems are detected around groups with a probability $\approx 2 - 3$ times higher compared to isolated galaxies, which is consistent with our findings at higher redshift. This result is proved to be valid even after matching the group and isolated samples in redshift, stellar mass and impact parameter, indicating that the environmental factors likely lead to the observed difference. Indeed, the Mg II covering fraction around the group and isolated LAEs at $z \approx 3$ is consistent within 1σ uncertainties with the $z \lesssim 2$ results around [O II] emitters. Several

mechanisms have been invoked at low redshift to explain this absorption excess. Modeling groups as a superposition of multiple haloes of isolated-like galaxies (see, e.g., Bordoloi et al., 2011) does not always succeed in reproducing the observations (Nielsen et al., 2018b; Dutta et al., 2020), suggesting that effects such as hydrodynamic and gravitational interactions or intragroup medium should not be neglected. Indeed, Dutta et al. (2021) interpreted the enhanced Mg II absorption in groups at $z \approx 1$ as possibly due to stripping processes driven by the intra-group medium. Mergers and tidal interactions are also able to perturb the cool gas, especially at $z \gtrsim 1$, when the occurrence of mergers is higher (see, e.g., Conselice et al., 2008; Conselice, 2014). However, we do not expect these processes to be equally efficient in the Universe at $z \approx 3$, when filamentary structures have not virialized yet. Although the results shown in Section 3.4.4 seems to suggest that the Mg II gas is more sensitive to galaxies' individual halos in groups compared to the C IV, what is responsible for boosting the covering fraction of cool gas at high redshift and sustaining its abundance for several billions of years still remains an open question.

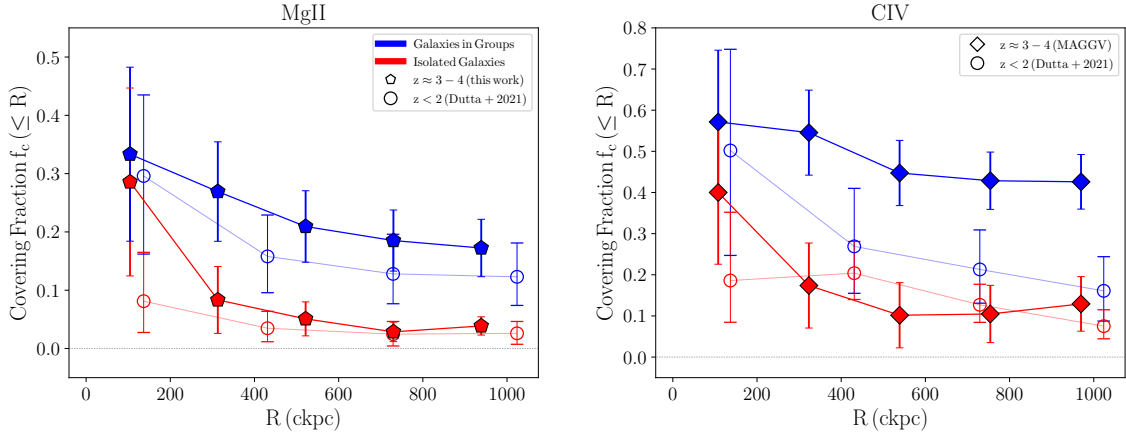


Figure 3.11: Comparison of the covering fraction of Mg II (left panel) and C IV (right panel) absorbers with $W_r \geq 0.1 \text{ \AA}$ detected around group and isolated galaxies in this work (filled points) with the results at $z \lesssim 2$ from Dutta et al. (2021) (empty points). The distance has been converted into co-moving units at the mean redshift of the samples ($z = 3.31$ in this work and $z = 0.95$ for Dutta et al., 2021). The covering fraction is estimated for the closest LAE at $z \approx 3-4$ and the most massive galaxy in groups at $z \lesssim 2$. The scales on the y-axis are different in the two panels.

A different picture emerges when considering the redshift evolution of the more ionized C IV gas in different galaxy environments. In MAGG V, we observed that the C IV absorbers at $z \gtrsim 3$ show a $\times 3$ higher probability to be found around group than around isolated LAEs (see also Muzahid et al., 2021; Banerjee et al., 2023). Conversely, at $z \lesssim 2$, this gas is much less sensitive to the environment. Moreover, at $z \lesssim 0.015$ Burchett et al. (2016) preferentially found $\log(N_{\text{CIV}}/\text{cm}^{-2}) > 13.5$ C IV gas around galaxies in low-density regions and a lack of detections in rich and massive groups. Similar results have been found at $z \lesssim 2$ by Dutta et al. (2021), who showed that the incidence and the rest-frame equivalent width of the C IV absorbers increase only weakly with galaxy overdensity. In addition, they do not detect any significant difference comparing the C IV covering fraction of groups and isolated galaxies, with only a marginal enhancement around groups at $R \lesssim 100 \text{ kpc}$. In Figure 3.11 (right panel), we show a direct comparison with the $W_r \geq 0.1 \text{ \AA}$ C IV covering fraction around group and isolated galaxies at $z \gtrsim 3$ in MAGG V and $z \lesssim 2$ by Dutta et al. (2021). While the covering fraction around isolated galaxies does not vary significantly with time, that of C IV gas around group galaxies drops from $z \approx 3$ to $z \lesssim 2$.

In conclusion, we do not observe significant redshift evolution for the excess of cool Mg II gas around LAEs in groups, but the enhancement of the more ionized gas phase around group galaxies reduces at lower redshifts. As also argued by Burchett et al. (2016) and Dutta et al. (2021), one interpretation of these results is that the Mg II absorbers arise from dense gas at small separation from the galaxies and are sensitive to the local environment, while the more diffuse and ionized gas phase is instead more extended (see details in Section 3.5.3) and thus more sensitive to the conditions of the large-scale environments. Moreover, the low incidence rate of C IV absorbers found by both Burchett et al. (2016) and Dutta et al. (2021) in massive haloes with $M_h \gtrsim 10^{12} - 10^{13} M_\odot$ may be indicative of the majority of the C IV gas to be shifted to higher ionization states. Based on this, the effect of the large-scale galaxy environment on the ionization state of the gas can be

responsible for the evolution of the covering fractions observed in Figure 3.10 for the different gas phases.

3.5.3 Radial distribution of different gas phases

As introduced in Section 3.5.2, the radial distribution of the different gas phases around the galaxies provides insights into the physical mechanisms behind the evolution of the effect of the large-scale galaxy environment. Indeed, at low redshift, it is well established that the strength of the Mg II (see e.g., Lanzetta & Bowen, 1990; Bergeron & Boissé, 1991; Chen et al., 2010a; Churchill et al., 2013b; Nielsen et al., 2013) and of the C IV (see e.g., Bordoloi et al., 2014; Burchett et al., 2016) absorption-line systems sharply decreases for increasing projected separations from the galaxies.

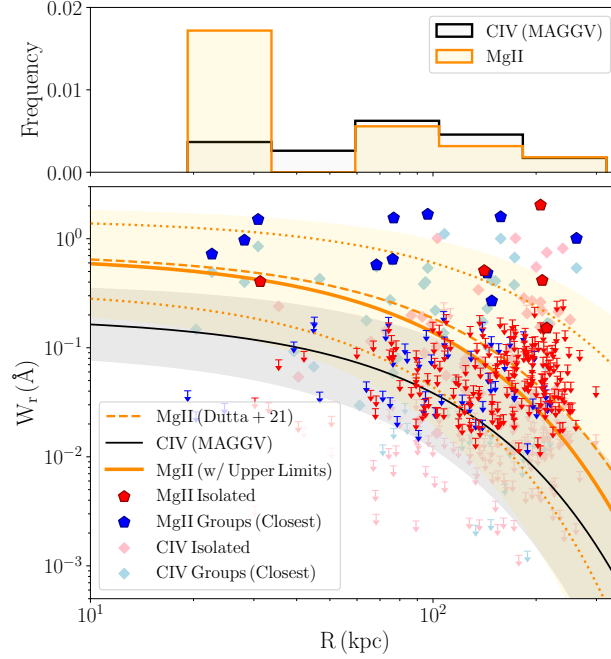


Figure 3.12: Radial profiles of the Mg II and C IV equivalent width. Upper panel: Impact parameter distribution of LAEs around Mg II (orange) and C IV (black) absorbers. Only the member closest to the line of sight is considered for each group. Lower panel: Rest-frame equivalent width of the Mg II and C IV absorbers around the group (closest member, blue and light blue, respectively) and isolated LAEs (red and pink), as a function of the galaxies’ projected separation. Arrows mark 3σ upper limits where no absorption systems are detected around LAEs. The best log-linear models fitted to the data and the upper limits are shown as solid lines, while the shaded area marks the 1σ uncertainties. As a comparison, we show the Mg II best-fit log-linear model at $z \lesssim 2$ from Dutta et al. (2021) (dashed orange line).

The picture is complicated when considering the effect of the large-scale galaxy environment. At $z \lesssim 2$ the Mg II rest-frame equivalent width drops beyond the galaxies’ virial radius, but the relation flattens at large projected separations for Mg II and C IV systems around groups (Chen et al., 2010a; Bordoloi et al., 2011; Nielsen et al., 2018b; Dutta et al., 2020, 2021; Nelson et al., 2021). Overall, when the radial distribution of the cool gas is compared to that of the more ionized phase, the first seems to be concentrated at small distances from the associated galaxies, while the latter appears more extended and reaches projected separations several times larger than the galaxies’ virial radius both in observations (see e.g., Dutta et al., 2021) and simulations (Ford et al., 2013; Ho et al., 2020, 2021; Weng et al., 2024). Moreover, Werk et al. (2014) found that the gas is ionized up to higher states with increasing distance from the host galaxy. How this condition evolves with redshift, especially in groups compared to isolated sources, may explain the evolution of the C IV covering fraction around multiple galaxies.

We have shown in MAGG V that at $z \gtrsim 3$ the C IV absorption strength does not decrease significantly for increasing projected distance from the associated LAEs and is indeed consistent with a flat radial profile (see also Muzahid et al., 2021; Banerjee et al., 2023). Compared to the C IV gas, we show the equivalent width of the Mg II absorbers as a function of the projected distance from the closest component of each group (blue) and the isolated LAEs (red) in the bottom panel in Figure 3.12. We estimate 3σ upper limits on the absorbers’ equivalent width in case any

Mg II systems are detected around LAEs within $\pm 500 \text{ km s}^{-1}$. The radial profile of the absorption strength is modeled by a log-linear function following the method detailed in Dutta et al. (2020). According to the best-fit parameters, $a = -0.061^{+0.188}_{-0.180}$ and $b = -0.001^{+0.001}_{-0.001}$, the radial profile of the Mg II absorbers shows a higher normalization compared to the C IV but decreases similarly at larger distances, while the strength of the detections remains flatter as a function of the projected separation. In addition, we do not observe any significant evolution across the cosmic time when comparing the Mg II equivalent width radial profile to the $z \approx 1$ result from Dutta et al. (2020) and at $z \lesssim 2$ from Dutta et al. (2021) (orange dashed line in Figure 3.12).

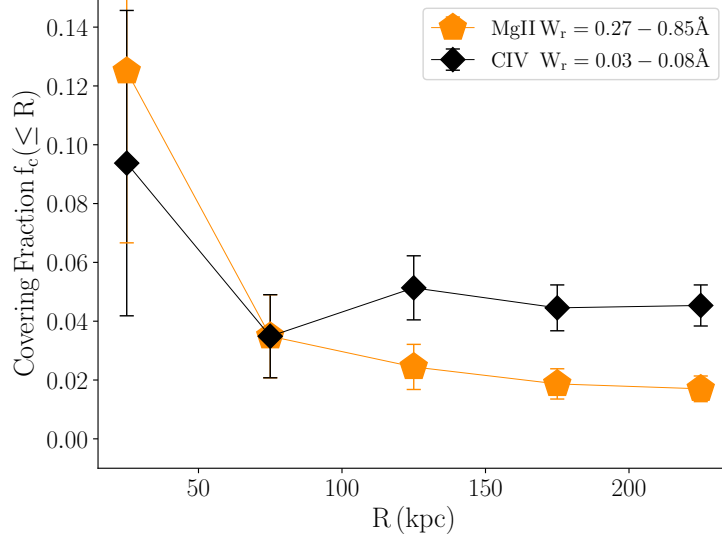


Figure 3.13: Covering fraction of Mg II and C IV systems with $12.80 < \log(N/\text{cm}^{-2}) \leq 13.30$ (filled markers). The corresponding equivalent width values in the linear regime are reported in the legend. Within $R \leq 100 \text{ kpc}$, the covering fraction of the Mg II gas is similar to the C IV while at larger distances, we observe a significant excess of C IV absorbers and a decline in the Mg II covering fraction.

In the upper panel in Figure 3.12, we show the distributions of the impact parameters of LAEs connected to the Mg II (orange) and C IV absorbers (black) considering only the closest component of each group. To assess the statistical difference between the radial distribution of the two gas phases, we perform a KS test and obtain a p -value ≈ 0.03 , which suggests that the two distributions are unlikely to arise from the same parent one. Combined with the excess of Mg II absorbers found at $R \lesssim 40 \text{ kpc}$, this result supports a picture in which different gas phases are possibly distributed differently around galaxies. We further explore this direction by measuring the Mg II and C IV covering fraction for two ranges of column density selected in the regime where the Mg II absorbers are not saturated (see Figure 3.3). This approach has the advantage of not being affected by the different selection functions of the two samples.

Since the two different ions trace different gas densities at a given equivalent width, we control for the effect of weak C IV systems ($W_{\text{C IV}} \lesssim 0.1 \text{ \AA}$) included in the analysis above by measuring the covering fractions of absorbers stronger than a threshold on the column density. This is chosen so that the corresponding lower limit in equivalent width is higher than the completeness limit of the samples. In Figure 3.13 we show the result for Mg II (orange) and C IV (black) absorbers with $12.80 < \log(N/\text{cm}^{-2}) \leq 13.30$ corresponding, in linear regime, to $W_{\text{Mg II}} = 0.26 - 0.85 \text{ \AA}$ and $W_{\text{C IV}} = 0.03 - 0.08 \text{ \AA}$ for the Mg II and C IV respectively. The Mg II gas is similar to the more ionized C IV at $R \leq 100 \text{ kpc}$ and steeply decreases at larger separations. On the contrary, the covering fraction of the C IV gas flattens at $R \geq 100 \text{ kpc}$ around a value ≈ 3 times larger than the Mg II. Altogether, this analysis provides evidence that the gas in a low ionization stage is more concentrated in close proximity to galaxies, while the more ionized gas phase traced by the C IV systems is more extended, consistently with the picture suggested at $z \lesssim 2$.

3.6 Summary and conclusions

In this paper, we have presented the results from a new 15.9h X-shooter program (PID 0109.A-0559; PI M. Galbiati), which complemented the MAGG survey with NIR spectroscopy of the background

quasars. This dataset allowed us to study 47 Mg II and compare them to 220 C IV absorption line systems from MAGG V as tracers of different gas phases around low-mass Ly α emitting galaxies at $z \approx 3 - 4$ as a function of the large-scale galaxy environment. The main results are:

- We computed the Ly α luminosity function and the projected cross-correlation function of LAEs identified within a velocity window of $|\Delta v| \leq 500 \text{ km s}^{-1}$ from the Mg II absorbers (Figure 3.6). LAEs cluster around the Mg II gas, as observed in previous MAGG analysis for other tracers (C IV and H I). Compared to C IV, the luminosity function appears elevated in normalization, similar to that of strong H I absorbers, and the projected cross-correlation function is better modeled by a steeper, $\gamma = 2.2$, power law. While C IV absorbers show a high rate of LAE detections for large equivalent widths ($\gtrsim 80$ percent), the detection rate of LAEs near Mg II remains around 40 percent, with a hint of decline toward the highest equivalent widths. An undetected (dustier) galaxy population may be responsible for this difference.
- The large-scale galaxy environment shapes the covering fraction and the properties of both the Mg II and the C IV bearing gas. The probability of detecting the absorbers is a factor of $\approx 3 - 4$ larger around LAEs in groups relative to isolated galaxies (Figure 3.7). The strength, probed by the rest-frame equivalent width, and the absorbers' kinematics, traced by the width in velocity space, positively correlate with the number of associated galaxies (Figure 3.8).
- Extending the analysis to $z \lesssim 2$, we observed no significant evolution in the covering fraction of absorbers around LAEs when considering the entire galaxy sample (Figure 3.10). However, after splitting between groups and isolated LAEs (Figure 3.11), we observed that the excess of Mg II absorbers around group galaxies persists from $z \approx 3 - 4$ to $z \lesssim 2$ without significant evolution. Conversely, the abundance of C IV gas around group LAEs seen at $z \approx 3 - 4$ drops at $z \lesssim 2$, matching the covering fraction observed around isolated galaxies.
- The radial profile of the Mg II equivalent width resembles the one of C IV up to distances of $\approx 250 \text{ kpc}$. However, when measuring the covering fraction above a fixed threshold of the column density, the more ionized C IV gas appears more extended than the cooler Mg II phase, which instead is more concentrated closer to the galaxies, as also observed at lower redshifts.

Complementing the MAGG survey with NIR quasar spectroscopy allowed us to trace simultaneously different gas phases around $z \approx 3$ galaxies and to extend the study of the cosmic gas to the surroundings of the low-mass end of the galaxies' mass function. The fundamental statistics (overdensity, luminosity function, and cross-correlation functions) of LAEs in regions selected by metals in different ionization stages (Mg II, C IV, and Si IV absorbers) and by neutral hydrogen share strong similarities, implying that galaxies are distributed inside cosmic structures that contain multiphase gas in composition and temperature. Yet, our analysis has also uncovered differences in the clustering of LAEs, allowing us to paint a more refined picture of the location of each gas component relative to galaxies. In MAGG V, by comparing the distributions of H I and C IV absorbers, we conclude that LAEs lie along gas filaments that contain the majority of the optically thick hydrogen seen at $z \approx 3$, as well as enriched gas traced by the strong C IV systems. Lower equivalent width absorbers are believed to trace instead a more diffuse gas phase extended around LAEs. Mg II gas shares similarities with the strong C IV absorbers and the H I gas, but with a steeper projected correlation function that suggests that this cool and enriched phase is more centrally concentrated near LAEs. However, with a 40 percent detection rate of LAEs for strong Mg II absorbers, an additional (star-forming and dusty) galaxy population is likely missing from our MUSE survey.

The MAGG survey has also added the dimension of the large-scale galaxy environment. Combining our previous results and this analysis, we find that the number of galaxies associated with the absorbers can shape their strength and kinematics for the Mg II and C IV gas. The covering fraction of these metal absorbers around group or isolated galaxies is the quantity that differs the most among the various gas phases. While the gravitational and hydrodynamic mechanisms able to enhance the absorbers' cross section or displace the enriched gas from the inner part of galaxies are better known at $z \lesssim 2$, the processes responsible for the excess of Mg II and C IV bearing gas observed at higher redshifts remain currently less clear. For instance, Muzahid et al. (2021) (see also MAGG V and Banerjee et al., 2023) propose that the abundance of gas around galaxies in groups might be due to the large-scale structure in which groups themselves are embedded.

Although both the Mg II and the C IV absorbers are found to be a factor $\approx 3 - 4$ times more abundant around groups of LAEs, we find a substantial difference in the distribution of the two gas phases, with a large fraction of Mg II absorbers at close separations from the galaxies compared to the C IV that is instead tracing a more ionized gas phase that is likely extended up to larger distances from the associated LAEs. This is consistent with the findings of [Werk et al. \(2014\)](#) at $z \approx 0.2$ and [Dutta et al. \(2021\)](#) at $z \approx 1$, according to which the gas is ionized to higher stages with increasing distance from the galaxies. This difference in the distribution and ionization state of the gas might drive the sharp decrease that we observed for the C IV covering fraction in groups from $z \approx 3$ to $z \lesssim 2$. In this case, the lack of difference in the covering fraction of C IV absorbers around lower-redshift groups finds a possible explanation if the gas is ionized to higher stages and is more sensitive to the large-scale, rather than the local, environment.

Further progress in completing our view of the multiphase gas near star-forming galaxies at cosmic noon requires extending our view of the galaxy population into the infrared and submillimeter to select, via optical or molecular lines, more massive and highly star-forming galaxies that are not easily detected in the UV rest frame. In parallel, a thorough exploration of the scenario provided above in cosmological simulations is also required to learn about plausible mechanisms at the origin of the differences uncovered in group and more isolated galaxies.

The MUSE Ultra Deep Field (MUDF) – VII: Probing high-redshift gas structures in the surroundings of ALMA-identified massive dusty galaxies

4.1 Introduction

Absorption lines identified in quasar spectra are excellent tracers of cosmic gas structures across a wide range of redshift and matter overdensities. Sensitive measurements of low-column-density gas against bright quasars have extensively mapped the redshift distribution of neutral and ionized hydrogen as well as that of metals inside the intergalactic medium (IGM) and around galaxies, within the circumgalactic medium (CGM), from the epoch of reionization to present days (see, e.g., [Oppenheimer et al. 2012](#); [Tumlinson et al. 2017](#), and [Péroux & Howk 2020](#) for a review). To date, advances in instrumentation, particularly with integral field spectrographs (IFSs), have enabled a direct view of hydrogen and metal-enriched gas external to the galaxy interstellar medium (ISM) in emission (see, e.g., [Hennawi & Prochaska, 2013](#); [Borisova et al., 2016](#); [Cantalupo, 2017](#); [Umehata et al., 2019](#); [Bacon et al., 2021](#); [Leclercq et al., 2022](#); [Dutta et al., 2024a](#); [Tornotti et al., 2024](#)). However, these studies are limited to the IGM and CGM densest portions, typically at close separations from galaxies. Therefore, absorption spectroscopy in quasar fields remains a crucial way to unravel the connection between the galaxy physical properties (e.g., stellar mass, star-formation rate, morphology) and those processes that regulate galaxy formation and evolution (e.g., gas accretion and outflows) within their larger-scale environment (see, e.g., [Weiner et al., 2009](#); [Chen et al., 2010a,b](#); [Steidel et al., 2010](#); [Zabl et al., 2019](#); [Dutta et al., 2020](#); [Galbiati et al., 2023](#); [Schroetter et al., 2024](#)).

At observed optical wavelengths, recent works in a wide range of redshifts, from $z \approx 0.5$ to $z \approx 4.5$, revealed a clear clustering of metal absorption lines, such as Mg II and C IV absorbers (see, e.g., [Dutta et al., 2021](#); [Banerjee et al., 2023](#); [Galbiati et al., 2023, 2024b](#)), or strong hydrogen absorbers, i.e., the Lyman limit systems (LLSs) and damped Ly α absorbers (DLAs; [Fumagalli et al., 2016](#); [Lofthouse et al., 2023](#)) around star-forming galaxies. These dense and complete spectroscopic surveys highlighted the importance of the galaxy environment in shaping the gas distribution at all redshifts, with a clear excess of cool/warm gas seen in absorption near group galaxies compared to isolated ones. With the clustering of multiple galaxies near absorption systems, a more complex picture emerges, where at least a fraction of the absorbers are probing the gas-rich cosmic structures within which galaxies reside. At longer wavelengths, surveys in the (sub-)millimeter, especially with the Atacama Large Millimeter Array (ALMA), provided a complementary view of the dusty and gas-rich galaxies associated with absorption line systems in quasar fields. Following the first detection of [CII] $_{158\mu\text{m}}$ line and dust-continuum emission from galaxies associated with $z \approx 4$ DLAs ([Neeleman et al., 2017](#)), systematic searches of molecular lines at $z \approx 0.5 - 2.5$ in the proximity of high-column density H I systems show that the highest-metallicity DLAs are associated with the most massive galaxies ([Møller et al., 2018](#); [Kanekar et al., 2018, 2020](#); [Neeleman et al., 2018, 2019](#); [Kaur et al., 2022a,b](#)). As in the case of optical searches, (sub-)millimeter observations uncover a wide range of environments, from more isolated galaxies to groups.

Based on the above results, multiple tracers are needed to complement our view of the gas distribution around cosmic structures populated by galaxies. Indeed, efforts to combine galaxy

observations in quasar fields across wavelengths, especially with the Multi Unit Spectroscopic Explorer (MUSE; [Bacon et al., 2010](#)) at the Very Large Telescope (VLT) and ALMA, are multiplying (e.g., [Klitsch et al., 2018](#); [Péroux et al., 2019](#); [Szakacs et al., 2021](#)). These combined observations reveal examples of groups composed of optically selected galaxies and dusty or molecular-rich galaxies near strong hydrogen absorbers ([Fynbo et al., 2018](#); [Klitsch et al., 2018, 2019](#)), particularly LLSs and DLAs. Conversely, multi-wavelength searches for galaxies associated with metal lines have been explored less (see, e.g., [Kashino et al., 2023](#)). However, results similar to those of LLSs and DLAs are expected. For instance, the search of Ly α emitting galaxies near Mg II absorbers with MUSE reveals an unexpected lack of detections near high-equivalent width absorbers at $z \approx 3-4$, perhaps hinting at the presence of an unseen dusty galaxy population ([Galbiati et al., 2024b](#)). Following this line of inquiry, in this paper, we present new ALMA observations of the MUSE Ultra Deep Field (MUDF; [Lusso et al., 2019](#); [Fossati et al., 2019](#); [Revalski et al., 2023](#)), a unique set of multiwavelength imaging and wide-field spectroscopic observations (from UV to near IR) at exceptional depths within a 2×2 arcmin² field hosting two bright $z \approx 3.22$ quasars (Q2139-4433 and Q2139-4434, hereafter QSO-NW and QSO-SE, respectively), which we use to map in absorption gaseous cosmic structures at $z \lesssim 3.22$. In this paper, we extend the MUDF dataset to longer wavelengths by mapping the 1.2-mm dust continuum with ALMA band 6 and spectral scans in band 3 to target the rotational transitions of carbon monoxide CO. With these data, we study the association of dusty and gas-rich galaxies with metal absorbers detected in the MUDF volume without preselection on the absorber properties. Comparisons with detections at optical and near-infrared (NIR) wavelengths, reaching the dwarf galaxy regime, are also presented to build a complete view of the galaxy environment surrounding absorbers.

This paper is structured as follows: in Sect. 4.2 we present the acquired data and illustrate the data processing. In Sect. 4.3, we describe the ancillary datasets used in this work. We devote Sect. 4.4 to analyzing the data, extracting the sources, and measuring the flux continuum and line luminosities. In Sect. 4.5, we discuss the properties of our galaxy sample in connection with those of the surrounding gas detected in absorption along the quasar sightlines. Finally, in Sect. 4.6 we summarize the results and draw our conclusions. In this work, we assume a standard Λ cold dark matter cosmology with $H_0 = 67.7 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.310$, and $\Omega_\Lambda = 1 - \Omega_m$ from [Planck Collaboration et al. \(2020b\)](#). We use AB magnitude and adopt 3σ limits when not stated otherwise.

4.2 ALMA observations and data reduction

4.2.1 Description of observations

In this work, we use ALMA band 6 mosaic observations from the program 2021.1.00285.S (Cycle 8, PI: M. Fumagalli), along with three ALMA band 3 spectral scans toward individual sources within the sky region covered by the band 6 mosaic (program ID: 2023.1.00461.S, Cycle 10, PI: A. Pensabene). The latter observations were designed to accurately measure the redshift of sources lacking robust spectroscopic constraints from MUSE (see Sect. 4.4.4). In what follows, we describe the acquisition and processing of the data.

The ALMA band 6 mosaic has been designed to cover the central part of the MUDF, where the MUSE observations achieve the maximum depth (see, [Tornotti et al., 2024](#)). The mosaic consists of 53 Nyquist-sampled pointings, each with a half-power beam width (HPBW) of $\approx 24''25$ at the reference frequency of 240.10 GHz, resulting in a covered rectangular sky area of $\approx 1'9 \times 1'6$ (Fig. 4.1). Observations were carried out in frequency division mode (FDM). The spectral tuning is composed of four 1.875 GHz-wide spectral windows (SPWs) centered at frequencies 254.498 GHz, and 257.998 GHz in the upper sideband (USB), and 240.102 GHz and 243.102 GHz in the lower sideband (LSB), respectively. The total effective bandwidth of the spectral setup is 7.5 GHz centered at wavelength 1.2 mm. Observations were subdivided into five execution blocks (EBs) carried out during the period from May 28 to June 3, 2022 employing 41 to 44 12-m antennas with baselines ranging between 15.1 – 783.5 m yielding a naturally-weighted synthesized beam FWHM of $0''.55 \times 0''.51$, and were conducted with a mean precipitable water vapor (PWV) of 0.3 – 1 mm. The total on-source integration time was $\simeq 5.86$ h achieving a sensitivity of $26 \mu\text{Jy beam}^{-1}$ over the total bandwidth. The native spectral resolution is $\simeq 7.8$ MHz, corresponding to $\simeq 9.5 \text{ km s}^{-1}$ at the reference frequency of 240.1 GHz. The J2126-4605 and J2258-2758 were used as phase and flux calibrators during observations.

The ALMA band 3 observations consist of three single-pointing spectral scans centered on three sources revealed in the band 6 mosaic (see, Fig. 4.1) to target their CO(3–2) line (rest-

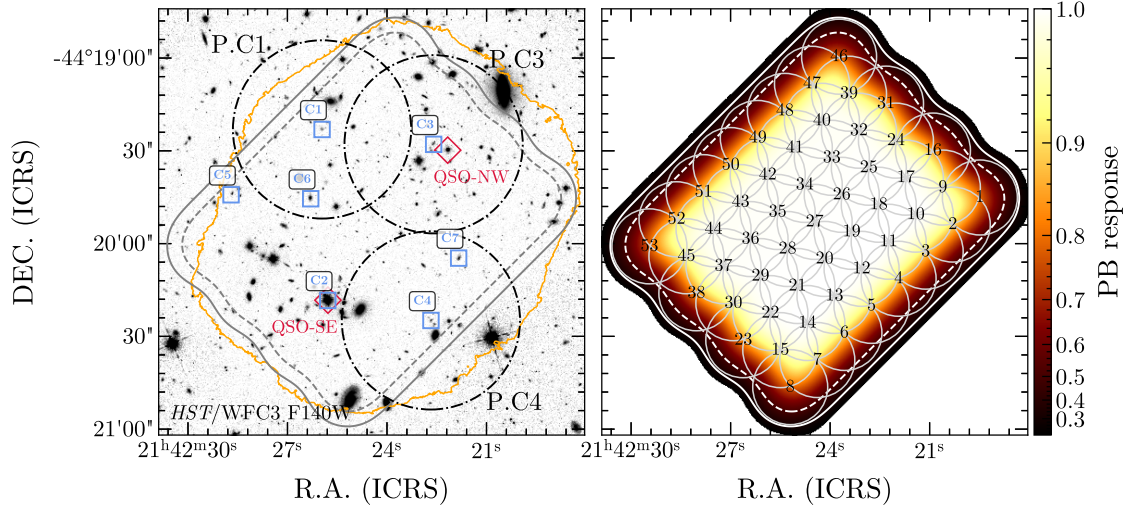


Figure 4.1: Footprints of ALMA observations toward the MUQF. *Left panel:* The background shows the HST/WFC3 F140W image, and the gray lines show the PB response of the ALMA band 6 mosaic at 30% and 50% (solid and dashed lines, respectively). The orange line indicates the region observed by VLT/MUSE with > 3 hours of exposure time. Red diamonds indicate the MUQF quasar pair. The blue squares with labels indicate the high-fidelity ALMA-selected sources extracted from the 1.2 mm continuum image. The dot-dashed black circles show the ALMA pointings of the band 3 spectral scans toward three individual sources (P.C1, P.C3, and P.C4). The diameter of the circles indicates the HPBW at 101 GHz. *Right panel:* PB response of the ALMA band 6 mosaic. The circles show the disposition of the 53 pointings with a diameter equal to the HPBW of the ALMA 12 m antennas at the reference frequency of the setup.

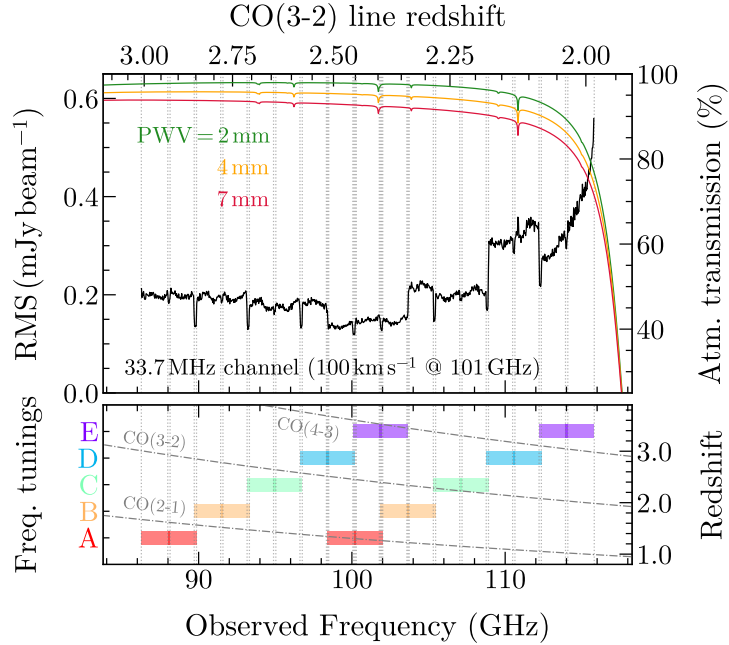


Figure 4.2: Sensitivity (RMS) of the ALMA band 3 spectral scans across the covered frequency range. *Top panel:* The colored curves and the right axis report the atmospheric transmission from the ALMA site at different levels of PWV obtained using `plotAtmosphere` task in the CASA `analysisUtils` package (Hunter et al., 2023a). The top axis reports the expected redshift of the CO(3-2) line. *Bottom panel:* Disposition of SPWs in the four frequency tunings. Vertical dotted lines indicate the edges of the SPWs. The maximum sensitivity is achieved within $z \approx 2.35 - 2.5$ where two SPWs in different frequency tunings overlap. The right axis reports the redshift of the various CO transitions (dash-dotted lines) covered by the spectral scan within $z \approx 1 - 4$.

frame frequency $\nu_{\text{rest}} = 461.041$ GHz) in the redshift range $z \simeq 2 - 3$. The frequency setups were tuned based on the photometric redshift probability distributions (pointing to the redshift range $1.8 \lesssim z \lesssim 2.8$) obtained from multiwavelength photometry available for galaxies in the MUDF (see, [Revalski et al., 2023](#)) using the EAZY code ([Brammer et al., 2008](#)). Observations were carried out in FDM employing five frequency tunings with two adjacent 1.875 GHz-wide SPWs per sideband designed to cover the spectral range 86.25 – 115.80 GHz contiguously. The total effective bandwidth is 29.55 GHz. Each frequency tuning was observed in two EBs, and executed between January 18th and March 4th 2024, employing 40 to 45 12-m antennas with baselines ranging in 15.1 – 483.9 m and a mean PWV = 2.7 – 7.1 mm. The uv coverage of these observations yields to a naturally-weighted synthesized beam size of $2''.5 \times 2''.3$ at the reference frequency of 101 GHz. The total on-source observing time was 8.1 h. Data were acquired with a native spectral resolution of 1.95 MHz, corresponding to $\simeq 5.8 \text{ km s}^{-1}$ at 101 GHz. The sensitivity achieved in the final dataset varies along with frequency due to the variation in atmospheric absorption and weather conditions across the EBs (Fig. 4.2) and has a median value of $0.20 \text{ mJy beam}^{-1}$ over 33.7 MHz, or $\simeq 100 \text{ km s}^{-1}$ at 101 GHz. Fig. 4.1 shows the footprints of the ALMA observations presented in this work, the pointing design, and the response of the primary beam (PB) of the band 6 mosaic. In Fig. 4.2, we show the spectral coverage of the acquired band 3 spectral scans and the sensitivity variation along the frequency range.

4.2.2 Data calibration and imaging

We calibrated the data using the Common Astronomy Software Application (CASA; [McMullin et al., 2007](#); [Hunter et al., 2023b](#)) by running the calibration pipeline `scriptForPI` provided with the raw measurement set using the CASA pipeline versions 6.2.1 and 6.5.4 for band 6 and 3 data, respectively. No self-calibration was performed on the datasets. We imaged band 6 visibilities by using the CASA task `tclean`. We first obtained a “dirty” continuum image by performing the Fourier transform of the visibilities (`niter=0`) using the multi-frequency synthesis mode (`specmode = ‘mfs’`) and adopting a natural weighting scheme to maximize the sensitivity per beam. In this procedure, we set “mosaic” as the gridding convolution function and put the phase center at the coordinates ICRS 21:42:24.4658 –44:19:53.007. We used a pixel scale of $0''.08$ to achieve the Nyquist sampling of the longest baselines. Similarly, we obtained “dirty” band 6 data cubes of the LSB and USB with 25 km s^{-1} channel width. Additionally, in order to improve the surface brightness sensitivity for faint extended emission, we obtained continuum image by applying a Gaussian taper in the uv plane with $\text{FWHM}_{uv} = 90 \text{ k}\lambda$ and setting the pixel size to $0''.4$, yielding a synthesized beam FWHM of $2''.19 \times 2''.39$. The “dirty” data preserve the intrinsic properties of the noise, and we used that to perform a blind search of continuum and line emitters in the field (see Sect. 4.4). We finally obtained “cleaned” continuum images and cubes by applying the cleaning algorithm down to 2σ , placing circular masks of $2''.0$ diameter around sources extracted with $S/N > 5$.

We imaged the ALMA band 3 pointings by obtaining “dirty” cubes of the combined dataset, including all the SPWs using `tclean` with `niter=0` and the “natural” weighting scheme of the visibilities. We set the pixel scale at $0''.3$ and the channel width at 100 km s^{-1} at the reference frequency of 101 GHz, which is expected to maximize the S/N per channel while resolving the line profile. In such datasets, sources are located at the phase center and appear faint and spatially unresolved and no bright continuum emitters are found in the cubes. For these reasons, the cleaning of cubes is not required.

4.3 Ancillary multi-wavelength observations

In this study, we use the catalog of MUDF galaxies from [Revalski et al. \(2023\)](#), which provides deep photometry and high-confidence redshifts constrained with HST and MUSE spectroscopy, to associate optical counterparts to the sources identified by ALMA. This field benefits from extremely deep MUSE observations (ESO PID 1100.A-0528, PI: M. Fumagalli) with an integration time of ≈ 142 hours in wide field mode and with extended wavelength coverage (4600 – 9350 Å, and spectral resolution $R \approx 2000 - 4000$). We refer to [Fossati et al. \(2019\)](#) for detailed descriptions of the survey design and the data reduction. The MUDF survey is specifically designed to study the connection between gas and galaxies across cosmic time and thus also includes high-resolution ($R \approx 40,000$) spectra of the two quasars (QSO-SE and QSO-NW with r -band magnitude 17.9 and 20.5, respectively) to probe the IGM and CGM in absorption (see, [Beckett et al., 2024](#)). The quasar

spectra were obtained using the Ultraviolet and Visual Echelle Spectrograph (UVES, [Dekker et al., 2000](#)) at the VLT (PIDs: 65.O-0299, 68.A-0216, 69.A-0204, and 102.A-0194, PI: V. D’Odorico) and cover the wavelength range 4100 – 9000 Å with a S/N per pixel of ≈ 25 and ≈ 10 for the bright and faint quasar, respectively. In HST Cycle 26, we also collected 90 orbits of WFC3/F140W imaging and WFC3/IR G141 grism spectroscopy (PID 15637, PI: M. Rafelski & M. Fumagalli), providing a deep and extensive sampling of the galaxy spectral energy distributions (SEDs) from UV to NIR wavelengths (see, [Revalski et al., 2024](#)). Custom calibration of HST observations is described by [Revalski et al. \(2023\)](#) together with a complete list of additional optical and near-UV (NUV) photometry¹.

4.3.1 Catalog of HST-selected galaxies

Galaxies detected in the deep WFC3/F140W image are included in a catalog that is 50% complete at 27.6 mag and covers the entire area observed with MUSE, as described by [Revalski et al. \(2023\)](#). To all the detected galaxies, a spectroscopic redshift is assigned together with a quality flag (see table 5 in [Revalski et al., 2023](#)), depending on the presence of high S/N lines in their spectra. The spectral fitting process used to identify and fit emission lines across a wide range of wavelengths, from UV to NIR, is detailed in [Revalski et al. \(2024\)](#). A high-quality redshift (redshift quality flag ≥ 3) was assigned to 25% of the sources with spectral coverage. The physical properties of these galaxies, such as stellar mass (M_*) and star formation rate (SFR), are obtained by fitting the multiwavelength photometry and the MUSE spectra simultaneously, as described in [Fossati et al. \(2018a, see, also, Revalski et al. 2024\)](#), to which we refer for full details about the method. Finally, we distinguished between isolated galaxies and those residing in groups. To this end, we applied a friends-of-friends algorithm to identify galaxies belonging to a group as all the sources that are not isolated within $\pm 500 \text{ km s}^{-1}$ along the line of sight (see, e.g., [Fossati et al., 2019](#), for a similar approach). Overall, we found that $\approx 86\%$ of the galaxies reside in 23 different groups across the redshift range $z = 1 - 3$, with five groups with more than five galaxies, and two with more than 15 members.

4.3.2 Catalog of quasar absorbers

The detailed procedure used to identify the absorption line systems in the spectra of the two quasars, as well as the complete list of the transitions and their properties, is illustrated in table 1 of [Beckett et al. \(2024\)](#), and is only briefly summarized below. The spectra are normalized by fitting a cubic spline to the continuum redward to the Ly α emission of the quasars by using the ESPRESSO data analysis software ([Cupani et al., 2016](#)). The spectra are then visually inspected to identify absorption line systems such as Mg I, Mg II, Fe II, Al II, Al III, C IV and Si IV. All these lines are then fitted using the Bayesian code Monte Carlo Absorption Line Fitter (MC-ALF, [Longobardi et al., 2023](#)) and decomposed into the minimum number of Voigt profiles required to reproduce the observed flux. The Doppler parameter, column density, and redshift of each Voigt component are the free parameters of the fit. Additional filler components account for blends with absorption features arising from different ions at different redshifts. We finally assign the redshift of the component with the largest column density to each absorber. The resulting sample includes 22 absorption systems in the QSO-SE sightline, and nine systems in the QSO-NW sightline. Only one absorption system is detected in both the sightlines at $z \approx 0.88$ (see, table 1 in [Beckett et al. 2024](#)). Multiple ions are often detected at the redshift of each system.

4.4 Extraction and characterization of ALMA sources

We performed a source search in the “dirty” ALMA band 6 continuum mosaic image using the Python-based code LINESEEKER². The complete details of the codes and performances are described in [González-López et al. \(2017, 2019\)](#). Here, we summarize the basic characteristics and

¹The final products can be retrieved as High Level Science Products (HLSPs) at doi:10.17909/81fp-2g44, together with the default data files at doi:10.17909/q67p-ym16 from the Mikulski Archive for Space Telescopes (MAST). The dedicated page to the MUDF program is available at <https://archive.stsci.edu/hlsp/mudf>.

²The code is publicly available at <https://github.com/jigonzal/LineSeeker>. Other similar source-finding codes are available in the literature, such as FINDCLUMPS ([Walter et al., 2016](#)) or MF3D ([Pavesi et al., 2018](#)), which differ mainly in details about the adopted spectral filter function. However, comparisons between codes have yielded very similar results (see [González-López et al., 2019](#)).

ID _{1.2mm} ⁽¹⁾	R.A. ⁽²⁾ (ICRS)	DEC. ⁽³⁾ (ICRS)	$S_{1.2\text{mm}}^{\text{peak}}$ ⁽⁴⁾ (mJy beam ⁻¹)	S/N ⁽⁵⁾	fidelity ⁽⁶⁾	z_{MUSE} ⁽⁷⁾	z_{ALMA} ⁽⁸⁾	ALMA line ⁽⁹⁾	m_{F140W} ⁽¹⁰⁾ (mag)
C1	21:42:25.957	-44:19:23.09	0.85 ± 0.07	24.3	1.00	—	2.2529	CO(3-2)	23.89
C2 (QSO-SE)	21:42:25.786	-44:20:18.37	0.75 ± 0.06	20.3	1.00	3.2236	3.2267	CO(4-3)	17.34
C3	21:42:22.595	-44:19:27.89	0.32 ± 0.06	8.4	1.00	—	2.2954	CO(3-2)	23.51
C4	21:42:22.676	-44:20:24.85	0.47 ± 0.08	8.4	1.00	—	2.3676	CO(3-2)	23.81
C5	21:42:28.701	-44:19:44.20	0.39 ± 0.09	5.6	1.00	2.4723 ^(†)	—	—	23.35
C6	21:42:26.300	-44:19:45.49	0.27 ± 0.06	5.1	0.96 ± 0.02	1.0541	1.0544	[Cl] _{609 μm}	22.21
C7	21:42:21.841	-44:20:04.69	0.13 ± 0.03	5.0	0.94 ± 0.02	1.7561	—	—	22.86

Table 4.1: ALMA 1.2-mm continuum-selected sources. High-fidelity ($F \geq 90\%$) candidate list extracted via blind search using LINESEEKER. ⁽¹⁾Identifier of ALMA 1.2-mm continuum-selected candidates. ⁽²⁾Right Ascension (ICRS). ⁽³⁾Declination (ICRS). ⁽⁴⁾Flux density extracted from the source peak pixel in the 1.2 mm tapered continuum image. ⁽⁵⁾Signal-to-noise ratio. ⁽⁶⁾Fidelity and its uncertainties estimated by LINESEEKER using negative detections. All sources have a NIR counterpart, as revealed in *HST* / WFC3 F140W image, therefore, the actual fidelity should be considered $F = 1$. ⁽⁷⁾High-confidence redshift measurement from MUSE spectroscopic data. ^(†)Tentative redshift estimate measured from low-quality MUSE spectrum. ⁽⁸⁾Redshift measurement derived from emission line detected with ALMA (see, also, Table 4.2). ⁽⁹⁾Emission line detected in either ALMA band 6 data cube or band 3 spectral scans. ⁽¹⁰⁾Apparent AB magnitude in *HST*/WFC3 F140W filter. All reported uncertainties correspond to 1σ .

operation of the algorithm. LINESEEKER has been designed to perform a blind search of line emitters in large data cubes from (sub-)millimeter surveys such as the ALMA Spectroscopic Survey in the *Hubble* Ultra Deep Field (ASPECS; see, e.g., Decarli et al., 2019; González-López et al., 2019, 2020), but can be equivalently employed to search sources in 2D continuum images. The code searches for the line emitter in the cube by adopting a matched filter approach. LINESEEKER convolves the data cube along the spectral axis with a Gaussian kernel with a range of input widths set based on a typical linewidth. For each convolution iteration, the code finds all the voxels above a given S/N threshold. The noise is estimated in the convolved cube through 5σ -clipped statistics, and the signal is measured from the peak flux density per beam. The sources are extracted by grouping connected voxels using the Density-Based Spatial Clustering of Application with Noise algorithm (DBSCAN; Ester et al., 1996)³. Then, the final list of sources is drawn by cross-matching the source catalogs obtained from different convolution kernels and selecting the maximum S/N . For each identified candidate, LINESEEKER estimates the probability of false positive detection by performing the source search on the negative cube. Since any negative peak in a (sub-)millimeter extragalactic survey is expected to be result of noise, the ratio between positive and negative detection distributions is used to estimate the fidelity (or reliability) of source candidates as a function of S/N as

$$F(S/N) = 1 - \frac{N_{\text{neg}}(S/N)}{N_{\text{pos}}(S/N)}, \quad (4.1)$$

where N_{neg} and N_{pos} are the number of negative and positive detections at a given S/N . The cumulative distribution of negative detections is then modeled assuming Gaussian-distributed noise by using a function of the form $N[1 - \text{erf}(S/N/\sqrt{2}\sigma)]$, where erf is the error function, and N , σ are free parameters. The best-fit model is finally used to estimate $F(S/N)$ for all positive source candidates. The search of continuum sources in 2D images can be performed similarly with LINESEEKER by skipping the convolution steps.

4.4.1 Continuum-selected sources in the 1.2-mm image

We selected source candidates in the ALMA band 6 mosaic by running LINESEEKER on the untapered “dirty” continuum image, excluding the region below 50% of the maximum PB response, where the low telescope sensitivity increases the probability of selecting spurious sources. The “dirty” data are preferred over the “cleaned” ones, since the former preserves the intrinsic properties of the noise. Furthermore, we input the “dirty” image into LINESEEKER without correcting for the PB response, thus preserving the spatial homogeneity of the noise across the field-of-view (FoV). With LINESEEKER we extracted all the detections above $S/N > 3$ and selected the source candidates with $F \geq 90\%$, corresponding to $S/N \geq 5.0$. We recovered a total of seven 1.2-mm continuum-selected candidates (labeled C1 to C7), including the host galaxy of the QSO-SE (C2). We report the location of these sources in Fig.4.1 and their characteristics in Table 4.1.

³This algorithm is included within the Python package Scikit-learn (Pedregosa et al., 2011).

We performed similar search on the tapered band 6 continuum image, which is more sensitive to faint extended emission, and found consistent result. To further investigate the reliability of faint candidates, we complemented the search for continuum sources by cross-matching the catalog of galaxies detected in the HST/WFC3 F140W image (see Sect. 4.3.1) with the sources selected by LINESEEKER with fidelity values within $20\% \leq F < 90\%$. To this purpose, we search HST counterparts around the sky position of ALMA continuum sources within a separation limit of $0''.5$, comparable to the beam size of the 1.2-mm continuum untapered image, and found no associations. For the goal of this study, we therefore include in our fiducial sample of ALMA continuum sources only the high-fidelity candidates selected as previously described. For all of these sources, we found associations in the HST catalog as described in Sect 4.4.4. We evaluated the completeness of our ALMA band 6 survey by injecting artificial point-like sources at random locations in the untapered dirty 1.2-mm continuum image. Then, we computed the rate of recovered sources by running LINESEEKER and searching for candidates with $F \geq 90\%$ encompassed within $0''.275$ from the injected position, corresponding to half of the beam major axis. We repeated the simulation by injecting 20 sources simultaneously into the image and repeating this procedure for different values of flux density in the range $[2, 10] \times \text{RMS}$. We finally iterated the entire simulation 30 times to achieve good statistics. As a result, we find that our source search is complete to $\approx 90\%$ at $S/N \approx 7.5$. During the simulation, we also evaluated the effect of flux boosting (see, e.g., Hogg & Turner, 1998; Scott et al., 2002; Coppin et al., 2006) on the recovered sources, and we determined that it is negligible at $S/N \gtrsim 5$. Thus, we do not correct the measured fluxes of our source sample for this effect.

4.4.2 Line emitters in band 6 data cubes

The band 6 data cubes cover multiple molecular and atomic lines which enter the LSB and USB at different redshifts. In particular, for $z < 3.5$ the carbon monoxide CO transitions are covered by our frequency setup from the CO(3–2) at $z \simeq 0.45$ in the USB, to the CO(9–8) at $z \simeq 3.34^4$, while the fine-structure lines of the atomic carbon [CII] $_{609,369\mu\text{m}}$ are observable from $z \approx 0.9$ up to 2.38. To complete the source search in the field, we employed LINESEEKER to detect line emitters in the cubes blindly. Similarly to the search for continuum emitters (see Sect. 4.4.1), we restricted the source extractions to the survey area with PB response $\geq 50\%$, and we run the line-search algorithm on both LSB and USB cubes using a Gaussian kernel with width ranging from 1 to 18 channels, which is adequate to match the typical linewidths of $\text{FWHM} \sim 50 - 1000 \text{ km s}^{-1}$ observed in high-redshift galaxies (see, e.g., Birkin et al., 2021). We extracted all line-emitting candidates with $S/N \geq 3.0$ and then selected those with $F \geq 90\%$. As a result, we identified one line emission in the spectrum of the C6 source, which we already identified through its continuum at 1.2 mm (see Table 4.1). This source has a robust spectroscopic redshift of $z = 1.05413$ measured with the MUSE data (see Sect. 4.4.4) that allowed us to identify the line as [CII] $_{609\mu\text{m}}$. We measured the line flux by fitting the spectrum extracted from the source peak pixel (see, Sect. 4.4.5), and obtained $F_{[\text{CII}]609\mu\text{m}} = 0.36 \pm 0.10 \text{ Jy beam}^{-1} \text{ km s}^{-1}$. The source redshift we estimated with this line is consistent with the one measured from the MUSE spectrum. The visual inspection of ALMA band 6 spectra at the location of all the 1.2-mm continuum-selected sources did not reveal additional emission lines, consistently with the spectroscopic redshift measurements of the sources. Finally, we complemented our line search by cross-matching the HST catalog of galaxies with the source catalog of line-emitting candidates delivered by LINESEEKER following the same approach described in Sect. 4.4.1. This procedure did not reveal additional line emitters.

4.4.3 Source detection in band 3 spectral scans

We analyzed ALMA band 3 spectral scans targeting sources C1, C3, and C4. We inspected the spectra at the center of the pointing around the nominal locations of the continuum detections revealed at 1.2 mm in the ALMA mosaic (see Table 4.1). The CO(3–2) line, which is identified based on the posterior distributions of the source photometric redshift⁵, is detected in all three sources, which appear spatially unresolved given the low angular resolution of these ALMA observations (see Sect. 4.2). We ran LINESEEKER on band 3 "dirty" cubes and continuum images similarly to what described in Sects. 4.4.1 and 4.4.2 for band 6 data. The search for line emitters within

⁴Higher- J CO transitions are covered for $z > 3.5$ by our ALMA band 6 spectral tunings.

⁵We derived photometric redshifts of C1, C3 and C4 from multiband photometry as detailed in Sect. 4.2, obtaining $z_{\text{phot}} = 2.21 \pm 0.19, 2.51 \pm 0.07, 2.44 \pm 0.17$, respectively, where nominal values and uncertainties are computed from the expected value and variance of the photo- z posteriors.

the cubes reveal the CO(4–3) line of the QSO-SE (C2) which is redshifted in the ALMA band 3 at $z \approx 3.22$. The analysis of this detection is beyond the scope of this work. Additionally, we performed visual inspection of the band 3 spectra in the location of secure continuum sources selected in the ALMA band 6 mosaic, and found no evident emission lines. Finally, in the 3-mm continuum images we did not find neither high-fidelity source candidates, nor HST counterparts associated to the low-fidelity ones.

4.4.4 HST counterparts of ALMA continuum sources

In Fig. 4.4, we show postage stamps of the seven ALMA continuum-selected sources and their NIR and optical/NUV counterparts in the field. As also described in Sect. 4.4.1, we cross-matched all the sources detected in the continuum at 1.2 mm with ALMA with the catalog of HST-selected galaxies. Within a projected separation of $0''.5$ in the sky, we found an unambiguous counterpart for all ALMA detections, except for C5. By visually inspecting the HST/WFC3 F140W image of this source, we found that the peak of the ALMA continuum appears to correspond to a faint and compact source located southeast relative to an extended object. Interestingly, the 1.2-mm continuum emission seems to be spatially associated with diffuse emission revealed by HST around the extended galaxy, which suggests irregular morphology, possibly due to the interaction with the nearby compact source. Alternatively, gravitational lensing may represent a viable explanation to the observed emission. However, given the limited S/N of the ALMA data, we cannot confidently assess the nature of this emission. The angular resolution of MUSE observations is insufficient for adequately deblending the spectra of these two closely spaced sources (see, Fig. 4.4). By inspecting the MUSE spectrum extracted at the location of the galaxies, we assigned to both the sources a low-quality redshift estimate of $z \approx 2.4723$ on the basis of only the tentative detection of C IV $\lambda\lambda 1548, 1550$ absorption line feature. Future deep spectroscopic observations using, for example, JWST, are needed to shed light on the characteristics of this system. A high-quality spectroscopic redshift is constrained from the MUSE spectra for C2, C6, and C7. The remaining ones (i.e., C1, C3, and C4), which are also the faintest in the F140W filter among the counterparts (see, Table 4.1), do not have a secure redshift measurement from MUSE or HST grism spectra. The redshifts of the latter, are accurately measured from the CO(3–2) lines detected in the band 3 spectral scans as previously described in Sect. 4.4.3.

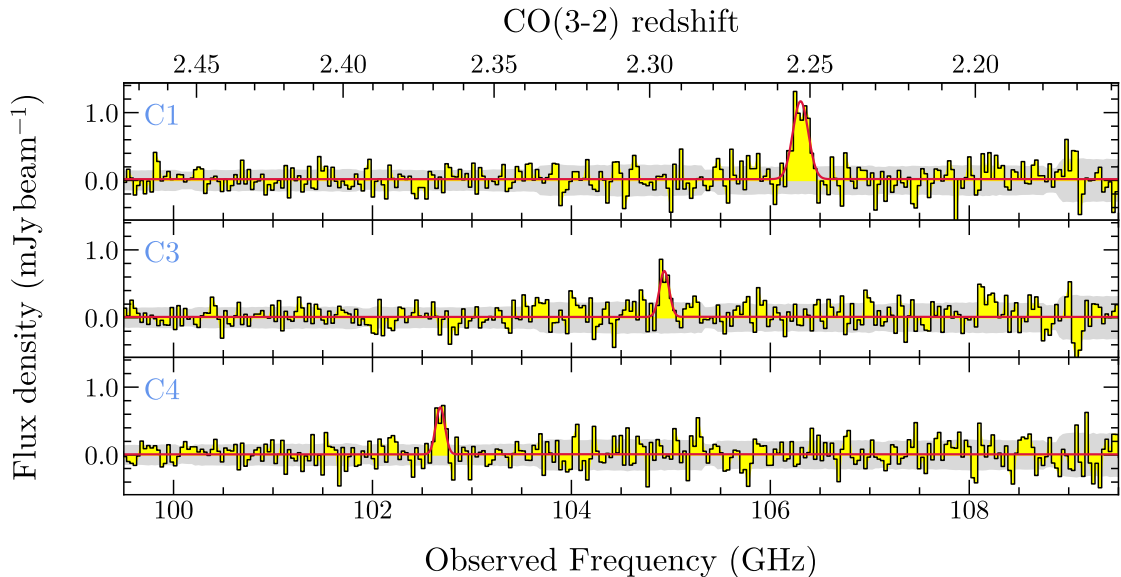


Figure 4.3: Spectra extracted from ALMA band 3 spectral scans toward 1.2-mm continuum-selected sources C1, C3 and C4 (from the top to the bottom panel). Data and the best-fit model are shown in yellow and red, respectively. The $\pm 1\sigma$ RMS noise range of the data is indicated by the gray shaded area. The top scale reports the expected CO(3–2) line redshift depending on the observed frequency.

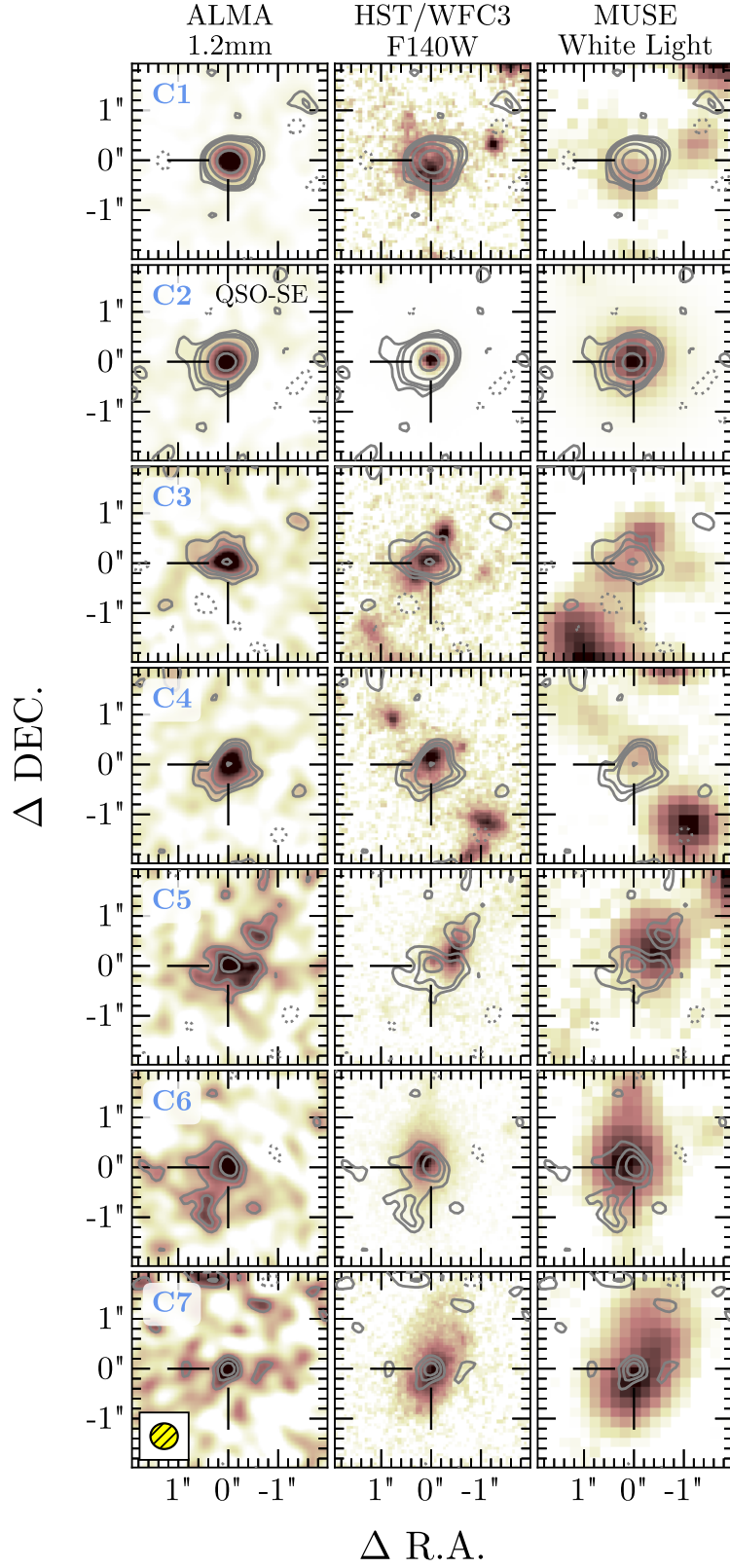


Figure 4.4: $5'' \times 5''$ postage stamps of NIR and optical/NUV counterparts of ALMA 1.2-mm continuum-selected sources in the MUDF. *From left to right:* 1.2-mm continuum map from ALMA band 6, HST/WFC3 F140W (central observed wavelength $1.4\mu\text{m}$), MUSE white light image (central observed wavelength 655nm). The solid contours show the $[2, 3, 2''] \times \text{RMS}$ (with $n \geq 2$) levels of the ALMA continuum. The yellow ellipse at the corner of the bottom left panel shows the synthesized ALMA beam.

4.4.5 Source fluxes and luminosities

All seven ALMA-selected sources detected in the continuum at 1.2 mm appear, at best, marginally resolved. To maximize the recovered signal, we estimated their flux density in the $\sim 2''$ tapered image through a single-pixel measurement of the continuum source flux peak that we then rescaled for the PB response at the corresponding location. We then applied the same PB rescaling to the RMS (in mJy beam $^{-1}$) measured on the 1.2-mm continuum image to estimate the 1σ uncertainty on $F_{1.2\text{ mm}}$. We report these quantities in Table 4.1. We measured the CO(3–2) line flux of C1, C3, and C4 detected in the band 3 cubes by fitting the spectra extracted from the source peak employing a single Gaussian plus a constant to model the line and the continuum. We then estimated the best-fit model parameters and uncertainties from the posterior probability distributions obtained through the Python Markov chain Monte Carlo (MCMC) ensemble sampler **emcee** package (Foreman-Mackey et al., 2013), assuming box-like priors based on a visual inspection of the data. We finally derived the line luminosities as (see, e.g., Solomon et al., 1997):

$$L_{\text{CO}}[L_{\odot}] = 1.04 \times 10^{-3} S\Delta\nu \nu_{\text{obs}} D_L^2, \quad (4.2)$$

$$L'_{\text{CO}}[\text{K km s}^{-1} \text{ pc}^2] = 3.25 \times 10^7 S\Delta\nu \frac{D_L^2}{(1z)^3 \nu_{\text{obs}}^2}, \quad (4.3)$$

where $S\Delta\nu$ is the velocity-integrated line flux in Jy km s $^{-1}$, ν_{obs} is the observed central frequency of the line in GHz, z is the source redshift measured from the centroid of the CO line, and D_L is the luminosity distance in Mpc. The relation between Eq. (4.2) and (4.3) is $L_{\text{CO}} = 3 \times 10^{-11} \nu_{\text{rest}}^3 L'_{\text{CO}}$, where ν_{rest} is the line rest frequency in GHz. In Fig. 4.3, we show the spectra and the best-fit models and report the line measurements in Table 4.2.

ID _{1.2mm} ⁽¹⁾	C1	C3	C4
z_{CO} ⁽²⁾	2.2529 $^{0.0003}_{-0.0003}$	2.2954 $^{0.0005}_{-0.0005}$	2.3676 $^{0.0003}_{-0.0003}$
$S\Delta\nu$ ⁽³⁾ (Jy beam $^{-1}$ km s $^{-1}$)	0.63 $^{0.06}_{-0.06}$	0.24 $^{0.06}_{-0.06}$	0.23 $^{0.04}_{-0.04}$
FWHM ⁽⁴⁾ (km s $^{-1}$)	510 $^{50}_{-50}$	330 $^{120}_{-80}$	320 $^{50}_{-40}$
L_{CO} ⁽⁵⁾ ($10^7 L_{\odot}$)	2.3 $^{0.2}_{-0.2}$	0.9 $^{0.2}_{-0.2}$	0.95 $^{0.14}_{-0.14}$
L'_{CO} ⁽⁶⁾ (10^{10} K km s $^{-1}$ pc ^{2})	1.77 $^{0.17}_{-0.17}$	0.70 $^{0.18}_{-0.18}$	0.72 $^{0.11}_{-0.10}$

Table 4.2: ALMA CO(3–2) line emitter. Derived properties from the CO(3–2) line. ⁽¹⁾Identifier of ALMA 1.2-mm continuum-selected galaxies. ⁽²⁾Redshift estimate for the CO(3–2) line. All reported uncertainties correspond to 1σ confidence interval as derived from the 16th and 84th percentile of the posterior probability distribution. ⁽³⁾Velocity-integrated line flux. ⁽⁴⁾Full Width at Half Maximum of the line. ^(5,6)Line luminosities as derived from Eq. (4.2) and (4.3).

4.5 Linking galaxies to absorbers in multiwavelength datasets

Absorption studies toward background quasars typically target galaxies in their rest-frame optical/UV. So far, (sub-)millimeter studies (e.g., with ALMA) have focused on specific classes of absorbers, such as high-column-density DLAs ($N_{\text{HI}} \geq 10^{20.3} \text{ cm}^{-2}$; see, e.g., Neeleman et al., 2017; Kanekar et al., 2018; Kaur et al., 2022a,b). However, little is known about the contribution of dusty and gas-rich galaxies to the cosmic metal enrichment across different epochs of the Universe from untargated observations (Dudzevičiūtė et al., 2021). In this work, we capitalize on our blind ALMA survey to study the association between cosmic gas traced by quasar absorption systems in the MUDF, and galaxies detected at millimeter wavelengths in their dust continuum.

4.5.1 Gas-galaxy association rate

With accurate redshift measurements ($\Delta\nu \lesssim 60 \text{ km s}^{-1}$) of our ALMA-selected galaxies from either MUSE spectra or ALMA CO(3–2) line (see Sect. 4.4.4, 4.4.5), we matched these galaxies with the absorbers in velocity space in either of the quasar sightlines (see, Sect. 4.3.2). Following the approach typically adopted in the literature for optical studies (see, e.g., Dutta et al., 2021; Schroetter et al., 2021; Galbiati et al., 2023; Beckett et al., 2024), we link the absorbing gas with galaxies by searching for association within $|\Delta\nu| \leq 500 \text{ km s}^{-1}$ with respect to the absorbers. We

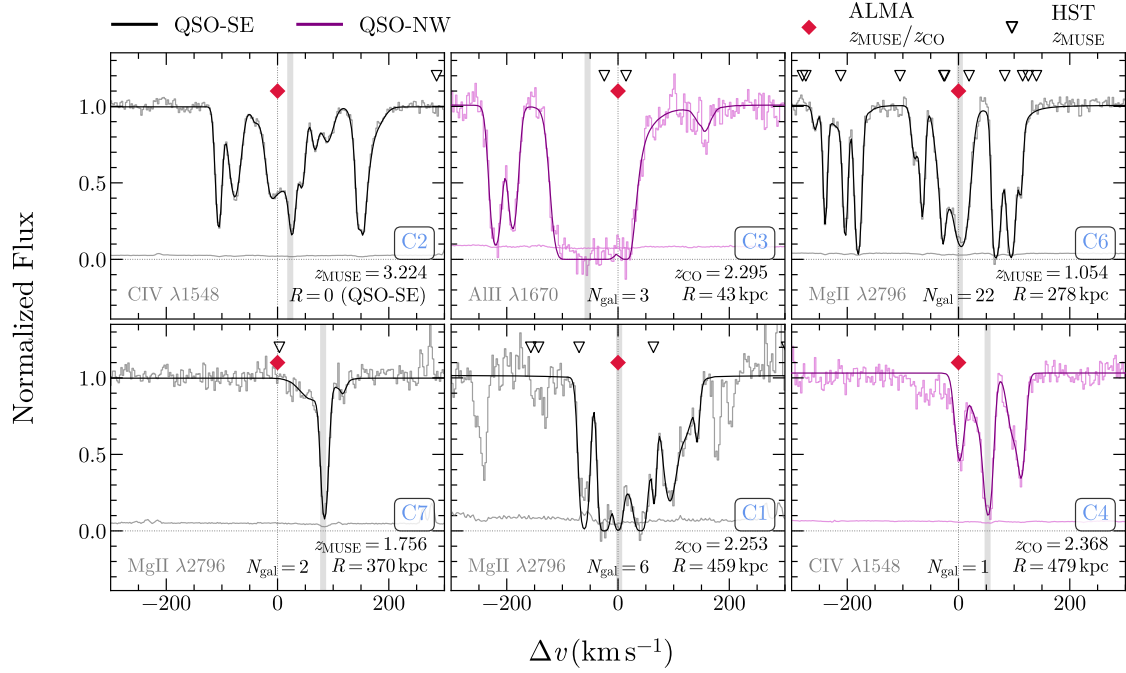


Figure 4.5: Connection between ALMA-selected galaxies and absorption systems in the spectra of the QSO-SE (black) and -NW (purple). The line-of-sight separations are centered on the redshift of the ALMA-selected galaxies. The vertical gray band indicates the velocity at which the optical depth of the absorption is 50% of the total (τ_{50}). The transitions shown (CIV λ1548, AlII λ1670, and MgII λ2796) are limited to the strongest lines among the multiple metals associated to each galaxy. Galaxies selected with ALMA are indicated by red diamonds and ordered (from the top left to the bottom right panel) for increasing sky-projected separation (R) from the QSO sightline along which the absorption is detected. HST-selected galaxies linked to the same absorbers are shown with black triangles. The number of galaxies (N_{gal}) associated to each group (see, Sect. 4.5.3) is also reported.

report our results in Fig. 4.5. We found that 100% of our ALMA-selected galaxies – for which robust spectroscopic redshifts are available – are unambiguously associated with metal absorbers. This detection rate decreases to $\geq 85\%$ when considering the uncertainties related to source C5 (see Sect. 4.4.4 for details). Indeed, if we adopt the tentative redshift of C5 measured from the low-quality MUSE spectrum ($z \approx 2.47$), we find no connected absorbers for this source. For comparison, we found that only $\approx 25\%$ of all spectroscopically confirmed HST-selected galaxies with $\log M_*/M_\odot \geq 8$ are associated with absorbers, increasing to $\approx 45\%$ considering galaxies with $\log M_*/M_\odot \geq 10$. Therefore, the association rate for galaxies selected in their rest-frame optical/UV is much lower than that observed for on-average more massive galaxies selected with ALMA in their dust continuum (see, Sect 4.5.2).

Interestingly, we note that three out of the six sources (i.e., C1, C3, C4) for which we find a robust association to absorbers would have been missed if only relying on MUSE and HST identifications. Indeed, such sources are the faintest among the NIR/optical counterparts (see Table 4.1) and lie in the MUSE redshift desert. This implies that studies limited to optical wavelengths may miss a significant fraction of the galaxy population around metal absorbers, as also highlighted in previous works (see, e.g., Neeleman et al., 2017, 2018, 2019; Fynbo et al., 2018; Galbiati et al., 2024b). Moreover, as the case of C4 exemplifies, ALMA is critical to identify the only galaxy we can associate with the QSO-NW absorber at $z = 2.368$, as no other galaxies were previously known at this redshift through MUSE and HST (see, Fig. 4.5). We evaluated the number density of our ALMA-selected galaxies by computing the completeness-corrected 1.2-mm number counts and found that these are consistent with that derived in blank field from ASPECS (González-López et al., 2020). In light of the above considerations, we conclude that, since obscured massive galaxies selected with ALMA have sufficient number density to be relevant in absorption line studies, blind (sub-)millimeter follow-ups of optical/NIR surveys are thus essential in providing the most complete view of the galaxy population residing around absorbers.

4.5.2 Properties of the ALMA galaxies around absorbers

In Fig. 4.6, we compare the stellar mass and SFR of the dusty galaxies detected with ALMA (see Appendix C.1 for full details) to the rest of the optical/UV-selected galaxy population in the MUDF and the sample of submillimeter galaxies (SMG) from [Dudzevičiūtė et al. \(2020\)](#) at $z \sim 1 - 2.5$. The ALMA sources lie at the high-mass end of the stellar mass function at those redshifts and are among the most massive members of the groups in which they reside (see Sect. 4.5.3). On the other hand, our ALMA galaxies appear to be less massive, but within in the scatter of the population of coeval SMGs, lying on the asymptotic part of the main sequence at $z \simeq 1 - 2$ (see, [Popesso et al., 2023](#)), although close to the starburst region.

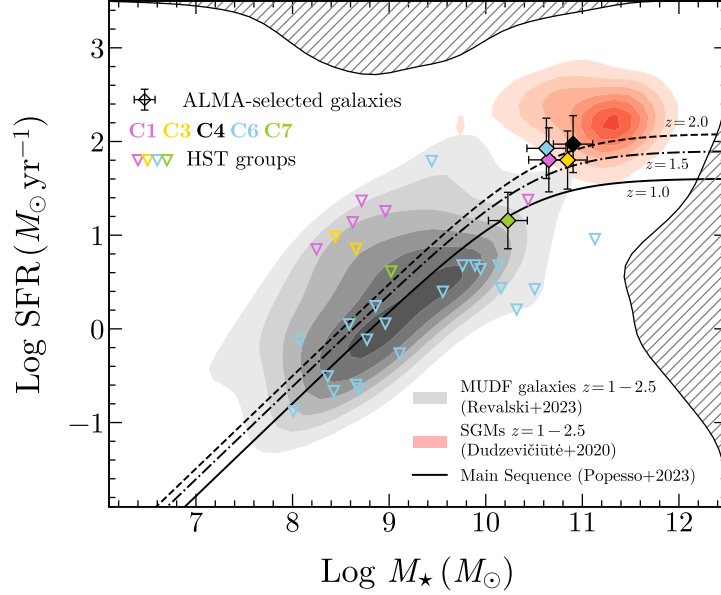


Figure 4.6: Star formation rate-stellar mass diagram of galaxies detected in the MUDF. The gray density plot shows the distribution of all galaxies selected with HST in the rest-frame optical/UV at $z \sim 1 - 2$. The marginalized distributions of $M_\star$ and SFR are also shown on the top and right axis, respectively. The red density plot report the sample of SMG with photometric redshift within $1 \leq z \leq 2.5$ from [Dudzevičiūtė et al. \(2020\)](#). The ALMA-selected galaxies with properties derived from SED modeling are reported with diamonds. Colored triangles mark the HST-selected group galaxies associated with the same absorption system of each ALMA-selected galaxy. The star-forming main sequence at $z \sim 1 - 2$ from [Popesso et al. \(2023\)](#) is also shown.

Using our measurements of the galaxy flux density at 1.2 mm on the Rayleigh-Jeans tail of the dust SED, we provide rough estimates of the dust masses (M_{dust}) of ALMA continuum-selected galaxies (excluding C5 for which we do not have reliable redshift measurement, and C2 that is the QSO-SE). By following a typical approach adopted in (sub-)millimeter studies of high- z galaxies (e.g. [Casey, 2012](#); [Riechers et al., 2013](#); [Faisst et al., 2020](#)), we employ a single modified black-body emission assuming a characteristic dust temperature of $T_{\text{dust}} = 32$ K, and emissivity index $\beta = 1.8$ typical of SMGs (see [Dudzevičiūtė et al., 2020](#)), and we find values in the range $\log(M_{\text{dust}}/M_\odot) \simeq 8.0 - 8.6$. Using the same assumptions, we find a lower value of $\log(M_{\text{dust}}/M_\odot) \simeq 7.8$ for C7, which appears to be more similar to typical main-sequence galaxies. By integrating the black-body models we computed the IR luminosities and converted them to SFR_{IR} using the calibration from [Murphy et al. \(2011\)](#), and find values in the range $\log(\text{SFR}_{\text{IR}}/M_\odot \text{ yr}^{-1}) \simeq 1.6 - 2.4$, comparable to those we derived from galaxy SED modeling (see Table C.1).

From the CO(3-2) emission line in C1, C3, C4, and the [CI] $_{609 \mu\text{m}}$ line detected in C6 (see, Sect. 4.4.2), we derive estimates of galaxy molecular gas fractions. For C1, C3, and C4 we adopt a CO(3-2)-to-CO(1-0) conversion factor of $r_{31} = 0.80 \pm 0.14$ from [Boogaard et al. \(2020\)](#) and a range of α_{CO} . At the lower end, we use a value of $0.8 M_\odot / (\text{K km s}^{-1} \text{ pc}^2)$, typical of SMGs and quasar hosts (see, e.g., [Downes & Solomon, 1998](#)). At the upper end, we adopt $4.3 M_\odot / (\text{K km s}^{-1} \text{ pc}^2)$ as measured in giant molecular clouds in the Milky Way (see, e.g., [Bolatto et al., 2013](#); [Carilli & Walter, 2013](#)). For C6, we first derive atomic carbon mass from equation 2 in [Weiß et al. \(2005\)](#) and we convert it to molecular gas mass by assuming a C-to-H $_2$ abundance ratio of $X[\text{CI}]/X[\text{H}_2] = M_{\text{CI}}/(6M_{\text{H}_2}) = (8.4 \pm 3.5) \times 10^{-5}$, as derived by [Walter et al. \(2011\)](#) for a sample of $z \sim 2 - 3$

FIR-bright sources. By employing such recipes, we find $f_{\text{gas}} = M_{\text{H}_2}/(M_{\text{H}_2}M_{\star}) = 51_{-17}^{14}\%$, $24_{-10}^{15}\%$, $20_{-9}^{12}\%$, and $65_{-12}^{10}\%$ for C1, C3, C4, and C6 respectively. These values are scattered in the typical range observed in both main-sequence galaxies and ULIRGs/SMGs at $z \gtrsim 1$ (see, e.g., Tacconi et al., 2018; Birkin et al., 2021). Interestingly, we note that C1 and C6 exhibit high values of the gas fraction and also live in rich groups (see, Sect. 4.5.3). This fact might suggest that efficient gas accretion is in place in these structures.

4.5.3 The link between the large-scale galaxy environment and the absorbing gas

In Sect. 4.3.1, we found that many HST-selected sources do not live in isolation but are associated with galaxy groups instead. In Fig. 4.7, we show that four (out of six) of our dusty galaxies (i.e., C1, C3, C6, and C7) detected with ALMA actually live in such overdense regions⁶. This is also evident from Fig. 4.5, in which multiple galaxies selected in HST are found to be associated with the same absorbers as those identified around ALMA galaxies. The high association rate between galaxies in groups and absorbers has also been found in previous works (see, e.g., Nielsen et al., 2018a; Dutta et al., 2020; Hamanowicz et al., 2020; Lofthouse et al., 2023; Qu et al., 2023), suggesting that the large-scale environment plays a significant role in shaping the absorption properties. Indeed, these absorbers exhibit multiple components and are observed to be strong, showing typical rest-frame equivalent widths of $\text{EW}_{\text{CIV}} \gtrsim 0.3 \text{ \AA}$, $\text{EW}_{\text{MgII}} \gtrsim 1 \text{ \AA}$, and with complex kinematics. In particular, we note that dust-mass-selected sources living in groups of more than two galaxies (i.e., C1, C3, and C6) are associated to absorbers exhibiting high column density of $N_{\text{AIII, MgII}} \gtrsim 10^{16} \text{ cm}^{-2}$, while C4, and C7 show significantly lower values of $N_{\text{CIV, MgII}} \simeq 10^{13} - 10^{14} \text{ cm}^{-2}$. Other works which studied quasar absorption around galaxy groups at different redshifts reported similar evidence (see, e.g., Lan et al., 2014; Fossati et al., 2019; Dutta et al., 2021; Nelson et al., 2021; Galbiati et al., 2023, 2024b).

We now turn to investigate the 3D separation between galaxies and the associated gas structures. The dusty galaxies selected with ALMA are more aligned in redshift space with the center of absorption features than the other connected galaxies. In particular, we find that C1 and C6, which are embedded in rich groups, are located in correspondence of the 50% of the optical depth (τ_{50}) of the associated absorbers (see, Fig. 4.5). Similarly, despite living in less overdense regions, we find that the remaining ALMA galaxies (i.e., C3, C4, and C7) lie within 100 km s^{-1} from τ_{50} . This fact, in combination with their massive nature (see, Sect. 4.5.2), suggests that these galaxies sit close to the center of the potential well of their large-scale structures. While our ALMA-selected galaxies show preferential alignment with the quasar absorbers along the line of sight, this is not necessarily the case if we consider the sky-projected distances from the location of the quasars. The impact parameters are typically large, of the order of $\sim 200 - 400 \text{ kpc}$ (see, Fig. 4.7), compared to the other galaxies associated to the absorbers. Therefore, while (sub-)millimeter observations are critical to gain a complete view of the galaxy population and possibly the group potential well, our findings suggest that the absorbing gas may not be located in the inner CGM of the dusty galaxies but is still retained within a few virial radii⁷. In this regard, massive dusty galaxies appear to be more tracers of the large scale structure than the actual hosts of the absorbers. However, we caution that this may also be a consequence of the low number density of these galaxies. More SMG-QSO pairs at small separations are needed to statistically assess our conclusions. We note that only one out of six detections, that is the source C3, lies both close to the center of the associated absorber in redshift space and at low impact parameter ($R = 43 \text{ kpc}$) from QSO-NW sightline, which points to a significant contribution from the CGM of this galaxy to the observed absorption profile. However, given that C3 belongs to a compact system of three sources all located within 100 kpc , we cannot exclude that hydrodynamical and gravitational interactions may play an important role in shaping the properties of the surrounding gas (Gunn & Gott, 1972; Merritt, 1983; Boselli & Gavazzi, 2006; Klitsch et al., 2019).

In light of the aforementioned considerations, the observed absorption profiles associated to groups could potentially arise from the superposition of the halos of multiple galaxies located around the massive sources identified with ALMA (see, e.g., Bordoloi et al., 2011; Nielsen et al.,

⁶From the total stellar mass of the groups we estimated the dark matter halo masses of the systems in which ALMA galaxies reside by employing the stellar-to-halo mass relation from Moster et al. (2010), and obtain an average halo mass of $\sim 10^{13} M_{\odot}$. This is comparable to previous estimates of the halo mass of SMGs at $z > 1$ from clustering analysis (Hickox et al., 2012; Wilkinson et al., 2017; Lim et al., 2020; Stach et al., 2021).

⁷The virial radius for a halo mass of $10^{13} M_{\odot}$, typical of the ALMA-selected galaxies (Stach et al., 2021) according to Moster et al. (2010), is ≈ 310 and 220 kpc at $z = 1$ and 2 , respectively.

▽▽▽ HST groups

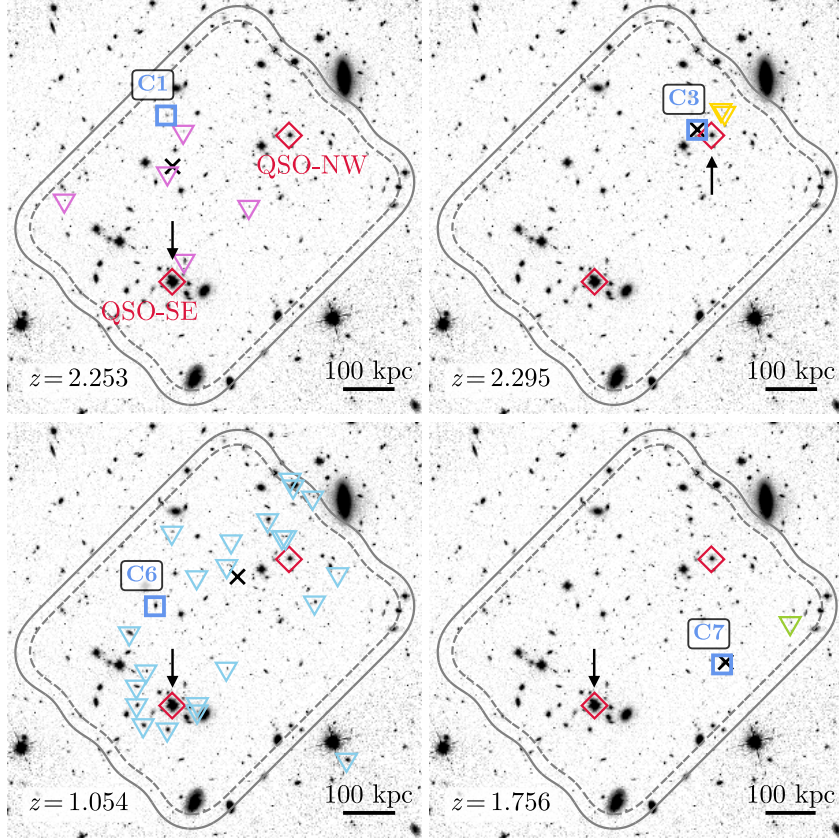


Figure 4.7: Groups of galaxies around ALMA sources (blue squares). The colored triangles indicate the HST-selected galaxies belonging to the same group found around the redshift of ALMA sources (bottom left corner). The two quasars in the MUDF are labeled with red diamonds, the black arrow indicates the QSO sightline in which absorption at the redshift of each group is detected. The black crosses mark the projected location of the galaxy stellar mass-weighted barycenter of the groups. The gray solid and dashed lines report the PB response at 0.3 and 0.5, respectively, of the ALMA band 6 mosaic.

2018a; Dutta et al., 2020). Therefore, one might expect the strength of the absorption to correlate with the projected galaxy surface density (see, e.g., Galbiati et al., 2024b). To test this hypothesis, we centered on the sky position and redshift of ALMA galaxies C1, C3, C6 and C7, and count the number of HST-selected sources within a velocity window of $|\Delta v| \leq 250 \text{ km s}^{-1}$ (chosen based on the broadest absorption system, see Fig. 4.5), and encompassed radius corresponding to the impact parameter from the QSO. We then divide the number of galaxies by the projected area corrected for the coverage of the MUSE FoV. In Fig. 4.8, we report such galaxy surface density against the EW of the associated absorbers⁸. As a result, we find a hint of positive correlation⁹, which is consistent with the scenario proposed above. In order to show the potential correlation, we fit the data with a linear model by employing the Python MCMC ensemble sampler `emcee` package (Foreman-Mackey et al., 2013) adopting a Poisson maximum likelihood estimator and flat priors on the free parameters. In Fig. 4.8, we report the best fit and 1σ confidence interval. However, we note that our analysis relies on a handful of sources and therefore is limited by the low statistics. Hence, surveys tailored to probe a broader range of galaxy overdensities around absorbers are required to draw more firm conclusions.

⁸We used Mg II for C1, C6, and C7, and Al II for C3, since we have no Mg II detection for this source. These are chosen as they have similar $f_{\text{osc}}\lambda_{\text{rest}}^2$, where f_{osc} is the oscillator strength, and λ_{rest}^2 is the rest-frame wavelength of the transitions. However, we caution that due to the differences in the relative abundances, at fixed EW different ions do not trace the same gas density.

⁹The Spearman’s rank correlation test (Spearman, 1904) yields a degree of monotonic correlation of $\rho = 0.8$ with a p -value 0.1, indicating that a positive correlation is in place at 80% confidence level.

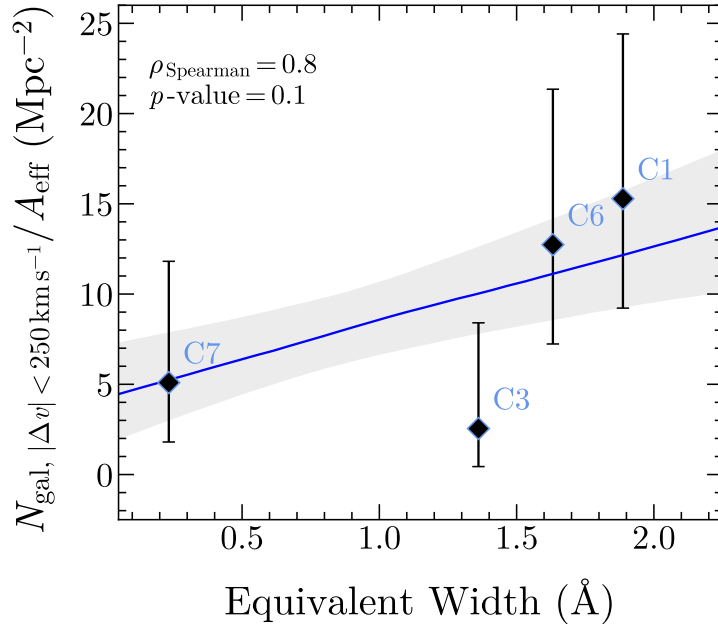


Figure 4.8: Galaxy surface density in groups as a function of EW of the associated absorbers (Mg II for C1, C6, and C7, and Al II for C3, respectively). Number density is computed by counting the number of galaxies (N_{gal}), and dividing for the effective area (A_{eff}) as described in Sect. 4.5.3. The associated uncertainties are obtained from Poisson statistics for low number counts (Gehrels, 1986). The blue solid line and gray shaded area are the best fit linear model ($N_{\text{gal}}/A_{\text{eff}} = a \times (\text{EW}/\text{\AA})b \text{ Mpc}^{-2}$, with $a = 4 \pm 2$, $b = 4 \pm 3$) and 1σ confidence interval, respectively. The Spearman’s rank correlation coefficient and the corresponding p -value are also reported.

4.6 Summary & Conclusions

In this work, we presented an ALMA blind dust continuum and molecular gas survey around the sightlines of the $z \approx 3.22$ quasar pair in the MUDF. By combining our (sub-)millimeter observations with available multiwavelength data from NIR to UV, we investigated the role of the dusty galaxy population in association to the intervening gas structures revealed in absorption in the quasar spectra. Below, we summarize our main findings:

- We identified a robust sample of seven sources detected at 1.2-mm in their dust continuum with ALMA. We determined their spectroscopic redshift from either the inspection of their MUSE spectra or via the CO(3–2) line detected with ALMA band 3 spectral scans.
- We found that the dust-selected galaxies and metal absorbers are associated within 500 km s^{-1} . This implies that at least $\geq 85\%$ of the dusty galaxies uncovered with ALMA are unambiguously linked to multiple absorption features arising from metal ions and observed in either of the quasar sightlines. This association rate is significantly higher than in previous surveys at NIR–UV wavelengths.
- By combining multiband photometry and information about dust and molecular gas gathered from our (sub-)millimeter data, we investigated the nature of the ALMA-selected galaxies by performing SED modeling. As a result, most of these galaxies appear massive ($M_{\star} \simeq 10^{10.2-10.7} M_{\odot}$), highly star forming ($\text{SFR} \simeq 10^{1.2-2.0} M_{\odot} \text{ yr}^{-1}$), and exhibit high dust and molecular gas content ($M_{\text{dust}} \simeq 10^{7.8-8.6} M_{\odot}$, $f_{\text{H}_2} \simeq 20 - 65\%$), resembling typical SMGs at similar redshifts.
- We discovered that a large fraction of ALMA-selected sources (4/6) live in galaxy groups in which they are among the most massive members. While the associated absorbers show complex kinematics, our galaxies are better aligned with the center of the absorption features than the other galaxies previously identified in the optical. This may suggest that these massive and dust-obscured galaxies sit close to the center of the gravitational potential well of their large-scale structures. However, the impact parameters of the ALMA galaxies are typically large, $\sim 200 - 400 \text{ kpc}$, indicating that absorption probed by the QSO sightlines might originate within a few virial radii.

This study highlights the crucial role of (sub-)millimeter surveys to obtain a complete census of the galaxy population around metal enriched cosmic gas structures revealing the presence of massive and dust-obscured galaxies. However, further efforts are needed to understand which processes are responsible for distributing metals into the IGM and how different galaxy populations contribute to the metal enrichment across the cosmic time. Furthermore, denser surveys are required to statistically investigate in absorption the properties of the inner CGM gas at the interface with the ISM of dusty galaxies.

Connecting the growth of galaxies to the large-scale environment in a massive node of the Cosmic Web at $z \sim 3$

5.1 Introduction

According to the current cosmological model, the majority of the inter-galactic gas is distributed in a network of filaments and sheets, called the Cosmic Web, formed through gravitational collapse of the initial density fluctuations (Bond et al., 1996). Structures form in the densest part of the Cosmic Web and grow hierarchically, merging into the massive virialized groups and clusters we observe in today’s universe at the intersection of filaments (see e.g., de Lapparent et al., 1986; Geller & Huchra, 1989; Alpaslan et al., 2014; Libeskind et al., 2018). This large-scale environment is thought to play a key role in regulating how the embedded galaxies form and evolve across the cosmic time. Indeed, galaxies found in virialized clusters at $z \lesssim 1$ are more massive and form stars at a lower rate compared to the galaxies living in the field (see e.g., Baldry et al., 2006; Bamford et al., 2009; Peng et al., 2010; Calvi et al., 2013; Tomczak et al., 2017; Lemaux et al., 2019; Tomczak et al., 2019; Old et al., 2020; Chartab et al., 2020; van der Burg et al., 2020; McNab et al., 2021). It is thus implied that the galaxies inhabiting extremely overdense regions assemble their mass rapidly and undergo intense episodes of star formation (see e.g., Thomas et al., 2005; Raichoor et al., 2011; Gu et al., 2018). Therefore, the $z \gtrsim 2$ progenitors of present-day clusters, often referred to as protoclusters (Sunyaev & Zeldovich, 1972; Cen & Ostriker, 2000; Overzier, 2016; Alberts & Noble, 2022), are an ideal laboratory within which we can study the fundamental link between galaxies and their large-scale environment.

Overdensities of Ly α and H α emitters, Lyman Break Galaxies (LBGs) or dusty star-forming galaxies are commonly used as a tracer for candidate clusters’ progenitors (see e.g., Steidel et al., 1998, 2000; Kurk et al., 2000, 2004; Hayashino et al., 2004; Chapman et al., 2009; Tanaka et al., 2011; Lee et al., 2014; Casey et al., 2015; Oteo et al., 2018; Shi et al., 2019a,b), although their spatial distribution does not necessarily follow fully virialized structures (Lovell et al., 2018; Toshikawa et al., 2020), unlike $z \approx 0$ clusters. Indeed, at $z \gtrsim 2$, Chiang et al. (2013) predicted protoclusters to extend over $\approx 10 - 20$ cMpc, while some well studied structures show multiple peaks in their galaxies’ overdensity and extend on even much larger scales (see e.g., Steidel et al., 1998; Pentericci et al., 2000; Cucciati et al., 2018). In similar $z \gtrsim 2$ overdense environments, galaxies are found to be sites of higher star formation activity and their stellar mass function shows an excess at the high-mass-end compared to the field (see e.g., Kodama et al., 2007; Koyama et al., 2013a,b; Shimakawa et al., 2018a,b; Greenslade et al., 2018; Miller et al., 2018; Ito et al., 2020; Lemaux et al., 2022; Taamoli et al., 2024; Toshikawa et al., 2024) as well as enhanced active galactic nuclei (AGN) activity (e.g., Lehmer et al., 2009; Tozzi et al., 2022b; Gatica et al., 2024). A high rate of mergers (Alonso et al., 2012; Mei et al., 2023), efficient gas accretion (e.g., D’Amato et al., 2020) and interactions with a *proto*-intracluster medium (Di Mascolo et al., 2023) are expected to play a key role in regulating galaxies growth and activity in these high-redshift overdensities. Though star forming galaxies are observed to be the dominant population in clusters’ progenitors (Kuiper et al., 2010; Spitler et al., 2012; Contini et al., 2016), the detections of quiescent galaxies has been reported in some structures (Kodama et al., 2007; Kubo et al., 2013; Shi et al., 2019b). These observations suggest that protoclusters could be ideal environments to study the onset of environmental quenching and the processes driving the suppression of the star formation activity

observed in $z \approx 0$ clusters' galaxies. Theoretical works (e.g., Muldrew et al., 2015; Chiang et al., 2017; Lovell et al., 2018) are consistent with these observational results, and both support a picture in which galaxies in protoclusters have assembled their masses and rapidly grown at earlier times than those in the field.

Complementary to blind searches, overdense structures of the Inter-Galactic Medium (IGM) itself (mapped via three dimensions Ly α tomography, see e.g., Lee et al., 2018; Newman et al., 2020, 2022), as well as the surroundings of highly biased objects, such as High-Redshift Radio Galaxies (HzRGs; e.g., Roettgering et al., 1994; Pentericci et al., 2000; Hayashi et al., 2012; Hatch et al., 2014; Rigby et al., 2014; Orsi et al., 2016; Pérez-Martínez et al., 2023), Hot Dust-Obscured Galaxies (Hot DOGs; Jones et al., 2014; Assef et al., 2015; Penney et al., 2019; Luo et al., 2022; Ginolfi et al., 2022; Zewdie et al., 2023) and luminous quasars (QSOs; e.g., Kashikawa et al., 2007; Venemans et al., 2007; Overzier et al., 2008; Priddey et al., 2008; Kim et al., 2009; Utsuni et al., 2010; Bañados et al., 2013; Husband et al., 2013; Adams et al., 2015; Mignoli et al., 2020; Kashino et al., 2023) have been typically explored as signposts of galaxy overdensities. In particular, high-redshift quasars are among the brightest objects in the $z \gtrsim 2$ universe and are thus expected to trace overdense structures. Indeed, a hyperluminous quasar at $z \approx 2.84$ was found at the center of a large overdensity of LBGs (Steidel et al., 2011), Ly α emitters (Kikuta et al., 2019) and sub-millimeter galaxies (Lacaille et al., 2019; Wang et al., 2024a), tracing a large-scale structure extending across ≈ 100 cMpc. Garcia-Vergara (2021) explored the galaxy clustering around $z \approx 4$ quasars and found a large overdensity of different galaxy populations, that is CO(4–3) line emitters and Ly α emitting galaxies. On the other hand, Falder et al. (2011); Trainor & Steidel (2012); Adams et al. (2015); Hennawi et al. (2015) reported the detection of galaxy overdensities in the environment of luminous quasars in only $\approx 10\%$ of the cases.

Despite the increasing number of $z \gtrsim 2$ galaxy overdensities and candidate protoclusters discovered in the literature, the heterogeneity of the tracers and the methods adopted to identify the galaxies in their surroundings makes it challenging to draw a unified picture and to consistently trace the assembly and the evolution of the overdense structures down to present-day clusters. Comparing the galaxy environment around radio-loud AGNs from the Clusters Around Radio-Loud AGNs (CARLA; Wylezalek et al. 2013) program and radio-quiet quasars with similar masses and redshifts, Hatch et al. (2014) found that the choice of the biased tracer itself has an effect on the large-scale galaxy environment identified in its surroundings. Indeed, on average, radio-loud AGN are observed to be surrounded by larger overdensities and to trace larger structures than radio-quiet quasars. Additionally, differences among overdensities of galaxies selected with a similar criterion around the same biased tracer may arise as a consequence of these structure being in a different evolutionary stage, as shown by, e.g., Shimakawa et al. (2018b).

We present in this work a multiwavelength survey of the large-scale galaxy environment around the $z \approx 3.25$ quasar CTS G18.01, which has been pre-selected for being surrounded by extended and filamentary-shaped Ly α emission from the cosmic gas (named MQN01), as discovered by Borisova et al. (2016) and as confirmed later by deeper observations covering a wider area (Cantalupo et al., in prep.). In particular, we targeted the field with a mosaic of observations taken with the Multi Unit Spectroscopic Explorer (MUSE; Bacon et al., 2010) at the Very Large Telescope (VLT), the FOcal Reducer and low dispersion Spectrograph 2 (FORS2; Appenzeller et al., 1998) and the High Acuity Wide field K-band Imager (HAWK-I; Pirard et al., 2004; Casali et al., 2006; Kissler-Patig et al., 2008; Siebenmorgen et al., 2011) at the VLT. The combination of these datasets aims at selecting star forming galaxies at the redshift of the quasar across 24×24 cMpc² and provides spectroscopic confirmation within an area of 4×4 cMpc² that is covered by MUSE. Additional HST/ACS, JWST/NIRCam and *Chandra* observation are used to derive star formation rates and stellar masses by modeling the galaxies' spectral energy distribution (SED) and to explore this peculiar environment in the context of galaxies' formation and evolution. The goal of this work is thus to exploit this rich multiwavelength dataset to explore an overdense region at $z \approx 3$ (Pensabene et al., 2024b) and unveil the roles of both the local and large-scale environment on the stellar mass assembly and star formation activity of the galaxies populating this unique field.

We describe the data in Section 5.2 and present in Section 5.3 the samples of galaxies identified around the target quasar. In Sections 5.4.1–5.4.3 we investigate the overdensity of star-forming galaxies within 4×4 cMpc² around the quasar and explore their properties, compared to galaxies in the field, in Sections 5.4.4–5.4.6. We then explore the overdensity of LBGs across 24×24 cMpc² in Section 5.4.7. Finally, we discuss this structure in the context of galaxies formation and evolution in Section 5.5. Throughout, unless otherwise noted, we quote magnitudes in the AB system, distances in physical units, and adopt the Planck 2015 cosmology ($\Omega_m = 0.307$, $H_0 = 67.7$ km s^{−1}

Mpc⁻¹; Planck Collaboration et al. (2016)).

5.2 Observations and data reduction

5.2.1 MUSE observations

The main dataset consists of a mosaic of four VLT/MUSE pointings around the quasar CTS G18.01. The observations are part of a MUSE Guaranteed Time of Observations (GTO) programme ID 0102.A-0448(A), PI S. Cantalupo. They were all conducted in October 2018 in visitor mode with an integration time of, on average¹, ≈ 10 hours per pointing using MUSE with AO in Wide-Field mode. The data were taken on clear nights, at airmass ≤ 2.8 , angular distance from the moon ≥ 30 deg, lunar illumination $\leq 20\%$ and requiring an image quality better than $\text{FWHM} \leq 0.8$ arcsec. The different frames of each pointing were observed with relative inclination of 90 degrees and small dithers ($\lesssim 1$ arcsec) to mitigate the differences in the performance of the 24 IFUs (see e.g., Borisova et al., 2016; Marino et al., 2018). Each of the pointing covers 1×1 arcmin², resulting in a final mosaic with a Field-of-View (FoV) of $\approx 2 \times 2$ arcmin² which corresponds to 4×4 cMpc² at the redshift of the central quasar, $z \approx 3.25$.

The raw data were initially reduced with the ESO Muse pipeline (Weilbacher et al., 2014, version 2.8.1). The SCIBASIC routine of the ESO Muse pipeline was used to subtract bias, apply flat fielding, perform twilight and illumination corrections, calibrate the wavelengths and the line spread function using arcs and reduce the sky flats. The same calibrations are applied to the field targeting the standard star that is then used to calibrate the flux. The known astrometric offsets (Bacon et al., 2015) are corrected by registering the datacubes using the point sources identified in each collapsed individual frame. Once aligned, all the individual frames are combined in a final datacube by using a 3D drizzling interpolation process. The final stacked datacube produced by the ESO pipeline are known (see e.g., Borisova et al., 2016; Bacon et al., 2017; Marino et al., 2018; Cantalupo et al., 2019) to be affected by imperfections in the flat-fielding process and instabilities of the detector. To correct for these effects and improve the quality of the datacubes, we post-process the final products using the CUBEXTRACTOR (CUBEX hereafter, see, de Beer et al., 2023 and Cantalupo et al., in prep.) package, version 1.8, and its subroutines (see figure 1 in Lofthouse et al. 2020b for an example of the post-processing results compared to the science products reduced by the ESO pipeline). In details, CUBEFIX applies the flat-fielding correction by using the observed sky to self-calibrate the exposures. This process is applied iteratively after masking the continuum-detected sources to avoid contamination. The result is an optimal re-alignment of the relative illumination between the stacks and slices of each IFU, which are not completely corrected by the daily flats. CUBESHARP performs flux-conserving sky subtraction by including empirical corrections of the line spread function. By means of the CUBECOMBINE tool, we then combined the individual exposures into a single datacube. Such cube is then employed to mask continuum sources detected in the collapsed white-light image to optimize a second iteration of CUBEFIX and CUBESHARP. Finally, CUBECOMBINE is used to produce the final and fully-reduced datacube.

5.2.2 FORS2 and HAWK-I mosaics

We followed-up the MUSE observations with a joint VLT/FORS2 and VLT/HAWK-I programme (ID 110.23ZX, PI S. Cantalupo). The FORS2 mosaic is designed to cover an area of $\approx 24 \times 24$ cMpc², which corresponds to a FoV that is $\approx 6 \times 6$ larger than the one of MUSE and to overlap in the same area covered by MUSE. The observations were executed during ESO Cycle 110-111 in both visitor and service mode, as detailed below and consists of imaging in three different filters: U_{High} (7h on source, $\lambda_{\text{eff}} = 361$ nm), B_{High} (5h on source, $\lambda_{\text{eff}} = 437$ nm) and R_{Special} (45 minutes on source, $\lambda_{\text{eff}} = 655$ nm). The U-band data were taken in visitor mode during two half-nights in October 2022 (Run ID 110.23ZX.001) with the blue-sensitive E2V FORS2 CCD. The B-band and R-band observations were taken in service mode with the MIT CCD. The imaging observations in all the three filters were executed with the default CCD read-out mode for imaging, 2×2 binning and low gain. Our observational constraints include clear nights, airmass ≤ 1.5 , $\leq 20\%$ of lunar illumination and a seeing better than 0.8 arcsec FWHM. We observed each mosaic position twice and applied

¹The final, combined cube consists on a 2×2 mosaic of 4 pointings. All of these overlaps in a small region ≈ 60 kpc away from the quasar, where the integrated exposure time is of 20h. In all the other regions it ranges from 8h, at the edges, to 12h.

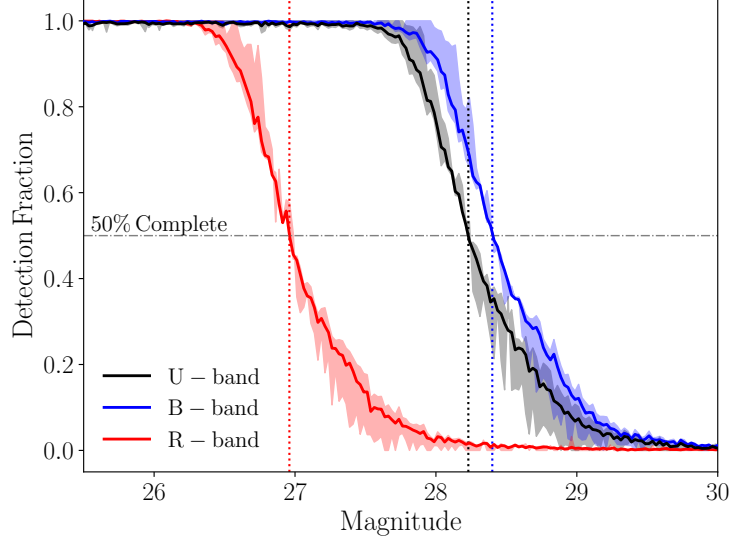


Figure 5.1: Completeness estimated for the VLT/FORS2 images in the U (blue), B (black) and R band (red). The solid line is the result across the entire FoV, while the shaded regions reproduce the variation from the shallower to the deepest regions of each mosaic. The vertical dotted lines mark the magnitudes corresponding to the 50% completeness limit (black horizontal line).

small dithers between the individual frames in each individual OB to improve contamination removal, background subtraction and to cover the gap between the two chips. To reduce the data, we used the ESO pipeline recipes for FORS2 (Izzo et al., 2010, version 5.6.2) to apply standard bias, flat and flux calibrations². The 2D background was locally modelled and subtracted using the python package PHOTUTILS (Bradley et al., 2023, version 1.9.0). Finally, we aligned the astrometry of each individual frame with stars from Gaia DR3 catalog (Gaia Collaboration et al., 2016, 2023) by using the SCAMP package (Bertin, 2006, version 1.20). The reduced and registered frames are sigma-clipped and combined into the large-scale mosaic using SWARP (Bertin, 2010, version 2.21). To characterize the depth of the final mosaics, we measured the 3σ limiting magnitude of each image (see Table 5.1) within apertures of radius $r = 4$ pixels ≈ 1 arcsec, corresponding to twice the maximum value of the PSF, and placed in 5000 random positions in blank regions, with all the sources masked to avoid contamination. To estimate the image quality of the science products we used the software PSFEX (Bertin, 2011) which extracts a sample of unsaturated stars to measure the Point Spread Function (PSF) across the entire FoV accounting for spatial variations. We report the mean value across the FoV in Table 5.1, which will also be used to perform accurate photometry later in the analysis: the variation between the minimum and maximum PSF across the FoV is $\lesssim 10\%$ for all the images.

Filter	$m_{3\sigma}^a$ (mag)	$m_{50\%}^b$ (mag)			PSF ^c (arcsec)
U	27.90	28.25	28.23	28.20	0.78
B	27.45	28.40	28.38	28.33	0.65
R	26.08	26.99	26.96	26.91	0.60

Table 5.1: Depth and PSF of the VLT/FORS2 imaging observations. ^(a) 3σ magnitude limit measured in apertures of radius $r = 4$ pixels ≈ 1 arcsec. ^(b) 50% completeness limit for the deep, medium and shallow part of the mosaic. ^(c) Mean PSF (FWHM).

As a final step, we performed injection/recovery experiments to estimate the completeness of the science mosaics. We first created background-only versions of the images by replacing each pixel belonging to known sources, as identified by SExtractor (Bertin & Arnouts 1996, version 2.24.2) using the same parameters listed in Section 5.3.1, with a value randomly extracted from a normal

²The photometric calibration of the B-band and R-band observations is done with the extinction-corrected instrumental zero-point available on the ESO data quality-control database (<http://archive.eso.org/qc1/qc1.cgi>). These measurements are processed using the ESO pipeline by combining standard star observations taken for at least 7 stable nights around the nights of the observations. For the U-band images, instead, we run the ESO pipeline on the daily standard frames to measure the zero-point.

distribution centered on the median value of the background. To create the artificial sources, we convolved point sources with the PSF and randomly injected 1500 mocks for each $\Delta m = 0.025$ mag bin of magnitude, from $m = 24$ mag to $m = 31$ mag, in the background-only mosaics. The number of mocks is carefully chosen to avoid overlap and deblending issues. We then run SEXTRACTOR and measured the fraction of sources we can recover as a function of the magnitude. We compute the results (see Figure 5.1 and Table 5.1) for the full mosaic, distinguishing three different regions (the deepest one is designed to overlap with the MUSE FoV at the center of the mosaic) depending on the number of exposures. The survey strategy described above for the FORS2 observations was employed to image the same area of $\approx 24 \times 24$ cMpc² around the target quasar with VLT/HAWK-I. We observed the field with a set of three filters: CH4 (1h on source, $\lambda_{\text{eff}} = 1.575 \mu\text{m}$), H (2h on source, $\lambda_{\text{eff}} = 1.620 \mu\text{m}$) and K_s (2h on source, $\lambda_{\text{eff}} = 2.146 \mu\text{m}$). The observations were executed in October-December 2022 requiring the same constraints listed for FORS2. The raw data were bias subtracted, flat fielded and flux calibrated by using the ESO Pipeline recipes for HAWK-I imaging (version 2.4.13) implemented in ESOREX (ESO CPL Development Team 2015, version 3.13.6). The 2MASS (Skrutskie et al., 2006) catalogue is used to measure the flux zero-point. After subtracting the background, we used the HAWKI_SCIENCE_POSTPROCESS recipe in the ESO Pipeline to combine the exposures into the final science mosaic. As a final step, we registered the astrometry of the final mosaic to Gaia by using the DRIZZLEPAC package³ (Gonzaga et al., 2012; Hoffmann et al., 2021).

5.2.3 Multiwavelength photometry and ancillary data

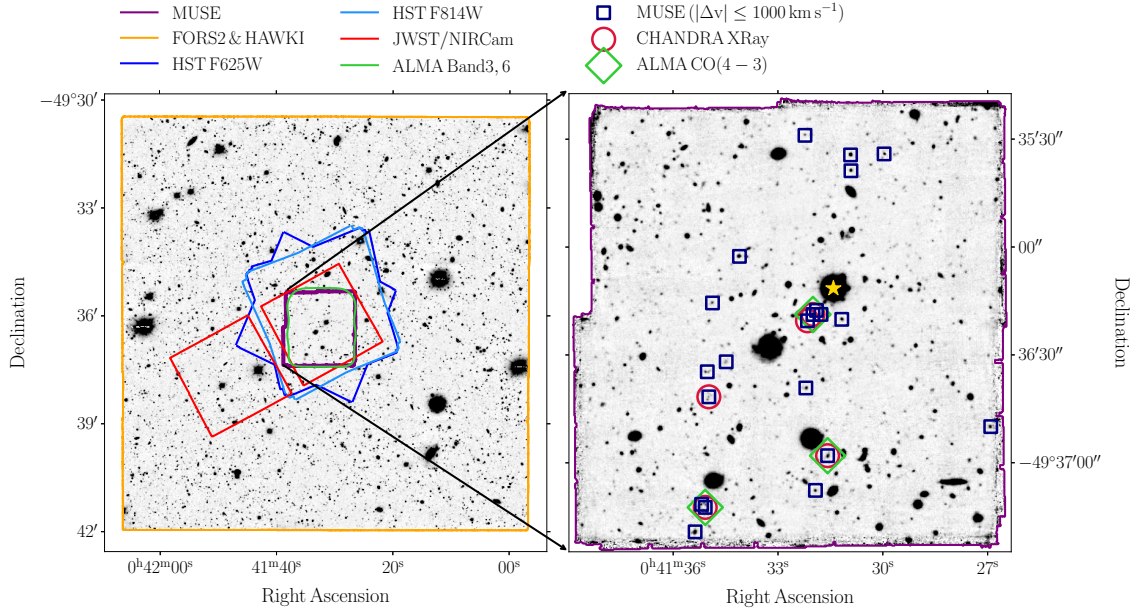


Figure 5.2: Visualisation of the multiwavelength data explored in this work. Left panel: FORS2/R-band image with overlaid the footprint of the MUSE (purple), FORS2 and HAWK-I (orange), HST/ACS F625W (blue), HST/ACS F814W (lightblue), JWST/NIRCam (red) and ALMA/Band3 and Band6 (green) observations. Right panel: the zoom-in shows the MUSE White-Light image in the background with highlighted the MQN01 high-confidence sample of galaxies found in the MUSE FoV within a line-of-sight separation $|\Delta v| \leq 1000 \text{ km s}^{-1}$ around the quasar (blue squares), the cross-matched CO(4–3) line-emitters found by ALMA (green diamonds, see Pensabene et al. 2024b) and X-Ray sources detected by *Chandra* (red circles, see Travascio et al., submitted). The target quasar is marked by a gold star.

This field has also been targeted by a multiwavelength campaign spanning from the X-Ray to the millimeter wavelengths (see the left panel in Figure 5.2), which enables us to perform a complete census of the different population of sources in the environment of the central quasar, including AGN and dusty star-forming galaxies, that may be elusive in optical observations. Travascio et al., (submitted) collected 600ks of broadband (0.57.0 KeV) observations with the Advanced CCD Imaging Spectrometer (ACIS) mounted onboard the *Chandra* spacecraft (Weisskopf et al., 2000) during Cycle 23 (PI: S.Cantalupo). The final FoV extends above the 24×24 cMpc² observed

³<https://www.stsci.edu/scientific-community/software/drizzlepac.html>

by FORS2 and HAWK-I and provides ≈ 1 arcsec resolution imaging and moderate resolution spectroscopy. We also observed the field with the Hubble Space Telescope (HST) for a total of 22 orbits in Cycle 30 (PID: 17065, PI: S. Cantalupo) and used the HST/ACS (Advanced Camera for Survey, Ford et al. 1998) instrument to image a $\approx 3.4 \times 3.4$ arcmin² FoV around the central quasar with the filters ACS F625W (25397 sec) and ACS F814W (30238 sec). We complemented our multiwavelength dataset with the *James Webb* Space Telescope (JWST, Rigby et al. 2023) pre-imaging observations associated to the GO Cycle 1 programme ID 1835 (Cantalupo et al. 2021, PI S. Cantalupo). With an on-source exposure time of $\approx 2 \times 27$ minutes, we pointed JWST/NIRCam to image a 2×5 arcmin² FoV with the extra-wide filters F150W2 (in the short wavelength channel, $0.6 - 2.3 \mu\text{m}$) and F322W2 (in the long wavelength channel, $2.4 - 5.0 \mu\text{m}$). Details on the data reduction are reported in Wang et al. (2024b). Finally, Pensabene et al. (2024b) targeted the field with a Cycle 8 ALMA programme (ID 2021.1.00793.S, PI S. Cantalupo) surveying the 4×4 cMpc² MUSE FoV with Band 3 and Band 6 mosaics to detect dusty star forming galaxies via their 1.2 mm continuum emission or CO(4–3) emission lines. The data and the reduction steps are detailed in Pensabene et al. (2024b).

5.3 Data analysis

We describe in this section the methods we adopted to search for galaxies around the target quasar. The rest-frame UV color selections used to identify the Lyman break galaxies in the FORS2 images are detailed in Section 5.3.1 and Appendix D.1. The sample of continuum-selected galaxies identified in the MUSE mosaic, as well as their spectroscopic redshift measurements, is described in Section 5.3.2.

5.3.1 Photometric selection of Lyman break galaxies

To identify the Lyman break galaxies at $z \approx 3 - 3.5$ across the FORS2 FoV, we calibrated a selection based on $U - B$ and $B - R$ colors (see e.g., Bielby et al., 2011). We carefully choose those photometric bands to avoid any possible contamination due to the presence of bright Ly α emission which may be arising from the recombination of the ionizing radiation emitted by the central quasar, rather than from the galaxies’ activity. To calibrate the selection region, we used BPASS SPS models (Stanway & Eldridge 2018, v2.2-100bin-t8.0-Z001) with a Salpeter IMF (Salpeter, 1955) up to a stellar mass limit of $M_{\star} = 100 M_{\odot}$, binary evolution and stellar metallicity $Z_{\star} = 0.001$ (corresponding to $Z_{\star} \approx 0.07 Z_{\odot}$ with solar abundances from Asplund et al. 2009). We assumed a constant star formation history and stellar age of $t = 10^8 \text{ Yr} = 100 \text{ Myr}$, chosen to be consistent with the median age of the stellar population of the galaxies identified in MUSE (see Section 5.3.4 and also Steidel et al. 2018 for a similar approach). We also considered discrete values of reddening: $E(B - V) = 0, 0.05, 0.10, 0.15, 0.20$, which are typical values for LBGs at $z \approx 3$ (see e.g., Shapley et al., 2003). The models are then post-processed to account for the attenuation due to the absorption from the inter-galactic medium. We thus generated 500 IGM realizations randomly extracted from 1000 sightlines from Steidel et al. (2018) and produced by inserting along each line-of-sight intervening H I absorption-line-systems, based on their column density distribution function reported in Rudie et al. (2013), to reproduce the stochastic nature of the IGM absorption. Finally, the color criterion requires:

1. $U > U_{\text{BKG}}^{3\sigma}$
2. $U - B > 1.89$
3. $B - R < 1.74 \times (U - B) - 1.24$
4. $0.31 < B - R < 2.20$

The first condition corresponds to requiring that the source is not detected in U-band above $3\sigma_{\text{BKG}}$, where σ_{BKG} is the RMS of the background measured within the same aperture used for colors, and is added to avoid contamination from sources at lower redshift and stars (see an application of this same criterion to select U-dropouts galaxies at $z \approx 3$ in van der Burg et al. 2010).

We choose the R-band as the detection image and run SExtractor with the following parameters: i) THRESHOLD=1.5 weighted on the mosaic exposure map to assure a uniform threshold

across the FoV; ii) MINAREA=5 pixels, corresponding to $\approx 1.3 \text{ arcsec}^2$; iii) default deblending parameters. The image is also filtered with a 3×3 Gaussian kernel to optimize the detection of faint objects. Colors are measured within a fixed circular aperture of $d = 5 \text{ pixels} \approx 1.26 \text{ arcsec}$, although we found that the results do not depend on the radius of the aperture by varying this within the range $0.5 - 2.0 \text{ arcsec}$. The radius is then re-scaled for the variations of the PSF between the different images, as listed in Table 5.1. The results are reported in Figure 5.3. All the candidates have been visually inspected to remove the objects that lie partially outside the edge of the FoV, the photometry of which cannot be performed on the full-size of the source. As a result, we identified 535 $z \approx 3.0 - 3.5$ LBGs across $24 \times 24 \text{ cMpc}^2$. The quality of the color selections is tested by cross-matching the catalogue of LBG candidates to the spectroscopic sample of galaxies identified in MUSE (Section 5.3.2), as reported in Table 5.2.

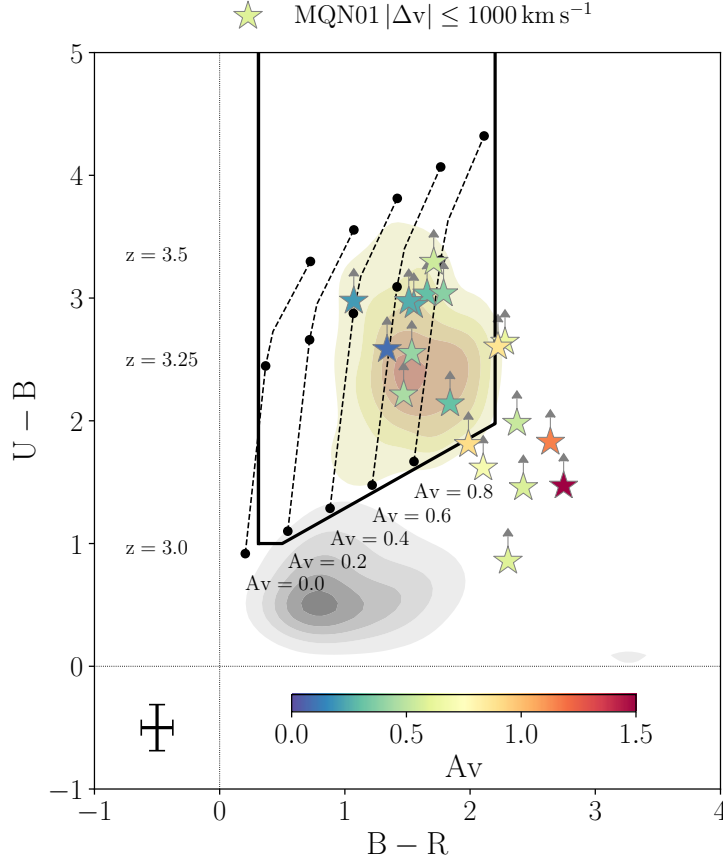


Figure 5.3: *UBR* color selection of LBGs. The selection region (delimited by solid black lines) is calibrated based on the redshift tracks of typical LBG templates (dashed lines). The orange density plot shows the LBGs candidate at $z \approx 3.0 - 3.5$ that are selected by our criterion. The sources with spectroscopic redshift from MUSE and found within $|\Delta v| \leq 1000 \text{ km s}^{-1}$ from the central quasar are shown as stars and are color-coded by their visual extinction, A_v , estimated by fitting their SEDs (see, Section 5.3.4). The gray density plot marks all the R-band detection within FORS2 FoV. Typical errors on the colors are shown in the lower-left corner.

5.3.2 MUSE spectroscopic sample of galaxies

By leveraging the deep MUSE observations, we searched i) for a spectroscopic confirmation of the color-selected Lyman break galaxies; ii) additional continuum-selected sources in close proximity to the central quasar that are not detected in the shallower FORS2/R-band nor color-selected as LBGs. To this end, we run SEXTRACTOR in the MUSE white-light image, obtained by collapsing the datacube along the wavelength direction, with the parameters: i) THRESHOLD=2.5 rescaled by the RMS of the background to guarantee a uniform threshold across the FoV; ii) MINAREA=6 pixels, corresponding to $\approx 1.2 \text{ arcsec}^2$; iii) minimum deblending parameter DEBLEND_MINCONT=0.0001. The spectra are extracted within an aperture $r = 2.5 \times R_{\text{Kron}}$ to collect the total flux of the sources.

We assign a spectroscopic redshift to the continuum-detected sources by using M. Fossati fork⁴ of MARZ tool (Hinton et al., 2016), which includes additional high-resolution templates suitable for high-redshift galaxies. We attribute confidence scores to the redshift measurements following the method illustrated in Bielby et al. (2019).

We are able to assign a high-confidence spectroscopic redshift to $\approx 38\%$ of the sources. We attribute low confidence to all the sources that exhibit bright Ly α emission at $z \approx z_{\text{QSO}}$ on top of a low signal-to-noise continuum and in absence of other spectral features. This choice is motivated by the presence of the bright target quasar, which prevents us from ruling out whether the Ly α emission is powered by a $z \approx z_{\text{QSO}}$ galaxy or it arises from recombination emission of the quasar radiation on top of a featureless continuum emitted by a foreground or background object. We thus require at least two of the following stellar and interstellar absorption or nebular emission lines to be clearly recognizable in the rest-frame UV spectra: Ly α 1216 Å, N V λ 1240 Å, Si II λ 1260 Å, O I+Si II λ 1303 Å, C II λ 1334 Å, Si II λ 1526 Å, C IV λ 1548, 1550 Å, Fe II λ 1608 Å, He II λ 1640 Å, Al II λ 1670, 1854, 1862 Å, C III λ 1906, 1908 Å.

With this method, we identify 21 galaxies with high-confidence spectroscopic redshifts lying within a line-of-sight separation $|\Delta v| \leq 1000 \text{ km s}^{-1}$ from the central quasar (right panel in Figure 5.2). To prove that our selection of galaxies associated with the QSO is robust against the choice of velocity window, we searched for galaxies within a larger range $|\Delta v| \leq 2500 \text{ km s}^{-1}$ in the MUSE cube and found that the number of sources increases only marginally, by $\lesssim 10\%$. In order to complete the characterization the galaxy population around the central quasar, we cross-match the catalogue of MUSE galaxies to the sample of X-Ray emitting AGN observed with *Chandra* (Travascio et al., submitted) and to that of dusty star-forming galaxies detected with ALMA by Pensabene et al. (2024b). We identified a total of five AGN detected in X-Ray, three of which are also CO(4–3) line emitters (see Figure 5.2).

As a final step, we cross-match the catalogue of color-selected Lyman break galaxies with the spectroscopic sample assembled in MUSE and report the results in Table 5.2. The majority of the LBG candidates ($\gtrsim 45\%$) is found within $|\Delta v| \leq 1000 \text{ km s}^{-1}$ from the quasar. Only $\approx 15\%$ has a redshift outside this narrow window, but within the range $2.8 < z < 3.5$ the color selection is calibrated on. The remaining sources have insufficient spectroscopic information to confidently constrain their redshifts. Compared with the spectroscopic sample, the galaxies that are not color-selected are on average $\times 1.2$ more massive (stellar mass $\log(M_*/M_\odot) \gtrsim 10$), and $\times 2.4$ more dust extincted (visual extinction $A_v \gtrsim 0.5$) than those classified as LBGs.

LBGs selection ^(a)	Area ^(b) [cMpc ²]	MUSE ^(c)
UBR	24×24	53% (10/19)
UGR	4×4	79% (15/19)

Table 5.2: Report of *UBR* and *UGR* color selections calibrated to identify $3.0 < z < 3.5$ Lyman break galaxies around the central quasar. ^(a) Color selections described in Section 5.3.1 and in Appendix D.1. ^(b) Size of the FoV on which the color-selections are performed. ^(c) Fraction of galaxies with high-confidence spectroscopic redshift within $|\Delta v| \leq 1000 \text{ km s}^{-1}$ from the quasar and detected in FORS2/R-band that are independently recovered by the color selections. The 3 sources that are not detected in FORS2/R-band are excluded from this statistics.

5.3.3 Control-field samples

We describe in this section the two independent control samples we built to properly compare the properties of the galaxies identified near the quasar to the expectations for random regions of the universe. To guarantee a consistent comparison, we make sure that these control samples have a similar selection function to the spectroscopic sample of galaxies identified in MUSE.

We first explored publicly available MUSE datacubes. We selected fields observed with 10 hours of exposure time and without any bright quasar at $z \approx 3.0 - 3.5$, which are known to be biased tracers of galaxies overdensities (see e.g., García-Vergara et al., 2017). In the final sample we included the most recent data release (Bacon et al., 2023) of the MUSE-Hubble Ultra Deep Field (UDF, Bacon et al., 2017), Q0055–269 and Q1317–0507 from the MUSEQuBES survey (Muzahid et al., 2020), which were all observed as part of the MUSE-GTO programme, and the archived J162116+004823 and J200324–325144 (Muzahid et al., 2020) from the MAGG survey (Lofthouse et al., 2020b; Fossati et al., 2021; Galbiati et al., 2023, 2024b). To be fully consistent with the

⁴<https://matteofox.github.io/Marz/>

selection function of this work, we searched these datacubes for continuum-detected galaxies and required high-confidence spectroscopic redshifts to be measured with the same criterion adopted in Section 5.3.2. As a result, we identify 43 galaxies (hereafter called “MUSE-FIELD” sample) with spectroscopic redshift in the range $2.5 < z < 4.5$ spread across an area over three times larger than the MQN01 MUSE FoV. As a complementary analysis, we also identified a control sample from our own MUSE datacube of the MQN01 including 11 continuum-detected galaxies within the range $2.5 \lesssim z \lesssim 4.5$, excluding those with velocity separation $|\Delta v| \leq 1500 \text{ km s}^{-1}$ from the quasar’s redshift to avoid contamination due to the proximity to the central QSO.

5.3.4 Galaxies’ physical properties and SED fitting

We aggregated all the photometric measurements from the multiwavelength data described in Section 5.2 to derive the properties of all the galaxies confidently identified within the redshift range $2.5 \lesssim z \lesssim 4.5$ in the MQN01 MUSE mosaic, that is, those around the central quasar as well as those in the control-field sample. Specifically, we aim at constraining the stellar mass (M_*), star formation rate (SFR) and dust extinction (A_v) of each source.

As five of the 21 sources included in the MUSE spectroscopic sample are X-Ray detected AGN, we use the publicly available code CIGALE (Burgarella et al. 2005; Noll et al. 2009; Boquien et al. 2019, version 2022.1, which includes X-Ray implementation by Yang et al. 2020, 2022) to separate the contribution of the AGN from that of the host and derive a reliable estimate of the galaxies’ SFR and stellar mass. We chose high-resolution Bruzual & Charlot (2003a) stellar population models, Chabrier (2003) initial mass function and Calzetti et al. (2000) dust law. We fixed both the stellar and the gas-phase metallicity to 0.20 (solar), 0.08 and 0.04 in separate runs. Similarly, we also performed different runs with an exponentially declining, a delayed exponentially declining or an effectively constant ($\tau = 10 \text{ Gyr}$) star formation history. We notice that such different assumptions on the metallicity and star formation histories lead to results that are consistent within 1σ with each other, and thus consider the output obtained assuming solar metallicity and exponentially declining star formation history across the paper. The AGN and the X-Ray emissions are modelled following Fritz et al. (2006) and Yang et al. (2022), respectively. The full input grid of parameters is adapted from Barrufet et al. (2020) and Riccio et al. (2023), and listed in Table D.1. To estimate realistic uncertainties on the best-fit stellar mass and SFR, we follow Pacifici et al. (2023) and combine the observed uncertainties inherited from photometry, with those related to the choice of the models (e.g., star formation history and stellar physics). We adopted an uncertainty value of 0.2 dex for the stellar mass (see also, Conroy, 2013) and of 0.3 dex for the SFR. Compared to the typical errors returned by the fitting code ($\approx 0.09 \text{ dex}$ for the stellar masses and $\approx 0.04 \text{ dex}$ for the SFR), the dispersion due to systematics is the dominant source of uncertainty. The properties estimated for each galaxy are listed in Table D.2. We also prove that these results are independent on the code used to fit the galaxies’ SED in the Appendix D.5.

5.4 Results

5.4.1 A large overdensity of star-forming galaxies clustered around the quasar

To characterize the large-scale environment around the central QSO, we first investigate the clustering of the galaxies relative to the quasar along the line-of-sight. We reproduce in Figure 5.4 the redshift distribution of the MQN01 high-confidence spectroscopic sample of galaxies identified in the MSUE FoV (black) and of the Lyman break galaxies selected via *UBR* colors (Section 5.3.1) for which a secure measurement of the spectroscopic redshift is available from MUSE (hatched gray histogram). The histogram is normalized by the different areas we surveyed to assemble the samples. An excess of galaxies peaked within the redshift interval $3.15 \lesssim z \lesssim 3.35$ clearly emerges. To roughly estimate the amplitude of the local clustering, we compare these results to expectations for the “MUSE-FIELD” sample and report that the surface density of galaxies identified in the redshift bin closest to the quasar is a factor of ≈ 8 larger than the field. Similarly, the surface density of *UBR*-selected galaxies is higher by a factor of ≈ 5.5 compared to the field. Zooming in around the quasar in terms of the line-of-sight velocity, we observe that the majority of the sources are clustered within an even narrower velocity range (see the inset in Figure 5.4), corresponding to $\pm 1000 \text{ km s}^{-1}$. We thus select it as a linking window to identify the galaxies connected to the central quasar throughout the following analysis.

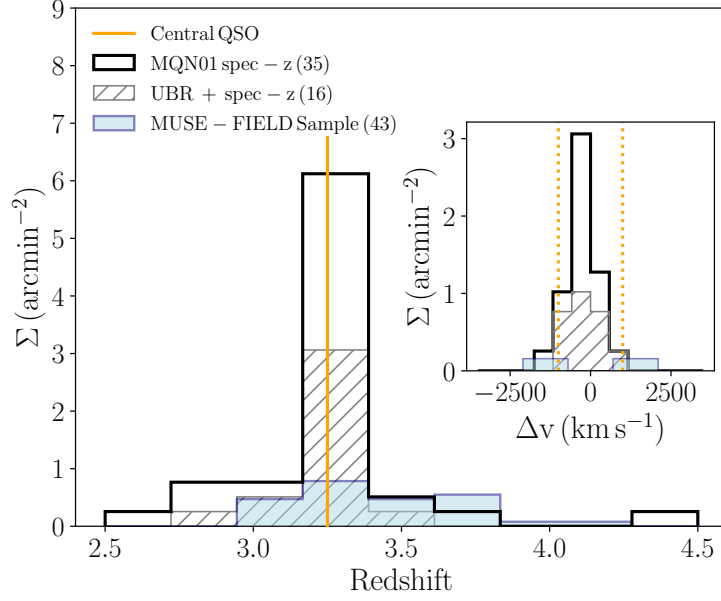


Figure 5.4: Redshift distributions of the MQN01 high-confidence spectroscopic sample (black histogram) and of the *UBR*-selected LBGs with secure spectroscopic redshift from MUSE (hatched gray histogram). These are compared to the “MUSE-FIELD” sample and normalized by the surveyed area. The redshift of the central quasar is marked by an orange vertical line. The inset shows the velocity distribution of the MUSE galaxies and of the LBGs relative to the redshift of the quasar. The line-of-sight velocity window corresponding to $|\Delta v| \leq 1000 \text{ km s}^{-1}$ is highlighted by dotted vertical lines.

5.4.2 Luminosity Function and amplitude of the overdensity

Once established that an excess of star-forming galaxies is found in proximity to the quasar, we now provide a more statistical argument to investigate the amplitude of the overdensity relative to blank fields. We thus compute the cumulative luminosity function (LF), $\phi_c(< M_R)$, defined as the number density of galaxies as a function of their luminosity. In order to account for the difference in redshift compared to the surveys available in the literature, we convert the flux measured in the R-band into absolute magnitude, M_R . To obtain the cumulative luminosity function, we first compute the differential number of galaxies at a given magnitude and per unit volume within nine intervals of $\Delta M_R = 0.175 \text{ mag}$ and associate uncertainties from Poisson statistics. The differential measure is then integrated along the magnitude axis. We apply this procedure for the sample of 21 galaxies spectroscopically identified to lie within $|\Delta v| \leq 1000 \text{ km s}^{-1}$ around the redshift of the quasar. The results are shown in Figure 5.5. We produce the cumulative luminosity function also for the “MUSE-FIELD” sample of galaxies in the range $3.0 \leq z \leq 3.5$ following the same steps described above. For these galaxies, the M_R magnitude is estimated by convolving their spectra with the transmission curve of the FORS2/ R_{special} filter. Similarly, we also produce the cumulative luminosity function for all the galaxies identified in the MUSE mosaic within $2.5 < z < 4.5$, masking the peak of the overdensity in velocity space around the quasar. The luminosity functions of these two independent field samples are consistent with each other, both in shape and normalization. The consistency of the comparison is achieved by having searched for the galaxies in the same MUSE datacube or in those with similar exposure time, $\approx 10 \text{ h}$, and applied the same method and constraints to select only the sources with a precise measure of their spectroscopic redshift. On these bases, we assume the selection function to be the same across the samples and do not apply any completeness corrections to the above luminosity functions. In addition, we prove the normalization of the luminosity function of the field sample to be consistent with results derived in $\mathcal{R}$ -band by Reddy & Steidel (2009) for a sample of color-selected LBGs with a spectroscopic redshift in the range $2.7 \leq z < 3.4$.

To statistically measure the amplitude of the overdensity around the quasar, we take the ratio⁵

⁵We employed the luminosity functions to measure the expected number density of galaxies for blank fields. For the sample of LBGs, we integrated R-band luminosity function of Reddy et al. (2008) and multiplied it by the completeness curve of the FORS2/R-band (see Figure 5.1) in order to resemble for the selection function of our sample. For the sample of spectroscopically-selected galaxies, instead, such value was directly taken from the luminosity function of the “MUSE-FIELD” sample (which is also consistent, within 1σ uncertainty, to the

between the cumulative luminosity function and the one from Reddy & Steidel (2009). As shown in Figure 5.5 (lower panel), we measure an overdensity $\rho/\bar{\rho} = 53 \pm 17$ at $M_R \leq -19.25$ mag. We also note that the overdensity is flat as a function of the magnitude threshold up to $M_R = -19.5$ mag and declines at fainter magnitudes due to the incompleteness of the sample. This suggests that, on average, we only find a shift in the normalization of the luminosity function and do not observe any excess of bright or faint galaxies in the environment around the quasar compared to the field. We also perform the same analysis for the sample of color-selected LBGs and show the derived rest-frame UV luminosity functions in Appendix D.2.

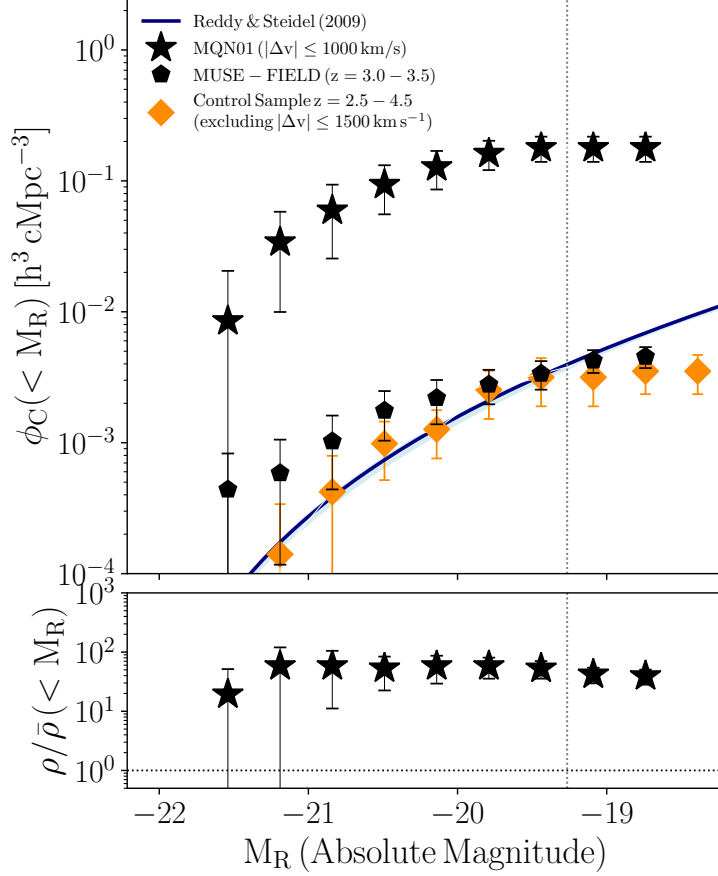


Figure 5.5: Luminosity functions and overdensity of galaxies identified around the central quasar. Upper panel: cumulative R-band luminosity function of the MQN01 galaxies found within $|\Delta v| \leq 1000 \text{ km s}^{-1}$ from the redshift of the quasar (black stars), compared to that of the “MUSE-FIELD” sample (black pentagons) and of the control sample of sources found in the MQN01 MUSE FoV within $2.5 < z < 4.5$, masking $|\Delta v| \leq 1500 \text{ km s}^{-1}$ around the quasar (orange diamonds). The vertical dotted line show the magnitude at which the luminosity function of the spectroscopic samples flattens relative to the completeness-corrected one from Reddy & Steidel (2009). Lower panel: overdensity as a function of the R-band magnitude.

5.4.3 Clustering analysis

Comparing the correlation length of the galaxies clustered around different quasars at similar redshifts encodes key information about the structures in which they are hosted. We follow the formalism from Trainor & Steidel (2012), TS12 hereafter, and derive the galaxy-galaxy auto-correlation function, w_{GG} , and the quasar-galaxies cross-correlation function, w_{QG} . In three dimensions, the correlation functions are typically modelled by a power law, $\xi = (r/r_0)^{-\gamma}$, where γ is the slope and r_0 the correlation length. Once integrated in the redshift direction, the projected correlation functions take a functional form that still depends on the parameters γ and r_0 (see, eq. 10 in TS12). In each radial bin, we compute $w_{GG,QG}(R) = (N_{\text{obs}}/N_{\text{exp}}) - 1$, where N_{obs} is the observed

estimates from the luminosity function of Reddy & Steidel (2009) and that of the control-field sample we built masking the overdensity around the quasar). As a result, we measured a co-moving density in blank fields to be $\bar{\rho}_{\text{LBGs}} = (2.4 \pm 0.2) \times 10^{-3} \text{ cMpc}^{-3}$ and $\bar{\rho}_{\text{spec-z}} = (2.5 \pm 0.8) \times 10^{-3} \text{ cMpc}^{-3}$ for the color-selected LBGs and the MUSE spectroscopic sample, respectively.

number of galaxies identified within a radial bin and N_{exp} is the expected one, derived from the luminosity functions of the field sample from Section 5.4.2.

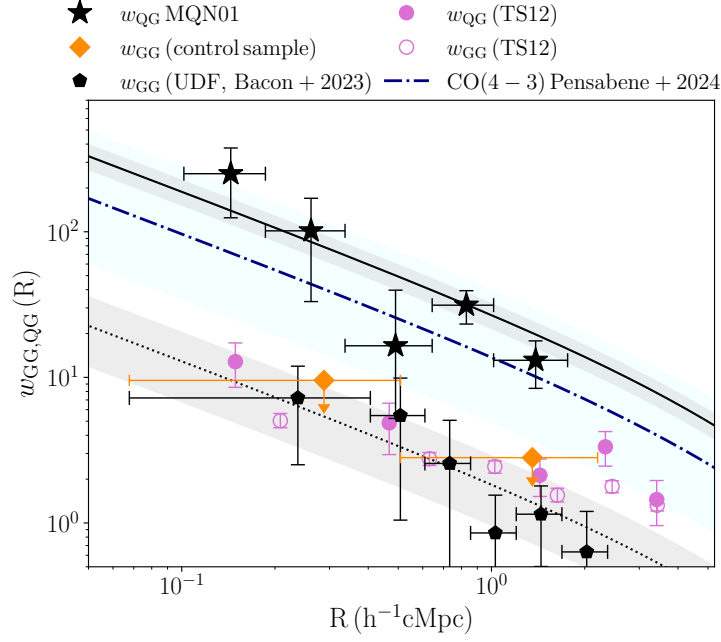


Figure 5.6: Projected correlation functions assuming fixed $\gamma = 1.8$ and linking window $|\Delta v| \leq 1000 \text{ km s}^{-1}$. The quasar-galaxy cross-correlation function for the MUSE spectroscopic sample (black stars) is similar to that of CO(4–3) line emitters identified by Pensabene et al. (2024b) in the same volume (blue line and shaded area), but significantly above that of Trainor & Steidel (2012) (purple, filled points). The galaxy-galaxy auto-correlation function is shown for the control-field sample of $z = 2.5 - 4.5$ galaxies identified in the UDF (black pentagons, Bacon et al., 2023), together with upper limits for the galaxies in our field within the same redshift range, but masking $|\Delta v| \leq 1500 \text{ km s}^{-1}$ around the redshift of the quasar (orange diamonds). Both the estimates are consistent with Trainor & Steidel (2012) (purple, empty points). The best-fit results are shown as dotted and solid lines for the auto and cross-correlation functions, respectively. All the uncertainties are derived from the 16th and 84th percentiles from bootstrap distributions.

Correlation function	$r_0 [h^{-1} \text{cMpc}]$ $ \Delta v \leq 1000 \text{ km s}^{-1}$	$r_0 [h^{-1} \text{cMpc}]$ $ \Delta v \leq 1500 \text{ km s}^{-1}$	γ
$w_{\text{QG}}^{\text{MQN01}}$	$23.06^{+3.49}_{-3.43}$	$22.39^{+3.72}_{-3.36}$	1.5
	$15.71^{+1.79}_{-1.83}$	$14.08^{+1.76}_{-2.01}$	1.8
$w_{\text{GG}}^{\text{UDF}}$	$4.25^{+1.56}_{-1.44}$	$2.65^{+1.28}_{-1.10}$	1.5
	$3.54^{+1.05}_{-1.08}$	$2.36^{+0.92}_{-0.83}$	1.8

Table 5.3: Quasar-galaxies cross-correlation and galaxy-galaxy auto-correlation lengths measured for the MUSE spectroscopic sample and for the field galaxies identified in the UDF-MOSAIC (Bacon et al., 2023), respectively.

To compute the galaxy-galaxy auto-correlation function without being affected by the presence of the quasar in our field, we took the $2.5 < z < 4.5$ galaxies with spectroscopic redshift measured from rest-frame UV absorption lines, that is, similar to our selection in Section 5.3.2, from the UDF catalog (Bacon et al., 2023). In addition, we proved this result to be consistent with upper limits for the auto-correlation of the galaxies identified in our field at $2.5 < z < 4.5$, masking $|\Delta v| \leq 1500 \text{ km s}^{-1}$ around the redshift of the quasar. The results are shown in Figure 5.6, with the horizontal errorbars marking the width of each radial bin and the vertical ones derived from the 16th and 84th percentile of the distributions obtained by bootstrapping over the sample of galaxies. We derive the correlation length by fitting the equation 10 from TS12 to the data. The

slope, γ , is degenerate with the correlation length and therefore fixed to different values to facilitate consistent comparisons with the literature. The best-fit results are shown in Figure 5.6 and listed in Table 5.3. The uncertainties are obtained from the 16th and 84th percentiles of the bootstrap distributions. The cross-correlation length estimated for spectroscopically-selected star forming galaxies in MQN01 is substantially higher than the galaxy-galaxy correlation lengths obtained using the same selection, and higher than the values found around other quasars in the literature at similar times, e.g. compared to TS12. In Section 5.5.1, we discuss the implications of these results.

5.4.4 Galaxies' star formation activity compared to the field

How the large-scale environment affects galaxies growth is a key question within the framework of galaxy evolution in a hierarchical universe, especially in overdense regions throughout cosmic time. A useful approach to gauge galaxies' ability to accrete cold gas and sustain star formation is by looking at their positioning in the stellar mass versus star formation rate (or specific star formation rate, sSFR) plane relative to the "main sequence" (see e.g., Brinchmann et al., 2004; Daddi et al., 2007; Elbaz et al., 2007; Noeske et al., 2007; Salim et al., 2007).

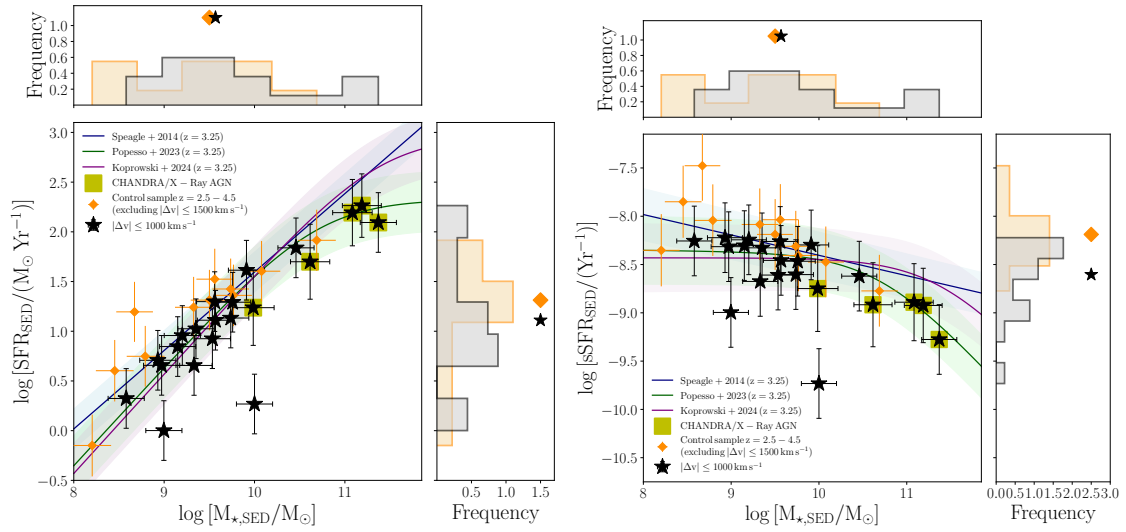


Figure 5.7: Star forming main sequence, in terms of SFR (left panel) and sSFR (right panel), of the MQN01 galaxies clustered around the quasar (black stars), compared to the control-field sample (orange diamonds). The data are compared to the main sequence at $z \approx 3.25$ derived from Speagle et al. (2014) (blue), Popesso et al. (2023) (green) and Koprowski et al. (2024) (purple). The X-Ray detected AGN are marked by yellow squares. The histograms on the top and right panels show the distribution of the SFR or sSFR and galaxies' stellar mass, respectively, of the MQN01 galaxies (black) and of the control-field sample, together with their median values on top.

In Figure 5.7 we directly compare the star formation rate (left panel) and the specific star formation rate (right panel) as a function of galaxies' stellar mass derived for the galaxies clustered around the quasar with the star forming main sequence at $z \approx 3.25$. We parametrized the main sequence with the functional forms from Speagle et al. (2014) (blue), Popesso et al. (2023) (green) and Koprowski et al. (2024) (purple) at the quasar redshift, allowing for a 0.3 dex dispersion. The method we used to identify the galaxies (Section 5.3.2) selects rest-frame UV bright sources, that is, actively star forming and not too heavily obscured by dust. In other words, our selection tends to exclude passive or extremely red galaxies. As a result, at a given stellar mass, we do not observe any significant offset from the main sequence calibrated on blank fields at $z \approx 3$ for the clustered galaxies, nor for the AGNs detected by *Chandra* (yellow squares in Figure 5.7). The evidence that galaxies form stars at main sequence rates seems to be a recurrent feature in some of the known $z \approx 2 - 3$ protoclusters (see e.g., Koyama et al., 2013a,b; Cooke et al., 2014; Shimakawa et al., 2018a,b; Shi et al., 2019a,b, 2020; Toshikawa et al., 2024).

Due to the differences in the selection function and in the SFR estimates between the sample of MQN01 galaxies and the ones used to calibrate the parametrizations of the main sequence, it is more informative to directly compare the properties of the galaxies in the overdensity to those included in a field sample with a similar selection function. To this end, we consider the control-

field sample of 11 galaxies identified in the MUSE FoV (see, Section 5.3.3). All these galaxies are scattered along the expected main sequence. In addition, we compare the distributions of stellar mass (top panels), SFR and sSFR (right panels) shown in Figure 5.7 by performing a KS test (significant at 2σ level for p -values ≤ 0.05). While the distributions of sSFR are unlikely drawn from the same parent one, p -value ≈ 0.01 , there is no statistical difference in those of stellar mass and SFR (p -value ≈ 0.35 and p -value ≈ 0.45 , respectively). Nonetheless, we notice that we find only one galaxy in the field sample located in the $\log(M_\star/M_\odot) \gtrsim 10.5$ region of the main sequence, and it is possibly the excess of galaxies in the high-mass tail of the distribution that is driving the difference in the sSFR. The direct comparison between the distributions of the two samples does not take into account the different encompassing co-moving volume ($\approx 2 \times 10^4 \text{ cMpc}^3$ for the field sample, which is more than 50 times larger compared to that of the sample of galaxies clustered around the quasar). Normalizing the number of sources by the co-moving volume, we find that the number density of massive ($M_\star \gtrsim 10^{10.5} M_\odot$) star forming galaxies is 217^{+200}_{-183} times higher around the quasar than in a blank field. Even accounting for the overdensity of ≈ 53 measured around the quasar in Section 5.4.2, the presence of massive galaxies exceeds by a factor of 4^{+2}_{-3} the expectations for the field (see, Section 5.8 for details). On the other hand, the incidence of $M_\star \lesssim 10^{10.5} M_\odot$ galaxies around the quasar is 92^{+80}_{-43} times larger than in the field sample, which corresponds to a factor of $1.7^{+1.6}_{-0.8}$ once accounting for the overdensity. This results also corresponds to a factor of 4^{+2}_{-3} more galaxies forming stars at a rate $\log[\text{SFR}/(M_\odot \text{Yr}^{-1})] \geq 1.7$ around the quasar compared to blank fields. In terms of sSFR, the massive star-forming galaxies discussed above reside at $\log(\text{sSFR}/\text{Yr}^{-1}) < -8.5$, leading to a large overdensity (11^{+7}_{-9}) in this sSFR range as well. This scenario is thus consistent with a picture where a consistent fraction of galaxies in this overdense environment is experiencing a phase of accelerated mass assembly and is rapidly consuming its gas content.

5.4.5 Galaxies' stellar mass function

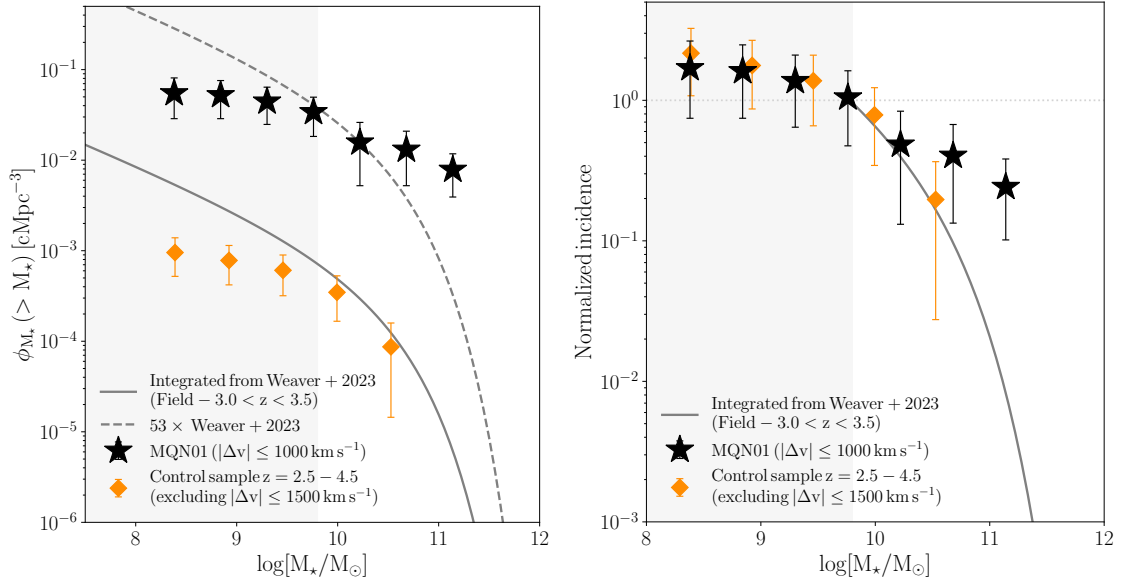


Figure 5.8: Excess of massive $\log(M_\star/M_\odot) \gtrsim 10.5$ galaxies around the MQN01 quasar. Left panel: cumulative stellar mass function of the MQN01 spectroscopic sample in proximity to the quasar (black stars) compared to the control-field sample (orange diamonds) and the expectations for blank fields from the COSMOS survey ($3.0 < z < 3.5$, Weaver et al. 2023, gray). Right panel: same as the left panel, but normalized to the value of the cumulative mass function at $M_\star = 10^{9.8} M_\odot$, corresponding to the 90% completeness limit of the MQN01 spectroscopic sample.

Motivated by the results obtained in Section 5.4.4, we now compute the galaxies' number density as a function of their stellar mass by integrating the differential stellar mass function and propagating the Poisson errors arising in each interval of masses. To also take into account the errors on the measure of the stellar mass itself, we bootstrap varying the fiducial values within the uncertainties derived from the galaxies' SED fitting for 1000 times. The standard deviation from the 16th and 84th percentiles of the resulting bootstrap distribution is then squared and

summed with errors propagated from the integral above. In the right-panel in Figure 5.8 we show the resulting cumulative stellar mass function, while in the left-panel we normalize it at the stellar mass completeness limit⁶ to better capture any difference in shape. The estimate for the galaxies around the quasar (black stars) is compared to the expectations for different field samples. We thus take as a reference the stellar mass function computed by Weaver et al. (2023) in the COSMOS2020 field within the redshift range $3.0 < z < 3.5$ and rescale it by the MQN01 overdensity $\rho/\bar{\rho} \approx 53$. This is also consistent with what we obtain for the galaxies included in the control-field sample and identified within the redshift range $2.5 < z < 4.5$, masking the peak of the overdensity in velocity space (orange diamonds). While both the stellar mass functions flatten and deviate from Weaver et al. (2023) field expectations at $\log(M_*/M_\odot) \lesssim 9.5$, likely due to the incompleteness of the samples, that of star-forming galaxies clustered with the quasar is also flatter at masses $\log(M_*/M_\odot) \gtrsim 10.5$. In other words, this change in the shape of the stellar mass function provides evidence that the surroundings of the quasar are richer in massive galaxies compared to the field, with 4 ± 2 and 24 ± 11 times more sources in the two bins at $\log(M_*/M_\odot) \gtrsim 10.5$, respectively, once accounted for the overdensity.

5.4.6 Connecting the local overdense environment to the galaxy properties

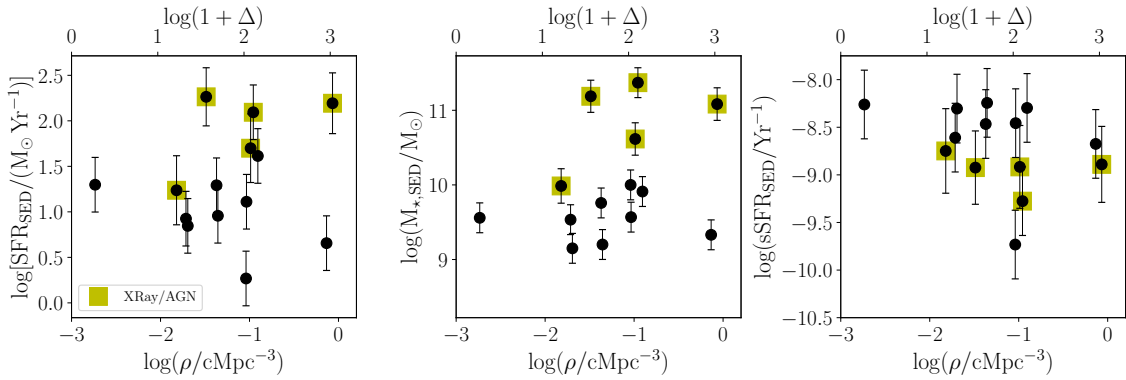


Figure 5.9: MQN01 galaxies’ star formation rate (1st panel), stellar mass (2nd panel) and specific star formation rate (3rd panel) as a function of the co-moving local galaxy density (bottom x-axis) and overdensity (top x-axis) with respect to the field. X-Ray emitting AGN detected in *Chandra* observations are highlighted as yellow squares.

We now turn to investigate how the properties of the galaxies depends on the local overdensity. We thus employ a scale-independent Voronoi Tessellation (one element per cell) to parametrize the local density of galaxies clustered around the quasar. According to this criterion, the size of each cell is, by construction, inversely proportional to the local galaxy surface density (i.e., small cells corresponds to crowded environments). Open cells at the edge of the FoV (7/21) are excluded from the analysis due to the loss of information about the objects located outside the MUSE mosaic. The overdensity is then defined as $\Delta = \rho/\bar{\rho} - 1$ (so that $\log(1 + \Delta) = 0$ corresponds to the case in which the galaxy density is the same as the field, i.e., $\rho = \bar{\rho}$). We show in Figure 5.9 how the SFR, stellar mass and sSFR correlate with the local galaxy density. To address this analysis on a more statistical basis, we also performed a Spearman-rank correlation test and report in Table 5.4 the strength of the correlation, ρ_{spearman} , and the probability that no monotonic relation exists between the two variables, p -value. We require $p\text{-value} \leq 0.05$ for a correlation to be significant at least at 2σ confidence level. As a result, we observe that the galaxies’ stellar mass and specific star formation rate are positively and negatively, respectively, correlated ($|\rho_{\text{spearman}}| \gtrsim 0.45$ and $p\text{-value} \approx 0.05$) with the local density at 2σ significance. On the other hand, we do not observe any significant environmental effect on the star formation rate itself ($p\text{-value} \approx 0.29$).

⁶To estimate an order-of-magnitude completeness for the MUSE spectroscopic samples, we adopt the empirical method from Pozzetti et al. (2010); Davidzon et al. (2013); Ilbert et al. (2013); Moustakas et al. (2013); Laigle et al. (2016), which has already been applied by Toshikawa et al. (2024) on U-dropout galaxies at $z \approx 3$. With a K-band magnitude limit of 25.5 mag (see Section 5.2.3), we estimate that our sample is 90% complete at $\log(M_*/M_\odot) \approx 9.8$.

Variable	ρ_{Spearman}	p -value
$\log[\text{SFR}/(\text{M}_{\odot} \text{Yr}^{-1})]$	+0.27	0.29
$\log(\text{M}_{\star}/\text{M}_{\odot})$	+0.46	0.05
$\log(\text{sSFR}/\text{Yr}^{-1})$	-0.50	0.04

Table 5.4: Results of the Spearman-rank test on the correlations between the properties of the MQN01 galaxies and the local overdensity.

5.4.7 Probing the structure on larger scales

The large-scale structures in which galaxies are embedded are expected to extend well beyond the $\approx 4 \times 4 \text{ cMpc}^2$ we are able to trace with the star-forming galaxies identified within the FoV of the MUSE mosaic. We thus model the 2D spatial distribution of the *UBR*-selected LBGs across the $\approx 24 \times 24 \text{ cMpc}^2$ FoV of FORS2 and compare it to that of the spectroscopically selected sources lying within $|\Delta v| \leq 1000 \text{ km s}^{-1}$ from the quasar in the $4 \times 4 \text{ cMpc}^2$ FoV of the MUSE mosaic. This is done by creating an evenly spaced grid of $7'' \times 7''$ and $5'' \times 5''$ cells covering the entire FORS2 and MUSE FoV, respectively. The number of galaxies identified within each cell is then normalized by the volume of the cell and compared to the co-moving density expected for the field sample (see Section 5.4.2). The grid is then oversampled (by a factor of 3 and 2, respectively) and smoothed with a Gaussian Kernel ($\sigma = 6, 5$, respectively, corresponding to $\approx 0.5''$). Results are shown in Figure 5.10.

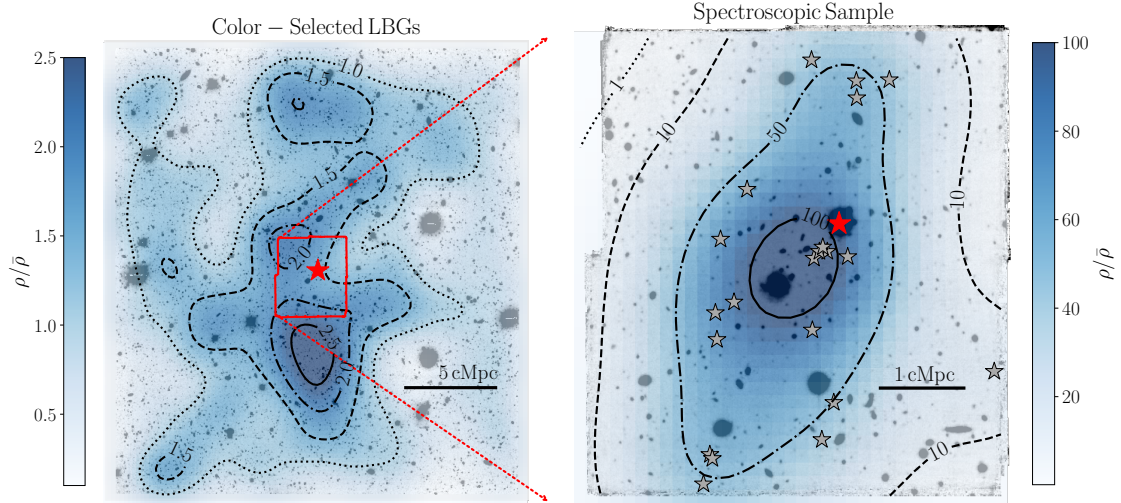


Figure 5.10: 2D overdensity maps of the LBGs and of the spectroscopically-selected galaxies found to be clustered with the quasar. Left panel: Overdensity map of the LBGs identified within the $24 \times 24 \text{ cMpc}^2$ FORS2 FoV. The amplitude of the overdensity is shown in shades of blue scaling with the colormap. Contours are shown for $\rho/\bar{\rho} = 2.5, 2.0, 1.5, 1.0$ as solid, dash-dotted, dashed and dotted lines respectively. The central quasar is marked as a red star, and the MUSE FoV is also displayed on top of the FORS2 R-band image. Right panel: zoom in the FoV of the MUSE mosaic, with the white-light image overlaid by the overdensity map of the spectroscopic sample of galaxies identified within $|\Delta v| \leq 1000 \text{ km s}^{-1}$ (gray stars) around the quasar (red star). Solid, dash-dotted, dashed and dotted contours correspond to $\rho/\bar{\rho} = 100, 50, 10, 1$. Note the different spatial scaling and that of the colormap in the two panels. North is up.

The LBGs appear to be overdense ($\rho/\bar{\rho} \gtrsim 1$) across the entire $\approx 24 \times 24 \text{ cMpc}^2$ FoV, even at large distances from the central quasar. The overdense regions $\rho/\bar{\rho} \gtrsim 1.5$ are concentrated closer to the center of the field and show a preferential vertical orientation, which corresponds to the North-South direction on the sky, and is also aligned with both the overdensity of spectroscopically-selected galaxies that we observe in MUSE (right panel) and the extended $\text{Ly}\alpha$ emitting gas observed by Borisova et al. (2016) and Cantalupo, et al. (in prep.). This evidence supports a picture in which the large-scale structure traced by the galaxies clustered with the quasar extends well beyond the FoV of the MUSE mosaic, for tens of cMpc. We reserve a more statistical treatment of the LBGs alignment as a function of the local overdensity to Appendix D.3. On the other hand, when looking at this structure on scales of $24 \times 24 \text{ cMpc}^2$, we observe that the most overdense regions, $\rho/\bar{\rho} \gtrsim 2$, are not found within the FoV of the MUSE mosaic, but are mostly concentrated

in the southern part of the field. A number of possible scenarios, or a combination of them, should be tested to explain this evidence: i) they are part of a foreground or background structure; ii) the structure surrounding the quasar itself extends in the southern direction; iii) despite being close to the peak of the overdensity in the MUSE FoV, the quasar is not the center of the structure, which is instead in the South; iv) they are tracing two merging or interacting structures. Spectroscopic confirmation of the LBGs would be essential to discriminate between these scenarios.

5.5 Discussion

In this section, we put the results presented above into a broader context by comparing them to previous results available in the literature. In particular, in Section 5.5.1 we discuss the overdensity and clustering measurements in MQN01 in light of other studies targeting quasars and overdense regions at $z > 2$. In Section 5.5.2, we focus on the galaxies' sSFR and stellar mass function in MQN01 comparing them to other works from the literature, both in the field and in overdense regions, in order to reveal the possible role of the local environment on galaxy formation and evolution.

5.5.1 MQN01 overdensity and clustering compared to other fields

The number of galaxies spectroscopically confirmed to be within $|\Delta v| \leq 1000 \text{ km s}^{-1}$ and in an area of $\approx 4 \times 4 \text{ cMpc}^2$ ($\approx 2 \times 2 \text{ arcmin}^2$) around the MQN01 quasar results in an overdensity of $\rho/\bar{\rho} = 53 \pm 17$ which corresponds to a surface density of $\approx 1.3 \text{ cMpc}^{-2}$. At a redshift similar to that of MQN01, several galaxy overdensities and protoclusters have been studied in the literature, such as SSA22 at $z \approx 3.08$ (Steidel et al., 1998, 2003), Hyperion at $z \approx 2.47$ (Cucciati et al., 2018) and Spiderweb at $z \approx 2.15$ (Roettgering et al., 1994; Pentericci et al., 1997). In particular, Topping et al. (2016) found 40 LBGs in SSA22 with spectroscopic redshift in the range $z = 3.05 - 3.12$ across $17.25 \times 17.25 \text{ cMpc}^2$, which result in a surface density of $\approx 0.13 \text{ cMpc}^{-2}$. Huang et al. (2022) discovered 53 Ly α emitters in a region extended over $\approx 30 \times 20 \text{ cMpc}^2$ in Hyperion, with a surface density of $\approx 0.08 \text{ cMpc}^{-1}$. In Spiderweb, Pérez-Martínez et al. (2023) found 39 spectroscopically confirmed H α emitters across $13.7 \times 19.3 \text{ cMpc}^2$, corresponding to a surface density of $\approx 0.15 \text{ cMpc}^{-1}$. Notably, the surface density measured in the MQN01 field is $\approx 10 - 20$ times higher with respect to these structures. A direct comparison, however, is not straightforward given the different galaxy populations selected in these works, the different levels of completeness reached and the surveyed volumes. A detail comparison would require, for instance, a MUSE (or KCWI) survey in these proto-cluster fields, similar to the one performed around MQN01.

Besides the large overdensity of star-forming galaxies in the field, a significant excess of galaxies in MQN01 obtained with different tracers has been reported. In particular, Pensabene et al. (2024b) found ≈ 18 times more CO(4-3) line emitters and higher 1.2mm source number counts compared to blank fields, with the latter result that is similar to the findings in SSA22. Exploring *Chandra* X-Ray data, an overdensity of $\rho/\bar{\rho} \approx 1000$ of AGNs has recently been unveiled by Travascio et al., (submitted) and found to exceed that observed in SSA22 and Spiderweb protoclusters. These studies support the peculiarity of the MQN01 field in terms of galaxy overdensity as suggested by our spectroscopic study in the rest-frame UV.

Stronger evidences about the rarity of the MQN01 structure are obtained by the cross-correlation function presented in Section 5.4.3. In particular, we obtained a quasar-galaxy cross-correlation length of $r_0 = 15.71^{+1.79}_{-1.83} h^{-1} \text{ cMpc}$, which is significantly higher than the galaxy-galaxy correlation length of $r_0 = 3.54^{+1.05}_{-1.08} h^{-1} \text{ cMpc}$ for fixed $\gamma = 1.8$. Despite the limited statistics, the cross-correlation lengths obtained using CO(4-3) line emitters around the MQN01 by Pensabene et al. (2024b), $r_0 = 10.8^{+8.8}_{-4.7} h^{-1} \text{ cMpc}$, are also consistent with these values. These values clearly indicate that galaxies in the MQN01 structure are more clustered around the quasar compared to both the average fields and other quasars at similar redshifts (see, e.g., Trainor & Steidel, 2012; Garcia-Vergara, 2021). For consistency and to control for the effects of different selection functions, we compare our results with works that studied the clustering of similar samples of galaxies, within similar volumes and at similar redshifts. TS12 measured an auto-correlation length $r_0 = (5.8 \pm 1.3) h^{-1} \text{ cMpc}$ and a cross-correlation length $r_0 = (7.1 \pm 1.3) h^{-1} \text{ cMpc}$ (once rescaled to the cosmology adopted in this paper, which corresponds to a correction factor of 1.03), assuming $\gamma = 1.5$, for a spectroscopic sample of galaxies identified within $\pm 1500 \text{ km s}^{-1}$ from 15 quasars at $z \approx 2.7$. The auto-correlation functions of galaxies selected as those found in MQN01, are consistent with that of TS12 (see, Table 5.3). On the other hand, the cross-correlation lengths of

the galaxies identified within $|\Delta v| \leq 1500 \text{ km s}^{-1}$ around the quasar is ≈ 3.2 times larger compared to TS12. We also measured the quasar-galaxy cross-correlation function in the MAGG survey (Lofthouse et al., 2020b), which 28 fields centered on quasars and observed with MUSE (see Appendix D.4 for more details). For consistency with the sample of galaxies used in this work, we considered only those that have been continuum-selected, as described in Lofthouse et al. (2020b), brighter than $m_R \leq 26.5 \text{ mag}$, with a secure measure of their spectroscopic redshifts and found within $|\Delta v| \leq 1500 \text{ km s}^{-1}$ from the redshift of each MAGG quasar (median $z \approx 3.7$). We derived a cross-correlation length of $r_0 = (3.8 \pm 0.8) h^{-1} \text{ cMpc}$ for fixed $\gamma = 1.5$. Interestingly, The one measured in the MQN01 field is ≈ 6 times larger. Altogether, these pieces of evidence suggest that the MQN01 quasar could possibly be associated with a much larger halo than the typical quasar host at these redshifts, which is estimated to be around $\log(M_h/M_\odot) \approx 12 - 12.5$ (TS12, see also, Fossati et al., 2021; Arrigoni Battaia et al., 2022; de Beer et al., 2023).

5.5.2 The role of the local environment on galaxy formation and evolution

One of the key results presented in Section 5.4.5 is the fact that the galaxy overdensity in the MQN01 galaxies depends on the galaxy stellar mass. In particular, Figure 5.8 shows that there is an elevated fraction of massive galaxies compared to the expectations from average regions of the universe, suggesting a possible link between the environment and galaxy growth. Indeed, despite the limited statistics, we find a hint that the local galaxy overdensity is positively correlated with the galaxies’ stellar mass (and negatively correlated with their sSFR) as shown in Figure 5.9. Our results reinforce previous findings in the literature (see e.g., Darvish et al., 2016; Chartab et al., 2020; Lemaux et al., 2022; Taamoli et al., 2024), which did not have the large dynamic range in overdensity as provided by our sample.

For instance, by using a large sample of protocluster candidates, Toshikawa et al. (2024) found that U-dropout galaxies living in overdense regions at $z \approx 3$ show a cumulative stellar mass function that is flatter at the massive end compared to the field. Similarly, Shimakawa et al. (2018a,b) found a statistically significant excess of massive galaxies in two $z \approx 2$ protoclusters, broadly consistent with other studies of continuum-selected galaxies and $\text{H}\alpha$ emitters in overdense regions at similar redshifts (see e.g., Steidel et al., 2005; Koyama et al., 2013a,b; Cooke et al., 2014). However, they also observe that such massive galaxies are mostly located in the densest regions of the protocluster, while sources in less dense areas have a stellar mass function distribution that is more similar to that of the field (see also, Forrest et al., 2024, at $3.20 < z < 3.45$).

The presence of an excess of massive galaxies in overdense regions, such as protoclusters, can be supported by different scenarios: i) the high density environment is able to trigger mergers as well as boost the rate of interactions between galaxies (e.g., Alonso et al., 2012; Cooke et al., 2014; D’Amato et al., 2020); ii) the access to a large gas reservoir that is efficiently accreted from the cosmic web onto the galaxies (see e.g., Shimakawa et al., 2018a). Indeed, Umehata et al. (2019) found that AGNs and sub-millimeter galaxies in the SSA22 protocluster are located within filaments of gas, which can potentially be accreted by these sources fuelling their star formation and driving their growth. On the other hand, there is also evidence that overdensities play a minor role in shaping the scaling relations of galaxies at $z > 2$ (see e.g., McGee et al., 2009; Fossati et al., 2017), which requires further investigation to be conducted in similar regions of the universe. Altogether, these possible scenarios suggest that these galaxies grew rapidly, either assembling their masses earlier or more efficiently than those in the field.

5.6 Summary and conclusions

We detected and studied a population of star-forming galaxies clustered within $\approx 4 \times 4 \text{ cMpc}^2$ around the $z \approx 3.25$ quasar CTS G18.01, known as MQN01 field. The quasar is surrounded by a giant $\text{Ly}\alpha$ emitting nebula discovered by Borisova et al. (2016) and Cantalupo, et al. (in prep.), and the field hosts a large overdensity of dusty star-forming galaxies (Pensabene et al., 2024b) and AGNs (Travascio, et al. submitted). By leveraging a rich set of multiwavelength data, including MUSE, HST, FORS2, HAWK-I, JWST and *Chandra*, we derived the star formation rates and the stellar masses of these galaxies and explored their connection to the local and large-scale overdense environment. The key results are summarized below:

- We discovered an overdensity of star forming galaxies of $\rho/\bar{\rho} = 53 \pm 17$ clustered within $4 \times 4 \text{ cMpc}^2$ and $|\Delta v| \leq 1000 \text{ km s}^{-1}$ around the central quasar. Measuring the amplitude of

the clustering as a function of the projected distance from the quasar, we derived a cross-correlation length that is ≈ 3.2 times larger than the average value found by [Trainor & Steidel \(2012\)](#) for $z \approx 2.7$ quasars. Galaxies in this structure are thus either more clustered or living in a more massive structure than those observed around average quasars at similar times.

- We found that the galaxies in the overdensity are forming stars at a rate consistent with the expected main sequence at $z \approx 3.25$. In addition, the ability of these galaxies to form stars appears to be more sensitive to the gas and baryon content of the galaxies, probed by their stellar masses, rather than to the local overdensities and the large-scale environment in which they reside.
- At the high mass end, the cumulative stellar mass function is flatter compared to the field. This suggests the presence of an excess of massive galaxies in the structure, with ≈ 4 times higher number density than in the field at $\log(M_*/M_\odot) \geq 10.5$. In addition, all the massive galaxies are also X-Ray AGNs. Such an excess of massive star forming galaxies may be due to a higher merger rate or efficient accretion that allow the galaxies to assemble their mass at earlier times or more efficiently than in the field and provide them of enough fuel to continue forming stars at main sequence rates.
- The distribution of color-selected Lyman break galaxies on scales $\times 36$ larger than the original FoV of the MUSE mosaic suggests that the overdensity extends along the North-South direction and is thus align with the $\text{Ly}\alpha$ emitting gas detected by [Borisova et al. \(2016\)](#) around the quasar and to the ≈ 4 cMpc filamentary structure observed in emission by Cantalupo et al., (in prep.) This may be a hint that the gaseous structure in which galaxies are embedded extends for tens of cMpc.

Overdense structures, such as MQN01, are ideal places to explore the possible connection between the environment and the properties of galaxies at high redshift. The advancement in the study of overdensities at $z > 2$ is finally making it possible to explore a wide range of environments, even at the cosmic noon. However, given their rarity, it is still challenging to build a large sample of such structures, and thus to increase the statistical significance of any possible connection between galaxies and environment. As such, even if based on a single field, our study provides important clues for future investigations to piece together the coevolution of galaxies and their large-scale environments from the very early epochs to the late times.

Conclusions and future outlook

In this thesis, we have explored the link between gas, galaxies and the large-scale environment at epochs close to cosmic noon, when the star formation rate density in the universe it is close to its peak. To do so, we have combined several methods to detect different galaxies populations, including spectroscopic surveys, color selections and submillimeter observations. At the same time, we have also detected the cosmic gas across a broad dynamic range of densities, composition and ionization state by combining observations in absorption and in emission.

In particular, in the MAGG survey we have found that at $z \approx 3 - 4.5$ multiple gas phases co-exist in the gas around low-mass galaxies, as it was previously observed only for more massive LBGs. The gas at different densities is also found to trace different structures, from the CGM, to filaments connecting galaxies, to more diffuse and faint component at larger distances. We have also found that the large-scale galaxy environment is able to affect the distribution and the properties of the gas, but the evolution of such effect across cosmic time differs depending on the gas phase. In the MUDF survey, we have discovered a surprisingly high association rate between the metal-enriched gas detected in absorption and a population of dusty, massive and star forming galaxies detected by ALMA, and in part missed by previous optical/NIR observations. This suggests that targeting also this population of galaxies in similar absorption-line studies is key to paint a complete picture of the galaxies-gas connection. In addition, these galaxies are typically massive and found close to the center of the potential well of the overdense structure in which they reside, suggesting that they are also crucial to fully characterize the effect of the large-scale environment on both the galaxies and the gas. Finally, we investigated the environmental effects on galaxy evolution in an extremely overdense structure identified around a $z \approx 3$ quasar. Within this structure, we identified overdensities of multiple populations, including star forming galaxies, LBGs, submillimeter galaxies, and AGNs. A significant excess of massive galaxies compared to the field suggests that the evolution is accelerated in this systems and galaxies assemble their mass at earlier times than in random regions of the universe.

These promising results contribute to shed light on the complex interplay between gas and galaxies. However, it is still challenging to draw a unified picture given the lack of a complete census of the galaxy populations, the field-to-field variations, and the limits in the detection of the cosmic gas. In particular, in the MAGG survey we have discovered that only a fraction of the strong Mg II absorbers are connected to galaxies. This is surprising, as they are tracers of overdense cool gas, which is an ideal site for galaxies to form. A possible caveat is that galaxies associated to these absorbers are located at such a small projected separation from the sightline to be hidden by the Point Spread Function of the quasar. Overcoming this limitation would require, for instance, observing these quasars at higher resolution. A second caveat is that these strong Mg II systems can be associated to a population of galaxies that does not emit bright Ly α radiation and is thus missed by our selection. This could be the case of dust-obscured galaxies. Therefore, we are currently obtaining follow-up 32.8 hours of ALMA observations (Program ID: 2024.1.00132.S, PI: A. Pensabene) to specifically target this galaxy population at the redshift of the strong Mg II absorbers. Similar surveys would be instrumental to improve the statistics and unveil the role of dusty galaxies in the connection with the multiphase cosmic gas. Additionally, we have found in the MUDF survey, that these galaxies are typically massive, highly star forming and located in overdensities, specifically close to the center of the gravitational potential well. This makes even more necessary to target these galaxies to have a comprehensive view of how both the properties

of the gas and the galaxy evolution depend on the large-scale environment.

On a similar fashion, also passive galaxies remains typically undetected in optical surveys. They are expected to play a leading role especially in extreme overdense regions, such as protoclusters, where environmental effects appear and become efficient in quenching star formation in galaxies. this is also the case of the structure identified in the MQN01 field and presented in Chapter 5, where we have performed a census of star forming and submillimeter galaxies, and AGNs, but we still miss the passive ones. to this end, we will take advantage of the photometric data taken by VLT/HAWK-I and sampling the 4000 Å Balmer Break at the redshift of the structure to search for quiescent galaxies. However, additional spectroscopic data would optimize this selection: with JWST it would be possible to sample the rest-frame wavelength range 400 – 700 nm, which covers the Balmer lines. Such spectra are already available for some galaxies in the field (Cantalupo et al., 2021; Wang et al., 2024b) from JWST/NIRSpec MSA observations, but grism spectroscopy would be necessary. In addition, we have also found that this structure is more extended than the field-of-view of our MUSE observations. Although we trace it by looking at the overdensity of color-selected LBGs, spectroscopic follow-up observations would be necessary to confirm their redshifts. We have thus obtained additional 84 hours of MUSE observations (Program ID: 112.25N3, PI: S. Cantalupo) to confirm the selection of these galaxies up to ≈ 14 cMpc and possibly even detect extended Ly α emission $\gtrsim 4$ cMpc away from the quasar. Finally, these findings are limited to an individual and rare field and cannot be generalised to entire population of protoclusters and overdensities. Furthermore, different studies tend to select and classify galaxies using different methods, which make it challenging to perform consistent comparisons. Also, these structures are though to evolve across cosmic time, and differences in their properties or galaxy populations may arise if observed at different stages of their evolution. Therefore, larger and more homogeneous samples, including structures similar to MQN01, would be necessary to unify the picture. Instruments such as MUSE or KCWI would be ideal to conduct deep and systematic wide surveys of extended quasar nebulae as signposts for these structures.

Bibliography

- Adams S. M., Martini P., Croxall K. V., Overzier R. A., Silverman J. D., 2015, [MNRAS](#), **448**, 1335
- 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
- Adler R. J., 1981, *The Geometry of Random Fields*
- Aggarwal K. M., 1983, [MNRAS](#), **202**, 15P
- Aguirre A., Hernquist L., Schaye J., Weinberg D. H., Katz N., Gardner J., 2001, [ApJ](#), **560**, 599
- Alberts S., Noble A., 2022, [Universe](#), **8**, 554
- Alberts S., et al., 2014, [MNRAS](#), **437**, 437
- Alonso S., Mesa V., Padilla N., Lambas D. G., 2012, [A&A](#), **539**, A46
- Alpaslan M., et al., 2014, [MNRAS](#), **438**, 177
- Anand A., Nelson D., Kauffmann G., 2021, [MNRAS](#), **504**, 65
- Anderson M. E., Bregman J. N., 2010, [ApJ](#), **714**, 320
- Anderson M. E., Bregman J. N., Dai X., 2013, [ApJ](#), **762**, 106
- Apostolovski Y., et al., 2024, [A&A](#), **683**, A64
- Appenzeller I., et al., 1998, *The Messenger*, **94**, 1
- Aragón-Calvo M. A., van de Weygaert R., Jones B. J. T., 2010, [MNRAS](#), **408**, 2163
- Arrigoni Battaia F., Yang Y., Hennawi J. F., Prochaska J. X., Matsuda Y., Yamada T., Hayashino T., 2015, [ApJ](#), **804**, 26
- Arrigoni Battaia F., Hennawi J. F., Prochaska J. X., Oñorbe J., Farina E. P., Cantalupo S., Lusso E., 2019a, [MNRAS](#), **482**, 3162
- Arrigoni Battaia F., Hennawi J. F., Prochaska J. X., Oñorbe J., Farina E. P., Cantalupo S., Lusso E., 2019b, [MNRAS](#), **482**, 3162
- Arrigoni Battaia F., et al., 2019c, [A&A](#), **631**, A18
- Arrigoni Battaia F., et al., 2022, [ApJ](#), **930**, 72
- Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, [ARA&A](#), **47**, 481
- Assef R. J., et al., 2015, [ApJ](#), **804**, 27
- Bañados E., Venemans B., Walter F., Kurk J., Overzier R., Ouchi M., 2013, [ApJ](#), **773**, 178
- Bacon R., et al., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III*. p. 773508, doi:10.1117/12.856027
- Bacon R., et al., 2015, [A&A](#), **575**, A75
- Bacon R., et al., 2017, [A&A](#), **608**, A1
- Bacon R., et al., 2021, [A&A](#), **647**, A107
- Bacon R., et al., 2023, [A&A](#), **670**, A4
- Bahcall J. N., Salpeter E. E., 1965, [ApJ](#), **142**, 1677
- Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, [MNRAS](#), **373**, 469
- Balogh M., et al., 2004, [MNRAS](#), **348**, 1355
- Bamford S. P., et al., 2009, [MNRAS](#), **393**, 1324
- Banerjee E., Muzahid S., Schaye J., Johnson S. D., Cantalupo S., 2023, [MNRAS](#), **524**, 5148
- Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, [ApJ](#), **304**, 15

- Barkana R., Loeb A., 2001, [Phys. Rep.](#), **349**, 125
- Barkana R., Loeb A., 2007, [Reports on Progress in Physics](#), **70**, 627
- Barnes L. A., Garel T., Kacprzak G. G., 2014, [Publications of the Astronomical Society of the Pacific](#), **126**, 969–1009
- Barrufet L., et al., 2020, [A&A](#), **641**, A129
- Becker G. D., Sargent W. L. W., Rauch M., Calverley A. P., 2011, [ApJ](#), **735**, 93
- Beckett A., et al., 2024, [arXiv e-prints](#), p. [arXiv:2408.11914](#)
- Beckmann R. S., et al., 2017, [MNRAS](#), **472**, 949
- Bergeron J., 1988, in Audouze J., Pelletan M. C., Szalay A., Zel’dovich Y. B., Peebles P. J. E., eds, Vol. 130, Large Scale Structures of the Universe. p. 343
- Bergeron J., Boissé P., 1991, [A&A](#), **243**, 344
- Bergeron J., Stasińska G., 1986, [A&A](#), **169**, 1
- Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, [Phys. Rep.](#), **367**, 1
- Bernstein R., Shectman S. A., Gunnels S. M., Mochnacki S., Athey A. E., 2003, in Iye M., Moorwood A. F. M., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes. pp 1694–1704, [doi:10.1117/12.461502](#)
- Bertin E., 2006, in Astronomical Data Analysis Software and Systems XV. p. 112
- Bertin E., 2010, SWarp: Resampling and Co-adding FITS Images Together, Astrophysics Source Code Library, record ascl:1010.068
- Bertin E., 2011, in Evans I. N., Accomazzi A., Mink D. J., Rots A. H., eds, Astronomical Society of the Pacific Conference Series Vol. 442, Astronomical Data Analysis Software and Systems XX. p. 435
- Bertin E., Arnouts S., 1996, [A&AS](#), **117**, 393
- Bialas D., Lisker T., Olczak C., Spurzem R., Kotulla R., 2015, [A&A](#), **576**, A103
- Bianconi M., Smith G. P., Haines C. P., McGee S. L., Finoguenov A., Egami E., 2018, [MNRAS](#), **473**, L79
- Bielby R. M., et al., 2011, [MNRAS](#), **414**, 2
- Bielby R. M., et al., 2016, [MNRAS](#), **456**, 4061
- 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
- Birkin J. E., et al., 2021, [MNRAS](#), **501**, 3926
- Black J. H., 1981, [MNRAS](#), **197**, 553
- Boera E., Murphy M. T., Becker G. D., Bolton J. S., 2014, [MNRAS](#), **441**, 1916
- Bolatto A. D., Wolfire M., Leroy A. K., 2013, [ARA&A](#), **51**, 207
- Bolmer J., et al., 2019, [A&A](#), **623**, A43
- Bolton J. S., Becker G. D., Haehnelt M. G., Viel M., 2014, [MNRAS](#), **438**, 2499
- Bond J. R., Kofman L., Pogosyan D., 1996, [Nature](#), **380**, 603
- Boogaard L. A., et al., 2020, [ApJ](#), **902**, 109
- Book L. G., Benson A. J., 2010, [ApJ](#), **716**, 810
- Boquien M., Burgarella D., Roehlly Y., Buat V., Ciesla L., Corre D., Inoue A. K., Salas H., 2019, [A&A](#), **622**, A103
- Bordoloi R., et al., 2011, [ApJ](#), **743**, 10
- Bordoloi R., et al., 2014, [ApJ](#), **796**, 136
- Bordoloi R., et al., 2023, [arXiv e-prints](#), p. [arXiv:2307.01273](#)
- Bordoloi R., et al., 2024, [ApJ](#), **963**, 28
- Borisova E., et al., 2016, [ApJ](#), **831**, 39
- Borthakur S., Heckman T., Strickland D., Wild V., Schiminovich D., 2013, [ApJ](#), **768**, 18
- Boselli A., Gavazzi G., 2006, [PASP](#), **118**, 517
- Boselli A., et al., 2021, [A&A](#), **646**, A139
- Boselli A., Fossati M., Sun M., 2022, [A&A Rev.](#), **30**, 3
- Bouché N., Lowenthal J. D., 2004, [ApJ](#), **609**, 513
- Bouché N., Murphy M. T., Péroux C., Csabai I., Wild V., 2006, [Monthly Notices of the Royal Astronomical Society](#), **371**, 495–512
- Bradley L., et al., 2023, astropy/photutils: 1.10.0, [doi:10.5281/zenodo.1035865](#)
- Brammer G. B., van Dokkum P. G., Coppi P., 2008, [ApJ](#), **686**, 1503
- Bregman J. N., 1980, [ApJ](#), **236**, 577

- Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, [MNRAS](#), **351**, 1151
- Bruzual G., Charlot S., 2003a, [MNRAS](#), **344**, 1000
- Bruzual G., Charlot S., 2003b, [MNRAS](#), **344**, 1000
- Budzynski J. M., Hewett P. C., 2011, [MNRAS](#), **416**, 1871
- Burbidge E. M., Lynds C. R., Burbidge G. R., 1966, [ApJ](#), **144**, 447
- Burchett J. N., et al., 2016, [ApJ](#), **832**, 124
- Burgarella D., Buat V., Iglesias-Páramo J., 2005, [MNRAS](#), **360**, 1413
- Byler N., Dalcanton J. J., Conroy C., Johnson B. D., Levesque E. M., Berg D. A., 2018, [ApJ](#), **863**, 14
- Cai Z., et al., 2016, [ApJ](#), **833**, 135
- Cai Z., et al., 2017, [ApJ](#), **839**, 131
- Calvi R., Poggianti B. M., Vulcani B., Fasano G., 2013, [MNRAS](#), **432**, 3141
- Calvi R., Castignani G., Dannerbauer H., 2023, [A&A](#), **678**, A15
- Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, [ApJ](#), **533**, 682
- Cantalupo S., 2017, in Fox A., Davé R., eds, *Astrophysics and Space Science Library* Vol. 430, *Gas Accretion onto Galaxies*. p. 195 ([arXiv:1612.00491](#)), [doi:10.1007/978-3-319-52512-9_9](#)
- Cantalupo S., Porciani C., Lilly S. J., Miniati F., 2005, [ApJ](#), **628**, 61
- Cantalupo S., Lilly S. J., Porciani C., 2007, [ApJ](#), **657**, 135
- Cantalupo S., Porciani C., Lilly S. J., 2008, [ApJ](#), **672**, 48
- Cantalupo S., Arrigoni-Battaia F., Prochaska J. X., Hennawi J. F., Madau P., 2014, [Nature](#), **506**, 63
- Cantalupo S., et al., 2019, [MNRAS](#), **483**, 5188
- Cantalupo S., et al., 2021, Unraveling the knots of gaseous Cosmic Web filaments at $z \sim 3$ through H-alpha emission observations, JWST Proposal. Cycle 1, ID. #1835
- Carilli C. L., Walter F., 2013, [ARA&A](#), **51**, 105
- Casali M., et al., 2006, in McLean I. S., Iye M., eds, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* Vol. 6269, *Ground-based and Airborne Instrumentation for Astronomy*. p. 62690W, [doi:10.1117/12.670150](#)
- Casey C. M., 2012, [MNRAS](#), **425**, 3094
- Casey C. M., et al., 2015, [ApJ](#), **808**, L33
- Cashman F. H., Kulkarni V. P., Kisieliński R., Ferland G. J., Bogdanovich P., 2017, [ApJS](#), **230**, 8
- Castignani G., Combes F., Salomé P., Benoist C., Chiaberge M., Freundlich J., De Zotti G., 2019, [A&A](#), **623**, A48
- Cen R., Ostriker J. P., 2000, [ApJ](#), **538**, 83
- Chabrier G., 2003, [PASP](#), **115**, 763
- Chapman S. C., Blain A., Ibatá R., Ivison R. J., Smail I., Morrison G., 2009, [ApJ](#), **691**, 560
- Chartab N., et al., 2020, [ApJ](#), **890**, 7
- 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., 2001, [ApJ](#), **556**, 158
- Chen H.-W., Helsby J. E., Gauthier J.-R., Shethman S. A., Thompson I. B., Tinker J. L., 2010a, [ApJ](#), **714**, 1521
- Chen H.-W., Wild V., Tinker J. L., Gauthier J.-R., Helsby J. E., Shethman S. A., Thompson I. B., 2010b, [ApJ](#), **724**, L176
- Chen S.-F. S., et al., 2017, [ApJ](#), **850**, 188
- Chen H.-W., Boettcher E., Johnson S. D., Zahedy F. S., Rudie G. C., Cooksey K. L., Rauch M., Mulchaey J. S., 2019, [ApJ](#), **878**, L33
- Chiang Y.-K., Overzier R., Gebhardt K., 2013, [ApJ](#), **779**, 127
- Chiang Y.-K., Overzier R., Gebhardt K., 2014, [ApJ](#), **782**, L3
- Chiang Y.-K., Overzier R. A., Gebhardt K., Henriques B., 2017, [ApJ](#), **844**, L23
- Christenson H. M., Jorgenson R. A., 2019, [ApJ](#), **883**, 17
- Churchill C. W., Nielsen N. M., Kacprzak G. G., Trujillo-Gomez S., 2013a, [The Astrophysical Journal](#), **763**, L42
- Churchill C. W., Trujillo-Gomez S., Nielsen N. M., Kacprzak G. G., 2013b, [ApJ](#), **779**, 87
- Cluver M. E., et al., 2020, [ApJ](#), **898**, 20
- Colberg J. M., Krughoff K. S., Connolly A. J., 2005, [MNRAS](#), **359**, 272
- Conroy C., 2013, [ARA&A](#), **51**, 393

- Conselice C. J., 2014, [ARA&A](#), **52**, 291
- Conselice C. J., Rajgor S., Myers R., 2008, [MNRAS](#), **386**, 909
- Contini E., De Lucia G., Hatch N., Borgani S., Kang X., 2016, [MNRAS](#), **456**, 1924
- Cooke J., O’Meara J. M., 2015, [ApJ](#), **812**, L27
- Cooke J., Wolfe A. M., Gawiser E., Prochaska J. X., 2006, [ApJ](#), **652**, 994
- Cooke E. A., Hatch N. A., Muldrew S. I., Rigby E. E., Kurk J. D., 2014, [MNRAS](#), **440**, 3262
- Cooksey K. L., Kao M. M., Simcoe R. A., O’Meara J. M., Prochaska J. X., 2013, [ApJ](#), **763**, 37
- Cooper M. C., et al., 2008, [MNRAS](#), **383**, 1058
- Coppin K., et al., 2006, [MNRAS](#), **372**, 1621
- Cornish T. M., et al., 2024, [MNRAS](#), **533**, 2399
- Cortese L., Gavazzi G., Boselli A., Franzetti P., Kennicutt R. C., O’Neil K., Sakai S., 2006, [A&A](#), **453**, 847
- Coziol R., Plauchu-Frayn I., 2007, [AJ](#), **133**, 2630
- Crighton N. H. M., et al., 2011, [MNRAS](#), **414**, 28
- Crighton N. H. M., Hennawi J. F., Simcoe R. A., Cooksey K. L., Murphy M. T., Fumagalli M., Prochaska J. X., Shanks T., 2015, [MNRAS](#), **446**, 18
- Cucciati O., et al., 2014, [A&A](#), **570**, A16
- Cucciati O., et al., 2018, [A&A](#), **619**, A49
- Cupani G., et al., 2016, in Chiozzi G., Guzman J. C., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 9913, Software and Cyberinfrastructure for Astronomy IV. p. 99131T, [doi:10.1117/12.2231379](#)
- D’Amato Q., et al., 2020, [A&A](#), **641**, L6
- D’Odorico V., et al., 2016, [MNRAS](#), **463**, 2690
- D’Odorico V., et al., 2018, [ApJ](#), **863**, L29
- D’Odorico V., et al., 2022, [MNRAS](#), **512**, 2389
- Daddi E., et al., 2007, [ApJ](#), **670**, 156
- Dale D. A., Helou G., Magdis G. E., Armus L., Díaz-Santos T., Shi Y., 2014, [ApJ](#), **784**, 83
- Dannerbauer H., et al., 2014, [A&A](#), **570**, A55
- Darvish B., Mobasher B., Sobral D., Rettura A., Scoville N., Faisst A., Capak P., 2016, [ApJ](#), **825**, 113
- Davé R., Tripp T. M., 2001, [ApJ](#), **553**, 528
- Davidzon I., et al., 2013, [A&A](#), **558**, A23
- Davis M., Peebles P. J. E., 1983, [ApJ](#), **267**, 465
- De Grandi S., Molendi S., 2002, [ApJ](#), **567**, 163
- Decarli R., et al., 2019, [ApJ](#), **882**, 138
- Dekel A., Sari R., Ceverino D., 2009, [ApJ](#), **703**, 785
- Dekker H., D’Odorico S., Kaufer A., Delabre B., Kotzlowski H., 2000, in Iye M., Moorwood A. F., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4008, Optical and IR Telescope Instrumentation and Detectors. pp 534–545, [doi:10.1117/12.395512](#)
- Di Mascolo L., et al., 2023, [Nature](#), **615**, 809
- Díaz C. G., Ryan-Weber E. V., Cooke J., Koyama Y., Ouchi M., 2015, [MNRAS](#), **448**, 1240
- Díaz C. G., Ryan-Weber E. V., Karman W., Caputi K. I., Salvadori S., Crighton N. H., Ouchi M., Vanzella E., 2021, [MNRAS](#), **502**, 2645
- Diener C., et al., 2017, [MNRAS](#), **471**, 3186
- Dijkstra M., 2014, [PASA](#), **31**, e040
- Dijkstra M., 2019, [Saas-Fee Advanced Course](#), **46**, 1
- Doroshkevich A. G., Shandarin S. F., 1978, *Soviet Ast.*, **22**, 653
- Downes D., Solomon P. M., 1998, [ApJ](#), **507**, 615
- Draine B. T., 2011, *Physics of the Interstellar and Intergalactic Medium*
- Dressler A., 1980, [ApJ](#), **236**, 351
- Dudzevičiūtė U., et al., 2020, [MNRAS](#), **494**, 3828
- Dudzevičiūtė U., et al., 2021, [MNRAS](#), **500**, 942
- Dutta R., et al., 2020, [MNRAS](#), **499**, 5022
- Dutta R., et al., 2021, [MNRAS](#), **508**, 4573
- Dutta R., et al., 2024a, [arXiv e-prints](#), p. [arXiv:2409.02182](#)
- Dutta R., Acebron A., Fumagalli M., Grillo C., Caminha G. B., Fossati M., 2024b, [MNRAS](#), **528**, 1895
- ESO CPL Development Team 2015, EsoRex: ESO Recipe Execution Tool, Astrophysics Source Code Library, record ascl:1504.003

- Eggen O. J., Lynden-Bell D., Sandage A. R., 1962, [ApJ](#), **136**, 748
- Eke V. R., et al., 2004, [MNRAS](#), **355**, 769
- Elbaz D., et al., 2007, [A&A](#), **468**, 33
- Ellison S. L., Songaila A., Schaye J., Pettini M., 2000, [AJ](#), **120**, 1175
- Erb D. K., Steidel C. C., Shapley A. E., Pettini M., Reddy N. A., Adelberger K. L., 2006, [ApJ](#), **646**, 107
- Ester M., Kriegel H.-P., Sander J., Xu X., 1996, in Second International Conference on Knowledge Discovery and Data Mining (KDD'96). Proceedings of a conference held August 2-4. pp 226–331
- Estrada N., et al., 2023, [A&A](#), **671**, A146
- Faisst A. L., Fudamoto Y., Oesch P. A., Scoville N., Riechers D. A., Pavesi R., Capak P., 2020, [MNRAS](#), **498**, 4192
- Falder J. T., et al., 2011, [ApJ](#), **735**, 123
- Farina E. P., et al., 2019, [ApJ](#), **887**, 196
- Felten J. E., 1976, [ApJ](#), **207**, 700
- Ferland G. J., Korista K. T., Verner D. A., Ferguson J. W., Kingdon J. B., Verner E. M., 1998, [PASP](#), **110**, 761
- Ferland G. J., et al., 2013, *Rev. Mexicana Astron. Astrofis.*, **49**, 137
- Field G. B., 1959, [ApJ](#), **129**, 525
- Finlator K., Davé R., 2008, [MNRAS](#), **385**, 2181
- Finn C. W., et al., 2016, [MNRAS](#), **460**, 590
- Fitzpatrick E. L., 1999, [PASP](#), **111**, 63
- Ford H. C., et al., 1998, in Bely P. Y., Breckinridge J. B., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 3356, Space Telescopes and Instruments V. pp 234–248, [doi:10.1117/12.324464](#)
- Ford A. B., Oppenheimer B. D., Davé R., Katz N., Kollmeier J. A., Weinberg D. H., 2013, [MNRAS](#), **432**, 89
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, [Publications of the Astronomical Society of the Pacific](#), **125**, 306
- Forrest B., et al., 2024, [arXiv e-prints](#), p. [arXiv:2405.18491](#)
- Fossati M., Fumagalli M., Boselli A., Gavazzi G., Sun M., Wilman D. J., 2016, [MNRAS](#), **455**, 2028
- Fossati M., et al., 2017, [ApJ](#), **835**, 153
- Fossati M., Mendel J. T., Boselli A., et al. 2018a, [A&A](#), **614**
- Fossati M., et al., 2018b, [A&A](#), **614**, A57
- Fossati M., et al., 2019, [MNRAS](#), **490**, 1451
- Fossati M., et al., 2021, [MNRAS](#), **503**, 3044
- Fraternali F., 2017, in Fox A., Davé R., eds, *Astrophysics and Space Science Library* Vol. 430, Gas Accretion onto Galaxies. p. 323 ([arXiv:1612.00477](#)), [doi:10.1007/978-3-319-52512-9_14](#)
- Fritz J., Franceschini A., Hatziminaoglou E., 2006, [MNRAS](#), **366**, 767
- Fujita Y., 2004, [PASJ](#), **56**, 29
- Fumagalli M., Prochaska J. X., Kasen D., Dekel A., Ceverino D., Primack J. R., 2011, [MNRAS](#), **418**, 1796
- Fumagalli M., O'Meara J. M., Prochaska J. X., Worseck G., 2013, [ApJ](#), **775**, 78
- Fumagalli M., O'Meara J. M., Prochaska J. X., 2016, [MNRAS](#), **455**, 4100
- Fumagalli M., et al., 2017, [MNRAS](#), **471**, 3686
- Fynbo J. P. U., et al., 2010, [MNRAS](#), **408**, 2128
- Fynbo J. P. U., et al., 2011, [MNRAS](#), **413**, 2481
- Fynbo J. P. U., et al., 2018, [MNRAS](#), **479**, 2126
- Gaia Collaboration et al., 2016, [A&A](#), **595**, A1
- Gaia Collaboration et al., 2023, [A&A](#), **674**, A1
- Galbiati M., Fumagalli M., Fossati M., Lofthouse E. K., Dutta R., Prochaska J. X., Murphy M. T., Cantalupo S., 2023, [MNRAS](#), **524**, 3474
- Galbiati M., et al., 2024a, [arXiv e-prints](#), p. [arXiv:2410.03822](#)
- Galbiati M., Dutta R., Fumagalli M., Fossati M., Cantalupo S., 2024b, [A&A](#), **690**, A7
- Gallagher John S. I., Ostriker J. P., 1972, [AJ](#), **77**, 288
- Gamow G., Teller E., 1939, [Physical Review](#), **55**, 654
- Garcia-Vergara C., 2021, in *Galaxy Cluster Formation II*. p. 49, [doi:10.5281/zenodo.5005051](#)
- García-Vergara C., Hennawi J. F., Barrientos L. F., Rix H.-W., 2017, [ApJ](#), **848**, 7
- García-Vergara C., Hennawi J. F., Barrientos L. F., Arrigoni Battaia F., 2019, [ApJ](#), **886**, 79
- Garzilli A., Bolton J. S., Kim T. S., Leach S., Viel M., 2012, [MNRAS](#), **424**, 1723

- Gatica C., Demarco R., Dole H., Polletta M., Frye B., Martinache C., Rettura A., 2024, [MNRAS](#), **527**, 3006
- Gauthier J.-R., Chen H.-W., Tinker J. L., 2009, [The Astrophysical Journal](#), **702**, 50–62
- Gawiser E., et al., 2007, [ApJ](#), **671**, 278
- Gehrels N., 1986, [ApJ](#), **303**, 336
- Geller M. J., Huchra J. P., 1989, [Science](#), **246**, 897
- Ghirardini V., et al., 2019, [A&A](#), **621**, A41
- Ginolfi M., et al., 2022, [Nature Communications](#), **13**, 4574
- Gnedin N. Y., Madau P., 2022, [Living Reviews in Computational Astrophysics](#), **8**, 3
- Gonzaga S., Hack W., Fruchter A., Mack J., 2012, *The DrizzlePac Handbook*
- González-López J., et al., 2017, [A&A](#), **597**, A41
- González-López J., et al., 2019, [ApJ](#), **882**, 139
- González-López J., et al., 2020, [ApJ](#), **897**, 91
- Gould A., Weinberg D. H., 1996, [ApJ](#), **468**, 462
- Greenslade J., et al., 2018, [MNRAS](#), **476**, 3336
- Gronke M., 2017, [A&A](#), **608**, A139
- Grützbauch R., et al., 2011, [MNRAS](#), **418**, 938
- Gu M., Conroy C., Brammer G., 2018, [ApJ](#), **862**, L18
- Guha L. K., Srianand R., 2023, [MNRAS](#), **519**, 3319
- Gunasekera C. M., van Hoof P. A. M., Chatzikos M., Ferland G. J., 2023, [Research Notes of the American Astronomical Society](#), **7**, 246
- Gunn J. E., Gott J. Richard I., 1972, [ApJ](#), **176**, 1
- Gunn J. E., Peterson B. A., 1965, [ApJ](#), **142**, 1633
- Gupta A., Mathur S., Krongold Y., Nicastro F., Galeazzi M., 2012, [ApJ](#), **756**, L8
- Haardt F., Madau P., 2012, [ApJ](#), **746**, 125
- Hafen Z., et al., 2019, [MNRAS](#), **488**, 1248
- Haines C. P., et al., 2015, [ApJ](#), **806**, 101
- Hamanowicz A., et al., 2020, [MNRAS](#), **492**, 2347
- Harshan A., et al., 2021, [ApJ](#), **919**, 57
- Hasan F., et al., 2020, [ApJ](#), **904**, 44
- Hasan F., Churchill C. W., Stemmock B., Nielsen N. M., Kacprzak G. G., Croom M., Murphy M. T., 2021, arXiv e-prints, p. [arXiv:2108.04924](#)
- Hatch N. A., et al., 2011, [MNRAS](#), **410**, 1537
- Hatch N. A., et al., 2014, [MNRAS](#), **445**, 280
- Hayashi M., Kodama T., Tadaki K.-i., Koyama Y., Tanaka I., 2012, [ApJ](#), **757**, 15
- Hayashino T., et al., 2004, [AJ](#), **128**, 2073
- Hazard C., 1963, [MNRAS](#), **126**, 489
- Heintz K. E., et al., 2019, [A&A](#), **621**, A20
- Heintz K. E., et al., 2023, [A&A](#), **679**, A91
- Hennawi J. F., Prochaska J. X., 2007, [ApJ](#), **655**, 735
- Hennawi J. F., Prochaska J. X., 2013, [ApJ](#), **766**, 58
- Hennawi J. F., et al., 2006, [AJ](#), **131**, 1
- Hennawi J. F., Prochaska J. X., Kollmeier J., Zheng Z., 2009, [ApJ](#), **693**, L49
- Hennawi J. F., Prochaska J. X., Cantalupo S., Arrigoni-Battaia F., 2015, [Science](#), **348**, 779
- Herenz E. C., et al., 2019, [A&A](#), **621**, A107
- Herrero Alonso Y., et al., 2021, [A&A](#), **653**, A136
- Hickox R. C., et al., 2012, [MNRAS](#), **421**, 284
- 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., Schaye J., 2020, [ApJ](#), **904**, 76
- Ho S. H., Martin C. L., Schaye J., 2021, [ApJ](#), **923**, 137
- Hoffmann S. L., Mack J., Avila R., Martlin C., Cohen Y., Bajaj V., 2021, in *American Astronomical Society Meeting Abstracts*. p. 216.02
- Hogan C. J., Weymann R. J., 1987, [MNRAS](#), **225**, 1P
- Hogg D. W., 1999, arXiv e-prints, pp [astro-ph/9905116](#)
- Hogg D. W., Turner E. L., 1998, [PASP](#), **110**, 727
- Horne K., 1986, [PASP](#), **98**, 609
- Hu E. M., Kim T.-S., Cowie L. L., Songaila A., Rauch M., 1995, [AJ](#), **110**, 1526
- Huang Y., et al., 2022, [ApJ](#), **941**, 134

Hung D., et al., 2024, [arXiv e-prints](#), p. [arXiv:2410.00237](#)

Hunter T. R., Petry D., Barkats D., Corder S., Indebetouw R., 2023a, *analysisUtils*, [doi:10.5281/zenodo.7502160](#)

Hunter T. R., et al., 2023b, *PASP*, **135**, 074501

Husband K., Bremer M. N., Stanway E. R., Davies L. J. M., Lehnert M. D., Douglas L. S., 2013, *MNRAS*, **432**, 2869

Hwang J.-S., Park C., Banerjee A., Hwang H. S., 2018, *ApJ*, **856**, 160

Ilbert O., et al., 2013, *A&A*, **556**, A55

Inoue A. K., 2011, *MNRAS*, **415**, 2920

Ito K., et al., 2020, *ApJ*, **899**, 5

Izzo C., de Bilbao L., Larsen J., Bagnulo S., Freudling W., Moehler S., Ballester P., 2010, in Silva D. R., Peck A. B., Soifer B. T., eds, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 7737, Observatory Operations: Strategies, Processes, and Systems III*. p. 773729, [doi:10.1117/12.858043](#)

Jin S., et al., 2021, *A&A*, **652**, A11

Johnston R., 2011, *A&A Rev.*, **19**, 41

Jones S. F., et al., 2014, *MNRAS*, **443**, 146

Kanekar N., et al., 2018, *ApJ*, **856**, L23

Kanekar N., Prochaska J. X., Neeleman M., Christensen L., Møller P., Zwaan M. A., Fynbo J. P. U., Dessauges-Zavadsky M., 2020, *ApJ*, **901**, L5

Kashikawa N., Kitayama T., Doi M., Misawa T., Komiyama Y., Ota K., 2007, *ApJ*, **663**, 765

Kashino D., Lilly S. J., Matthee J., Eilers A.-C., Mackenzie R., Bordoloi R., Simcoe R. A., 2023, *ApJ*, **950**, 66

Kauffmann G., White S. D. M., 1993, *MNRAS*, **261**, 921

Kaur B., Kanekar N., Rafelski M., Neeleman M., Prochaska J. X., Revalski M., 2022a, *ApJ*, **933**, L42

Kaur B., Kanekar N., Revalski M., Rafelski M., Neeleman M., Prochaska J. X., Walter F., 2022b, *ApJ*, **934**, 87

Kelson D. D., 2003, *PASP*, **115**, 688

Kereš D., Katz N., Weinberg D. H., Davé R., 2005, *MNRAS*, **363**, 2

Kikuta S., et al., 2019, *PASJ*, **71**, L2

Kim T.-S., Hu E. M., Cowie L. L., Songaila A., 1997, *AJ*, **114**, 1

Kim S., et al., 2009, *ApJ*, **695**, 809

Kirkman D., Tytler D., 1997, *ApJ*, **484**, 672

Kirshner R. P., Oemler A. J., Schechter P. L., Shectman S. A., 1981, *ApJ*, **248**, L57

Kissler-Patig M., et al., 2008, *A&A*, **491**, 941

Klitsch A., Péroux C., Zwaan M. A., Smail I., Oteo I., Biggs A. D., Popping G., Swinbank A. M., 2018, *MNRAS*, **475**, 492

Klitsch A., et al., 2019, *MNRAS*, **482**, L65

Kodama T., Tanaka I., Kajisawa M., Kurk J., Venemans B., De Breuck C., Vernet J., Lidman C., 2007, *MNRAS*, **377**, 1717

Kollmeier J. A., Zheng Z., Davé R., Gould A., Katz N., Miralda-Escudé J., Weinberg D. H., 2010, *ApJ*, **708**, 1048

Koprowski M. P., Wijesekera J. V., Dunlop J. S., McLeod D. J., Michałowski M. J., Lisiecki K., McLure R. J., 2024, *arXiv e-prints*, p. [arXiv:2403.06575](#)

Koyama Y., Kodama T., Tadaki K.-i., Hayashi M., Tanaka M., Smail I., Tanaka I., Kurk J., 2013a, *MNRAS*, **428**, 1551

Koyama Y., et al., 2013b, *MNRAS*, **434**, 423

Koyama Y., et al., 2021, *MNRAS*, **503**, L1

Kubo M., et al., 2013, *ApJ*, **778**, 170

Kuiper E., et al., 2010, *MNRAS*, **405**, 969

Kurk J. D., et al., 2000, *A&A*, **358**, L1

Kurk J. D., Pentericci L., Röttgering H. J. A., Miley G. K., 2004, *A&A*, **428**, 793

Kusakabe H., et al., 2022, *A&A*, **660**, A44

Lacaille K. M., et al., 2019, *MNRAS*, **488**, 1790

Lacey C., Cole S., 1993, *MNRAS*, **262**, 627

Laigle C., et al., 2016, *ApJS*, **224**, 24

Lammers C., Hill R., Lim S., Scott D., Cañameras R., Dole H., 2022, *MNRAS*, **514**, 5004

Lan T.-W., 2020, *ApJ*, **897**, 97

- Lan T.-W., Ménard B., Zhu G., 2014, [ApJ](#), **795**, 31
- Langen V., Cantalupo S., Steidel C. C., Chen Y., Pezzulli G., Gallego S. G., 2023, [MNRAS](#), **519**, 5099
- Lanzetta K. M., Bowen D., 1990, [ApJ](#), **357**, 321
- Larson R. B., Tinsley B. M., Caldwell C. N., 1980, [ApJ](#), **237**, 692
- Le Fèvre O., et al., 2015, [A&A](#), **576**, A79
- Leclercq F., et al., 2017, [A&A](#), **608**, A8
- Leclercq F., et al., 2022, [A&A](#), **663**, A11
- Lee K.-G., et al., 2013, [AJ](#), **145**, 69
- Lee K.-S., Dey A., Hong S., Reddy N., Wilson C., Jannuzi B. T., Inami H., Gonzalez A. H., 2014, [ApJ](#), **796**, 126
- Lee M. M., et al., 2017, [ApJ](#), **842**, 55
- Lee K.-G., et al., 2018, [ApJS](#), **237**, 31
- Lehmer B. D., et al., 2009, [ApJ](#), **691**, 687
- Lemaux B. C., et al., 2014, [A&A](#), **572**, A41
- Lemaux B. C., et al., 2018, [A&A](#), **615**, A77
- Lemaux B. C., et al., 2019, [MNRAS](#), **490**, 1231
- Lemaux B. C., et al., 2022, [A&A](#), **662**, A33
- Lewis I., et al., 2002, [MNRAS](#), **334**, 673
- Liang C. J., Chen H.-W., 2014, [MNRAS](#), **445**, 2061
- Libeskind N. I., et al., 2018, [MNRAS](#), **473**, 1195
- Lifshitz E. M., 1946, *Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki*, **16**, 587
- Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, [ApJ](#), **772**, 119
- Lim C.-F., et al., 2020, [ApJ](#), **895**, 104
- Lin X., et al., 2023, [ApJ](#), **956**, 148
- Liu Y., Gao L., Liao S., Zhu K., 2024, [arXiv e-prints](#), p. [arXiv:2409.11088](#)
- Lofthouse E. K., et al., 2020b, [MNRAS](#), **491**, 2057
- Lofthouse E. K., et al., 2020a, [MNRAS](#), **491**, 2057
- Lofthouse E. K., et al., 2023, [MNRAS](#), **518**, 305
- Longobardi A., et al., 2023, [RAS Techniques and Instruments](#), **2**, 470
- Lopes P. A. A., Ribeiro A. L. B., Brambila D., 2024, [MNRAS](#), **527**, L19
- López S., et al., 2016, [A&A](#), **594**, A91
- Lovell C. C., Thomas P. A., Wilkins S. M., 2018, [MNRAS](#), **474**, 4612
- Luo Y., et al., 2022, [ApJ](#), **935**, 80
- Lusso E., et al., 2019, [MNRAS](#), **485**, L62
- Mackenzie R., et al., 2019, [MNRAS](#), **487**, 5070
- Macuga M., et al., 2019, [ApJ](#), **874**, 54
- Madau P., Dickinson M., 2014, [ARA&A](#), **52**, 415
- Maio U., Khochfar S., Johnson J. L., Ciardi B., 2011, [MNRAS](#), **414**, 1145
- Marino R. A., et al., 2018, [ApJ](#), **859**, 53
- Martin C. L., Bouché N., 2009, [The Astrophysical Journal](#), **703**, 1394–1415
- Martin C., Moore A., Morrissey P., Matuszewski M., Rahman S., Adkins S., Epps H., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III*. p. 77350M, [doi:10.1117/12.858227](#)
- Martin D. C., Chang D., Matuszewski M., Morrissey P., Rahman S., Moore A., Steidel C. C., Matsuda Y., 2014, [ApJ](#), **786**, 107
- Martinache C., et al., 2018, [A&A](#), **620**, A198
- Martizzi D., Vogelsberger M., Torrey P., Pillepich A., Hansen S. H., Marinacci F., Hernquist L., 2020, [MNRAS](#), **491**, 5747
- Mas-Ribas L., Dijkstra M., Hennawi J. F., Trenti M., Momose R., Ouchi M., 2017, [ApJ](#), **841**, 19
- Matejek M. S., Simcoe R. A., 2012, [ApJ](#), **761**, 112
- Mathur S., Das S., Gupta A., Krongold Y., 2023, [MNRAS](#), **525**, L11
- McGee S. L., Balogh M. L., Bower R. G., Font A. S., McCarthy I. G., 2009, [MNRAS](#), **400**, 937
- McKee C. F., Ostriker E. C., 2007, [ARA&A](#), **45**, 565
- McMullin J. P., Waters B., Schiebel D., Young W., Golap K., 2007, in *Astronomical Data Analysis Software and Systems XVI*. p. 127
- McNab K., et al., 2021, [MNRAS](#), **508**, 157
- McNamara B. R., Nulsen P. E. J., 2012, [New Journal of Physics](#), **14**, 055023

- McQuinn M., 2016, [ARA&A](#), **54**, 313
- McQuinn M., Oh S. P., Faucher-Giguère C.-A., 2011, [ApJ](#), **743**, 82
- Mehra V., et al., 2015, [ApJ](#), **811**, 141
- Mei S., et al., 2023, [A&A](#), **670**, A58
- Meiksin A. A., 2009, [Reviews of Modern Physics](#), **81**, 1405
- Ménard B., Fukugita M., 2012, [ApJ](#), **754**, 116
- Merritt D., 1983, [ApJ](#), **264**, 24
- Merritt D., 1984, [ApJ](#), **276**, 26
- Meyer R. A., et al., 2020, [MNRAS](#), **494**, 1560
- Mignoli M., et al., 2020, [A&A](#), **642**, L1
- Miley G. K., et al., 2006, [ApJ](#), **650**, L29
- Miller T. B., et al., 2018, [Nature](#), **556**, 469
- Mo H. J., White S. D. M., 1996, [MNRAS](#), **282**, 347
- Møller P., Warren S. J., Fall S. M., Fynbo J. U., Jakobsen P., 2002, [ApJ](#), **574**, 51
- Møller P., Fynbo J. P. U., Fall S. M., 2004, [A&A](#), **422**, L33
- Møller P., et al., 2018, [MNRAS](#), **474**, 4039
- Momose R., Shimizu I., Nagamine K., Shimasaku K., Kashikawa N., Kusakabe H., 2021, [ApJ](#), **911**, 98
- Moon J.-S., An S.-H., Yoon S.-J., 2019, [ApJ](#), **882**, 14
- Moore B., Katz N., Lake G., Dressler A., Oemler A., 1996, [Nature](#), **379**, 613
- Morrissey P., et al., 2018, [ApJ](#), **864**, 93
- Moster B. P., Somerville R. S., Maubetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, [ApJ](#), **710**, 903
- Moustakas J., et al., 2013, [ApJ](#), **767**, 50
- Muldrew S. I., Hatch N. A., Cooke E. A., 2015, [MNRAS](#), **452**, 2528
- Muldrew S. I., Hatch N. A., Cooke E. A., 2018, [MNRAS](#), **473**, 2335
- Murphy E. J., et al., 2011, [ApJ](#), **737**, 67
- Muzahid S., et al., 2020, [MNRAS](#), **496**, 1013
- Muzahid S., et al., 2021, arXiv e-prints, [p. arXiv:2105.05260](#)
- Ménard B., Wild V., Nestor D., Quider A., Zibetti S., Rao S., Turnshek D., 2011, [Monthly Notices of the Royal Astronomical Society](#), **417**, 801–811
- Neeleman M., Prochaska J. X., Wolfe A. M., 2015, [ApJ](#), **800**, 7
- Neeleman M., Kanekar N., Prochaska J. X., Rafelski M., Carilli C. L., Wolfe A. M., 2017, [Science](#), **355**, 1285
- Neeleman M., Kanekar N., Prochaska J. X., Christensen L., Dessauges-Zavadsky M., Fynbo J. P. U., Møller P., Zwaan M. A., 2018, [ApJ](#), **856**, L12
- Neeleman M., Kanekar N., Prochaska J. X., Rafelski M. A., Carilli C. L., 2019, [ApJ](#), **870**, L19
- Nelson D., Byrohl C., Peroux C., Rubin K. H. R., Burchett J. N., 2021, [MNRAS](#), **507**, 4445
- Nestor D. B., Johnson B. D., Wild V., Ménard B., Turnshek D. A., Rao S., Pettini M., 2011, [Monthly Notices of the Royal Astronomical Society](#), **412**, 1559–1572
- Newman A. B., et al., 2020, [ApJ](#), **891**, 147
- Newman A. B., et al., 2022, [Nature](#), **606**, 475
- Nicastro F., et al., 2023, [ApJ](#), **955**, L21
- Nielsen N. M., Churchill C. W., Kacprzak G. G., 2013, [ApJ](#), **776**, 115
- Nielsen N. M., Kacprzak G. G., Pointon S. K., Churchill C. W., Murphy M. T., 2018a, [ApJ](#), **869**, 153
- Nielsen N. M., Kacprzak G. G., Pointon S. K., Churchill C. W., Murphy M. T., 2018b, [ApJ](#), **869**, 153
- Nilsson K. K., Fynbo J. P. U., Møller P., Sommer-Larsen J., Ledoux C., 2006, [A&A](#), **452**, L23
- Noeske K. G., et al., 2007, [ApJ](#), **660**, L47
- Noll S., Burgarella D., Giovannoli E., Buat V., Marcillac D., Muñoz-Mateos J. C., 2009, [A&A](#), **507**, 1793
- Nomoto K., Kobayashi C., Tominaga N., 2013, [ARA&A](#), **51**, 457
- North P. L., Courbin F., Eigenbrod A., Chelouche D., 2012, [A&A](#), **542**, A91
- O’Meara J. M., et al., 2015, [AJ](#), **150**, 111
- Old L. J., et al., 2020, [MNRAS](#), **493**, 5987
- Oppenheimer B. D., Davé R., 2006, [MNRAS](#), **373**, 1265
- Oppenheimer B. D., Davé R., 2008, [MNRAS](#), **387**, 577
- Oppenheimer B. D., Davé R., Katz N., Kollmeier J. A., Weinberg D. H., 2012, [MNRAS](#), **420**, 829

- Orsi Á. A., Fanidakis N., Lacey C. G., Baugh C. M., 2016, *MNRAS*, **456**, 3827
- Osterbrock D. E., 1962, *ApJ*, **135**, 195
- Oteo I., et al., 2018, *ApJ*, **856**, 72
- Ouchi M., et al., 2010, *ApJ*, **723**, 869
- Ouchi M., Ono Y., Shibuya T., 2020, *ARA&A*, **58**, 617
- Overzier R. A., 2016, *A&A Rev.*, **24**, 14
- Overzier R. A., et al., 2008, *ApJ*, **673**, 143
- Pacifici C., et al., 2023, *ApJ*, **944**, 141
- Park C., Hwang H. S., 2009, *ApJ*, **699**, 1595
- Patel S. G., Holden B. P., Kelson D. D., Illingworth G. D., Franx M., 2009, *ApJ*, **705**, L67
- Patnaik A. R., Browne I. W. A., Walsh D., Chaffee F. H., Foltz C. B., 1992, *MNRAS*, **259**, 1P
- Patrício V., et al., 2016, *MNRAS*, **456**, 4191
- Pavesi R., et al., 2018, *ApJ*, **864**, 49
- Pawlik A. H., Schaye J., van Scherpenzeel E., 2009, *MNRAS*, **394**, 1812
- Pedregosa F., et al., 2011, *Journal of Machine Learning Research*, **12**, 2825
- Peebles P. J. E., 1965, *ApJ*, **142**, 1317
- Peebles P. J. E., 1980, The large-scale structure of the universe
- Peebles P. J. E., Yu J. T., 1970, *ApJ*, **162**, 815
- Peng Y.-j., et al., 2010, *ApJ*, **721**, 193
- Peng Y., Maiolino R., Cochrane R., 2015, *Nature*, **521**, 192
- Penney J. I., et al., 2019, *MNRAS*, **483**, 514
- Pensabene A., et al., 2024a, *arXiv e-prints*, p. arXiv:2410.06249
- Pensabene A., et al., 2024b, *A&A*, **684**, A119
- Pentericci L., Roettgering H. J. A., Miley G. K., Carilli C. L., McCarthy P., 1997, *A&A*, **326**, 580
- Pentericci L., et al., 2000, *A&A*, **361**, L25
- Pérez-Martínez J. M., et al., 2023, *MNRAS*, **518**, 1707
- Pérez-Martínez J. M., et al., 2024, *MNRAS*, **527**, 10221
- Péroux C., Howk J. C., 2020, *ARA&A*, **58**, 363
- Péroux C., Dessauges-Zavadsky M., D’Odorico S., Kim T.-S., McMahon R. G., 2003, *MNRAS*, **345**, 480
- Péroux C., et al., 2017, *MNRAS*, **464**, 2053
- Péroux C., Rahmani H., Arrigoni Battaia F., Augustin R., 2018, *MNRAS*, **479**, L50
- Péroux C., et al., 2019, *MNRAS*, **485**, 1595
- Petitjean P., Webb J. K., Rauch M., Carswell R. F., Lanzetta K., 1993, *MNRAS*, **262**, 499
- Pettini M., Madau P., Bolte M., Prochaska J. X., Ellison S. L., Fan X., 2003, *ApJ*, **594**, 695
- Pieri M. M., Schaye J., Aguirre A., 2006, *ApJ*, **638**, 45
- Pilbratt G. L., et al., 2010, *A&A*, **518**, L1
- Pirard J.-F., et al., 2004, in Moorwood A. F. M., Iye M., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 5492, Ground-based Instrumentation for Astronomy. pp 1763–1772, doi:10.1117/12.578293
- Planck Collaboration et al., 2016, *A&A*, **594**, A1
- Planck Collaboration et al., 2020a, *A&A*, **641**, A1
- Planck Collaboration et al., 2020b, *A&A*, **641**, A6
- Polletta M., et al., 2021, *A&A*, **654**, A121
- Polletta M., Dole H., Martinache C., Lehnert M. D., Frye B. L., Kneissl R., 2022, *A&A*, **662**, A85
- Popesso P., et al., 2023, *MNRAS*, **519**, 1526
- Pozzetti L., et al., 2010, *A&A*, **523**, A13
- Prescott M. K. M., Kashikawa N., Dey A., Matsuda Y., 2008, *ApJ*, **678**, L77
- Prescott M. K. M., et al., 2012, *ApJ*, **752**, 86
- Press W. H., Schechter P., 1974, *ApJ*, **187**, 425
- Priddey R. S., Ivison R. J., Isaak K. G., 2008, *MNRAS*, **383**, 289
- Primack J. R., 2024, *Annual Review of Nuclear and Particle Science*
- Prochaska J. X., Wolfe A. M., 1997, *ApJ*, **487**, 73
- Prochaska J. X., Wolfe A. M., 2009, *ApJ*, **696**, 1543
- Prochaska J. X., Chen H.-W., Howk J. C., Weiner B. J., Mulchaey J., 2004, *ApJ*, **617**, 718
- Prochaska J. X., O’Meara J. M., Worseck G., 2010, *ApJ*, **718**, 392
- Prochaska J. X., Lau M. W., Hennawi J. F., 2014, *ApJ*, **796**, 140
- Prochaska J. X., et al., 2019, *ApJS*, **243**, 24
- Prochaska J., et al., 2020, *The Journal of Open Source Software*, **5**, 2308

Prochter G. E., Prochaska J. X., Burles S. M., 2006, [ApJ](#), **639**, 766

Qu Z., et al., 2023, [MNRAS](#), **524**, 512

Rafelski M., Wolfe A. M., Prochaska J. X., Neeleman M., Mendez A. J., 2012, [ApJ](#), **755**, 89

Rafelski M., Neeleman M., Fumagalli M., Wolfe A. M., Prochaska J. X., 2014, [ApJ](#), **782**, L29

Raichoor A., et al., 2011, [ApJ](#), **732**, 12

Rakic O., Schaye J., Steidel C. C., Rudie G. C., 2011, [MNRAS](#), **414**, 3265

Ramakrishnan V., et al., 2023, [ApJ](#), **951**, 119

Reddy N. A., Steidel C. C., 2009, [ApJ](#), **692**, 778

Reddy N. A., Steidel C. C., Pettini M., Adelberger K. L., Shapley A. E., Erb D. K., Dickinson M., 2008, [ApJS](#), **175**, 48

Revalski M., et al., 2023, [ApJS](#), **265**, 40

Revalski M., et al., 2024, [ApJ](#), **966**, 228

Ribaud J., Lehner N., Christopher Howk J., Werk J. K., Tripp T. M., Xavier Prochaska J., Meiring J. D., Tumlinson J., 2011, [The Astrophysical Journal](#), **743**, 207

Riccio G., et al., 2023, [A&A](#), **678**, A164

Richstone D. O., 1976, [ApJ](#), **204**, 642

Riechers D. A., et al., 2013, [Nature](#), **496**, 329

Rigby E. E., et al., 2014, [MNRAS](#), **437**, 1882

Rigby J., et al., 2023, [PASP](#), **135**, 048001

Robert P. F., Murphy M. T., O’Meara J. M., Crighton N. H. M., Fumagalli M., 2019, [MNRAS](#), **483**, 2736

Roettgering H. J. A., Lacy M., Miley G. K., Chambers K. C., Saunders R., 1994, [A&AS](#), **108**, 79

Rubiño-Martín J. A., Chluba J., Sunyaev R. A., 2006, [MNRAS](#), **371**, 1939

Rubin K. H. R., Weiner B. J., Koo D. C., Martin C. L., Prochaska J. X., Coil A. L., Newman J. A., 2010, [The Astrophysical Journal](#), **719**, 1503–1525

Rubin K. H. R., Xavier Prochaska J., Koo D. C., Phillips A. C., 2012, [The Astrophysical Journal](#), **747**, L26

Rubin K. H. R., Prochaska J. X., Koo D. C., Phillips A. C., Martin C. L., Winstrom L. O., 2014, [The Astrophysical Journal](#), **794**, 156

Rubin K. H. R., Diamond-Stanic A. M., Coil A. L., Crighton N. H. M., Moustakas J., 2018, [ApJ](#), **853**, 95

Rudie G. C., et al., 2012, [ApJ](#), **750**, 67

Rudie G. C., Steidel C. C., Shapley A. E., Pettini M., 2013, [ApJ](#), **769**, 146

Rudie G. C., Steidel C. C., Pettini M., Trainor R. F., Strom A. L., Hummels C. B., Reddy N. A., Shapley A. E., 2019, [ApJ](#), **885**, 61

Salim S., et al., 2007, [ApJS](#), **173**, 267

Salpeter E. E., 1955, [ApJ](#), **121**, 161

Sánchez S. F., et al., 2022, [ApJS](#), **262**, 36

Scannapieco E., Ferrara A., Madau P., 2002, [ApJ](#), **574**, 590

Schaye J., 2001a, [ApJ](#), **559**, 507

Schaye J., 2001b, [ApJ](#), **562**, L95

Schaye J., Aguirre A., Kim T.-S., Theuns T., Rauch M., Sargent W. L. W., 2003, [ApJ](#), **596**, 768

Schechter P., 1976, [ApJ](#), **203**, 297

Scheuer P. A. G., 1965, [Nature](#), **207**, 963

Schlaflly E. F., Finkbeiner D. P., 2011, [ApJ](#), **737**, 103

Schmidt M., 1963, [Nature](#), **197**, 1040

Schmidt M., 1968, [ApJ](#), **151**, 393

Schneider R., Ferrara A., Natarajan P., Omukai K., 2002, [ApJ](#), **571**, 30

Scholz T. T., Walters H. R. J., Burke P. J., Scott M. P., 1990, [MNRAS](#), **242**, 692

Schroetter I., et al., 2016, [ApJ](#), **833**, 39

Schroetter I., et al., 2019, [MNRAS](#), **490**, 4368

Schroetter I., et al., 2021, [MNRAS](#), **506**, 1355

Schroetter I., Bouché N. F., Zabl J., Wendt M., Cherrey M., Langan I., Schaye J., Contini T., 2024, [A&A](#), **687**, A39

Scott J., Bechtold J., Dobrzycki A., Kulkarni V. P., 2000, [ApJS](#), **130**, 67

Scott S. E., et al., 2002, [MNRAS](#), **331**, 817

Scoville N., et al., 2013, [ApJS](#), **206**, 3

Shapiro P. R., Field G. B., 1976, [ApJ](#), **205**, 762

Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, [ApJ](#), **588**, 65

Sheinis A. I., Bolte M., Epps H. W., Kibrick R. I., Miller J. S., Radovan M. V., Bigelow B. C., Sutin B. M., 2002, *PASP*, **114**, 851

Shen S., Madau P., Guedes J., Mayer L., Prochaska J. X., Wadsley J., 2013, *ApJ*, **765**, 89

Shi K., et al., 2019a, *ApJ*, **871**, 83

Shi K., et al., 2019b, *ApJ*, **879**, 9

Shi K., Toshikawa J., Cai Z., Lee K.-S., Fang T., 2020, *ApJ*, **899**, 79

Shi K., Toshikawa J., Lee K.-S., Wang T., Cai Z., Fang T., 2021, *ApJ*, **911**, 46

Shi K., Malavasi N., Toshikawa J., Zheng X., 2024, *ApJ*, **961**, 39

Shimakawa R., et al., 2018a, *MNRAS*, **473**, 1977

Shimakawa R., et al., 2018b, *MNRAS*, **481**, 5630

Siebenmorgen R., Carraro G., Valenti E., Petr-Gotzens M., Brammer G., Garcia E., Casali M., 2011, *The Messenger*, **144**, 9

Silk J., 1968, *ApJ*, **151**, 459

Simcoe R. A., Sargent W. L. W., Rauch M., 2004, *ApJ*, **606**, 92

Simcoe R. A., et al., 2011, *ApJ*, **743**, 21

Skrutskie M. F., et al., 2006, *AJ*, **131**, 1163

Solomon P. M., Downes D., Radford S. J. E., Barrett J. W., 1997, *ApJ*, **478**, 144

Songaila A., 2005, *AJ*, **130**, 1996

Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, *ApJS*, **214**, 15

Spearman C., 1904, *The American Journal of Psychology*, **15**, 72

Spitler L. R., et al., 2012, *ApJ*, **748**, L21

Srianand R., Hussain T., Noterdaeme P., Petitjean P., Krühler T., Japelj J., Pâris I., Kashikawa N., 2016, *MNRAS*, **460**, 634

Stach S. M., et al., 2021, *MNRAS*, **504**, 172

Stanway E. R., Eldridge J. J., 2018, *MNRAS*, **479**, 75

Steidel C. C., Sargent W. L. W., 1992, *ApJS*, **80**, 1

Steidel C. C., Gialavalisco M., Pettini M., Dickinson M., Adelberger K. L., 1996, *ApJ*, **462**, L17

Steidel C. C., Adelberger K. L., Dickinson M., Gialavalisco M., Pettini M., Kellogg M., 1998, *ApJ*, **492**, 428

Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., Gialavalisco M., 2000, *ApJ*, **532**, 170

Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., Gialavalisco M., 2003, *ApJ*, **592**, 728

Steidel C. C., Adelberger K. L., Shapley A. E., Erb D. K., Reddy N. A., Pettini M., 2005, *ApJ*, **626**, 44

Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, *ApJ*, **717**, 289

Steidel C. C., Bogosavljević M., Shapley A. E., Kollmeier J. A., Reddy N. A., Erb D. K., Pettini M., 2011, *ApJ*, **736**, 160

Steidel C. C., Bogosavljević M., Shapley A. E., Reddy N. A., Rudie G. C., Pettini M., Trainor R. F., Strom A. L., 2018, *ApJ*, **869**, 123

Steinhauser D., Schindler S., Springel V., 2016, *A&A*, **591**, A51

Stockton A. N., Lynds C. R., 1966, *ApJ*, **144**, 451

Stott J. P., et al., 2020, *MNRAS*, **497**, 3083

Sunyaev R. A., Zeldovich Y. B., 1972, *A&A*, **20**, 189

Szakacs R., et al., 2021, *MNRAS*, **505**, 4746

Taamoli S., et al., 2024, *ApJ*, **966**, 18

Tacconi L. J., et al., 2018, *ApJ*, **853**, 179

Tadaki K.-i., et al., 2019, *PASJ*, **71**, 40

Tanaka I., et al., 2011, *PASJ*, **63**, 415

Tejos N., et al., 2014, *MNRAS*, **437**, 2017

Thom C., Werk J. K., Tumlinson J., Prochaska J. X., Meiring J. D., Tripp T. M., Sembach K. R., 2011, *The Astrophysical Journal*, **736**, 1

Thomas D., Maraston C., Bender R., Mendes de Oliveira C., 2005, *ApJ*, **621**, 673

Tomczak A. R., et al., 2017, *MNRAS*, **472**, 3512

Tomczak A. R., et al., 2019, *MNRAS*, **484**, 4695

Topping M. W., Shapley A. E., Steidel C. C., 2016, *ApJ*, **824**, L11

Tornotti D., et al., 2024, *arXiv e-prints*, p. arXiv:2406.17035

Toshikawa J., et al., 2014, *ApJ*, **792**, 15

- Toshikawa J., Malkan M. A., Kashikawa N., Overzier R., Uchiyama H., Ota K., Ishikawa S., Ito K., 2020, *ApJ*, **888**, 89
- Toshikawa J., et al., 2024, *MNRAS*, **527**, 6276
- Tozzi P., et al., 2022a, *A&A*, **662**, A54
- Tozzi P., et al., 2022b, *A&A*, **667**, A134
- Trainor R. F., Steidel C. C., 2012, *ApJ*, **752**, 39
- Travascio A., et al., 2024, *arXiv e-prints*, p. [arXiv:2410.03933](#)
- Tremonti C. A., et al., 2004, *ApJ*, **613**, 898
- Tremonti C. A., Moustakas J., Diamond-Stanic A. M., 2007, *The Astrophysical Journal*, **663**, L77–L80
- Tsai S.-H., Chen K.-J., Smith A., Chiang Y.-K., 2024, *arXiv e-prints*, p. [arXiv:2410.00450](#)
- 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., Crain R. A., Rudie G., Steidel C. C., Strom A., Theuns T., 2017, *MNRAS*, **471**, 690
- Umehata H., et al., 2019, *Science*, **366**, 97
- Utsumi Y., Goto T., Kashikawa N., Miyazaki S., Komiyama Y., Furusawa H., Overzier R., 2010, *ApJ*, **721**, 1680
- Veilleux S., Cecil G., Bland-Hawthorn J., 2005, *ARA&A*, **43**, 769
- Venemans B. P., et al., 2007, *A&A*, **461**, 823
- Verhamme A., et al., 2018, *MNRAS*, **478**, L60
- Vernet J., et al., 2011, *A&A*, **536**, A105
- Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, *ApJ*, **640**, 691
- Vogt S. S., et al., 1994, in Crawford D. L., Craine E. R., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 2198, Instrumentation in Astronomy VIII. p. 362, [doi:10.1117/12.176725](#)
- Walter F., Weiß A., Downes D., Decarli R., Henkel C., 2011, *ApJ*, **730**, 18
- Walter F., et al., 2016, *ApJ*, **833**, 67
- Wang T., et al., 2016, *ApJ*, **828**, 56
- Wang T., et al., 2018, *ApJ*, **867**, L29
- Wang G. C. P., et al., 2024a, *arXiv e-prints*, p. [arXiv:2406.16637](#)
- Wang W., et al., 2024b, *arXiv e-prints*, p. [arXiv:2409.17956](#)
- Weaver J. R., et al., 2023, *A&A*, **677**, A184
- Weilbacher P. M., Streicher O., Urrutia T., Pécontal-Rousset A., Jarno A., Bacon R., 2014, in Manset N., Forshay P., eds, Astronomical Society of the Pacific Conference Series Vol. 485, Astronomical Data Analysis Software and Systems XXIII. p. 451 ([arXiv:1507.00034](#)), [doi:10.48550/arXiv.1507.00034](#)
- Weiner B. J., et al., 2009, *The Astrophysical Journal*, **692**, 187–211
- Weiß A., Downes D., Henkel C., Walter F., 2005, *A&A*, **429**, L25
- Weisskopf M. C., Tananbaum H. D., Van Speybroeck L. P., O’Dell S. L., 2000, in Truemper J. E., Aschenbach B., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4012, X-Ray Optics, Instruments, and Missions III. pp 2–16 ([arXiv:astro-ph/0004127](#)), [doi:10.1117/12.391545](#)
- Welsh L., Cooke R., Fumagalli M., 2019, *MNRAS*, **487**, 3363
- Weng S., Péroux C., Ramesh R., Nelson D., Sadler E. M., Zwaan M., Bollo V., Casavecchia B., 2024, *MNRAS*, **527**, 3494
- Werk J. K., Prochaska J. X., Thom C., Tumlinson J., Tripp T. M., O’Meara J. M., Meiring J. D., 2012, *ApJS*, **198**, 3
- 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
- Werner M. W., et al., 2004, *ApJS*, **154**, 1
- White S. D. M., Frenk C. S., 1991, *ApJ*, **379**, 52
- White S. D. M., Rees M. J., 1978, *MNRAS*, **183**, 341
- Wilkinson A., et al., 2017, *MNRAS*, **464**, 1380
- Wise J. H., Turk M. J., Norman M. L., Abel T., 2012, *ApJ*, **745**, 50
- Wisotzki L., et al., 2016, *A&A*, **587**, A98

Wolfe A. M., Gawiser E., Prochaska J. X., 2005, [ARA&A](#), **43**, 861
 Worseck G., et al., 2014, [MNRAS](#), **445**, 1745
 Wu Y., et al., 2023, [ApJ](#), **956**, L40
 Wylezalek D., et al., 2013, [ApJ](#), **769**, 79
 Xu X., et al., 2023, CLASSY VI: Density, Structure and Size of Galactic Outflows ([arXiv:2301.11498](#))
 Yamada T., Nakamura Y., Matsuda Y., Hayashino T., Yamauchi R., Morimoto N., Kousai K., Umemura M., 2012, [AJ](#), **143**, 79
 Yang Y., Zabludoff A., Tremonti C., Eisenstein D., Davé R., 2009, [ApJ](#), **693**, 1579
 Yang G., et al., 2020, [MNRAS](#), **491**, 740
 Yang G., et al., 2022, [ApJ](#), **927**, 192
 Zabl J., et al., 2019, [MNRAS](#), **485**, 1961
 Zel'dovich Y. B., 1970, [A&A](#), **5**, 84
 Zeng X., Ao Y., Zhang Y., 2024, [MNRAS](#), **528**, 2964
 Zewdie D., et al., 2023, [A&A](#), **677**, A54
 Zhang C., et al., 2021, [ApJ](#), **911**, 57
 Zhang Y. H., et al., 2024, [arXiv e-prints](#), p. [arXiv:2410.10169](#)
 Zheng Z., Miralda-Escudé J., 2002, [ApJ](#), **578**, 33
 Zhu G., Ménard B., 2013, [ApJ](#), **770**, 130
 Zwicky F., 1933, *Helvetica Physica Acta*, **6**, 110
 de Beer S., et al., 2023, [MNRAS](#), **526**, 1850
 de La Vieuville G., et al., 2019, [A&A](#), **628**, A3
 de Lapparent V., Geller M. J., Huchra J. P., 1986, [ApJ](#), **302**, L1
 de Vos K., Merrifield M. R., Hatch N. A., 2024, [MNRAS](#), **531**, 4383
 van Dokkum P. G., 2001, [PASP](#), **113**, 1420
 van der Burg R. F. J., Hildebrandt H., Erben T., 2010, [A&A](#), **523**, A74
 van der Burg R. F. J., et al., 2020, [A&A](#), **638**, A112

Appendix to chapter 2

A.1 Theoretical comparison between C iv and Si iv absorbers as tracers of the ionized gas

We complemented the comparison between C iv and Si iv absorbers shown in Section 2.4.6 with a more theoretical approach: indeed, one way to test the assumption that C iv and Si iv absorbers trace the same gas phase is to examine whether absorbers arising at the same redshift are composed of the same number of components that are aligned in redshift. By searching for Si iv lines detected within $|\Delta v| \leq 150 \text{ km s}^{-1}$ from a C iv absorber, we identify 63 matched systems (corresponding to the $\approx 29\%$ of the full C iv sample, and $\approx 58\%$ of the Si iv systems). We model the absorption of these profiles by fitting the lines with the MC-ALF code, following the method presented in Rudie et al. (2019) to identify the number of aligned components. Based on the alignment of each component, we classify the matched systems into three classes (see Figure A.1 for examples): i) systems that are clearly misaligned, suggesting that C iv and Si iv absorbers are not tracers of the same gas structure. This class contains $\approx 10\%$ (6/63) of the matched systems; ii) systems in which the strong components are offset or only partially aligned, i.e. separated by $\Delta v \gtrsim 10 - 20 \text{ km s}^{-1}$ in velocity space ($\approx 13\%$, 8/63, of the matched systems belong to this class, plus the 157 C iv systems that do not show a Si iv counterpart); iii) systems with the same number of components that are clearly aligned (this class contains $\approx 77\%$ (49/63) of the matched systems). The rate of matches we observe is qualitatively in line with the recent statistical measures from D’Odorico et al. (2022). With a sample of over 600 Si iv absorbers detected in 147 quasar sightlines at redshift $2.1 \lesssim z_{\text{QSO}} \lesssim 6.5$, they studied the column density distribution function and found that the abundance of C iv and Si iv ions trace each other across a large redshift range. However, the Si iv number density per unit redshift path is a factor ≈ 2 lower than the C iv absorbers at any redshift. This difference implies a fraction of $\lesssim 50\%$ of the C iv systems with a detected Si iv counterpart, as we observe. Having identified the aligned components in each transition, we measure the rest-frame equivalent width for class 2 and 3 systems and recover how the Si iv equivalent width, $W_{\text{Si iv}}$, scales as a function of the equivalent width of the matched C iv absorber, $W_{\text{C iv}}$. This estimate allow us to better understand the nature of the class 2 systems and test whether the partial alignment or the offset of the strong components is a sign that systems are not tracing the same gas structures. To complete the picture, we also compare the observed trend with the theoretical predictions from a set of ionization models using CLOUDY (Ferland et al., 2013). The results from the observations are shown in Figure A.2, where the shaded region and its limits are derived from CLOUDY using the “minimal grid” from Fumagalli et al. (2016). Assuming the metallicity to be fixed at $\log(Z/Z_{\odot}) = -2.0$, we show the mean equivalent width ratio over the redshift range $z = 3.0 - 4.5$ as a function of the neutral hydrogen column density and the density in the left-panel of Figure A.3.

As it can be seen from this grid, the equivalent width ratio is $\log(W_{\text{Si iv}}/W_{\text{C iv}}) \approx -1.0$ at densities $\log(n_{\text{H}}/\text{cm}^{-3}) \lesssim -4$ (for reference, the mean cosmic density at redshift $z \sim 3$ is $\log(n_{\text{H}}/\text{cm}^{-3}) \approx -5.0$), increases to $\log(W_{\text{Si iv}}/W_{\text{C iv}}) \approx 0$ at $-4 \lesssim \log(n_{\text{H}}/\text{cm}^{-3}) \lesssim -2.5$ and about $\log(W_{\text{Si iv}}/W_{\text{C iv}}) \approx 0.5$ at higher densities $\log(n_{\text{H}}/\text{cm}^{-3}) \gtrsim -2.5$. Thus, for common CGM/IGM densities of $\log(n_{\text{H}}/\text{cm}^{-3}) \lesssim -2.5$, it is reasonable to expect that only a fraction of C iv absorbers is detected also in Si iv, which will typically appear as weaker line. Indeed, the

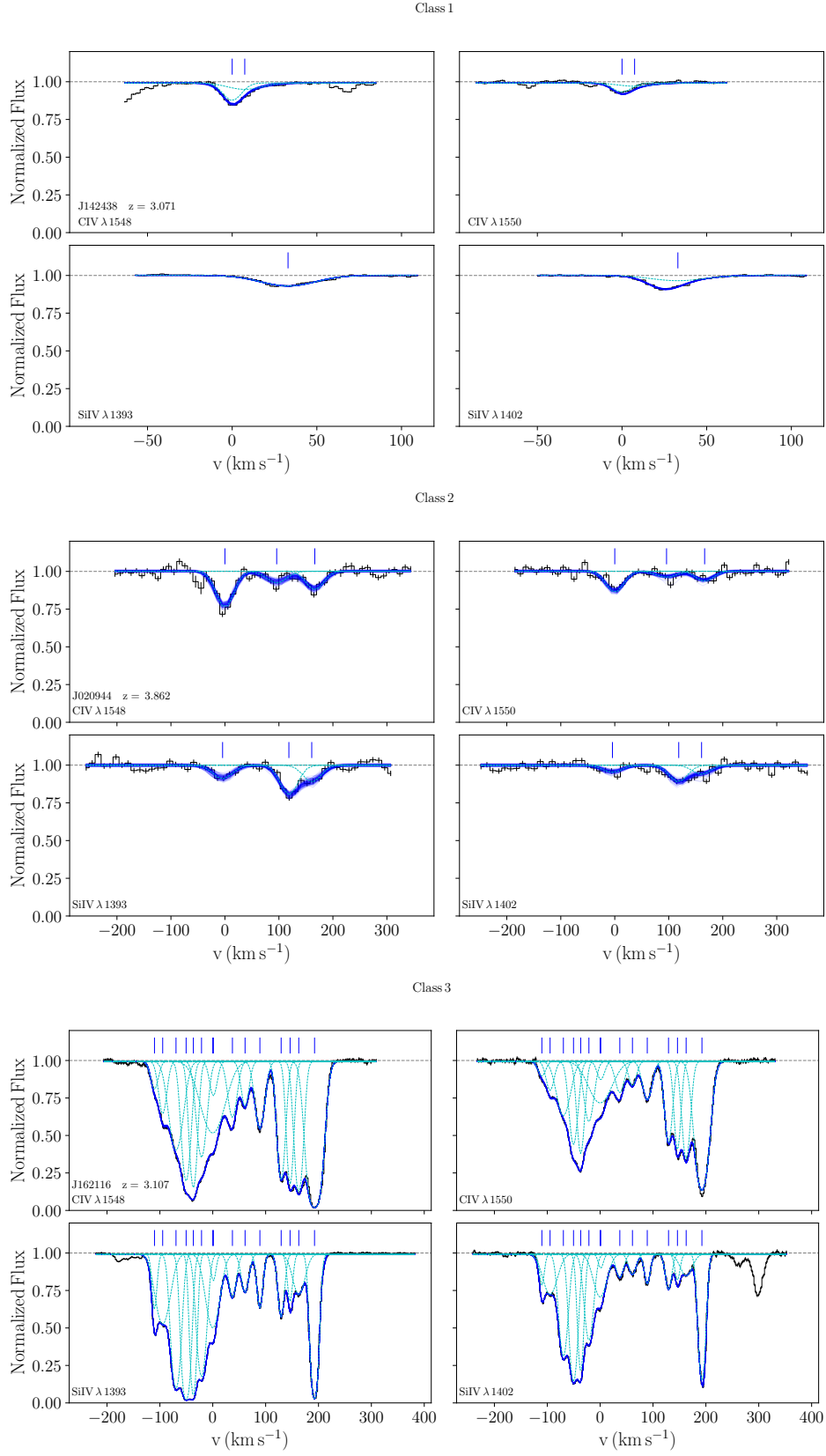


Figure A.1: Examples of CIV and SiIV systems matched within ± 150 km s⁻¹ in velocity space and classified according to the number and the alignment of Voigt components (class 1-3 from top to bottom, see text for details).

right panel of Figure A.3, which shows the distribution of the C IV-Si IV detections as a function of the C IV rest frame equivalent width, provides clear evidence that detection of Si IV counterparts decline with decreasing strength of the C IV absorbers, significantly dropping close to the sensitivity limits of our spectra at $W_{\text{Si IV}} \lesssim 10^{-2} \text{ \AA}$.

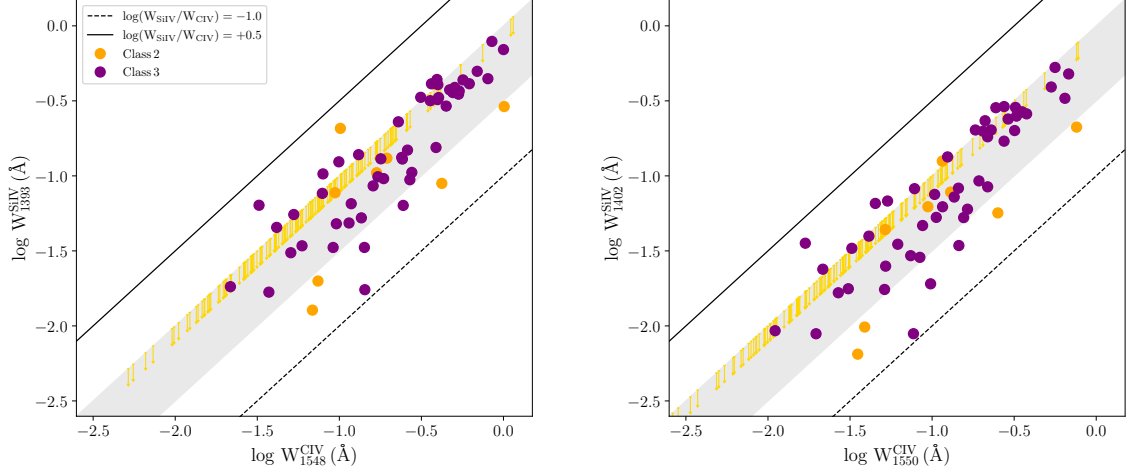


Figure A.2: Comparison between the rest-frame equivalent width of Si IV absorbers, $W_{\text{Si IV}}$, as a function of the equivalent width of the matched C IV system, $W_{\text{C IV}}$. Aligned absorbers from class 3 (purple) are distinguished from class 2 systems (orange). The predictions from CLOUDY ionization models (see text for details) are shown as solid and dashed lines, marking $\log(W_{\text{Si IV}}/W_{\text{C IV}}) = +0.5$ and $\log(W_{\text{Si IV}}/W_{\text{C IV}}) = -1.0$ respectively. The shaded region highlights the range occupied by the $\approx 2/3$ of the distribution of aligned detections, corresponding to $-0.5 < \log(W_{\text{Si IV}}/W_{\text{C IV}}) < 0$. Arrows show upper limits corresponding to $\log W_{\text{Si IV}} \leq \log W_{\text{C IV}}$ for Si IV non detections.

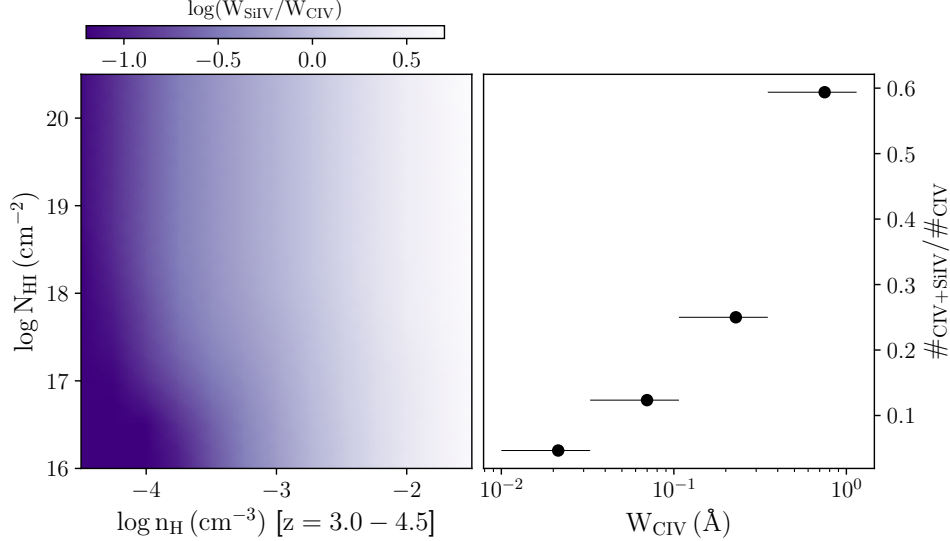


Figure A.3: Left panel: mean equivalent width ratio over the redshift range $z = 3.0 - 4.5$ as predicted from a grid of ionization models in the $\log N_{\text{HI}} - \log n_{\text{H}}$ plane. The metallicity is fixed to $\log(Z/Z_{\odot}) = -2.0$. Right panel: distribution of the matching C IV-Si IV detections, normalized to the number of C IV systems in each W_r interval, as a function of the C IV rest frame equivalent width. Horizontal error-bars account for the width of each bin.

Figure A.2 provides evidence that our observations of the aligned components are consistent with the ionization models predictions since the class 3 systems are scattered between $-1.0 \lesssim \log(W_{\text{Si IV}}/W_{\text{C IV}}) \lesssim 0.5$. We observe most of the matched systems at $\log W_{\text{Si IV}} \lesssim \log W_{\text{C IV}}$, with only few Si IV absorbers stronger than their C IV counterpart. About 38% (3/8) of the partially-aligned systems (class 2) follow the trend observed for the aligned systems, while the remaining sample is scattered toward a lower ratio $\log(W_{\text{Si IV}}/W_{\text{C IV}}) \approx -1.0$ pinpointing to the presence of

C IV absorption that is significantly stronger or with a higher number of components compared to the Si IV counterpart. We show upper limits (gold arrows in Figure A.2) for the C IV absorbers without Si IV counterpart within $\pm 150 \text{ km s}^{-1}$, assuming $\log W_{\text{Si IV}} \leq \log W_{\text{C IV}}$. As expected from the analysis above, we observe that a large fraction of non-detections corresponds to weak C IV systems with $\log(W_{\text{C IV}}/\text{\AA}) \lesssim -1.5$, with only few strong absorbers ($\log(W_{\text{C IV}}/\text{\AA}) \gtrsim -0.5$) lacking associations. The last result supports the hypothesis that weak C IV systems would be associated to aligned Si IV absorbers below the detection limit of the survey. In the end, this analysis supports the idea that C IV and Si IV absorbers are, to first approximation, two tracers of the same gas phase and are expected to lead to comparable results if used to trace an intermediate ionized phase of the CGM and IGM near LAEs.

A.2 Kinematics of C iv and Mg ii absorbers

As detailed in Section 3.3.3, we found 28 Mg II absorbers matching C IV lines from MAGG V in velocity space within $\pm 150 \text{ km s}^{-1}$. Among these, we identify 19 cases in which the Mg II and C IV absorption line profiles are matched in velocity space (see an example in the upper panel in Figure A.4), while the remaining systems are mis-matched and only partially aligned (lower panel in Figure A.4). The former systems are likely arising from the same gas clouds and are thus direct evidence of the multiphase nature of the CGM of Ly α emitting galaxies. Rudie et al. (2019) probed the gas clouds in star-forming galaxies' CGM by tracing a variety of absorbers, from low to high ions, at $z \approx 3$. They observed different kinematics for the elements in different ionization stages: indeed, O VI and C IV systems show broader absorption features compared to singly, doubly and triply ionized species.

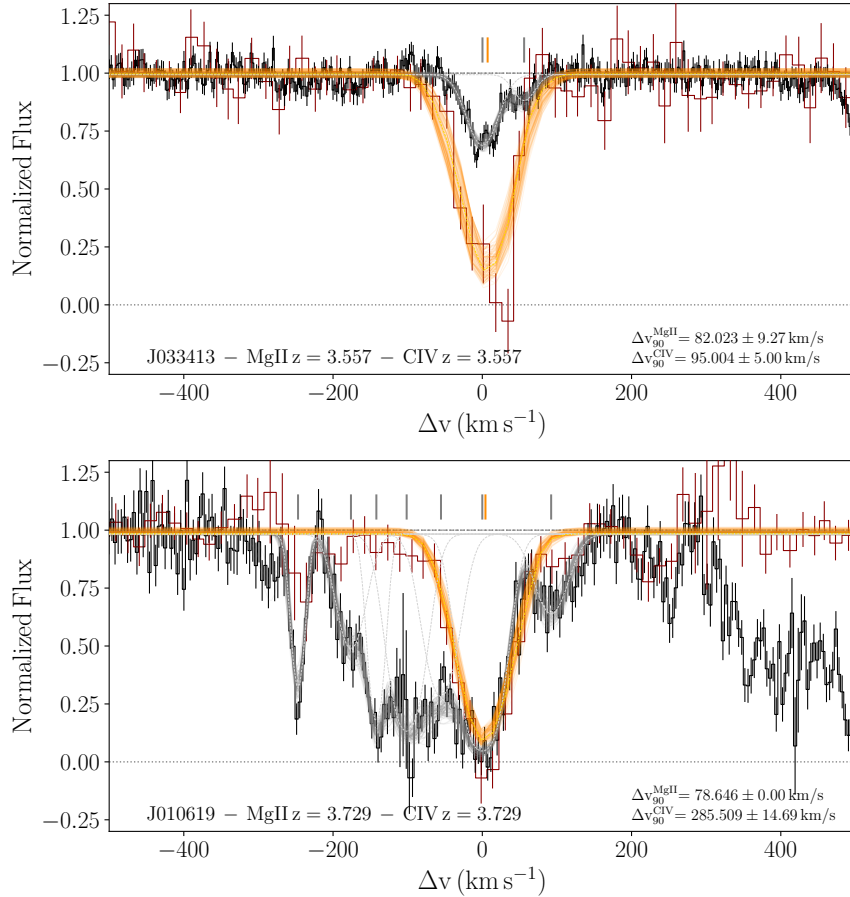


Figure A.4: Examples of Mg II (dark-red spectrum and orange best-fit profile) and C IV (black spectrum and gray profile) absorbers matched in velocity space. The redshift of the C IV component with the largest column density corresponds to $\Delta v = 0 \text{ km s}^{-1}$. We show in the upper panel the case $\Delta v_{90}^{\text{C IV}} \approx \Delta v_{90}^{\text{Mg II}}$ (matched system) and in the lower panel the case $\Delta v_{90}^{\text{C IV}} > 2 \times \Delta v_{90}^{\text{Mg II}}$ with only a partial alignment between the two ions (mis-matched system). The sightline, redshifts and measure of the velocity width are reported in the figures.

These aligned systems identified in MAGG allow us to extend this analysis at $z \gtrsim 3$. However, the different resolution of the optical and NIR quasar spectra in which we identified the C IV and Mg II absorbers, respectively, prevents from directly compare the individual Voigt components. We thus show in Figure A.5 an alternative tracer of the gas kinematics, that is, the width of the lines in velocity space, for the Mg II absorbers (x-axis) relative to that of the matching C IV lines (y-axis). We find that a large fraction of the data-points are scattered around the one-to-one relation, corresponding to $\Delta v_{90}^{\text{C IV}} = \Delta v_{90}^{\text{Mg II}}$. A few systems show kinematics that is consistent within $1\sigma = \pm 10 \text{ km s}^{-1}$, the typical uncertainty on the measurements of the velocity width for the two ions. In more details, $\approx 71\%$ ($\approx 58\%$) of the C IV systems (matched C IV systems) show, on average, a ≈ 1.2 times larger Δv_{90} and is scattered around $\Delta v_{90}^{\text{C IV}} = \Delta v_{90}^{\text{Mg II}} \pm 2.5\sigma$ (see an example in the upper panel in Figure A.4). About 30 percent of the systems populates the upper-left corner of Figure A.5, corresponding to the region of $\Delta v_{90}^{\text{C IV}} > 2 \times \Delta v_{90}^{\text{Mg II}}$. In these cases the Mg II absorption profile significantly differs from that of the C IV and is partially aligned (see the lower panel in Figure A.4). Indeed, the mild difference in the kinematics is washed out when only the matched systems are taken into account. In the inset in Figure A.5 we show a complementary tracer of the kinematics of the two ions: the velocity shift between the peak of the Mg II optical depth relative to that of the C IV. The distribution (black histogram for the aligned Mg II and C IV systems) is centered around a median value of $v_{\text{peak}}^{\text{Mg II}} - v_{\text{peak}}^{\text{C IV}} \approx 2.7 \text{ km s}^{-1}$, which is indicative of only a mild difference between the Mg II and C IV kinematics and consistent with the analysis above. Including also the partially aligned systems (lightgray histogram) results in a median value $v_{\text{peak}}^{\text{Mg II}} - v_{\text{peak}}^{\text{C IV}} \approx 5.4 \text{ km s}^{-1}$ and does not change the outcome. In conclusion, $\approx 58\%$ of the Mg II-C IV systems matched along the line-of-sight shows a velocity width that is consistent within $\approx 2.5\sigma$. While in mis-matched systems we typically observe $\Delta v_{90}^{\text{C IV}} \gtrsim 2 \times \Delta v_{90}^{\text{Mg II}}$, we do not find any Mg II absorber significantly broader than the corresponding C IV line.

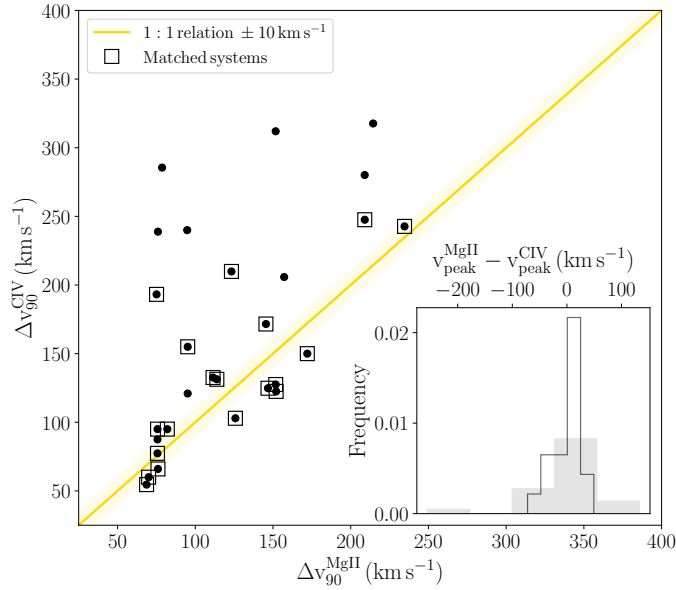


Figure A.5: Comparison between the velocity width of the Mg II (x-axis) and the C IV (y-axis) systems found to be matched within $\pm 150 \text{ km s}^{-1}$ along the line of sight. We show the $\Delta v_{90}^{\text{Mg II}} = \Delta v_{90}^{\text{C IV}}$ relation (gold line) with the shaded area marking the typical 1σ uncertainty on the kinematics. Open squares highlight the Mg II and C IV systems whose absorption line profiles are matched in velocity space. The inset shows the velocity shift between the peak of the Mg II optical depth relative to the C IV for the full sample (lightgray) and restricted to only the matched systems (black).

Appendix to chapter 3

B.1 Kinematics of C iv and Mg ii absorbers

As detailed in Section 3.3.3, we found 28 Mg II absorbers matching C IV lines from MAGG V in velocity space within $\pm 150 \text{ km s}^{-1}$. Among these, we identify 19 cases in which the Mg II and C IV absorption line profiles is matched in velocity space (see an example in the upper panel in Figure B.1), while the remaining systems are mis-matched and only partially aligned (lower panel in Figure B.1). The former systems are likely arising from the same gas clouds and are thus direct evidence of the multiphase nature of the CGM of Ly α emitting galaxies. Rudie et al. (2019) probed the gas clouds in star-forming galaxies' CGM by tracing a variety absorbers, from low to high ions, at $z \approx 3$. They observed different kinematics for the elements in different ionization stages: indeed, O VI and C IV systems show broader absorption features compared to singly, doubly and triply ionized species.

These aligned systems identified in MAGG allow us to extend this analysis at $z \gtrsim 3$. However, the different resolution of the optical and NIR quasar spectra in which we identified the C IV and Mg II absorbers, respectively, prevents from directly compare the individual Voigt components. We thus show in Figure B.2 an alternative tracer of the gas kinematics, that is, the width of the lines in velocity space, for the Mg II absorbers (x-axis) relative to that of the matching C IV lines (y-axis). We find that a large fraction of the data-points are scattered around the one-to-one relation, corresponding to $\Delta v_{90}^{\text{C IV}} = \Delta v_{90}^{\text{Mg II}}$. A few systems show kinematics that is consistent within $1\sigma = \pm 10 \text{ km s}^{-1}$, the typical uncertainty on the measurements of the velocity width for the two ions. In more details, $\approx 71\%$ ($\approx 58\%$) of the C IV systems (matched C IV systems) show, on average, a ≈ 1.2 times larger Δv_{90} and is scattered around $\Delta v_{90}^{\text{C IV}} = \Delta v_{90}^{\text{Mg II}} \pm 2.5\sigma$ (see an example in the upper panel in Figure B.1). About 30 percent of the systems populates the upper-left corner of Figure B.2, corresponding to the region of $\Delta v_{90}^{\text{C IV}} > 2 \times \Delta v_{90}^{\text{Mg II}}$. In these cases the Mg II absorption profile significantly differs from that of the C IV and is partially aligned (see the lower panel in Figure B.1). Indeed, the mild difference in the kinematics is washed out when only the matched systems are taken into account. In the inset in Figure B.2 we show a complementary tracer of the kinematics of the two ions: the velocity shift between the peak of the Mg II optical depth relative to that of the C IV. The distribution (black histogram for the aligned Mg II and C IV systems) is centered around a median value of $v_{\text{peak}}^{\text{Mg II}} - v_{\text{peak}}^{\text{C IV}} \approx 2.7 \text{ km s}^{-1}$, which is indicative of only a mild difference between the Mg II and C IV kinematics and consistent with the analysis above. Including also the partially aligned systems (lightgray histogram) results in a median value $v_{\text{peak}}^{\text{Mg II}} - v_{\text{peak}}^{\text{C IV}} \approx 5.4 \text{ km s}^{-1}$ and does not change the outcome. In conclusion, $\approx 58\%$ of the Mg II-C IV systems matched along the line-of-sight shows a velocity width that is consistent within $\approx 2.5\sigma$. While in mis-matched systems we typically observe $\Delta v_{90}^{\text{C IV}} \gtrsim 2 \times \Delta v_{90}^{\text{Mg II}}$, we do not find any Mg II absorber significantly broader than the corresponding C IV line.

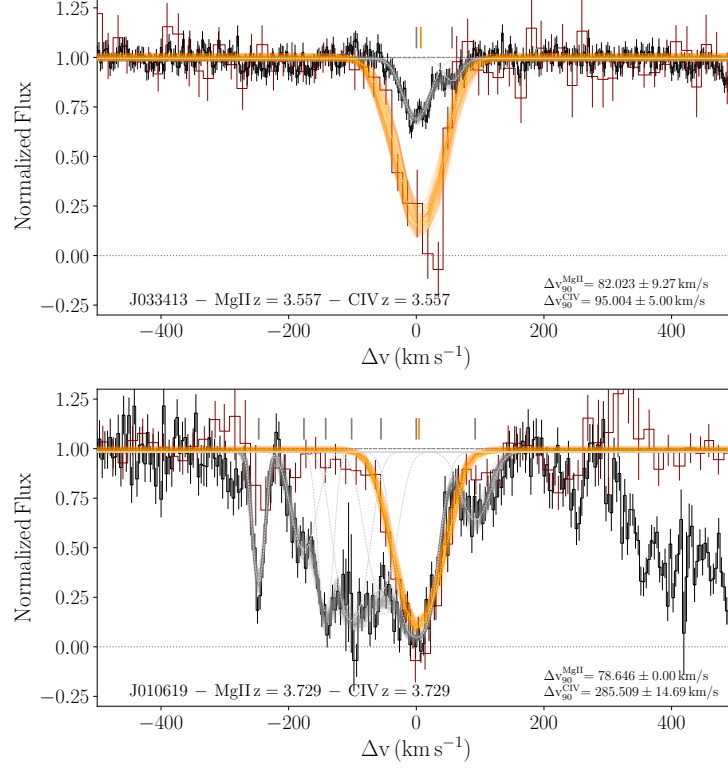


Figure B.1: Examples of Mg II (dark-red spectrum and orange best-fit profile) and C IV (black spectrum and gray profile) absorbers matched in velocity space. The redshift of the C IV component with the largest column density corresponds to $\Delta v = 0$ km s⁻¹. We show in the upper panel the case $\Delta v_{90}^{\text{C IV}} \approx \Delta v_{90}^{\text{Mg II}}$ (matched system) and in the lower panel the case $\Delta v_{90}^{\text{C IV}} > 2 \times \Delta v_{90}^{\text{Mg II}}$ with only a partial alignment between the two ions (mis-matched system). The sightline, redshifts and measure of the velocity width are reported in the figures.

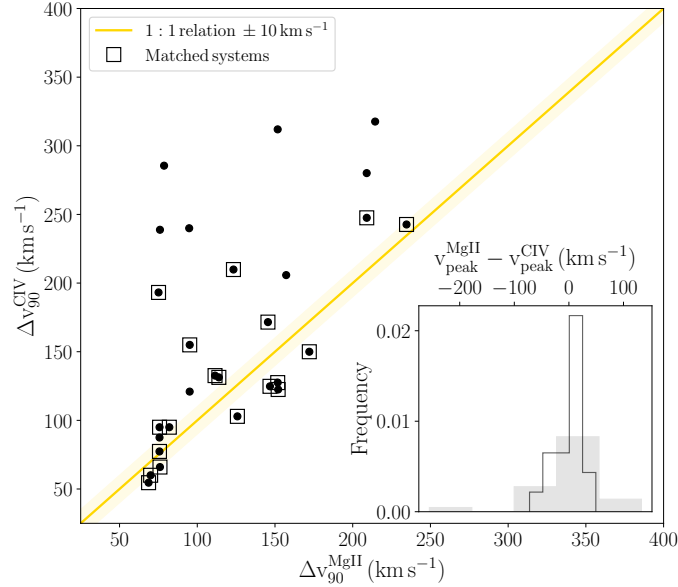


Figure B.2: Comparison between the velocity width of the Mg II (x-axis) and the C IV (y-axis) systems found to be matched within ± 150 km s⁻¹ along the line of sight. We show the $\Delta v_{90}^{\text{Mg II}} = \Delta v_{90}^{\text{C IV}}$ relation (gold line) with the shaded area marking the typical 1 σ uncertainty on the kinematics. Open squares highlight the Mg II and C IV systems whose absorption line profiles are matched in velocity space. The inset shows the velocity shift between the peak of the Mg II optical depth relative to the C IV for the full sample (lightgray) and restricted to only the matched systems (black).

Appendix to chapter 4

C.1 SED of the ALMA–selected galaxies in the MUDF field

We derived properties of our ALMA-selected galaxies with secure photometric measurements (i.e., excluding C2 that is QSO-SE, and C5; see Sect. 4.4.4), such as M_* , SFR, and visual extinction (A_V) by performing simultaneous fit of multiwavelength photometry and MUSE spectra. The photometry includes custom MUSE top-hat filters, covering 4750 Å to 9300 Å (as described in table 2 in Fossati et al. 2019), multiple UV to NIR HST filters (WFC3/F140W, F125W, and WFC2/F450W, F702W, see Revalski et al. 2023), VLT/HAWK-I Ks filter (Revalski et al., 2024), and 1.2-mm flux from ALMA band 6 observations.

We used the Monte Carlo Spectro-Photometric Fitter (MC-SPF, Fossati et al., 2018a, 2019) code. We built a grid of stellar spectra based on the updated version of the Bruzual & Charlot (2003b) high-resolution models at 40% solar metallicity, assuming a Chabrier (2003) initial mass function and a delayed exponentially declining star formation history¹. We scaled a grid of nebular emission lines models from Byler et al. (2018) to the number of Lyman continuum photons from the stellar models, converted it into flux at the redshift of the galaxies and then added to templates. To include the effect of dust attenuation on both the stellar continuum and the nebular emission lines, we assumed the double attenuation law from Calzetti et al. (2000), which accounts for extra extinction of the young stars compared to those older than 10 Myr. During the fitting procedure, we impose an energy balance between the UV flux from young stars that is absorbed by dust grains and re-emitted at FIR wavelengths. We used the dust emission model from Dale et al. (2014). We report best-fit results from SED modeling in Fig. C.1 and Table C.1.

ID _{1.2mm} ⁽¹⁾	$\log M_*$ ⁽²⁾ ($M_\odot$)	$\log \text{SFR}$ ⁽³⁾ ($M_\odot \text{ yr}^{-1}$)	A_V ⁽⁴⁾ (mag)
C1	10.7 ± 0.2	1.8 ± 0.3	1.86 ± 0.18
C3	10.8 ± 0.2	1.8 ± 0.3	1.77 ± 0.08
C4	10.9 ± 0.2	2.0 ± 0.3	1.99 ± 0.06
C6	10.6 ± 0.2	1.9 ± 0.3	3.25 ± 0.14
C7	10.2 ± 0.2	1.2 ± 0.3	0.58 ± 0.02

Table C.1: Results from SED fitting of the ALMA-selected galaxies. ⁽¹⁾Identifier of ALMA 1.2-mm continuum-selected galaxies. ⁽²⁾Logarithm of the stellar mass. ⁽³⁾Logarithm of the star-formation rate. ⁽⁴⁾Visual extinction. Statistical uncertainties on measurements are estimated from the 16th and 84th percentiles of the parameter posterior probability distribution. Uncertainties on M_* and SFR are reported after adding in quadrature 0.3 and 0.2 dex, respectively, to account for systematics (see, Pacifici et al., 2023; Beckett et al., 2024).

¹The set of models is publicly available at <http://www.bruzual.org/CB19/>. A detailed description is provided in appendix A of Sánchez et al. (2022).

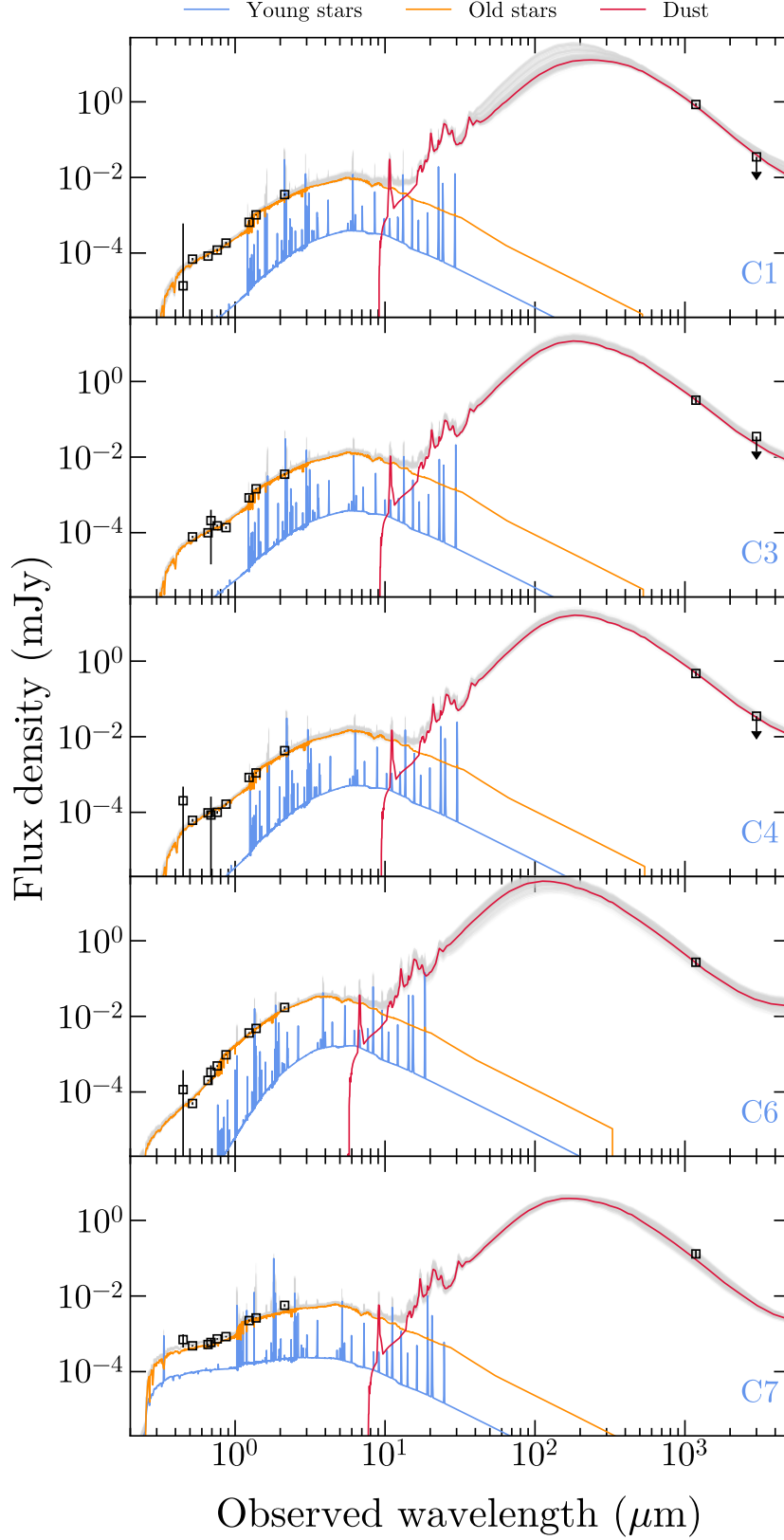


Figure C.1: Best fit SED models of ALMA-selected galaxies. The black squares report the multiband photometric data. Flux upper limits at $5\text{-}\sigma$ for C1, C3, and C4, obtained from 3-mm continuum images are also reported. Red, blue, and orange lines represent the dust emission, and the young (star age < 10 Myr) and old (> 10 Myr) stellar components, respectively. The gray lines show random extraction from the posterior probability distribution of the total galaxy SED.

Appendix to chapter 5

D.1 Custom G-band and UGR selection

As described in Section 5.3.1, we measured $U - B$ and $B - R$ colors to select Lyman break galaxies within the redshift range $z \approx 3 - 3.5$ across a much larger FoV compared to that of the MUSE mosaic. In order to be able to compare our results with other sample in the literature, it is essential to keep the selection function between different studies as similar as possible. Taking advantage of the MUSE data, we restrict to the $4 \times 4 \text{ cMpc}^2$ FoV of the MUSE mosaic and identify the LBG candidates using a UGR (that is, based on $U - G$ and $G - R$ colors) color selection (Steidel et al., 1996, 2003). To this end, we construct a custom-made G-band image covering the FoV of the MUSE mosaic. We therefore combine the B-band image taken with FORS2 with a top-hat medium-band image built out of the MUSE data, producing a custom G-band image with effective wavelength $\lambda_{\text{eff}} = 4727.3 \text{ \AA}$. The MUSE top-hat filter spans a wavelength range from $4920 \text{ \AA}$, corresponding to a transmission of 5% for the FORS2 B_{High} filter, to $5510 \text{ \AA}$, which corresponds to the maximum wavelength of the FORS2 G_{High} filter. To calibrate the color selection, we used the same SPS models described in Section 5.3.1 convolved for the transmission functions of the FORS2 filters and of the custom G-band. As a result, the selection criterion reads as following:

1. $U > U_{\text{BKG}}^{3\sigma}$
2. $U - G > 1.22$
3. $G - R < 0.88 \times (U - G) - 0.81$
4. $0.21 < G - R < 1.75$

Finally, the results of this alternative color selection are shown in Figure 5.3, while the success rate for the sample of spectroscopically-selected galaxies clustered around the quasar is reported in Table 5.2. In details, within the $\approx 4 \times 4 \text{ cMpc}^2$ FoV of the MUSE mosaic, we are able to select 33 candidate Lyman break galaxies at $z \approx 3.0 - 3.5$, 23 of them also selected by the UBR criterion.

D.2 The UV luminosity function of LBGs

We here estimate the overdensity of UBR color-selected LBGs by applying the same method described in Section 5.4.2 and thus deriving the rest-frame UV cumulative luminosity functions in the FoV of MUSE and FORS2 (see Figure D.2). These are corrected for the completeness of the R-band, that is used for the detection of the galaxies (see Figure 5.1). The highest normalization is observed for the LBGs with spectroscopic redshift within $|\Delta v| \leq 1000 \text{ km s}^{-1}$ around the quasar in the FoV of the MUSE mosaic. This sample mimics a selection based on photometry and a spectroscopic follow-up, but fails to select massive and dusty galaxies with $\log(M_*/M_\odot) \gtrsim 10$ and $A_V \gtrsim 0.5$ (as reported in Section 5.3.1). In this case, the overdensity estimated relative to the completeness corrected luminosity function from Reddy & Steidel (2009), is $\rho/\bar{\rho} = 24 \pm 14$ at $M_R \leq -19.25 \text{ mag}$ and thus consistent within the uncertainties with the results we obtained for the main spectroscopic sample. An overdensity $\rho/\bar{\rho} = 3 \pm 1$ is observed even when considering a $\times 6$ larger volume in redshift. However, the luminosity function becomes fully consistent with Reddy

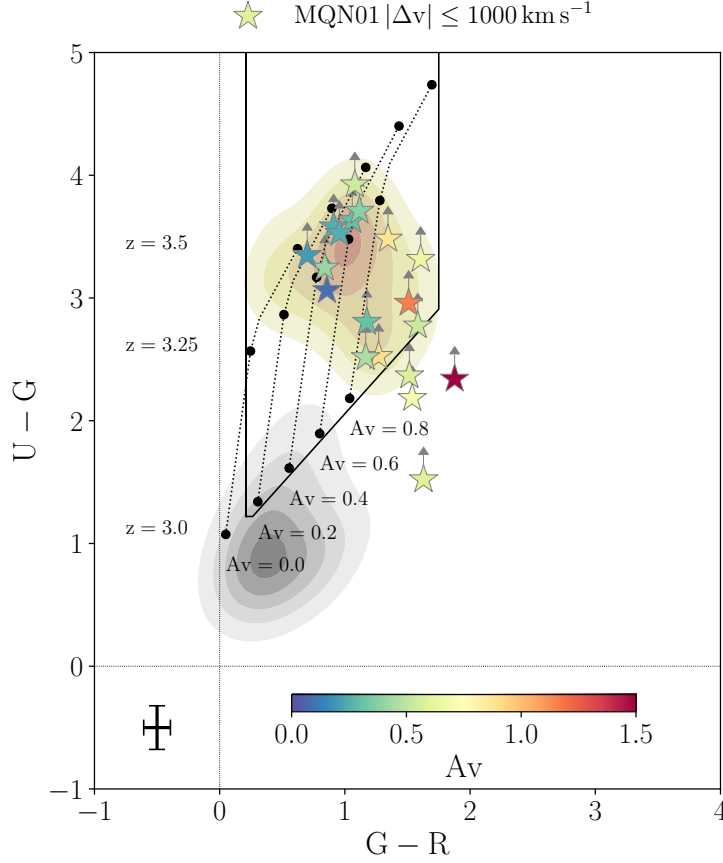


Figure D.1: Same as Figure 5.3, but showing the results of the UGR color selection of LBGs.

& Steidel (2009) once computed for all the LBGs detected across the $\times 36$ larger area sampled by FORS2, suggesting that the overdensity is diluted in such a large volume. This provides the evidence that the 3D distribution of the star-forming galaxies is significantly peaked not only along the line of sight but also spatially around the quasar.

D.3 Alignment of the LBGs with the large-scale overdensity

Assuming the LBGs span the entire redshift range $3.0 < z < 3.5$ (see Section 5.3.1), we measured the local co-moving galaxy density by using the N_{th} neighbor estimator with the same formalism presented by Shimakawa et al. (2018a). We choose $N=5$. Although we expect this method not to be efficient at the edges of the FoV, these galaxies should not affect significantly the statistics and are mostly limited to $R \gtrsim 11$ cMpc. Furthermore, we validate the choice of the estimator by comparing the N_{th} neighbor results to the co-moving density obtained by the scale independent the Voronoi Tessellation, described in Section 5.4.6. We observe that the 3rd neighbor systematically over-estimate the density by a factor of ≈ 1.5 compared to the Voronoi Tessellation, but the two methods turn out to be consistent for $N \geq 5$. The local density is finally converted into a co-moving overdensity $\rho/\bar{\rho}$ against the expected density of galaxies in a blank field (see Section 5.4.2).

We then measured the orientation of each galaxy with respect to the North-South direction and the central quasar. In bins of co-moving overdensity defined to contain roughly the same number of galaxies, we measured the fraction of LBGs oriented along the North-South direction around the quasar as a function of the environment (see Figure D.3). We require a galaxy for being aligned to the North-South direction to be offset by an angle within $\pm 25^\circ$ (black, corresponding to $0^\circ - 25^\circ$ or $155^\circ - 180^\circ$) or $\pm 45^\circ$ (magenta, corresponding to $0^\circ - 45^\circ$ or $135^\circ - 180^\circ$), chosen so that $\gtrsim 40\%$ and 100% , respectively, of the MUSE spectroscopic sample is found to be aligned.

Figure D.3 shows that the fraction of LBGs oriented along the North-South direction steeply increases with the local co-moving density. In other words, the regions that overdense in LBGs are not sparsely distributed across the large area we surveyed, but instead do preferentially follow the same orientation as the spectroscopically selected galaxies within 4×4 cMpc² and $\Delta v = \pm 1000$ km s⁻¹

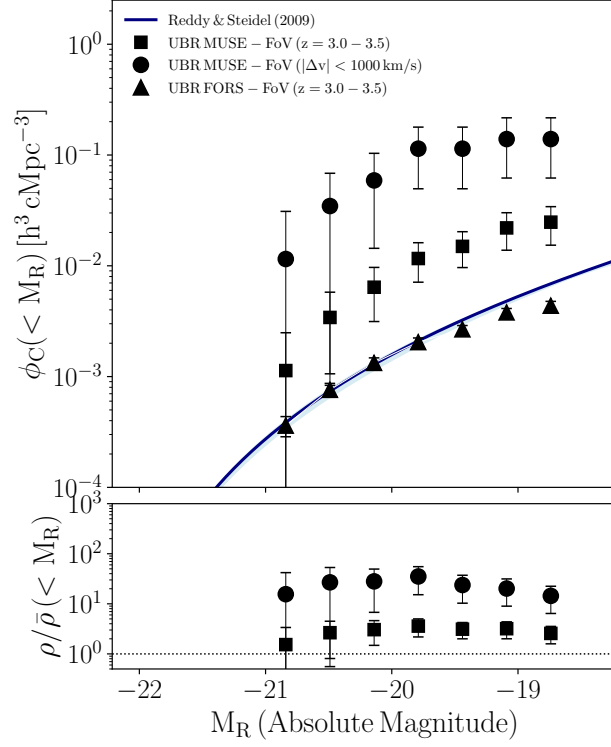


Figure D.2: Rest-frame UV luminosity function for the sample of *UBR* color-selected LBGs. Upper panel: cumulative luminosity function of the LBGs identified within the FoV of MUSE (squares) and FORS2 (triangles), and of those galaxies with spectroscopic redshift within $|\Delta v| \leq 1000 \text{ km s}^{-1}$ around the central quasar (points). The LBGs' luminosity functions are corrected for the 1D completeness of the R-band that is used for the detection (see Figure 5.1). Lower panel: overdensity as a function of the R-band magnitude.

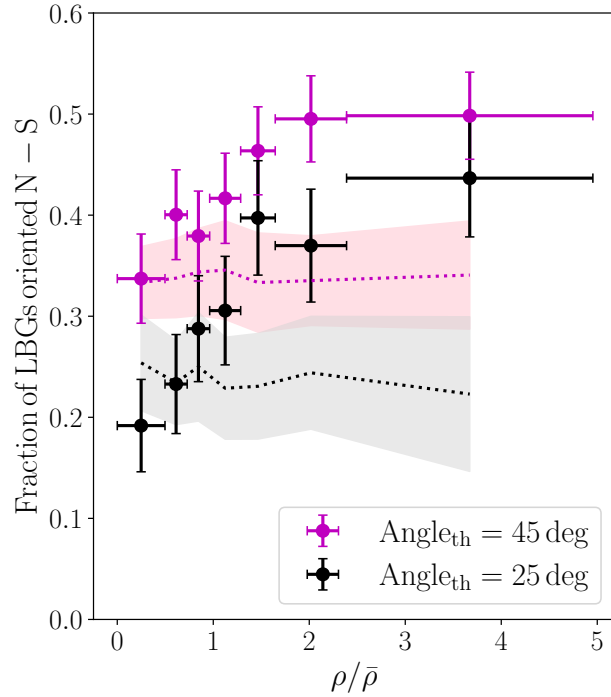


Figure D.3: Fraction of LBGs aligned with respect to the North-South direction, relative to the quasar, as a function of the local overdensity. We require a galaxy to be aligned if found within $\pm 25^\circ$ (black, corresponding to $0^\circ - 25^\circ$ or $155^\circ - 180^\circ$) or $\pm 45^\circ$ (magenta, corresponding to $0^\circ - 45^\circ$ or $135^\circ - 180^\circ$) the North-south direction. The shaded area show the 16th and 84th percentiles of the distribution expected for random samples of 1,000 bootstrap realizations.

from the central quasar. To further support this result, we reproduce the expectations for a sample of randomly-distributed galaxies by injecting the same number of LBGs of the real data in random positions across the FoV. We repeat the injection 1,000 times and show, despite the large uncertainties derived from the 16th and 84th percentile of the bootstrap distribution, that the fraction of aligned galaxies is flat as a function of the co-moving galaxy overdensity. To support this argument on a more statistical ground, we performed a Spearman's rank correlation test on both the data and the bootstrap results. For the data we found a positive correlation that is significant well above 2σ , as given by $\rho_{\text{spearman}} = 0.93$, $p\text{-value} = 0.003$ for the threshold $\pm 45^\circ$, and $\rho_{\text{spearman}} = 0.97$, $p\text{-value} = 0.0005$ for $\pm 25^\circ$. On the other hand, for the random-like samples we find $\rho_{\text{spearman}} = 0.01$, $p\text{-value} = 0.78$ for the threshold $\pm 45^\circ$, and $\rho_{\text{spearman}} = 0.02$, $p\text{-value} = 0.86$ for $\pm 25^\circ$. Altogether, this analysis provides a strong evidence for the distribution of massive LBGs to be preferentially aligned along the direction of the overdensity of star-forming galaxies observed in a much smaller volume around the central quasar and thus further supports a picture in which the gaseous structure embedding these galaxies extends for several tens of cMpc.

D.4 MQN01 clustering compared to the MAGG survey

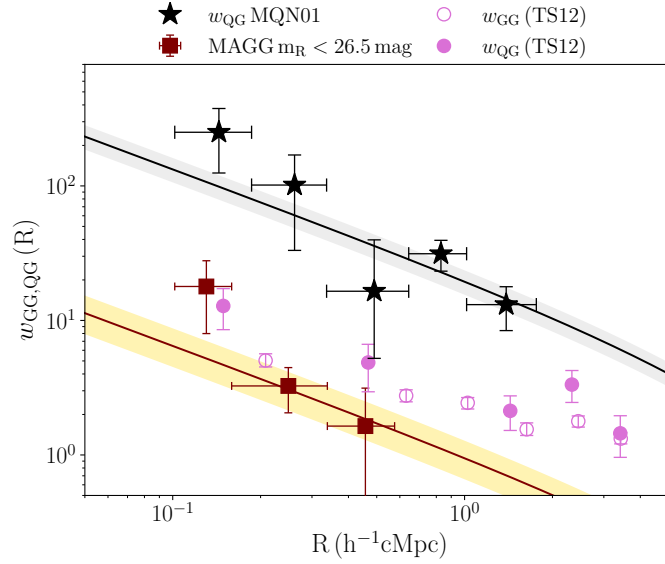


Figure D.4: Projected quasar-galaxy cross-correlation functions for the continuum-selected galaxies identified in the MAGG survey (darkred and gold-shaded region). Results for the MQN01 galaxies and from TS12 are shown as in Figure 5.6.

The clustering analysis presented in Section 5.4.3 and discussed in Section 5.5.1 provides hints that the MQN01 quasar can possibly be associated to a halo that is much larger compared to those of average quasars at similar redshifts. To further support this conclusion, we here give the details of the quasar-galaxy cross-correlation function we derived for continuum-selected galaxies clustered around the 28 quasars included in the MAGG survey. Such galaxies have been identified by Lofthouse et al. (2020b) running SEXTRACTOR (Bertin & Arnouts, 1996) on the MUSE white-light images obtained for each of the 28 cubes centered on a $3.2 \lesssim z \lesssim 4.5$ quasar (median $z \approx 3.7$). This software outputs a segmentation map that is used to extract the spectra of all the sources, which are finally inspected using MARZ Hinton et al. (2016) to assign a reliable spectroscopic redshift. As a final step, to produce a sample with a selection function that is as similar as possible to that of the MQN01 galaxies, we selected 20 galaxies with R-band magnitude $m_R \lesssim 26.5$ mag, high-quality spectroscopic redshift, and found within a line-of-sight separation $|\Delta v| \leq 1500 \text{ km s}^{-1}$ from the central quasar of the respective field. We finally derived the projected quasar-galaxy cross-correlation function by using the same procedure described in Section 5.4.3. The results are shown in Figure D.4. We measured a cross-correlation length $r_0 = (3.8 \pm 0.8) h^{-1} \text{ cMpc}$ for fixed $\gamma = 1.5$. Although the different sizes of the FoV, as well as the different selection and average redshift of the samples, make it difficult to compare this result with TS12, the fact that the cross-correlation in MAGG is ≈ 6 times smaller than that of MQN01 remains an important indicator of

the rare and massive nature of the quasar host halo, as discussed in Section 5.5.1.

D.5 Results from SED fitting

We provide in this section further details and results on the galaxies' SED fitting process. The parameters we set in the initial file of CIGALE are listed below in Table D.1, while Table D.2 shows the properties estimated for each of the 21 galaxies included in the main spectroscopic sample, as well as their redshift and results of the color-selections employed to identify the LBGs. For the sources detected in X-Ray, the derived SFR and stellar mass may be significantly affected by the contribution of the AGN emission to the total SED. We have thus investigated the morphology of these sources in the JWST/NIRCam data and verified that none of them shows compact bright emission in the center, suggesting that the contribution of the AGN emission does not outshine that of the stellar continuum. Furthermore, according to the best-fit model obtained by running CIGALE, the contribution of the AGN emission to the total galaxies' SED is, in terms of flux density, $\gtrsim 1$ dex smaller compared to the stellar continuum. Based on these considerations, we conclude that the properties we derived for these sources are reliable and are not dominated by presence of an AGN.

We also tested the results of the SED modeling to be independent of the code we used. To this end, we fit the observed galaxies' photometry using the Monte Carlo Spectro-Photometric Fitter (MC-SPF, Fossati et al. 2018b, 2019). The grid of stellar spectra is built upon the updated version¹ of the Bruzual & Charlot (2003a) high-resolution models at solar metallicity assuming an exponentially declining star formation history and the Chabrier (2003) initial mass function. A grid of nebular emission lines models from Byler et al. (2018) is scaled to the number of Lyman continuum photons from the stellar models, converted into flux at the redshift of the galaxies and then added to templates. To model the dust attenuation, we assume the double attenuation law from Calzetti et al. (2000), which includes extra extinction of the young stars compared to those older than 10 Myr. Comparing the two SED fitting codes, the median difference we observe in the galaxies' properties is 0.048 ± 0.050 dex for the stellar mass and 0.054 ± 0.055 dex for the star formation rate, which is in both the cases substantially smaller than other uncertainties associated to these parameters.

¹The set of models is publicly available at <http://www.bruzual.org/CB19/>. A detailed description is provided in Appendix A of Sánchez et al. (2022).

Parameter	Value
Exponential or Delayed star formation History	
e-folding time of the main stellar population model (Myr)	50-15000 in step $\Delta\tau = 50$ Myr
Age of the main stellar population (Myr)	25-2000 in step $\Delta\text{age} = 5$ Myr
Stellar population models: Bruzual & Charlot (2003a)	
Initial mass function	Chabrier (2003)
Metallicities (in separate runs)	0.008, 0.004, 0.020
Age of the separation between the young and the old star population (Myr)	10
Dust attenuation law: Calzetti et al. (2000)	
E(B-V)l: color excess of the nebular lines	0.0-0.75 in step $\Delta E(B-V)l = 0.015$
E(B-V)f: reduction factor to compute the E(B-V) for the stellar continuum attenuation	0.44
Nebular Emission: Inoue (2011)	
Ionization parameter	-2.0
Gas metallicity (set to be same as the stellar metallicity)	0.008, 0.004, 0.020
Fraction of Lyman continuum photons escaping the galaxy	0.0
AGN emission ^(a) : Fritz et al. (2006)	
AGN fraction: contribution of the AGN to the total L_{IR}	0.0, 0.01, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9
Opening angle of the dust torus	10, 30, 60
Angle between AGN axis and line of sight	0.001, 89.9
Optical depth at 9.7	1.0, 6.0
X-ray emission ^(a) : Yang et al. (2022)	
Photon index of the AGN intrinsic X-ray spectrum	1.8
Power-law slope connecting L_ν at rest-frame 2500Å and 2keV (α_{ox})	-1.9, -1.7, -1.5, -1.3, -1.1, -0.9
Maximum allowed deviation of α_{ox} from the empirical $\alpha_{\text{ox}} - L_\nu$ (2500Å)	0.2
Deviation from the expected LMXB scaling relation (δ_{LMXB})	0.0
Deviation from the expected HMXB scaling relation (δ_{HMXB})	0.0

Table D.1: Input parameters for the SED fitting code CIGALE. ^(a) The AGN and X-ray emission are included only for the five sources detected with *Chandra*.

Coordinates	spec-z	$\log [\text{SFR}/(\text{M}_\odot \text{Yr}^{-1})]$	$\log [\text{sSFR}/\text{Yr}^{-1}]$	$\log [M_\star/\text{M}_\odot]$	A_V	AGN	UBR	UGR
00:41:35.39 -49:37:19.41	3.254	1.84±0.30	-8.62±0.36	10.46±0.20	0.64±0.03			✓
00:41:35.21 -49:37:11.76	3.246	0.92±0.30	-8.61±0.36	9.53 ±0.20	0.70±0.09			
00:41:31.95 -49:37:07.89	3.253	1.02±0.30	-8.33±0.36	9.35 ±0.20	0.41±0.03		✓	✓
00:41:31.60 -49:36:58.25	3.245	2.26±0.32	-8.92±0.39	11.19±0.22	1.15±0.05	✓		✓
00:41:26.93 -49:36:50.15	3.254	1.13±0.30	-8.60±0.36	9.74 ±0.20	0.30±0.04		✓	✓
00:41:34.99 -49:36:41.87	3.241	1.24±0.38	-8.75±0.44	9.99 ±0.23	0.90±0.06	✓		✓
00:41:32.22 -49:36:39.41	3.246	1.29±0.30	-8.47±0.36	9.76 ±0.20	0.53±0.06			✓
00:41:35.04 -49:36:34.90	3.236	0.32±0.30	-8.26±0.36	8.58 ±0.20	0.08±0.02		✓	✓
00:41:34.50 -49:36:32.13	3.243	1.11±0.30	-8.46±0.36	9.57 ±0.20	0.57±0.06			
00:41:32.18 -49:36:20.74	3.243	1.70±0.38	-8.92±0.43	10.62±0.22	1.70±0.13	✓		
00:41:31.19 -49:36:20.33	3.254	0.85±0.30	-8.30±0.36	9.15 ±0.20	0.25±0.02		✓	✓
00:41:31.79 -49:36:18.97	3.248	0.66±0.30	-8.67±0.36	9.33 ±0.20	0.46±0.04			
00:41:32.02 -49:36:18.98	3.243	2.19±0.33	-8.89±0.40	11.08±0.22	0.88±0.06	✓		✓
00:41:31.92 -49:36:17.84	3.250	0.27±0.30	-9.73±0.36	10.0 ±0.20	0.31±0.25			
00:41:34.89 -49:36:15.77	3.247	0.66±0.30	-8.32±0.36	8.97 ±0.20	0.30±0.02		✓	✓
00:41:34.12 -49:36:02.81	3.247	1.30±0.30	-8.26±0.36	9.56 ±0.20	0.38±0.02		✓	✓
00:41:30.93 -49:35:39.02	3.239	0.96±0.30	-8.24±0.36	9.20 ±0.20	0.25±0.02		✓	✓
00:41:30.94 -49:35:34.61	3.248	1.61±0.30	-8.30±0.36	9.91 ±0.20	0.54±0.03		✓	✓
00:41:29.98 -49:35:34.32	3.258	0.71±0.30	-8.22±0.36	8.93 ±0.20	0.20±0.03		✓	✓
00:41:32.25 -49:35:29.16	3.244	0.01±0.30	-8.96±0.36	9.00 ±0.20	0.45±0.15		✓	✓
00:41:35.13 -49:37:12.81	3.245	2.09±0.30	-9.28±0.36	11.37±0.20	0.58±0.03	✓		

Table D.2: Detailed results of the galaxies' SED fitting. The properties listed in the table refer to the sample of spectroscopically selected galaxies that have been identified in the FoV of the MUSE mosaic within a line-of-sight separation $|\Delta v| \leq 1000 \text{ km s}^{-1}$ from the central quasar. Uncertainties on the SFR, sSFR and $M_\star$ are computed as described in Section 5.3.4.

List of publications

This thesis is based on the publications listed below, and includes three first-author papers (two of them already published and one recently submitted to a journal) and one paper in which I appear as a second author, but with equal contribution. Reported below are also the other papers to which I have contributed during the PhD and are related to the results presented in this thesis.

Publications included in this thesis

Chapter 2, Galbiati et al. (2023):

“MUSE Analysis of Gas around Galaxies (MAGG)-V: Linking ionized gas traced by CIV and SiIV absorbers to Ly α emitting galaxies at $z \approx 3-4$ ” M. Galbiati, M. Fumagalli, M. Fossati, E.K. Lofthouse, R. Dutta, J.X. Prochaska, M.T. Murphy, S. Cantalupo [MNRAS, 524, 3474 doi:10.1093/mnras/stad2087] 2023

Chapter 3, Galbiati et al. (2024b):

“MUSE Analysis of Gas around Galaxies (MAGG)-VI. The cool and enriched gas environment of $z \gtrsim 3$ Ly α emitters” M. Galbiati, R. Dutta, M. Fumagalli, M. Fossati, S. Cantalupo [A&A, 690, A7 doi:10.1051/0004-6361/202450741] 2024

Chapter 4, Pensabene et al. (2024a):

“The MUSE Ultra Deep Field (MUDF) VII. Probing high-redshift gas structures in the surroundings of ALMA-identified massive dusty galaxies” A. Pensabene, M. Galbiati (second author with equal contribution), et al. [Submitted to A&A 10.48550/arXiv.2410.06249] 2024

Chapter 5, Galbiati et al. (2024a):

“Connecting the growth of galaxies to the large-scale environment in a massive node of the Cosmic Web at $z \sim 3$ ” M. Galbiati, S. Cantalupo, C. Steidel, A. Pensabene, A. Travascio, W. Wang, M. Fossati, M. Fumagalli, G. Rudie, A. Fresco, T. Lazeyras, N. Ledos, G. Quadri [Submitted to A&A 10.48550/arXiv.2410.03822] 2024

Other relevant publications

1) “X-ray view of a massive node of the Cosmic Web at $z \sim 3$ I. An exceptional overdensity of rapidly accreting SMBHs” A. Travascio, S. Cantalupo, P. Tozzi, F. Vito, A. Paggi, G. Pezzulli, M. Elvis, G. Fabbiano, F. Fiore, M. Fossati, A. Fresco, M. Fumagalli, M. Galbiati, et al. [Submitted to A&A 10.48550/arXiv.2410.03933] 2024

2) “A Giant Disk Galaxy Two Billion Years After The Big Bang” W. Wang, S. Cantalupo, A. Pensabene, M. Galbiati, et al. [Submitted 10.48550/arXiv.2409.17956] 2024

3) “The MUSE Ultra Deep Field (MUDF). VI. The relationship between galaxy properties and metals in the circumgalactic medium” A. Beckett, M. Rafelski, M. Revalschi, M. Fumagalli, M. Fossati, K. Nedkova, R. Dutta, R. Bielby, S. Cantalupo, P. Dayal, V. D’Odorico, M. Galbiati, et al. [ApJ in press doi:10.3847/1538-4357/ad702d] 2024

- 4) “ALMA survey of a massive node of the Cosmic Web at $z \sim 3$. I. Discovery of a large overdensity of CO emitters“ A. Pensabene, S. Cantalupo, C. Cicone, R. Decarli, **M. Galbiati**, et al. [A&A, 684, A119 [doi:10.1051/0004-6361/202348659](https://doi.org/10.1051/0004-6361/202348659)] 2024
- 5) “Resolving the physics of quasar Ly α nebulae (RePhyNe): I. Constraining quasar host halo masses through circumgalactic medium kinematics“ S. de Beer, S. Cantalupo, A. Travascio, G. Pezzulli, **M. Galbiati**, et al. [MNRAS, 526, 1850 [doi:10.1093/mnras/stad2682](https://doi.org/10.1093/mnras/stad2682)] 2023
- 6) “Towards an automatic approach to modelling the circumgalactic medium: new tools for mock making and fitting of metal profiles in large surveys“ A. Longobardi, M. Fossati, M. Fumagalli, B. Agarwal, E.K. Lofthouse, **M. Galbiati**, et al. [RAS Techniques and Instruments, 2, 470 [doi:10.1093/rasti/rzad031](https://doi.org/10.1093/rasti/rzad031)] 2023
- 7) “MUSE Analysis of Gas around Galaxies (MAGG)–IV. The gaseous environment of $z \sim 3$ –4 Ly α emitting galaxies“ E.K. Lofthouse, M. Fumagalli, M. Fossati, R. Dutta, **M. Galbiati**, et al. MNRAS, 518, 305. [doi:10.1093/mnras/stac3089](https://doi.org/10.1093/mnras/stac3089)] 2023