

---

# Multi-phase gas reservoirs in and around high- $z$ galaxies

Nahir Muñoz Elgueta

---



München 2023



---

# Multi-phase gas reservoirs in and around high- $z$ galaxies

Nahir Muñoz Elgueta

---

Dissertation  
an der Fakultät für Physik  
der Ludwig-Maximilians-Universität  
München

vorgelegt von  
Nahir Muñoz Elgueta  
aus La Serena, Chile

München, den 25.07.2023

Erstgutachter: Prof. Dr. Guinevere Kauffmann

Zweitgutachter: PD Dr. Klaus Dolag

Tag der mündlichen Prüfung: 21.09.2023

*To my uncle Mene,  
who will always live in my heart*

*“Lift up your face, child, and receive the stars.  
When you first look, they all pierce and freeze you,  
and then the sky begins to sway like a cradle they’re rocking,  
and you give yourself up wholly to be carried away, away.”  
Gabriela Mistral*



# Contents

<b>Zusammenfassung</b>	<b>xv</b>
<b>Abstract</b>	<b>xvii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Overview of galaxy formation and evolution . . . . .	2
1.1.1 The Lambda Cold Dark Matter ( $\Lambda$ CDM) cosmological model . . . . .	2
1.1.2 Formation of cosmic structures . . . . .	3
1.1.3 Fueling galaxy growth . . . . .	3
1.1.4 Active Galactic Nuclei . . . . .	5
1.2 The Circumgalactic Medium (CGM) . . . . .	6
1.2.1 CGM as traced by emission lines . . . . .	8
1.2.2 Cool phase of CGM in observations . . . . .	10
1.2.3 Cool dense CGM from theoretical perspective . . . . .	13
1.3 Cold gas reservoirs in high- $z$ galaxies, from galactic to halo scales . . . . .	15
1.3.1 The Interstellar Medium (ISM) . . . . .	15
1.3.2 Cold gas tracers . . . . .	17
1.3.3 Main sequence, depletion time and molecular gas fraction . . . . .	23
1.3.4 Observations of cold gas in high- $z$ galaxies . . . . .	24
1.3.5 Cold gas reservoirs from theoretical perspective . . . . .	26
1.4 Structure and goals of this thesis . . . . .	28
<b>2 APEX at the QSO MUSEUM</b>	<b>31</b>
2.1 Introduction . . . . .	32
2.2 Sample and observational data . . . . .	35
2.3 Observational results . . . . .	41
2.3.1 Emission line measurements . . . . .	42
2.3.2 Redshifts from molecular lines . . . . .	47
2.4 Molecular mass estimates . . . . .	50
2.4.1 Atomic carbon mass . . . . .	50
2.4.2 LVG Models and CO constraint . . . . .	52
2.4.3 Molecular mass constraints from a joint CO and [C $\text{I}$ ] analysis . . . . .	54
2.5 Comparison with the large-scale Ly $\alpha$ emission . . . . .	57

2.5.1	Molecular lines versus Ly $\alpha$ line properties . . . . .	57
2.5.2	Molecular tracers vs Ly $\alpha$ line profiles . . . . .	59
2.6	Discussion . . . . .	63
2.6.1	Quasar host properties . . . . .	63
2.6.2	Ly $\alpha$ nebulae kinematics with respect to molecular gas . . . . .	67
2.6.3	Relation between large molecular reservoirs and Ly $\alpha$ nebulae . . . . .	68
2.7	Summary . . . . .	69
2.8	Appendix A: Stability of data . . . . .	71
2.9	Appendix B: Reliability of line detections . . . . .	72
<b>3</b>	<b>Modelling [CII] emission in star-forming galaxies</b>	<b>75</b>
3.1	Introduction . . . . .	75
3.2	Observations . . . . .	78
3.3	Methods . . . . .	79
3.3.1	IllustrisTNG simulations . . . . .	79
3.3.2	Selection of simulated galaxies . . . . .	80
3.3.3	[CII] cooling rate and flux estimation . . . . .	81
3.3.4	[CII] surface brightness maps and profiles . . . . .	84
3.4	Results . . . . .	86
3.4.1	Relation between [CII] luminosity and star formation rate . . . . .	86
3.4.2	[CII] Surface brightness profiles . . . . .	87
3.5	Discussion . . . . .	94
3.5.1	Origin of large-scale star-forming gas . . . . .	94
3.5.2	Role of satellite galaxies in producing [CII] halos . . . . .	96
3.5.3	Role of outflows in [CII] emission of high and low-SFR galaxies . . . . .	98
3.5.4	Caveats . . . . .	102
3.6	Summary . . . . .	104
<b>4</b>	<b>Summary and Future outlook</b>	<b>107</b>
4.1	Summary and implications of the main results . . . . .	107
4.1.1	Link between molecular gas reservoirs and extended Ly $\alpha$ emission in $z \sim 3$ quasars . . . . .	107
4.1.2	[CII] emission in simulated TNG50 galaxies at $z \sim 4-6$ . . . . .	108
4.2	Future directions and projects extensions . . . . .	109
4.2.1	Multi-phase gas reservoirs in $z \sim 3$ quasars . . . . .	109
4.2.2	Extension of [CII] predictions . . . . .	114
	<b>Bibliography</b>	<b>117</b>
	<b>Acknowledgements</b>	<b>146</b>

# List of Figures

1.1	Large scale structure of the Universe derived from the Millenium Simulation.	4
1.2	Gas density projection centred on a quasar host galaxy at $z \sim 6$ , extracted from cosmological simulations. . . . .	7
1.3	Artistic representation of the CGM. . . . .	9
1.4	Examples of Ly $\alpha$ surface brightness maps for extended Ly $\alpha$ nebulae around high- $z$ quasars. . . . .	12
1.5	Neutral hydrogen column density map in a 600 kpc column, of a Milky Way-mass galaxy at $z=0$ . . . . .	14
1.6	Representation of the structure of a PDR. . . . .	17
1.7	CO SLED for different sources. . . . .	20
1.8	Representation of fine-structure levels of CI and C $^+$ . . . . .	21
1.9	Atmospheric transmission for ALMA, between 180 and 380 GHz. . . . .	23
2.1	The peak specific luminosity at the Ly $\alpha$ line of the quasar versus $M_{1450\text{\AA}}$ for the QSO MUSEUM sample . . . . .	36
2.2	$40 \times 40$ (or $\sim 312 \text{ kpc} \times 312 \text{ kpc}$ ) MUSE Ly $\alpha$ surface brightness maps of the nebulae surrounding the 9 targets of this study . . . . .	38
2.3	Histograms of the PWV values during the observations . . . . .	39
2.4	The seven APEX CO(6-5) spectral line observations . . . . .	43
2.5	Same as Figure 2.4, but for the eight APEX CO(7-6)-CI(2-1) spectral line observations . . . . .	44
2.6	CO SLEDs obtained from LVG modelling, for $T_{\text{kin}} = 30 \text{ K}$ . . . . .	53
2.7	CO SLEDs obtained with the joint CO and CI constraints, for the sources with IDs 8, 43 and 50 . . . . .	55
2.8	CO SLEDs for the sources with IDs 8, 43 and 50 compared to the SLEDs for the quasars Cloverleaf, MG0751 and PSS14 . . . . .	56
2.9	FWHM of the detected molecular lines versus Ly $\alpha$ properties . . . . .	58
2.10	Molecular gas mass $M_{\text{H}_2}$ derived from the joint CO and CI constraints versus Ly $\alpha$ properties . . . . .	60
2.11	Normalized APEX CO(6-5) emission compared with the nebulae and quasar Ly $\alpha$ emission from MUSE . . . . .	61
2.12	Same as Figure 2.11, but for the APEX CI(2-1) detections . . . . .	62

2.13	Velocity shifts between $z_{\text{sys}}$ (from CIV) and $z_{\text{mol,sys}}$ (blue), and redshift of Ly $\alpha$ nebulae and $z_{\text{mol,sys}}$ . . . . .	62
2.14	APEX CI(2-1) spectral line observation for Q-0115-30 . . . . .	64
2.15	MUSE white-light image for the field around ID 29 . . . . .	66
2.16	RMS of the final combined spectra as a function of the bin size for each of the nine sources . . . . .	73
2.17	RMS values (for each scan) versus PWV relation for the source SDSS J0100+2105 (ID 39) . . . . .	74
3.1	Comparison of the distribution of the SFRs and $M_*$ values between the simulated and observed samples. . . . .	82
3.2	Phase-diagram for one of the selected TNG50 galaxies at $z=5$ . . . . .	83
3.3	Example of different input spectra used in the CLOUDY calculations. . . . .	85
3.4	[CII] emission surface-brightness map for one of the selected TNG50 galaxies at $z=5$ . . . . .	86
3.5	[CII] luminosity versus SFR for the TNG50 galaxies and the ALPINE survey. . . . .	88
3.6	Averaged [CII] SB profiles obtained for our sample of 72 TNG50 halos at $z=4, 5$ and $6$ . . . . .	89
3.7	Left: Averaged contribution of the star-forming gas and non star-forming gas to the total [CII] emitting gas. Right: SFR density map for a TNG50 galaxy at $z=5$ . . . . .	90
3.8	Averaged [CII] SB profiles assuming $C^+$ fractions from the UBV and local young stars case. . . . .	91
3.9	Averaged contribution of the infalling gas and outflowing gas to the total [CII] emitting gas. . . . .	92
3.10	Left: Gas radial velocity versus radial distance, weighted by [CII] luminosity, for a stacking of all the selected TNG50 galaxies. Right: the histograms shows the total distribution of inflowing and outflowing gas, at different radial bins. . . . .	93
3.11	Averaged contribution of the satellites gas and central galaxies gas to the total [CII] emitting gas. . . . .	94
3.12	Same as in Figure 3.10, but for a stacking of all satellites of all the selected TNG50 halos. . . . .	95
3.13	Average SFR-weighted radial velocity map for a galaxy at $z=5$ . . . . .	96
3.14	Left: Average metallicity (weighted by SFR) radial profile of the star-forming gas for one galaxy at $z=5$ . Right: Average SFR-weighted metallicity map for a galaxy at $z=5$ . . . . .	97
3.15	Averaged [CII] SB profiles obtained for a sample of 169 TNG50 galaxies at $z=4, 5$ and $6$ which include close satellites. . . . .	98
3.16	Averaged [CII] SB profiles for the outflowing (inflowing) gas only, represented by the red (blue) color. . . . .	99

---

3.17	Left: Averaged density profile, weighted by temperature, obtained for the outflowing gas only. Right: Averaged temperature profile, weighted by density, obtained for the outflowing gas only. . . . .	101
4.1	PDR models of <a href="#">Meijerink et al. (2007)</a> , as a function of $n$ and $G$ , for ID 3.	110
4.2	SEDs for J 0525-233 (ID 3). . . . .	111
4.3	$L_{[\text{CII}]}/L_{\text{FIR}}$ versus $L_{\text{CO}(1-0)}/L_{\text{FIR}}$ , adapted from <a href="#">Venemans et al. (2017)</a> . . .	112
4.4	Left: $30'' \times 30''$ Ly $\alpha$ SB maps of the five targets proposed compared to the one of ID 29. Right: Quasars' peak specific luminosity at the Ly $\alpha$ line versus $M_{1450}$ for the QSO MUSEUM sample. . . . .	113
4.5	Left: SEDs for the two targets of the proposal. Right: $60'' \times 60''$ MUSE Ly $\alpha$ surface brightness maps of the nebulae surrounding the 2 proposed targets.	114



# List of Tables

1.1	Excitation potentials, rest wavelengths, rest frequencies, Einstein $A$ coefficients and critical densities for different cold gas tracers. . . . .	19
2.1	Sample, observations setup and log . . . . .	40
2.2	CO(6-5), CI(2-1), CO(7-6) fluxes, FWHM, and integrated S/N ratios. . . .	45
2.3	Redshifts, central frequencies and velocity shifts. . . . .	49
2.4	Molecular, carbon and dynamical mass estimations. . . . .	51
2.5	Numbers of spurious lines in the negative and jack-knife tests. . . . .	74



# Zusammenfassung

Das zirkumgalaktische Medium (CGM) spielt eine entscheidende Rolle bei der Gestaltung der Prozesse der Galaxienentstehung und -entwicklung und bietet wertvolle Einblicke in das komplexe Zusammenspiel zwischen Galaxien und ihrer Umgebung. In dieser Arbeit konzentriere ich mich auf die Untersuchung von Gasreservoirs in Galaxien mit hoher Rotverschiebung, vom Halo bis zur Galaxienskala.

Zunächst untersuchte ich die molekularen Gasreservoirs anhand der APEX CO(6-5), CO(7-6) und [CI](2-1) Emissionslinien in neun  $z \sim 3$  Quasaren, die bereits mit MUSE/VLT für die Ly $\alpha$  Emission auf CGM-Skalen beobachtet wurden. Durch die Verknüpfung der CO- und [CI]-Informationen konnte ich die Molekulargasmassen der Wirtsgalaxien der Quasare bestimmen. Außerdem habe ich die Beobachtungen des molekularen Gases mit den Eigenschaften der diffusen Ly $\alpha$ -Emission verglichen, um die Verbindung zwischen den kühlen ( $T \sim 10^4$  K) und den molekularen Gasphasen zu verstehen. Ich fand große Geschwindigkeitsverschiebungen zwischen der Mehrheit der Ly $\alpha$  Emission und den geschätzten systemischen Rotverschiebungen des Molekulargases in vier Quellen. Die Quellen mit den größten Abweichungen haben die größten Ly $\alpha$ -Linienbreiten in der Stichprobe, was stark auf turbulenter Gasbedingungen und/oder großräumige Ein- und Ausflüsse um diese Quasare hindeutet. Ich habe auch herausgefunden, dass die hellsten und breitesten molekularen Linien für die kleinsten und schwächsten Ly $\alpha$ -Nebel in der Stichprobe entdeckt werden, was wahrscheinlich darauf hindeutet, dass die Verdunkelung der Wirtsgalaxie eine wichtige Rolle bei der Verringerung der ionisierenden Photonen spielen kann, die in den Halo entkommen können. Diese Studie ist eine der ersten Beobachtungsarbeiten, die Einschränkungen und Korrelationen zwischen der kalten Gasphase, die durch die Ly $\alpha$ -Emission aufgespürt wird, und dem Gehalt an molekularem Gas in stark rotverschobenen Quasaren liefert.

Anschließend untersuchte ich kalte ( $T < 100$  K) Gasreservoirs aus theoretischer Sicht. Dazu verwendete ich die kosmologischen Illustris-TNG50-Simulation, auf die ich ein Nachbearbeitungsmodell zur Vorhersage der [CII]-Emissionen von  $z \sim 4-6$  sternbildenden Galaxien anwendete. Mein Ziel war es, den Ursprung der diffusen [CII]-Emission in Galaxien mit hoher Rotverschiebung zu untersuchen und unsere Ergebnisse mit den Beobachtungen der ALPINE-Durchmusterung zu vergleichen. Unter Einbeziehung von C<sup>+</sup>-Anteilen, die auf Annahmen über den UV-Hintergrund und junge Sterne (Alter  $< 10$  Myr) als Strahlungsquellen beruhen, habe ich [CII]-Emissionen geschätzt, künstliche Beobachtungen durchgeführt und [CII]-SB-Profile extrahiert. Beim Vergleich mit den Beobachtungen stellte ich fest, dass das Modell auf galaktischen Skalen ( $r < 10$  kpc bei  $z=5$ ) ähnliche Werte für

die [CII]-Emission vorhersagt, auf größeren Skalen jedoch die Emission um einen Faktor von  $\sim 10$  unterschätzt. Es werden mehrere Effekte diskutiert, die zu dieser Unstimmigkeit führen können. Darüber hinaus stelle ich beim Vergleich der SB-Profile von Galaxien mit niedriger und hoher SFR eine weitere Diskrepanz zu den Beobachtungen fest: Es gibt keinen erkennbaren Unterschied in der ausgedehnten [CII]-Emission zwischen den beiden simulierten Teilproben. Diese Diskrepanz könnte auf die begrenzte Auflösung der Simulation und das Feedback-Modell zurückzuführen sein. Schließlich legt diese Analyse nahe, dass die ausgedehnte [CII]-Emission aus einer Kombination von Gas aus Satellitengalaxien (die derzeit in den Modellen im Bereich von  $0,6 < r/R_{200} < 1$  dominieren) und Ausflüssen aus zentralen Galaxien stammt. Eine genaue Abschätzung der Bedeutung der einzelnen Beiträge erfordert jedoch die Entwicklung weiterer Subgrid-Beschreibungen. Diese Studie ist der erste Versuch, die [CII]-Emission in Galaxien mit Hilfe des IllustrisTNG-Simulationsprojektes zu simulieren.

Durch diesen kombinierten Ansatz aus Beobachtungen und theoretischen Untersuchungen liefert diese Arbeit wertvolle Erkenntnisse über die physikalischen Prozesse, die den Gasreservoirs in Galaxien mit hoher Rotverschiebung und ihren Verbindungen mit dem CGM zugrunde liegen.

# Abstract

The circumgalactic medium (CGM) plays a crucial role in shaping the processes of galaxy formation and evolution, containing valuable insights into the complex interplay between galaxies and their surrounding environment. In this thesis, I focus on the study of gaseous reservoirs in high-redshift galaxies, from halo to galaxy scales.

First, I investigated the molecular gas reservoirs by using APEX CO(6-5), CO(7-6) and [CI](2-1) emission lines in nine  $z \sim 3$  quasars already observed with MUSE/VLT for the Ly $\alpha$  emission on CGM scales. By joining the CO and [CI] information, I constrained the molecular gas masses of the quasars' host galaxies. I also compared the molecular gas observations with the extended Ly $\alpha$  properties to understand the link between the cool ( $T \sim 10^4$  K) and the molecular gas phases. I found large velocity shifts between the bulk of Ly $\alpha$  and the estimated molecular gas systemic redshifts in four sources. The sources with the largest offsets have the largest Ly $\alpha$  line widths in the sample, strongly suggesting more turbulent gas conditions and/or large-scale inflows/outflows around these quasars. I also found that the brightest and the widest molecular lines are detected for the smallest and dimmest Ly $\alpha$  nebulae in the sample, likely indicating that host galaxy obscuration can play an important role in reducing the ionizing photons able to escape to halo scales. This study stands out as one of the first observational works providing constraints and correlations between the cool gas phase, traced by Ly $\alpha$  emission, and the molecular gas content in high-redshift quasars.

Then, I investigated cold ( $T < 100$  K) gas reservoirs from a theoretical perspective. For this, I used the Illustris-TNG50 cosmological simulation to apply a post-processing model to predict [CII] emission levels for  $z \sim 4-6$  star-forming galaxies. I aimed to investigate the origin of extended [CII] emission in high-redshift galaxies, comparing our results with the observational ALPINE survey. By incorporating C<sup>+</sup> fractions based on assumptions about the UV background and young stars (ages  $< 10$  Myr) as radiation sources, I estimated [CII] emissivities, produced mock observations and extracted [CII] SB profiles. Comparing with observations, I found that the model predicts similar values of the [CII] emission on galactic scales ( $r < 10$  kpc at  $z=5$ ), however at larger scales it underestimates the emission by a factor of  $\sim 10$ . Several caveats that may lead to this disagreement are discussed. Additionally, I find further tension with observations when comparing the SB profiles of low and high-SFR galaxies, finding no discernible difference in the extended [CII] emission between the two simulated subsamples. This discrepancy could be attributed to the limited resolution of the simulation and the feedback model. Finally, this analysis suggests that

the extended [CII] emission originates from a combination of gas from satellite galaxies (currently dominating in the models in the range the range of  $0.6 < r/R_{200} < 1$ ) and outflows from central galaxies. However, a firm estimate of the importance of each contribution requires the development of further subgrid prescriptions. This study represents the first effort to simulate [CII] emission in galaxies using the IllustrisTNG simulations suite.

Through this combined approach of observational and theoretical investigations, this thesis provides valuable knowledge on the physical processes underlying the gaseous reservoirs in high-redshift galaxies and their connections with the CGM.

# Chapter 1

## Introduction

The study of the high-redshift<sup>1</sup> ( $z$ ) Universe plays an essential role in advancing our understanding of galaxy evolution and the physical processes that shape the cosmos. By probing galaxies in the early stages of their formation, we gain insights into the fundamental mechanisms that drive their growth and transformation over cosmic time. Central to this investigation is the examination of gas reservoirs and their interplay with various astrophysical phenomena.

Gas reservoirs, spanning from galactic to halo scales, are key to this understanding. The interplay between galaxies and their gaseous reservoirs influence their physical properties, such as their star formation rates, chemical enrichment, morphologies, and overall evolution. High- $z$  galaxies offer a unique window into these early stages, where the interplay between gas inflows, outflows, and feedback processes is particularly significant. The state-of-the-art observations and simulations, have revolutionized our ability to study galaxy formation and evolution. In particular, cosmological simulations provide a powerful tool to explore the underlying physical processes and make predictions that can be compared with observational data. However, bridging the gap between theory and observations requires a comprehensive understanding of the observational techniques and instruments used to probe the high- $z$  universe.

In this context, the present thesis aims to contribute to our understanding of galaxy evolution by investigating the properties of cold ( $\lesssim 10^4$  K) gas reservoirs in and around high- $z$  galaxies. Through a combination of observations and cosmological simulations, we explore the nature of these gas reservoirs, their connection with the host galaxies, as well as their impact on key observables at large scales, such as Lyman alpha and [CII]  $158\mu\text{m}$  emission. This introductory Chapter sets the stage for the subsequent chapters by providing a general overview of galaxy formation and evolution, the circumgalactic medium, cold gas tracers and cold gas reservoirs, from galactic to halo scales. It also presents the motivation

---

<sup>1</sup>The redshift ( $z$ ) of an astronomical object is expressed as:

$$z = \frac{\lambda_{\text{obs}} - \lambda_{\text{rest}}}{\lambda_{\text{rest}}}$$

where  $\lambda_{\text{obs}}$  and  $\lambda_{\text{rest}}$  represent the observed and rest wavelengths, respectively.

for the thesis work and establishes the research objectives that will be addressed throughout the thesis.

## 1.1 Overview of galaxy formation and evolution

### 1.1.1 The Lambda Cold Dark Matter ( $\Lambda$ CDM) cosmological model

The formation of galaxies and the underlying mechanisms involved have been central questions in the field of astronomy ever since Edwin Hubble’s discovery of extragalactic nebulae in 1929. Astronomers have dedicated decades of research to gather information in order to uncover the processes that govern the growth of galaxies and the formation of large-scale structures. As a result, a comprehensive framework known as the “standard cosmological model” has emerged, offering a theoretical description of the Universe and its evolution. According to the current cosmological model, so-called  $\Lambda$  Cold Dark Matter ( $\Lambda$ CDM) model (e.g., [Spergel et al. 2007](#)), at early times quantum fluctuations amplified by cosmic inflation caused certain regions of the Universe to become denser than others. These density fluctuations served as the seeds for the formation of structures in the Universe, creating substructures due to gravitational instability ([Jeans, 1902](#)). Over time, under the influence of gravity, these denser regions accumulated more matter, leading to the emergence of large-scale structures such as galaxies, galaxy clusters, and cosmic filaments.

In the  $\Lambda$ CDM model, the density distribution of the Universe consists of three main components: 1) Dark energy, associated with the accelerated expansion of the Universe, 2) Cold dark matter, which refers to non-relativistic matter that interacts with other matter predominantly through gravity, acting as the gravitational glue that binds structures together, and 3) Baryonic matter, which refers to ordinary matter composed by protons and neutrons, including all the visible matter in the Universe. A combination of observations from different sources has allowed to robustly constrain the so-called cosmological parameters: dark energy density ( $\Omega_\Lambda$ ), the matter density ( $\Omega_m$ ), the baryonic density ( $\Omega_b$ ), the Hubble constant<sup>2</sup> ( $H_0$ ), the power-law index of the primordial power spectrum ( $n_s$ ) and the linear variance of the matter distribution ( $\sigma_8$ ). The most recent estimates ([Planck Collaboration et al., 2020](#)) have estimated the following values:  $\Omega_\Lambda = 0.685$ ,  $\Omega_m = 0.315$ ,  $\Omega_b = 0.049$ ,  $h = 0.674$ <sup>3</sup> = km s<sup>-1</sup> Mpc<sup>-1</sup>,  $n_s = 0.966$ ,  $\sigma_8 = 0.812$ .

---

<sup>2</sup>The Hubble constant, denoted as  $H_0$ , represents the current rate of expansion of the Universe. It quantifies the relationship between the distance and recessional velocity of galaxies, as described by Hubble’s law. The Hubble’s law is expressed as  $v = H_0 D$ , with  $v$  the recessional velocity of a galaxy and  $D$  its proper distance from us.

<sup>3</sup> $h = H_0/(100 \text{ km s}^{-1} \text{ Mpc}^{-1})$ , this parameter is the dimensionless equivalent of the Hubble constant. It is often employed instead of  $H_0$  because it simplifies comparisons and interpretations of cosmological data and models across different studies.

### 1.1.2 Formation of cosmic structures

The current cosmological model suggests that the Universe began with a hot and dense state 13.7 billion years ago, expanding and cooling thereafter. This expansion, that started from a singularity known as the “Big Bang”, was followed by a brief ( $\sim 10^{-30}$  s) inflationary phase where quantum fluctuations grew to cosmological scales. Dark matter, interacting only gravitationally, allowed these small perturbations to collapse and grow. This growth of structures can be then explained by the evolution of density perturbations. Assuming that fluctuations occur via gravitational instability and dark matter is the dominant component, we can consider the dark matter and baryons as a single fluid. Then, the “density contrast”  $\delta$  at some point  $r$  and time  $t$  of the Universe, can be expressed as:

$$\delta(r, t) = \frac{\rho(r, t) - \bar{\rho}(t)}{\bar{\rho}(t)} \quad (1.1)$$

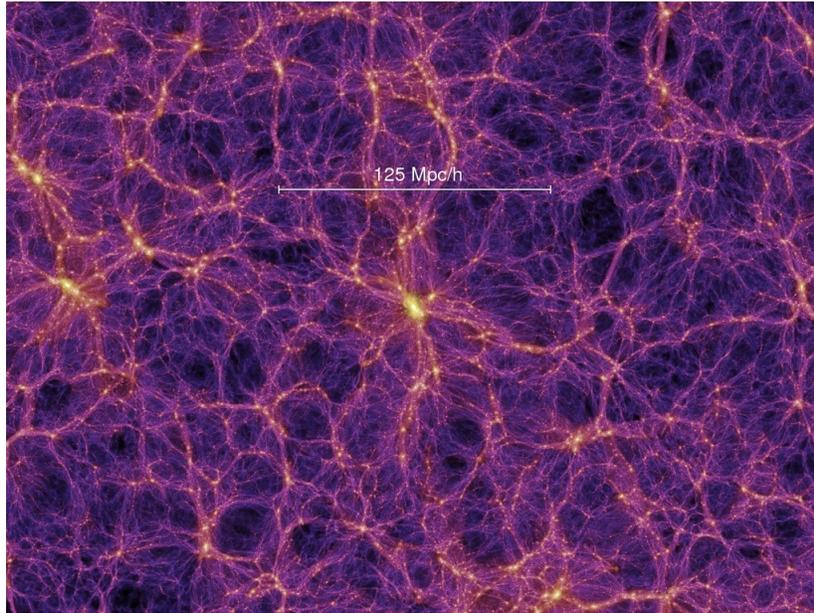
where  $\rho$  the density at some specific  $r$  and  $t$ , and  $\bar{\rho}$  is the mean density of the Universe. As a result of self-gravity, a region of the Universe will undergo expansion at a slower pace compared to the overall expansion of the Universe. This evolution can be described within the linear regime if  $\delta(r, t) \ll 1$  (e.g., Lifshitz 1946; Zel’dovich 1970; Peebles 1980). In contrast, if the initial density of another region is significantly higher, after a certain time the density becomes non-linear, causing the region to collapse under its own gravitational pull and form a virialized system (e.g., Tomita 1967; Sunyaev & Zeldovich 1972).

In other words, the initial evolution of the Universe can be explained with linear perturbation theory, but nonlinearity became dominant over time. N-body simulations like the Millenium Simulation (Springel et al., 2005) aided in studying the nonlinear regime. These simulations revealed the so-called cosmic web (Bond et al., 1996), a filamentary structure formed by matter collapsing under gravity. The cosmic web represents the large-scale structure of the Universe, with galaxies and clusters found along interconnected filaments. The diffuse medium that fills the space between galaxies is known as intergalactic medium (IGM). Figure 1.1 shows the cosmic web, as simulated by the Millenium Simulation. This cosmic web has also been observed through measurements of galaxy and clustering in big galaxy surveys (e.g., Tegmark et al. 2004; Cole et al. 2005; Percival et al. 2007; Reid et al. 2010).

### 1.1.3 Fueling galaxy growth

White & Rees (1978) made a significant contribution by introducing a model that connected the growth of structures with the formation of luminous galaxies. They built upon the formalism presented in Press & Schechter (1974) for the non-linear regime, extending the arguments given by Binney (1977), Rees & Ostriker (1977), and Silk (1977). This model proposes that as gas from the IGM falls into the gravitational potential well of dark matter haloes, it undergoes shock heating, reaching a temperature approximately equal to the virial<sup>4</sup> temperature of the halo (of the order of  $10^6$  K), denoted as  $T_{\text{vir}}$ . The virial

<sup>4</sup>The virial theorem states that for a system in equilibrium, the time-averaged total kinetic energy (K)



**Figure 1.1:** Large scale structure of the Universe derived from the Millennium Simulation. It illustrates a projected density field of a  $15 \text{ Mpc h}^{-1}$  thick slice at  $z=0$ . Credit: <https://www.mpa-garching.mpg.de/galform/virgo/millennium/>

temperature characterizes the average temperature of a system in virial equilibrium. For a self-gravitating system, it is expressed as:

$$T_{\text{vir}} = \left( \frac{\mu m_p}{2k} \right) \left( \frac{GM}{R} \right) \quad (1.2)$$

where  $\mu$  is the mean molecular weight of the gas,  $m_p$  is the proton mass,  $k$  is the Boltzmann constant,  $G$  is the gravitational constant,  $M$  is the total mass of the system, and  $R$  is its characteristic size or radius.

The above process occurs as the gas accretes onto the growing haloes through hierarchical clustering. Within these haloes, gas cooled and fragmented, giving rise to the first stars and galaxies, as discussed in Rees & Ostriker (1977). This pioneering model laid the groundwork for modern galaxy formation theories.

In this context, more recent works have extensively studied the deposition of gas into galaxies from the surrounding medium (e.g. Kereš et al. 2005; Dekel & Birnboim 2006; Faucher-Giguère et al. 2011; van de Voort et al. 2011; Anglés-Alcázar et al. 2017; Stevens et al. 2017; Martin et al. 2019). Two main modes of galaxy fueling have been identified: the “hot accretion mode” and the “cold accretion mode”. In the hot mode, gas is first shock-heated to the  $T_{\text{vir}}$  of the halo, and then it dissipates its gravitational and thermal energy before being accreted onto the galaxy. This mode dominates for massive haloes ( $\gtrsim 10^{12} M_{\odot}$ ) at low redshifts ( $z \lesssim 1$ ). In the case of cold-mode accretion, gas is directly

---

is related to the time-averaged total potential energy ( $U$ ) by:  $2K + U = 0$ .

deposited into the galaxy without undergoing virialization. In other words, the gas accreted along filaments with low circular velocities is not expected to undergo shock heating nor to reach the  $T_{\text{vir}}$  of the halo. Instead, this gas is anticipated to flow in a cold state, with temperatures  $\sim 10^4$  K. This mode of accretion is common in lower-mass haloes ( $< 10^{12} M_{\odot}$ ), where the cooling time<sup>5</sup> ( $t_{\text{cool}}$ ) of the gas is shorter than the gravitational free fall time<sup>6</sup> ( $t_{\text{ff}}$ ).

However, simulations and theoretical models have demonstrated the importance of the cold accretion mode even for high mass haloes that contain hot, hydrostatic gas, indicating that both modes can coexist. The fact that the gas preferentially accretes along filaments, where relatively dense gas has low cooling times, enables the cold streams to penetrate the hot haloes. This process is expected to be the dominant mode of mass inflow at high- $z$  ( $z \gtrsim 2$ , e.g. Kereš et al. 2005; Dekel et al. 2009; Kereš et al. 2009; van de Voort et al. 2012; Faucher-Giguère et al. 2011).

In summary, the dominant accretion mode depends on both the halo mass and redshift. At higher redshifts the  $t_{\text{cool}}$  of gas is shorter mainly due to the higher density ( $\rho$ ), as  $t_{\text{cool}} \propto \rho^{-1}$ . Conversely, higher mass haloes have higher  $T_{\text{vir}}$ , resulting in longer  $t_{\text{cool}}$ , especially for temperatures above  $10^6$  K (e.g., Birnboim & Dekel 2003; Katz et al. 2003; Kereš et al. 2005; Crain et al. 2010).

Regardless of the dominant accretion mode of haloes, cold gas clouds within dark matter haloes will eventually undergo gravitational instabilities, leading to its collapse and fragmentation. These fragmented clouds are converted into stars. The ratio between the gas mass converted into stars and the time that has elapsed is commonly referred to as the star formation rate (SFR). As galaxies form stars, several physical mechanisms act on different scales. For instance, star formation process contributes to the growth of galactic bulges.

### 1.1.4 Active Galactic Nuclei

In the central regions of galaxies, the gas may also be accreted onto a supermassive black hole (SMBH), triggering the formation of an active galactic nucleus (AGN) powered by the energy released in the accretion process. It is now widely accepted that the majority of galaxies harbor SMBHs with masses  $> 10^6 M_{\odot}$  (e.g., Magorrian et al. 1998)

---

<sup>5</sup>The  $t_{\text{cool}}$  represents the characteristic timescale for the gas in the halo to radiatively lose its thermal energy. It is expressed as

$$t_{\text{cool}} = \left( \frac{3nkT}{2n_{\text{H}}^2\Lambda} \right)$$

with  $n$  the number density of gas particles,  $n_{\text{H}}$  the number density of hydrogen,  $T$  the gas temperature and  $\Lambda$  the cooling function.

<sup>6</sup>The  $t_{\text{ff}}$  represents the timescale for the gas in the halo to collapse under the influence of gravity, assuming no other forces are significant. It is expressed as

$$t_{\text{ff}} = \sqrt{\frac{3\pi}{32G\rho}}$$

where  $\rho$  is the volumetric gas density of the gas.

AGNs can release an enormous amount of energy and form winds and jets which could have an impact on their host galaxies. These processes are known as “AGN feedback”. Importantly, current models of galaxy evolution require the presence of AGNs to reproduce the stellar mass and luminosity function found in observations, and prevent the formation of over-massive galaxies (e.g., Bower et al. 2006; Schaye et al. 2015; Pillepich et al. 2019). Overall, AGN feedback can influence its host galaxy through two primary mechanisms. Firstly, AGN outflows can have a “positive feedback” effect by enhancing the star formation. This happens by compressing the interstellar medium<sup>7</sup> and molecular clouds (e.g, Ishibashi & Fabian 2012). However, direct evidence of this positive feedback is rare (e.g., Cresci et al. 2015; Shin et al. 2019). On the contrary, AGN can also exert “negative feedback” by employing mechanical energy from winds, outflows, or jets to heat the surrounding interstellar medium. This prevents molecular gas from radiatively cooling or expels gas from the host galaxy, thereby suppressing the star formation (e.g. Binney & Tabor 1995; Ciotti & Ostriker 2001; Croton et al. 2006; Cattaneo et al. 2009; Nesvadba et al. 2010; Heckman & Best 2014). Jets can further heat the halo gas, preventing future star-formation by inhibiting gas cooling (e.g. Ciotti & Ostriker 2001; McNamara & Nulsen 2007). These feedback mechanisms play crucial roles in regulating the growth and evolution of galaxies hosting AGN, and their CGM.

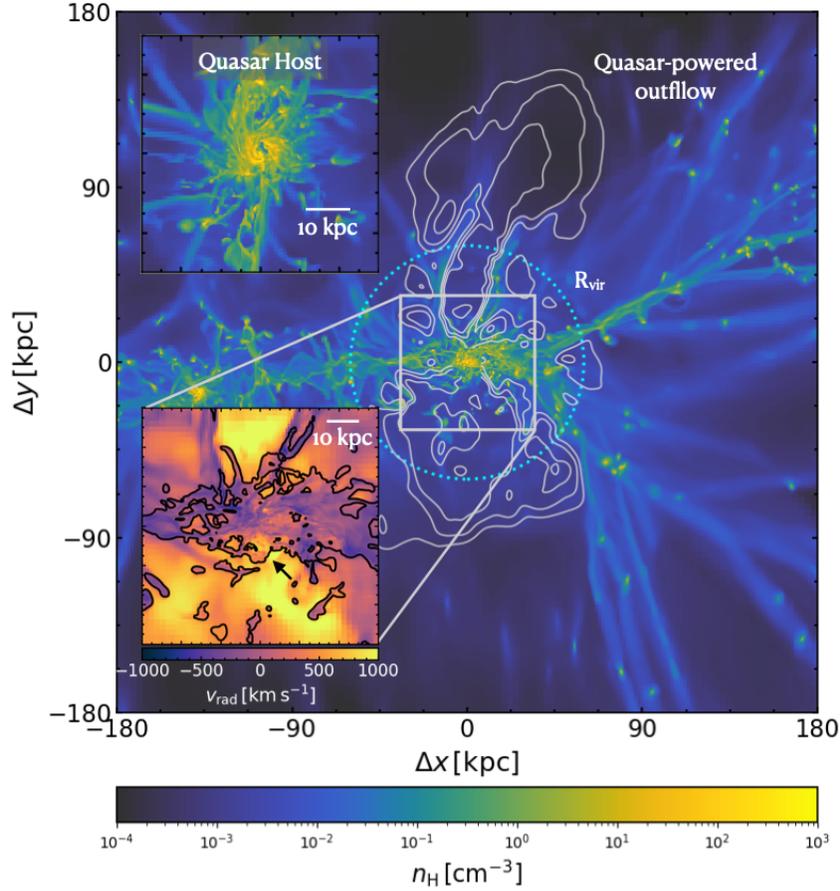
Specifically, part of this thesis focuses on *quasars*, which refers to the most luminous AGN (bolometric luminosities<sup>8</sup>  $L_{\text{bol}}$  up to  $10^{50} \text{erg s}^{-1}$ , Shen et al. 2020) in which we can directly observe the nuclear emission (e.g., Antonucci 1993; Elvis 2000). They inhabit massive halos ( $\sim 10^{12}$ - $10^{13} M_{\odot}$ , e.g., Porciani et al. 2004; Shen et al. 2007; White et al. 2012; Timlin et al. 2018), and their host galaxies appear more massive and star-forming than typical galaxies at their cosmic time. A quasar is able to expel and highly ionize the gas reservoir within the host and its surroundings, thus regulating its own growth (e.g., Di Matteo et al. 2005). Considering the masses of their halos, they are expected to accrete gas in cold mode at  $z \gtrsim 2$ , therefore having both a cool and a warm/hot halo gas phases (e.g., Kereš et al. 2005). Observations of the gas reservoirs surrounding quasars from galaxy to CGM scales are thus key in understanding the role of quasars in galaxy formation and their co-evolution. Figure 1.2 shows a simulated gas density map of a quasar host galaxy and its surrounding medium. It highlights the complex interplay between quasar-driven outflows (white contours) and inflows from filaments (purple regions in left inset plot).

## 1.2 The Circumgalactic Medium (CGM)

The circumgalactic medium (CGM; e.g., Tumlinson et al. 2017) is the extended gaseous halo surrounding galaxies, encompassing the region between the interstellar medium of a

<sup>7</sup>The interstellar medium is the vast expanse of gas, dust, and plasma that fills the space between stars within a galaxy. It consists of various components, including molecular clouds, atomic gas, ionized gas, and dust particles.

<sup>8</sup>Bolometric luminosity refers to the total luminosity emitted by an astronomical object over all wavelengths.



**Figure 1.2:** Gas density projection centred on a quasar host galaxy ( $M_{\text{vir}}=2.4\times 10^{12}M_{\odot}$ ) at  $z\sim 6$ , extracted from cosmological simulations. The white contours are levels of radial velocity (100, 300 and 500  $\text{km s}^{-1}$ ), representing a quasar-driven outflow. The top left inset plot shows the gas density around the host galaxy, in a face-on configuration. The bottom left inset plot shows the radial velocity of gas within the central 12 arcsec (or  $\sim 70$  kpc), revealing large-scale outflows (orange regions) and inflows (purple regions). Figure extracted from [Costa et al. \(2022\)](#).

galaxy and the IGM, extending up to the virial radius ( $R_{\text{vir}}$ ). The virial radius of a dark matter halo is a characteristic radius within which the gravitational effects dominate and the halo is in approximate dynamical equilibrium. It is defined in relation to the critical density ( $\rho_{\text{crit}}$ ) of the Universe (e.g., [Mo et al. 2010](#)):

$$\rho_{\text{crit}} = \frac{3H(z)^2}{8\pi G} \quad (1.3)$$

with  $H(z)$  the Hubble parameter. Then,  $R_{\text{vir}}$  is given by:

$$R_{\text{vir}} = \left( \frac{3M_{\text{vir}}}{4\pi\rho_{\text{crit}}\Delta} \right)^{1/3} \quad (1.4)$$

where  $\Delta$  is the overdensity parameter and  $M_{\text{vir}}$  is the total mass within  $R_{\text{vir}}$ . For purposes of this thesis, we assume  $\Delta=200$  (Bryan & Norman, 1998), therefore we use the notation  $R_{\text{vir}}$  and  $R_{200}$  interchangeably:

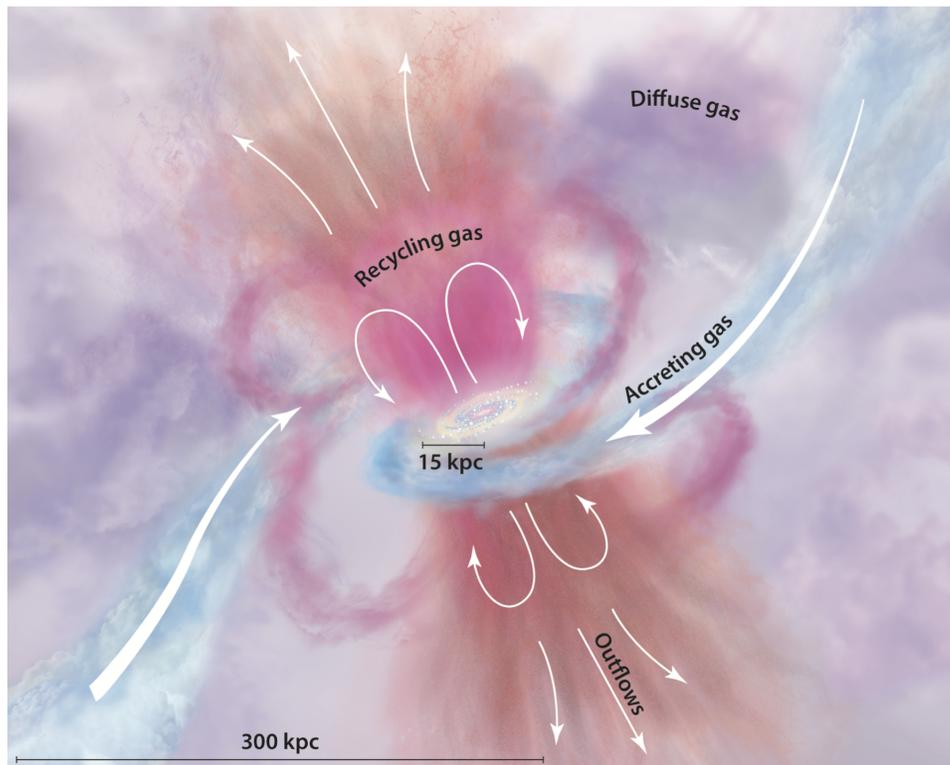
$$R_{200} = R_{\text{vir}} = \left( \frac{3M_{200}}{4\pi\rho_{\text{crit}}200} \right)^{1/3} \quad (1.5)$$

Several processes are relevant for galaxy formation and evolution on CGM scales. These processes include gas inflows from the IGM and/or satellite galaxies, gas outflows from the galaxy, stellar and AGN feedback, interactions with satellites that can lead to tidal stripping and other effects, and mixing processes (see Tumlinson et al. 2017 and references therein). Therefore, CGM facilitates the exchange of baryons between galaxies and their surroundings at larger scales (see Fig. 1.3, e.g. Burchett et al. 2018; Dutta et al. 2020, 2021). Evidence suggests that the CGM gas exhibits a complex multi-phase structure. This includes at least a hot virialized gas phase with  $T\sim 10^6\text{-}10^7$  K (e.g., Anderson et al. 2016; Kaaret et al. 2020; Das et al. 2021) and a cold gas phase with  $T<10^5$  K (e.g., Prochaska et al. 2013; Meiksin et al. 2017; Fossati et al. 2021).

The CGM has been a subject of extensive study around quasars, with investigations primarily focused on the cool phase ( $T\sim 10^4$  K). This has been accomplished through both absorption and emission techniques. Given the diffuse nature expected for the CGM ( $n_{\text{H}} \lesssim 0.1 \text{ cm}^{-3}$ ), in the past this medium has been typically studied by analysing absorption features based on one bright background quasar sightline per foreground halo to gather statistical insights into the physical properties of the CGM surrounding quasars (e.g., Bergeron et al. 2002; Hennawi et al. 2006; Prochaska et al. 2013; Tumlinson et al. 2011; Farina et al. 2013, 2014; Turner et al. 2014; Lau et al. 2018). Making use of the absorption technique with projected quasar pairs at  $z \sim 2\text{-}3$ , studies have provided important results. Firstly, the anisotropic clustering of HI systems around quasars suggests that the ionizing radiation from quasars escapes on CGM scales in a non-uniform or intermittent manner (Hennawi & Prochaska 2007; Jalan et al. 2019). Secondly, these studies have revealed the presence of substantial reservoirs ( $>10^{10} M_{\odot}$ ) of cool, metal-enriched ( $Z \gtrsim 0.1 Z_{\odot}$ ) gas within the CGM, indicating the significant contribution of this CGM phase to the overall gas content and metal budget of halos (Prochaska et al. 2013, 2014; Lau et al. 2016). Lastly, investigations into the kinematics of the CGM suggest that this gas is in virial equilibrium with the dark matter halo, i.e internal motions of the gas (such as inflows or outflows) are balanced by the gravitational potential energy of the halo (Prochaska et al. 2014; Lau et al. 2018). However, Lau et al. (2018) found evidence for the presence of outflowing gas, indicated by an offset of  $\sim 200 \text{ km s}^{-1}$  between the mean absorption centroids and the systemic.

### 1.2.1 CGM as traced by emission lines

Absorption measurements offer limited information on gas surface density, as they provide projected and narrow measures along a single sightline per galaxy due to the rarity of



**Figure 1.3:** Artistic representation of the CGM. Gas that is being accreted from the IGM onto the central galaxy is represented by blue. Gas that is outflowing from the galactic disk is represented by pink and orange. The recycling gas corresponds to gas that has been previously ejected. The purple color represents the complex gas reservoir that consists of a mixture of all these processes. Figure extracted from [Tumlinson et al. \(2017\)](#).

background quasars. For this reason, it is essential to complement these studies with direct emission from the CGM. Theoretical studies on CGM have primarily focused on absorption measurements, while the number of studies predicting CGM line emission is relatively limited (e.g., [Bertone & Schaye 2012](#); [van de Voort & Schaye 2013](#); [Sravan et al. 2016](#); [Augustin et al. 2019](#); [Byrohl et al. 2021](#); [Nelson et al. 2021](#); [Piacitelli et al. 2022](#)).

Lines from different ions serve as tracers of gas in various temperature regimes. Some promising lines that can potentially be detected by current and upcoming optical instruments at high- $z$  are  $\text{Ly}\alpha$  (1216 Å),  $\text{HeII}$  (1640 Å),  $\text{CIII}$  (977 Å),  $\text{Si III}$  (1207 Å) and  $\text{Si IV}$  (1394, 1403 Å) that predominantly trace gas with  $T < 10^5$  K near galaxies, making them valuable tools for studying cold accretion flows and colder regions of galactic winds. On the other hand, lines like  $\text{CIV}$  (1548, 1551 Å),  $\text{Nv}$  (1238, 1243 Å),  $\text{OVI}$  (1032, 1038 Å) and  $\text{NeVIII}$  (770, 780 Å) can trace the warmer gas ( $10^5 \lesssim T \lesssim 10^6$  K), providing insights into the properties of gas in haloes and filaments (e.g., [Bertone & Schaye 2012](#)). Additionally, X-ray emission lines such as  $\text{OVII}$  (22 Å) and  $\text{OVIII}$  (19 Å), are promising to probe the hot-CGM ( $T > 10^6$  K) at low redshift (e.g., [van de Voort & Schaye 2013](#); [Nelson et al. 2023](#)).

Through the analysis of hydrodynamic cosmological simulations, [van de Voort & Schaye](#)

(2013) studied the emission of rest UV metal lines at  $z = 3$ , finding that CIII is the strongest emission line. Also based in cosmological simulations, Bertone & Schaye (2012) investigated the intensity of rest UV lines in the IGM and CGM at  $z \sim 2-5$ , and predicted that the brightest emission line would be from HI Ly $\alpha$ , and the strongest metal line would be CIII. This study emphasized the significant dependence of the emission on the gas density and temperature, particularly in relation to the cooling curves specific to the emitting ion. Corlies et al. (2020) studied emission lines in the CGM at  $z=3$  based on cosmological simulations that forces higher spatial resolution on the CGM. Their results highlight that increasing the resolution of the CGM has a significant impact on the total luminosity of the considered lines (up to two orders of magnitude for CIII and CIV).

In the next sections, I focus on the cool ( $\sim 10^4$  K) gas phase of the CGM, given its relevance to the scope of this thesis.

## 1.2.2 Cool phase of CGM in observations

In the last decade, significant advancements in observational capabilities have revolutionized the study of the CGM gas surrounding quasars, particularly through the use of sensitive integral field unit spectrographs. Instruments such as the Multi-Unit Spectroscopic Explorer (MUSE; Bacon et al. 2010), the Keck Cosmic Web Imager (KCWI; Morrissey et al. 2012) and the Palomar Cosmic Web Imager (PCWI; Matuszewski et al. 2010) have enabled deeper observations of the CGM through emission lines in a reasonable amount of time. In particular, several studies are currently routinely reporting the presence of extended Lyman alpha (hereafter Ly $\alpha$ ) emission around  $z \gtrsim 2$  quasars (e.g., Borisova et al. 2016; Arrigoni Battaia et al. 2019a; Cai et al. 2019; Fossati et al. 2021). Furthermore, the sensitivity of these instruments has allowed for the detection of Ly $\alpha$  haloes around typical star-forming galaxies (e.g., Wisotzki et al. 2016; Leclercq et al. 2017, 2020; Kusakabe et al. 2022). In the next paragraphs I focus on the brightest tracer of the cool gas phase ( $T \sim 10^4$  K) of the CGM, its Ly $\alpha$  emission.

### Ly $\alpha$ line

The Ly $\alpha$  emission line is a spectral line of the hydrogen atom, corresponding to the transition of an electron (e-) from the first excited state ( $2p$ ) to the ground state<sup>9</sup>( $1s$ ). It is placed in the ultraviolet portion of the electromagnetic spectrum, with a rest wavelength ( $\lambda_{\text{rest}}$ ) of 121.56 nm (or frequency  $\nu_{\text{rest}} = 2.45 \times 10^{15}$  Hz), and an Einstein coefficient<sup>10</sup>  $A$  of  $\sim 6.3 \times 10^8 \text{ s}^{-1}$ .

Under typical astrophysical conditions, it is anticipated that Ly $\alpha$  photons (with an energy of  $h\nu = 10.2$  eV) will undergo significant resonant scattering (e.g., Gould & Weinberg 1996). The optical depth of Ly $\alpha$  at the line center, which serves as an indicator of the

<sup>9</sup>The ground state corresponds to the lowest energy level, which has a principal quantum number of  $n = 1$ .

<sup>10</sup>The Einstein coefficient  $A$  describes the rate at which an atom or molecule undergoes a transition from an upper energy level to a lower energy level through spontaneous emission.

number of scatterings a photon encounters, can be mathematically expressed as follows:

$$\tau_0(Ly\alpha) = 8.02 \times 10^4 \frac{15 \text{ km s}^{-1}}{b} \tau(\text{Ly cont}) \quad (1.6)$$

where  $\tau(\text{Ly cont}) = 6.3 \times 10^{-18} \text{ cm}^2 \text{N(H)}$  corresponds to the Lyman continuum optical depth, i.e., the optical depth due to photoionization just above the hydrogen ionization energy<sup>11</sup>, with  $\text{N(H)}$  being the column density of hydrogen (e.g., [Draine 2011](#)). The parameter  $b$  corresponds to the Doppler broadening parameter that includes both thermal and turbulent gas motions:  $b = (v_{\text{thermal}}^2 + v_{\text{turb}}^2)^{-2}$ . Under Case B recombination<sup>12</sup>, it is expected  $\tau(\text{Ly cont}) > 1$ , then it is evident that  $\tau_0(Ly\alpha)$  is very large<sup>13</sup>. Therefore, a  $Ly\alpha$  photon generally undergoes multiple scatterings, being absorbed and re-emitted in different directions, before it eventually escapes the system. This process of resonant scattering contributes to the complex radiative transfer of  $Ly\alpha$  photons in astrophysical environments.

### Ly $\alpha$ nebulae

As mentioned at the beginning of this Section, extended  $Ly\alpha$  emission can be observed nowadays around  $z \gtrsim 2$  quasars. These so-called  $Ly\alpha$  nebulae extend up to hundreds of kiloparsecs and have diverse morphologies ([Husband et al. 2015](#); [Fumagalli et al. 2016](#); [Borisova et al. 2016](#); [Arrigoni Battaia et al. 2019a](#); [Cai et al. 2019](#); [O’Sullivan et al. 2020](#); [Mackenzie et al. 2021](#); [Fossati et al. 2021](#)). The bulk of the extended emission traces gas on a few tens of kpc near the quasars, while large-scale structures extending to  $R \gtrsim 80$  kpc are seen at lower surface brightness ( $\text{SB}_{Ly\alpha} \sim 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-2}$ ). Figure 1.4 shows an example of observed  $Ly\alpha$  nebulae at different  $z$ . In most of the cases, the  $Ly\alpha$  kinematics in the extended nebulae is consistent with gravitational motions in haloes with masses consistent with typical quasar hosts ( $\sim 10^{12.5} M_{\odot}$ , e.g., [Timlin et al. 2018](#)), with full width at half maximum<sup>14</sup> (FWHM) of  $\sim 600 \text{ km s}^{-1}$  (e.g., [Arrigoni Battaia et al. 2019a](#)).

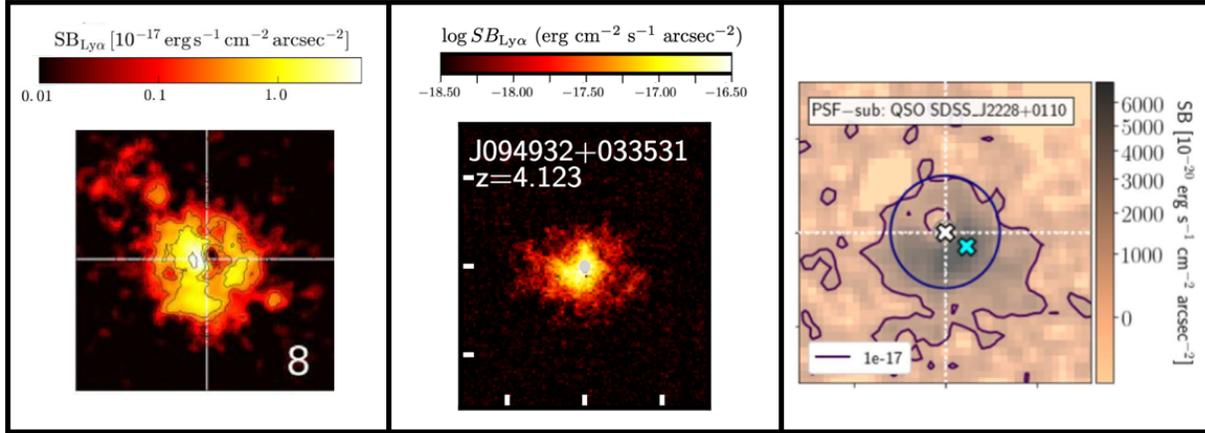
The origin of  $Ly\alpha$  emission in these nebulae remains a topic of ongoing discussion. While in previous studies it has been commonly assumed that this  $Ly\alpha$  emission is primarily a result of photoionizing radiation from the quasar (e.g., [Heckman et al. 1991](#); [Cantalupo et al. 2014](#)), the relative contribution of different mechanisms is still debated. The following list of possible mechanisms has been proposed:

<sup>11</sup>The ionization energy of hydrogen corresponds to the energy required to remove an electron from a hydrogen atom, which is approximately 13.6 eV or a wavelength of 912 angstroms

<sup>12</sup>Recombination refers to the radiative decay of an e<sup>-</sup> from an unbound to a bound state of an atom. For hydrogen, [Baker & Menzel \(1938\)](#) proposed two limiting cases: *Case A*, where ionizing photons (above 13.6 eV) emitted during the recombination escape, and *Case B*, where the ionizing photons are immediately reabsorbed. For Case B conditions, all of the Lyman series transitions are optically thick.

<sup>13</sup>The optical depth  $\tau=1$  represents the threshold to define a medium as optically thin ( $\tau \ll 1$ ) and optically thick ( $\tau > 1$ ). In the optically thin regime, photons can easily pass through the medium without significant attenuation. In the optically thick medium, the material strongly absorbs or scatters radiation, causing a significant reduction in the intensity or flux of the photons.

<sup>14</sup>The FWHM is commonly used to characterize the intrinsic linewidth of an emission or absorption feature, indicating the range of velocities or frequencies over which the line is spread.



**Figure 1.4:** Examples of Ly $\alpha$  surface brightness maps for extended Ly $\alpha$  nebulae around high- $z$  quasars. Left:  $z \sim 3$ , image of 30 arcsec  $\times$  30 arcsec (or  $\sim 230$  kpc  $\times$  230 kpc). Center:  $z \sim 4$ , image of  $\sim 43$  arcsec  $\times$  43 arcsec (or  $\sim 300$  kpc  $\times$  300 kpc). Right:  $z \sim 6$ , image of  $\sim 6$  arcsec  $\times$  6 arcsec (or  $\sim 35$  kpc  $\times$  35 kpc). From left to right, figures adapted from Arrigoni Battaia et al. (2019a), Fossati et al. (2021) and Drake et al. (2019), respectively.

- Photoionization; it refers to ionizing radiation from the AGN and/or star formation that excites neutral hydrogen atoms, which then recombine and emit Ly $\alpha$  photons (e.g., Heckman et al. 1991; Smith et al. 2009; Geach et al. 2009; Cantalupo et al. 2014). Photoionization from other sources, such as active companions, is also expected to contribute to the Ly $\alpha$  emission (e.g., Husemann et al. 2018; Arrigoni Battaia et al. 2019b).
- Shocks powered by outflows; in this scenario, superwinds driven by starburst supernovae or the AGN would produce Ly $\alpha$  emission in the swept up material (e.g., Taniguchi & Shioya 2000; Ohyama et al. 2003; Wilman et al. 2005; Mori & Umemura 2006).
- Gravitational cooling radiation; which is produced in cool ( $\sim 10^4$ - $10^5$  K) accretion streams (e.g., Haiman et al. 2000; Dijkstra et al. 2006; Faucher-Giguère et al. 2010; Rosdahl & Blaizot 2012). This scenario is supported by results from simulations that show that cold accretion mode is important in fueling galaxy formation (e.g., Kereš et al. 2005).
- Resonant scattering of Ly $\alpha$ ; where centrally produced Ly $\alpha$  photons undergo multiple scatterings within the gas distribution surrounding the quasar (e.g., Dijkstra & Loeb 2008; Hayes et al. 2011; Cen & Zheng 2013; Cantalupo et al. 2014).

Disentangling each role of these mechanisms by using Ly $\alpha$  only is difficult, considering that the presence of Ly $\alpha$  photons within the CGM is known to be influenced by resonant scattering. Resonant scattering redistributes Ly $\alpha$  photons both in space and frequency,

making it difficult to interpret the origin of this transition without the use of additional tracers (e.g., [Dijkstra 2017](#); [Gronke et al. 2017](#)).

In a recent work conducted by [Costa et al. \(2022\)](#), they used zoom-in simulations of a quasar host halo at  $z=6$  and found a significant connection between AGN feedback and Ly $\alpha$  nebula. The findings suggest that quasar-powered outflows play a crucial role in creating a pathway for Ly $\alpha$  flux to escape and remain in resonance with the gas in the halo. Their results also showed that the inclination of the host galaxy affects how we observe the extended Ly $\alpha$  emission, because of dust absorption and scattering effects. When the galaxy is seen edge-on, we expect to see fainter and more asymmetric Ly $\alpha$  nebulae due to quasar-powered outflows along the galaxy’s minor axis. This is because the AGN outflows create a pathway with least resistance for the propagation of Ly $\alpha$  emission along the host galaxy minor axis. Conversely, when the galaxy is seen face-on, we expect brighter and more symmetric Ly $\alpha$  nebulae.

At large scales ( $\gtrsim 50$  kpc of projected distances), it has been found in some systems that resonant scattering does not affect the gas (e.g., [Cai et al. 2017](#); [Arrigoni Battaia et al. 2018b](#)). While Ly $\alpha$  emission alone is not sufficient to determine the amount and density of gas in nebulae associated with quasars, follow-up observations in additional line diagnostics such as HeII (1640 Å) and CIV (1548, 1551 Å) (e.g., [Arrigoni Battaia et al. 2015](#); [Travascio et al. 2020](#); [Zhang et al. 2023](#)), along with constraints on  $N_{\text{H}}$  from absorption studies and detailed photoionization models<sup>15</sup>, provide valuable information. In particular, HeII is used to constrain the density of the emitting gas, while the ratio HeII/CIV is used as a metallicity indicator. These additional measurements (that neglect resonant scattering) revealed that both the brightest and largest nebulae, as well as the more typical Ly $\alpha$  nebulae discovered around  $z \sim 3$  quasars, indicate the presence of a significant reservoir ( $\gtrsim 10^{10} M_{\odot}$ ) of cool and dense ( $n_{\text{H}} \gtrsim 1 \text{ cm}^{-3}$ ) gas within these massive halos.

### 1.2.3 Cool dense CGM from theoretical perspective

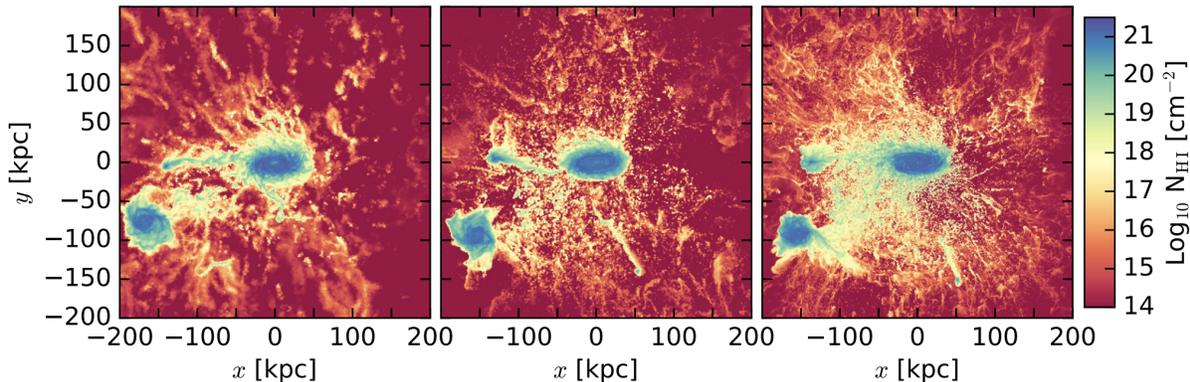
According to simulations, the expected densities in the CGM are generally low, with values on the order of  $n_{\text{H}} \lesssim 0.1 \text{ cm}^{-3}$  (e.g., [Oppenheimer et al. 2016](#); [Nelson et al. 2020](#)). Therefore, the presence of dense CGM gas found in observations was unexpected, as such high densities had been typically found in the interstellar medium of galaxies. In addition, the presence of cool halo gas is not completely understood, and cosmological simulations in the past used to under-predict its abundance (e.g., [Fielding et al. 2017](#); [Liang et al. 2016](#)). In general, determining the underlying physical processes in the CGM is challenging mainly due to the complexity of simulating small-scale physics (e.g., [Scannapieco & Brügggen 2015](#); [Mandelker et al. 2019](#)).

However, currently cosmological simulations are forcing higher resolutions in the CGM, following the formation and survival of such dense gas (e.g., [Hummels et al. 2019](#); [Peeples et al. 2019](#); [van de Voort et al. 2019](#); [Nelson et al. 2020](#)). For instance, [Nelson et al. \(2020\)](#)

<sup>15</sup>Photoionization models are computational tools used to study the ionization and excitation state of the gas in the presence of radiation sources. They can provide information about the ionization fractions, emission line spectra, and other observable properties of the gas.

used IllustrisTNG massive halos ( $\sim 10^{13}$ - $10^{13.5} M_{\odot}$ ) at  $z = 0.5$ , to study the abundance and origin of cool gas in the CGM. Despite these massive halos containing a hot ( $\sim 10^7$  K) virialized plasma, their results show also a cool gas component, which consist in thousands of small ( $\sim$ kpc scales) clouds. On the other hand, [van de Voort et al. \(2019\)](#) applied a refinement technique to simulate the CGM in zoom-in simulations of a Milky Way-mass galaxy, at uniform spatial resolution of up to  $\sim 1$  kpc. They found a CGM with significantly more small-scale structure at higher resolutions, also with more dense clumps and smaller turbulent eddies (see Fig. 1.5). Despite these advancements, the physics of the dense CGM gas is still largely unresolved by cosmological simulations

Finally, “cloud-crushing” simulations at resolution  $< 1$  pc (e.g., [Gronke & Oh 2018, 2020](#); [Kanjilal et al. 2021](#)) have studied this cool dense gas that shatters in smaller structures, entrains and survives in the hot medium, providing theoretical support to the observations. In brief, the acceleration of a cool cloud in a hot ( $\gtrsim 10^6$  K) wind is known as the cloud crushing problem (e.g., [Cowie & McKee 1977](#); [Klein et al. 1994](#); [Xu & Stone 1995](#)). The cloud-crushing or destruction time of a cool cloud with radius  $r_{\text{cl}}$  and density  $\rho_{\text{cl}}$ , which is affected by a wind with  $r_{\text{wind}}$  and  $\rho_{\text{wind}}$ , is  $t_{\text{cc}} = \chi^{1/2} r_{\text{cl}} / v_{\text{wind}}$ , with  $\chi = \rho_{\text{cl}} / \rho_{\text{wind}}$  the overdensity of the cloud. On the contrary, the acceleration time-scale of the cloud is  $t_{\text{drag}} \sim \chi^{1/2} t_{\text{cc}}$ . In the case of overdensities of  $\chi \sim 100$ - $1000$ , which correspond to cool clouds embedded within a hot wind,  $t_{\text{cc}} \ll t_{\text{drag}}$ . This implies that the cool gas present in these clouds will be destroyed before it has the chance to be accelerated. To solve this problem, a criteria based on efficient radiative cooling, specifically in the cooling time of the mixed gas  $t_{\text{cool,mix}}$  (which depends on the temperature and density of the mixed gas), has been proposed. When  $t_{\text{cool,mix}} < t_{\text{cc}}$ , the mixed gas cools quickly and increases the total cool gas mass. When  $t_{\text{cool,mix}} > t_{\text{cc}}$ , the survival of the cool gas will depend of the cloud size. In this way, the surviving clouds can grow by an order of magnitude in mass, all the while gaining momentum from the surrounding hot wind.



**Figure 1.5:** Neutral hydrogen column density map in a 600 kpc column, of a Milky Way-mass galaxy at  $z=0$ . From left to right, the resolution of the simulation increases, being the highest resolution (third panel) 1 kpc. We can note that higher resolution in the halo leads in more small-scale structure. Figure adapted from [van de Voort et al. \(2019\)](#).

## 1.3 Cold gas reservoirs in high- $z$ galaxies, from galactic to halo scales

In the previous section, I focused mostly in the cool gas phase of the CGM, as traced by Ly $\alpha$  emission. As already mentioned, the presence and survival of this dense halo gas seems plausible from cloud crushing simulations. Therefore, questions about the possible presence of even molecular gas on such large scales naturally arise. In this context, of importance is also to investigate the link between the CGM and galactic scales. This can be done through the study of the cold phase ( $T < 100$  K), which serves as the fuel for star formation. This gas phase can be detected by using (sub)millimeter observations of bright carbon monoxide  $^{12}\text{CO}$  (hereafter CO) emission lines or fine structure lines such as [CII]  $158\mu\text{m}$  (hereafter [CII]).

In this thesis, I also study these cold gas reservoirs in and around high- $z$  galaxies. In this context, in this Section I describe the different phases of the interstellar medium, cold gas tracers and works that are relevant for this thesis.

### 1.3.1 The Interstellar Medium (ISM)

The interstellar medium (ISM) is a relevant component of galaxies, since it is where star-formation occurs. Broadly speaking, it includes the gas, dust and plasma between stars. In the ISM is where most baryons and electrons ( $\sim 10$  per cent of the total mass) that determine the visible appearance of galaxies are located (Draine, 2011). For purposes of this thesis, here I will describe the characteristic phases of the ISM, which are important to understand the cold gas reservoirs in and around high- $z$  galaxies studied here. These phases, that account for most of the mass and volume of the ISM, consist of:

- Coronal gas: corresponds to gas that has been shock-heated by supernova explosions. It has temperatures of  $T \gtrsim 10^{5.5}$  K and low hydrogen densities of  $n_{\text{H}} \sim 4 \times 10^{-3} \text{ cm}^{-3}$ . It is also known as “hot ionized medium”.
- HII gas: corresponds to gas mainly composed by ionized hydrogen. Specifically, here the hydrogen has been photoionized by UV photons from young hot stars (mostly O-type stars). This gas can be classified in two types: “HII regions”, where the photoionized gas is dense material from a nearby cloud, and “diffuse HII”, which refers to low density intercloud medium. HII gas has typical values of  $T \sim 10^4$  K, and  $n_{\text{H}} \sim 0.2 - 10^4 \text{ cm}^{-3}$
- Warm HI: mostly composed by atomic gas. Also referred to as “warm neutral medium”. It is characterized by  $T \sim 5000$  K, and  $n_{\text{H}} \sim 0.6 \text{ cm}^{-3}$ .
- Cool HI: mostly composed by atomic gas, with  $T \sim 100$  K, and  $n_{\text{H}} \sim 30 \text{ cm}^{-3}$ . It is also known as “cold neutral medium”.

- Diffuse molecular gas: this phase is in principle similar to the cool HI, but with larger densities that promote the abundance of H<sub>2</sub> molecules through self-shielding<sup>16</sup>. It has typical values of  $T \sim 50$  K and  $n_{\text{H}} \sim 100 \text{ cm}^{-3}$ .
- Dense molecular gas: corresponds to gravitationally bound dense clouds. Within these regions, star formation takes place. Often, these molecular clouds are “dark”, with high visual extinction<sup>17</sup> of  $A_V \gtrsim 3$ <sup>18</sup>. They are characterized by  $T \sim 10 - 50$  K and  $n_{\text{H}} \sim 10^3 - 10^6 \text{ cm}^{-3}$ .
- Stellar outflows: corresponds to mass loss from stars. In particular, evolved stars lead to high density outflows, with low outflows velocities of  $\lesssim 30 \text{ km s}^{-1}$ . On the other hand, hot stars can produce less dense and faster outflows.

The temperature and density of the ISM are influenced by the surrounding environment and are regulated by a balance between heating and cooling processes. Various mechanisms contribute to the heating of molecular gas (e.g., [Goldsmith & Langer 1978](#)). Within dark molecular clouds, cosmic rays, which consist of high-energy protons and electrons, ionize H<sub>2</sub> molecules. The resulting free e- transfer their excess kinetic energy to other H<sub>2</sub> molecules, thereby heating the gas.

In molecular clouds where active star formation occurs, UV heating is an additional process that contributes to the heating of molecular gas. In this scenario, the intense radiation emitted by O and B stars, which are prevalent in star-forming regions, creates “photodissociation regions” (hereafter PDRs, [Hollenbach & Tielens 1999](#)). These PDRs correspond to the interface zone between an HII region and a dense molecular cloud. A PDR is characterized by an ionization front, which marks the boundary where hydrogen is 50 per cent ionized, and a photodissociation front, where hydrogen is 50 per cent atomic and 50 per cent molecular. In this region, molecular gas flows towards the photodissociation front, where it is dissociated, and the atomic gas flows away from the photodissociation front towards the ionization front ([Draine, 2011](#)). Figure 1.6 shows the structure of a PDR.

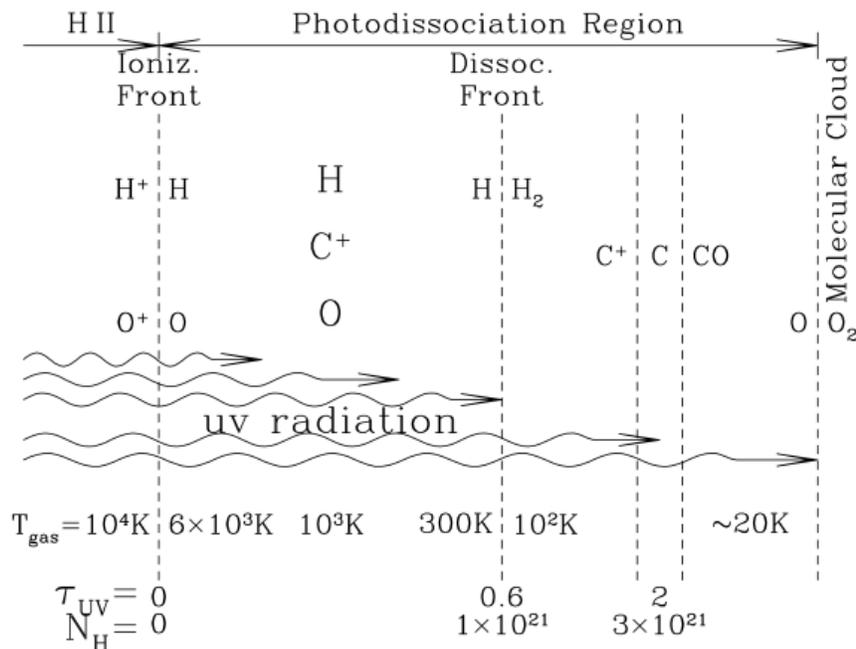
In addition to UV heating in PDRs, there is another type of heating that occurs through X-ray emission. This is known as an X-ray-dominated region (XDR) and is associated with the presence of AGN or hot plasmas heated by supernovae. X-ray radiation can penetrate deeper into the molecular clouds, compared to UV radiation (e.g., [Meijerink et al. 2006](#)), contributing to the heating of the molecular gas.

Regarding gas cooling, atomic fine-structure lines (FSLs) play a significant role in cooling interstellar gas in star-forming galaxies ([Spitzer, 1978](#)). They act as dominant coolants, particularly in cooler regions where hydrogen’s permitted lines cannot be excited ( $T < 10^4$  K). These FSLs are forbidden transitions, which means they are typically optically

<sup>16</sup>The self-shielding effect refers to the phenomenon where a gas medium, illuminated by a light source, exhibits a protective mechanism against dissociation of molecules. This occurs when specific wavelengths of light are absorbed by the gas, allowing the molecules to remain intact.

<sup>17</sup>Extinction refers to the process of absorption and scattering of electromagnetic radiation by dust and gas present between an emitting source and the observer.

<sup>18</sup> $A_V$  is the total extinction in the photometric  $V$  (550 nm) band.



**Figure 1.6:** Representation of the structure of a PDR. Figure extracted from [Draine \(2011\)](#).

thin and do not experience line trapping due to resonant scattering in regions with high column densities. Some FSLs have ionization potentials higher than hydrogen and primarily cool the ionized medium, such as [NII] and [OIII]. Other FSLs have lower ionization potentials in the far-infrared range, allowing them to also trace the neutral ISM. Examples of these FSLs include [CII], [OI], and [CI].

### 1.3.2 Cold gas tracers

#### Emission line luminosities

The line luminosities for the cold gas tracers at high- $z$  are usually expressed in two different ways: as the source luminosity in units of  $L_{\odot}$  ( $L_{\text{line}}$ ), or as the integrated source brightness temperature in units of  $\text{K km s}^{-1} \text{ pc}^2$  ( $L'_{\text{line}}$ ). Following [Solomon et al. \(1992\)](#):

$$L'_{\text{line}} = 3.25 \times 10^7 \times S_{\text{line}} \Delta v \frac{D_L^2}{(1+z)^3 \nu_{\text{obs}}^2} (\text{K km s}^{-1} \text{ pc}^2) \quad (1.7)$$

$$L_{\text{line}} = 1.04 \times 10^{-3} \times S_{\text{line}} \Delta v D_L^2 \nu_{\text{obs}} (L_{\odot}) \quad (1.8)$$

where  $S_{\text{line}} \Delta v$  corresponds to the velocity integrated flux of the line (in  $\text{Jy km s}^{-1}$ ),  $D_L$  is the luminosity distance<sup>19</sup> (in Mpc), and  $\nu_{\text{obs}}$  corresponds to the observed frequency (in

<sup>19</sup>The luminosity distance is defined as the distance at which an object with known intrinsic luminosity would have the same observed flux as the object being observed. The luminosity distance ( $D_L$ ) is related to the observed flux ( $F$ ) and the intrinsic luminosity ( $L$ ) of an object through:  $F = L/(4\pi D_L^2)$ .

Ghz).

The luminosity  $L_{line}$  is usually used when calculating the cooling capability, for example in relation to the far infrared (FIR) luminosity ( $L_{line}/L_{FIR}$ ) (e.g., [Hailey-Dunsheath et al. 2010](#); [Stacey et al. 2010](#)). On the other hand,  $L'_{line}$  is frequently used to estimate CO luminosities, which then are used to obtain molecular masses through a conversion factor that relates CO(1-0) luminosity to the total molecular gas mass (e.g., [Bolatto et al. 2013](#)). A comprehensive review of various methods for estimating cold gas masses and the challenges associated with them has been provided in [Tacconi et al. \(2020\)](#).

### Carbon monoxide CO rotation transitions

Molecular gas in galaxies is primarily composed of molecular hydrogen ( $H_2$ ), which lacks a permanent dipole moment and has high excitation requirements for its lowest transitions. Therefore, directly studying  $H_2$  emission is limited to a very small fraction of the molecular gas. Instead, emission from tracer molecules has been commonly used. In this context, Carbon monoxide is the preferred tracer due to several reasons. Firstly, it is the most abundant molecule after  $H_2$  in galaxies, therefore experiences frequent collisions with  $H_2$ . Additionally, CO has low excitation requirements, with its first excited state being accessible at temperatures of approximately 5 K. Moreover, CO can be conveniently observed from the ground using the 3-mm band at low- $z$  ( $z < 0.3$ ), particularly in its ground transition ([Carilli & Walter, 2013](#)).

The transitions of CO are described by quantum mechanical principles, specifically the laws governing the rotational energy levels of molecules. The rotational energy levels of a diatomic molecule like CO can be approximated by the rigid rotor model, which assumes that the molecule behaves as a rigid rotor with a fixed moment of inertia.

The energy of a rotational state ( $E(J)$ ) of a CO molecule can be expressed as (e.g., [Hollas 2002](#)):

$$E(J) = BJ(J + 1) \quad (1.9)$$

where  $E(J)$  is the energy of the rotational state with quantum number  $J$ , and  $B$  is the rotational constant of the molecule. The rotational constant depends on the mass of the molecule and the bond length between the carbon and oxygen atoms. The transition between two rotational states is accompanied by the emission or absorption of a photon with a specific frequency or wavelength. The frequency of the transition is given by:

$$\nu(J, J') = 2B(J' - J), \quad J' = 0, 1, 2, 3, \dots \quad (1.10)$$

where  $\nu(J, J')$  is the frequency of the transition between rotational states  $J$  and  $J'$ .

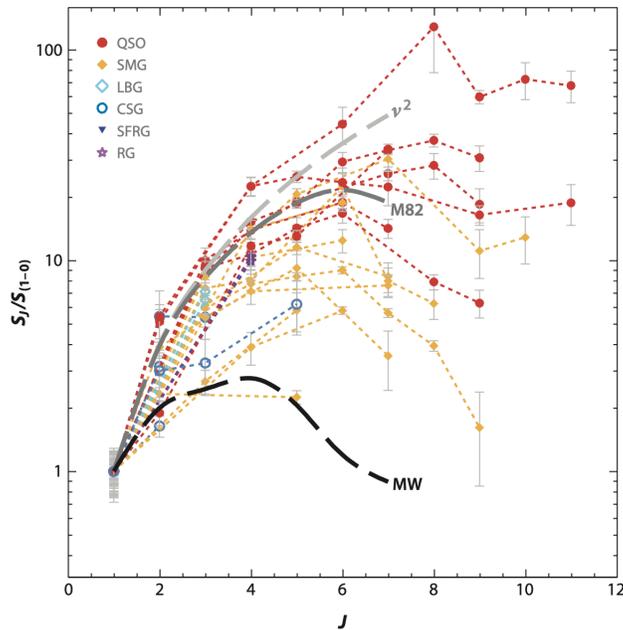
Rotational transitions in molecules are commonly characterized by the critical density ( $n_{crit}$ ) which represents the density at which collisional excitation and spontaneous radiative deexcitation reach a balance. It is expressed as  $n_{crit} = A/\gamma$ , with  $A$  the Einstein coefficient for spontaneous emission, and  $\gamma$  is the collision rate coefficient.  $A$  is solely determined by the molecular properties and is proportional to the cube of the transition frequency. This means that transitions with higher  $J$  values have higher deexcitation rates. On the other

Species	Transition	Excitation potential (K)	$\lambda_{\text{rest}}$ ( $\mu\text{m}$ )	$\nu_{\text{rest}}$ (GHz)	A ( $\text{s}^{-1}$ )	$n_{\text{crit}}$ ( $\text{cm}^{-3}$ )
CO	J = 1-0	5.5	2601	115.27	$7.2 \times 10^{-8}$	$2.1 \times 10^3$ <sup>a</sup>
CO	J = 2-1	16.6	1300	230.54	$6.9 \times 10^{-7}$	$2.1 \times 10^4$ <sup>a</sup>
CO	J = 3-2	33.2	867	345.80	$2.5 \times 10^{-6}$	$4.0 \times 10^4$ <sup>a</sup>
CO	J = 4-3	55.3	650.3	461.04	$6.1 \times 10^{-6}$	$6.0 \times 10^5$ <sup>a</sup>
CO	J = 5-4	83.0	520.2	576.27	$1.2 \times 10^{-5}$	$2.4 \times 10^5$ <sup>a</sup>
CO	J = 6-5	116.2	433.6	691.47	$2.1 \times 10^{-5}$	$3.1 \times 10^5$ <sup>a</sup>
CO	J = 7-6	154.9	371.7	806.65	$3.4 \times 10^{-5}$	$7.2 \times 10^6$ <sup>a</sup>
[CI]	$^3\text{P}_1 - ^3\text{P}_0$	23.5	609.14	492.16	$7.9 \times 10^{-8}$	492.5 <sup>b</sup>
[CI]	$^3\text{P}_2 - ^3\text{P}_1$	62.5	370.42	809.34	$2.7 \times 10^{-7}$	$2.8 \times 10^3$ <sup>b</sup>
[CII]	$^2\text{P}_{1/2} - ^2\text{P}_{3/2}$	91.2	157.74	1900.5	$2.3 \times 10^{-6}$	$4.8 \times 10^3$ <sup>a</sup>
						$2.4 \times 10^3$ <sup>b</sup>
						16 <sup>c</sup>

**Table 1.1:** Excitation potentials, rest wavelengths ( $\lambda_{\text{rest}}$ ) and rest frequencies ( $\nu_{\text{rest}}$ ) were extracted from the Leiden Atomic and Molecular Database (LAMDA) as accessed on 2023 July 02 (Schöier et al., 2005). The CO and [CI] critical densities were calculated using the Einstein A coefficients and collision rates from the LAMDA database, considering a temperature of 100 K. The [CII] critical densities were extracted from Goldsmith et al. (2012), for a temperature of 500 K. **Notes:** [a] Colliding partners are H<sub>2</sub> molecules; [b] Colliding partners are H atoms; [c] Colliding partners are electrons.

hand,  $\gamma$  is dependent on the temperature of the gas (Carilli & Walter, 2013). Table 1.1 shows different CO transitions (up to CO(7-6), which covers the range of interest for this thesis) and their main parameters.

Importantly, by observing multiple CO transitions (i.e., the CO spectral line energy distribution - SLED), different physical properties of the molecular gas, such as temperatures and densities can be constrained (e.g., Weiß et al. 2005). As an example, Fig. 1.7 shows the CO SLED for different type of objects (see caption). In particular (and important for this thesis), quasars (red) show the most excited CO transitions, with an average peak around  $J \sim 6$ . Recent works have found quasars even more excited, with a peak in the CO SLED around  $J \sim 15$  (e.g., Pensabene et al. 2021).



**Figure 1.7:** CO SLED for different sources. The CO line flux (y-axis) has been normalized to the CO(1-0) line. The black line represents the results for the Milky Way, the yellow line and dots the results for submillimeter galaxies (SMGs), while the red line and dots represent the values for quasars. Figure extracted from Carilli & Walter (2013).

### Atomic Carbon [CI] fine structure lines

An alternative molecular gas tracer is atomic carbon (CI). Different studies mainly focused on molecular clouds, galactic disk and center, and nearby galaxies have found that CI and CO emission are closely linked (e.g., Stutzki et al. 1997; Israel & Baas 2002). CI has three fine structure levels, denoted as  $^3P_0$ ,  $^3P_1$ , and  $^3P_2$ , which correspond to different total angular momentum values and have different energies associated with them (see Fig. 1.8). These fine structure levels arise from the splitting of energy levels due to the interaction between the electron spin and the orbital angular momentum of the atom.

The FSL of atomic carbon, namely  $[\text{CI}]({}^3\text{P}_1-{}^3\text{P}_0)$  (at  $609.13 \mu\text{m}$ ) and  $[\text{CI}]({}^3\text{P}_2-{}^3\text{P}_1)$  (at  $370.41 \mu\text{m}$ ) (hereafter  $[\text{CI}](1-0)$  and  $[\text{CI}](2-1)$ , respectively), are generally optically thin (e.g., Ikeda et al. 2002). They have excitation energies of 23.6 K and 62.5 K respectively, being the critical density of  $[\text{CI}](1-0)$  is similar to that of  $\text{CO}(1-0)$ , with a value of approximately  $10^3 \text{ cm}^{-3}$  (see Table 1.1).

The  $[\text{CI}]$  fine structure lines are of particular interest as they provide valuable information about the excitation temperature ( $T_{ex}$ ) of the gas in which they are observed. The  $T_{ex}$  is a measure of the relative population of atoms or ions in different energy states within a given system. The intensity of the line is related to the population difference between the upper and lower energy levels involved in the transition.

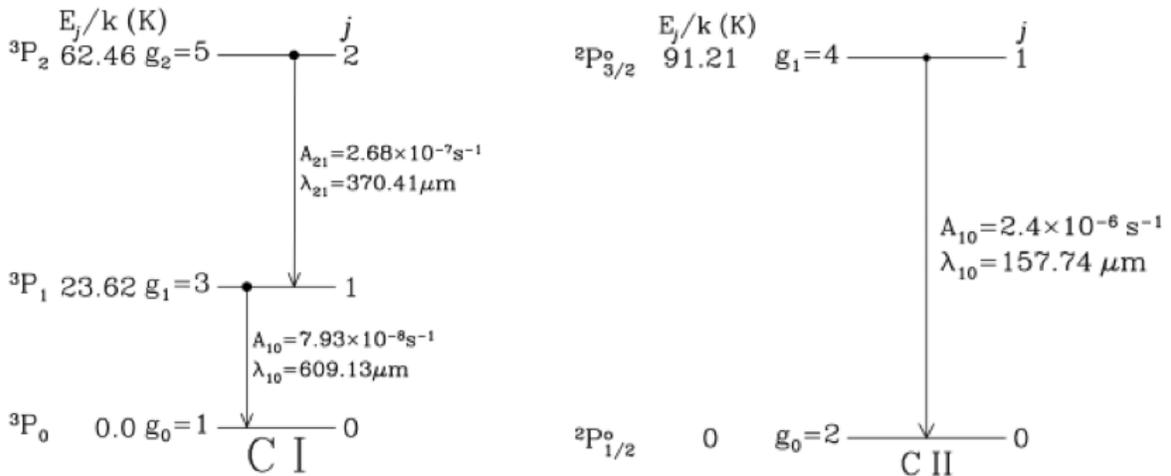
The population difference can be quantified using the Boltzmann equation:

$$\frac{N_u}{N_l} = \frac{g_u}{g_l} \cdot \exp\left(-\frac{\Delta E}{kT_{ex}}\right) \quad (1.11)$$

where  $N_u$  and  $N_l$  are the number densities of atoms or ions in the upper and lower energy levels,  $g_u$  and  $g_l$  are the statistical weights of the respective levels,  $\Delta E$  is the energy difference between the levels ( $= h\nu$ ),  $k$  is the Boltzmann constant. Then, following Schneider et al. (2003),  $T_{ex}$  is simplified to:

$$T_{ex} = 38.8 \times \ln\left(\frac{2.11}{R}\right)^{-1} \quad (1.12)$$

with  $R = L'_{[\text{CI}](2-1)}/L'_{[\text{CI}](1-0)}$ .



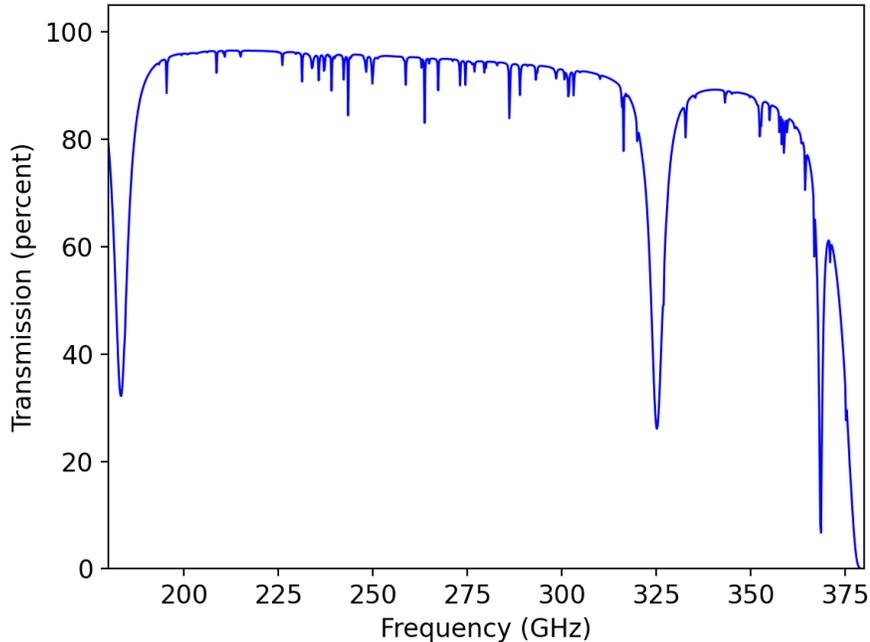
**Figure 1.8:** Representation of fine-structure levels of CI (left panel) and  $\text{C}^+$  (right panel). Figure adapted from Draine (2011).

## Ionized Carbon [CII] fine structure line

The  $C^+$  ion contains only two fine-structure levels, denoted as  $^2P_{1/2}$  and  $^2P_{3/2}$  (see Fig. 1.8, right panel). The ionization potential of atomic carbon is 11.3 eV, which is lower than the one of hydrogen (13.6 eV). The [CII] FSL at  $157.7\mu\text{m}$  ( $\nu_{rest} = 1900.54$  GHz) is the main coolant of the cool ( $\sim 100$  K) interstellar gas, and has become promising to investigate the gas physical conditions in the distant Universe (e.g., [Maiolino et al. 2009](#); [Hailey-Dunsheath et al. 2010](#); [Gallerani et al. 2012](#); [Ferkinhoff et al. 2014](#)). One of the reasons for its popularity is that the [CII] line is easily accessible from the ground. The Earth's atmosphere allows us to access at best the [CII] transition when it is redshifted in the frequency range 211.2–345.5 GHz, which corresponds to the redshift range  $4.5 < z < 8.5$ . This can be appreciated in Figure 1.9, which shows the atmospheric transmission for ALMA, between 180 and 380 GHz.

The [CII] transition is a result of the interaction between the ISM gas and far-ultraviolet (FUV) photons. The emission arises from the collisional excitation of  $C^+$  by electrons (e-), H atoms and  $H_2$  molecules, and its critical densities vary over three orders of magnitude depending on the temperature (see Table 1.1 for values of  $n_{crit}$  and other parameters). Considering that  $C^+$  is the dominant form of carbon under a wide range of conditions, the colliding partners previously mentioned can all be important, allowing it to originate from different phases of the ISM, such as HII regions, PDRs and diffuse molecular gas. Thus, the interpretation of the origin of the observed [CII] emission is challenging, and it will depend on various environmental factors like FUV radiation intensity, metallicity, cosmic-ray ionization rate (e.g., [Bisbas et al. 2021](#)), and X-ray intensity (e.g., [Mackey et al. 2019](#)). Furthermore, the analysis of [CII] is complicated considering that the determination of the optical depth of the line is difficult (e.g., [Neri et al. 2014](#)).

For the galactic plane of the Milky Way, [Pineda et al. \(2014\)](#) estimated that  $\sim 80$  per cent of the [CII] emission comes from atomic and molecular regions, while 20 per cent originates from ionized gas. In local star-forming galaxies, [Croxall et al. \(2017\)](#) found that approximately 60–80 per cent of the [CII] emission originates from neutral gas. At high- $z$ , it is often assumed that the [CII] emission arises predominantly from PDRs, based on observational and theoretical constraints. For instance, [Stacey et al. \(2010\)](#) studied a sample of twelve  $z \sim 1$ –2 galaxies, finding that the bulk of the [CII] line (i.e., a 70 per cent of the emission) originates from PDRs. On the other hand, [Gullberg et al. \(2015\)](#) investigated a sample of 20 dusty star-forming galaxies at  $z \sim 2$ –6, finding that their [CII] emission is consistent with being produced by PDRs. Finally, by using a model based on cosmological simulations of a  $z \sim 7$  galaxy, [Vallini et al. \(2015\)](#) found that the [CII] emission arises from PDRs, regardless of the galaxy properties. [CII] has also been suggested as a more reliable tracer of molecular gas compared to CO in regions where CO may be photodissociated or destroyed by cosmic rays (e.g., [Madden et al. 2020](#); [Vizgan et al. 2022](#)), however some works have reported that [CII] emission arising from molecular gas is low (e.g., [Franeck et al. 2018](#)).



**Figure 1.9:** Atmospheric transmission for ALMA, between 180 and 380 GHz. This transmission is obtained by assuming a precipitable water vapour of 0.6 mm, an elevation angle of 90 degrees, and a site altitude of 5059 m. The strongest absorption features at  $\sim 180$ , 325, 365 and 376 GHz are produced by H<sub>2</sub>O vapor. This figure has been generated with the *plotAtmosphere* routine from CASA (McMullin et al., 2007) version 6.2.

### 1.3.3 Main sequence, depletion time and molecular gas fraction

The so-called star-formation main sequence (MS) is a near-linear correlation between the stellar mass ( $M_*$ ) of a galaxy and its corresponding SFR (with a scatter of  $\pm 0.3$  dex) (see Tacconi et al. 2020 and references therein). It is believed that around a 90 per cent of cosmic star formation occurring between  $z=0$  and  $z=5$  takes place in galaxies located in the MS (e.g., Rodighiero et al. 2011). Different MS prescriptions have been proposed (e.g., Whitaker et al. 2012, 2014; Speagle et al. 2014). A prescription commonly assumed in the literature (e.g., Sanders et al. 2023; Pantoni et al. 2021; Tacconi et al. 2018) is the one from Speagle et al. (2014), which in terms of specific SFR (sSFR; with  $\text{sSFR} = \text{SFR}/M_*$ ) is expressed as:

$$\begin{aligned} \log[\text{sSFR}(M_S, z, M_*) (\text{Gyr}^{-1})] &= [-0.6 - 0.026 \times t_c (\text{Gyr})] \times [\log M_* (M_\odot) + 0.025] \\ &\quad - [6.51 - 0.11 \times t_c (\text{Gyr})] + 9 \end{aligned} \quad (1.13)$$

with  $\log(t_c) = 1.143 - 1.026 \times \log(1+z) - 0.599 \times \log^2(1+z) + 0.528 \times \log^3(1+z)$ , where  $t_c$  is the cosmic time. This MS prescription applies for the ranges of  $z = 0 - 5$ , and  $M_* = 10^9 - 10^{11.8} M_\odot$ .

An essential quantity that reflects the link between the evolution of molecular gas and the star-formation MS (and star formation history), is the ratio of molecular gas mass to the SFR. This ratio is known as the molecular depletion timescale, represented as  $t_{\text{depl}} = M_{\text{H}_2}/\text{SFR}$ , with  $M_{\text{H}_2}$  the molecular gas mass. It provides an estimate of the time required for the current star formation activity to deplete the available molecular gas reservoir, without accounting for any mass replenishment from stellar winds and supernovae (Tacconi et al., 2020). The studies of Genzel et al. (2015), Scoville et al. (2017) and Tacconi et al. (2018), have shown that  $t_{\text{depl}}$  can be expressed as products of power laws, that transformed in log-log space is:

$$\log(t_{\text{depl}}) \text{ (Gyr)} = A_t + B_t \times \log(1+z) + C_t \times \log \left[ \frac{\text{sSFR}}{\text{sSFR}(MS, z, M_*)} \right] + D_t \times [\log(M_*) - 10.7] \quad (1.14)$$

where  $\text{sSFR}/\text{sSFR}(MS, z, M_*)$  represents the offset from the MS line, which is a function of  $z$  and  $M_*$ , and  $A_t$ ,  $B_t$ ,  $C_t$  and  $D_t$  are fit parameters that determine the relationship between these variables.

By estimating molecular masses and SFRs of a large sample of 1444 star-forming galaxies at  $z=0-4$  with  $M_* = 10^9 - 10^{11.8} M_\odot$ , Tacconi et al. (2018) updated previous scaling relations of gas depletion time and molecular gas fractions ( $f_{\text{H}_2}$ ). For this, their approach was to analyze separately the dependencies with  $(1+z)$ ,  $\text{sSFR}/\text{sSFR}(MS, z, M_*)$  and  $M_*$ . The best fitting functions led to the following parameters for equation 1.14:  $A_t = -0.53$ ,  $B_t = 0.95$ ,  $C_t = -0.45$  and  $D_t = -0.08$ .

After establishing the parameters dependence of  $t_{\text{depl}}$ , it is in principle straightforward to determine the evolution of the gas-to-stellar mass ratio ( $\mu_{\text{H}_2} = M_{\text{H}_2}/M_*$ ). This is because  $\mu_{\text{H}_2} = t_{\text{depl}} \times \text{sSFR}$  (i.e., multiplication between 1.14 and 1.13), and then the molecular gas fraction is  $f_{\text{H}_2} = \mu_{\text{H}_2}/(1 + \mu_{\text{H}_2})$ . Alternatively, Tacconi et al. (2018) followed a similar approach than the one described above for  $t_{\text{depl}}$ , obtaining that  $\mu_{\text{H}_2}$  can be expressed as:

$$\log(\mu_{\text{H}_2}) = A_\mu + B_\mu \times [\log(1+z) - F]^2 + C_\mu \times \log \left[ \frac{\text{sSFR}}{\text{sSFR}(MS, z, M_*)} \right] + D_\mu \times [\log(M_*) - 10.7] \quad (1.15)$$

with  $A_\mu = 0.07$ ,  $B_\mu = -3.8$ ,  $C_\mu = 0.53$ ,  $D_\mu = -0.33$ , and  $F_\mu = 0.63$ . This expression will be used in Section 2.5.1 of this thesis, to derive molecular gas fractions. An update of these scaling relations has been presented in Tacconi et al. (2020).

### 1.3.4 Observations of cold gas in high- $z$ galaxies

In this subsection I briefly summarize a few representative studies of cold gas reservoirs in and around high- $z$  galaxies (star-forming galaxies and quasar host galaxies) from the literature, to provide context for this thesis.

Several studies have investigated the reliability of [CII] as a tracer of star formation in nearby and high- $z$  galaxies (e.g., Stacey et al. 1991; Boselli et al. 2002; Stacey et al. 2010; Sargsyan et al. 2012; Pineda et al. 2014; De Looze et al. 2014; Herrera-Camus et al. 2015).

For example, [De Looze et al. \(2014\)](#) studied a sample of different types of nearby galaxies. Their results showed that [C II] is a good tracer for the SFR except for low-metallicity sources. On the other hand, [Pineda et al. \(2014\)](#) studied this correlation in the Milky Way, finding that [CII] emission from various phases of the ISM correlates well with the SFR. By using *Herschel* data, [Herrera-Camus et al. \(2015\)](#) found that [CII], in absence of strong AGNs, is a good tracer of SFR in global and kpc scales.

In the context of quasar host galaxies, several studies have studied their cold gas reservoirs (e.g., [Barvainis et al. 2002](#); [Weiß et al. 2003, 2007](#); [Weiss et al. 2007](#); [Riechers et al. 2009](#); [Schumacher et al. 2012](#); [Wang et al. 2016](#); [Venemans et al. 2017](#); [Decarli et al. 2018](#); [Hill et al. 2019](#); [Novak et al. 2019](#); [Bischetti et al. 2021](#)). For instance, by using ALMA, NOEMA and JVLA observations of the far-infrared continuum CO and [CII] lines, [Bischetti et al. \(2021\)](#) studied eight luminous quasars at  $z \sim 2.4-4.7$ , reporting a 100 per cent emission line detection. They found that these quasars exhibit lower (factor of 4) CO luminosities than  $z \sim 0-3$  star-forming galaxies with the same infrared luminosities. They also inferred molecular gas fractions in the host galaxies of four quasars, finding lower (factor of 10-100) values than star-forming galaxies with the same stellar mass. In another study, [Venemans et al. \(2017\)](#) explored three  $z \sim 7$  quasar host galaxies with CO(6-5), CO(7-6) and [CI] ALMA observations. They detected CO in the three objects, and [CI] in one. They derived molecular gas reservoirs of the order of  $\sim 10^{10} M_{\odot}$ . The obtained ratios between CO, [CII] and [CI] were consistent with expectations from PDRs.

Recent observational studies have provided compelling evidence for the existence of cold gas extending to CGM scales at high- $z$ . Most of these studies have predominantly utilized transitions of CO (e.g., [Cicone et al. 2014](#); [Emonts et al. 2016](#); [Ginolfi et al. 2017](#); [Li et al. 2021](#)) and [CII] (e.g., [Cicone et al. 2015](#); [Fujimoto et al. 2020](#); [Herrera-Camus et al. 2021](#)). For example, by using ALMA [CII] observations, [Fujimoto et al. \(2019\)](#) claimed to find the first evidence of extended [CII] emission, the so-called [CII] halos, in a sample 18 star-forming galaxies at  $z \sim 5-7$ . In brief, they obtained a stacked image and extracted a radial profile that showed [CII] emission out to  $\sim 10$  kpc, indicating an extended component. Great progress in the study of [CII] emission at high- $z$  was done thanks to the ALPINE survey ([Le Fèvre et al., 2020](#)). This ALMA survey targeted [CII] emission for 118 star-forming galaxies at  $z \sim 4-6$ , at  $\sim 1$  arcsec of resolution. Based on this survey, different studies have found evidence of [CII] halos through the stacking technique (e.g., [Ginolfi et al. 2020b](#)) and in individual galaxies (e.g., [Fujimoto et al. 2020](#)). These results provide support for a star-formation driven outflows scenario as the main mechanism behind the origin of these [CII] halos.

A more recent ALMA survey is REBELS, which targeted [CII] emission in a sample of 40 star-forming galaxies at  $z \sim 6.5-9$  ([Bouwens et al., 2022](#)). Based on this survey, [Fudamoto et al. \(2022\)](#) analyzed 28 [CII] emitting galaxies, stacked their [CII] flux maps and extracted radial profiles. They found that the [CII] emission is  $\sim 2$  times more extended than the stellar component, as traced by the rest-frame UV emission. This is consistent with previous results obtained with the ALPINE survey.

An interesting work focused on one star-forming galaxy at  $z \sim 5.5$  was performed by [Herrera-Camus et al. \(2021\)](#). With ALMA [CII] observations of high resolution ( $\sim 0.3$  arc-

sec), they detected a diffuse component of [CII] emission, extending beyond the star-forming disk ( $r \sim 6$  kpc). On the other hand, [Akins et al. \(2022\)](#) studied a lensed star-forming galaxy at  $z \sim 7$ . They used ALMA [CII] and [OIII]  $88\mu\text{m}$  data, and found that while the [OIII] emission is compact ( $r \sim 4$  kpc), the [CII] line extends up to  $r \sim 12$  kpc. They suggest that the possible origin of this extended [CII] emission is related to the cooling of hot ionized outflows.

In studies at lower redshifts, the investigation of extended cold gas reservoirs has predominantly relied on the utilization of CO and [CI] transitions. For instance, [Riechers et al. \(2006\)](#) discussed CO(1-0) detections in three quasars at  $z \sim 4$ . They did not find extended molecular gas, however their measurement’s accuracy cannot rule out the possibility of an extended component contributing up to 30 per cent of the overall CO(1-0) luminosity. On the other hand, an extended molecular gas reservoir across  $\sim 50$ -70 kpc has been found around the massive Spiderweb Galaxy at  $z = 2.2$ , using the CO(1-0) and [CI](1-0) transitions ([Emonts et al. 2016, 2018](#)).

Recently, [Jones et al. \(2023\)](#) found CO(3-2) ALMA emission up to  $\sim 13$  kpc by a stacking analysis of seven AGN host galaxies at  $z \sim 2$ -2.5. Similarly, [Scholtz et al. \(2023\)](#) used ALMA CO(7-6) and [CI](1-0) observations to examine the radial surface brightness profiles of both individual sources and stacked emission from extremely red quasars at  $z \sim 2$ . They found extended cold gas and dust emission up to 14 kpc scales. Another promising and debated result comes from the study of a high-redshift radio galaxy where an extended bridge of [CI](1-0) has been reported ([Emonts et al., 2023](#)). Finally, [Li et al. \(2023\)](#) studied a sample of ten quasars with enormous Ly $\alpha$  nebulae at  $z \sim 2$ , with ALMA+ACA CO(4-3) observations. They reported extended (15-100 kpc) reservoirs of molecular gas in six sources.

It is worth noting that many of these observations, particularly those utilizing [CII], seem to primarily trace outflows driven by the AGN or star formation, rather than an extended gas reservoir directly feeding the central galaxy. Moreover, interferometers tend to resolve out a significant portion of the extended CGM emission, especially for [CII] that is observed at high frequencies (e.g., [Carniani et al. 2020](#); [Novak et al. 2020](#); [Decarli et al. 2021](#)). For this reason, more observations based on single-dish submillimetre telescopes are essential to study the CGM that covers a large spatial scale.

### 1.3.5 Cold gas reservoirs from theoretical perspective

In the last years, there has been an increased number of attempts to reproduce cold gas reservoirs on galactic scales in cosmological simulations, through post-processing<sup>20</sup> analysis (e.g., [Lagos et al. 2015](#); [Tomassetti et al. 2015](#); [Marinacci et al. 2017](#); [Diemer et al. 2018](#); [Stevens et al. 2021](#)). This strategy is necessary because currently large-scale cosmological simulations are not able to resolve the high densities needed to form molecular hydrogen. The majority of these models employ recipes taking into account factors such

<sup>20</sup>“Post-processing” refers to the analysis and additional calculations performed on a simulation data after the main simulation run has been completed. It involves applying various algorithms, models, and observational comparisons to extract additional meaningful information from the simulation results.

as the metallicity of the gas or the pressure conditions to determine the fraction of gas that is in atomic or molecular form. The pressure-based approach relies on the established empirical correlation between the pressure within a galaxy’s disk and the ratio of atomic to molecular hydrogen (e.g., Blitz & Rosolowsky 2004, 2006; Leroy et al. 2008). On the other hand, the metallicity-based methods utilize the metallicity of the gas as an indicator of the presence of dust grains, which play a crucial role in catalyzing the formation of  $H_2$  (e.g., Krumholz et al. 2008; Krumholz 2013; Gnedin & Kravtsov 2011).

Lagos et al. (2015) estimated the galactic  $H_2$  abundance in the EAGLE cosmological hydrodynamic simulations. Through postprocessing, they assign  $H_2$  masses to each gas particle based on two models (Gnedin & Kravtsov 2011 and Krumholz 2013), which rely on the local dust-to-gas ratio and the interstellar radiation field. They found that their  $H_2$  results match the observations, at low and high- $z$ . In particular, the predictions from the simulations are able to reproduce the observed scaling relations between the mass of  $H_2$  and the stellar mass, star formation rate and stellar surface density.

On the other hand, Diemer et al. (2018) developed a postprocessing framework to estimate the atomic and molecular gas abundances in the IllustrisTNG simulations. For this, they applied five different models for the molecular fraction, including empirical, simulation-based, and theoretical prescriptions, which rely mostly on the neutral hydrogen surface density and UV flux. Their estimations were done in two dimensions. They found that these five models show a general consensus on average, however the specific details of the modeling have an impact on individual galaxies and the spatial distribution of molecular hydrogen. In this way, they highlight that the estimated molecular fractions are only approximations due to the presence of significant systematic uncertainties. A similar approach but with estimations in three dimensions was performed by Popping et al. (2019), who found several tensions between their  $H_2$  predictions and observations at  $z > 1$ . For example, the  $H_2$  masses they predicted are lower than the observations by a factor of 2-3.

Olsen et al. (2017) presented a postprocessing framework to model FIR/millimeter emission lines, including CO(1-0), CO(2-1), CO(3-2) and [CII] in cosmological hydrodynamics simulations. In a posterior work, Olsen et al. (2021) updated this framework by introducing a sub-grid model<sup>21</sup> to estimate the distribution of dense gas up to densities of  $\sim 10^7 \text{ cm}^{-3}$ , that otherwise is not present in cosmological simulations. They did that by applying a probability density distribution for the gas fragmentation. After these densities and radiation field (from radiative transfer calculations) and density were determined, they used a photoionization code to model the emission lines. They found that most of the emission lines are in good agreement with observations of local galaxies (on galaxy scales). For the case of CO(3-2), its emission was overestimated by  $\sim 1$  dex. For the particular case of [CII], they found good agreement with respect to the [CII]-SFR relation for nearby galaxies.

Regarding the cold gas phase on CGM scales, to my knowledge currently there are no cosmological simulations reproducing this phase. However, there have been efforts to

---

<sup>21</sup>A sub-grid model refers to a modeling approach used to capture physical processes that occur on scales smaller than the resolution of the simulation.

compare predictions from galaxy-scale models with observations of cold gas in the CGM.

For instance, a set of high-resolution zoom-in simulations of  $z \sim 6$  ( $M_* \sim 10^{10} M_\odot$ ) star-forming galaxies was presented in [Pallottini et al. \(2017\)](#). These simulations include dust radiative transfer simulations as postprocessing step, and [CII] emission calculated from a photoionization code. They found that these galaxies are under-luminous with respect to the locally observed [CII]-SFR relation. Their results also indicate that  $\sim 95$  per cent of the [CII] emission comes from the dense gas in the disk. Furthermore, [Fujimoto et al. \(2019\)](#) used one galaxy from these simulations to perform mock observations and compare with ALMA observations of extended (CGM scales) [CII] emission around  $z \sim 6$  galaxies. They did not find such extended [CII] emission in the simulations.

In order to address the molecular phase of the CGM, an interesting approach was presented in the recent work of [Farber & Gronke \(2022\)](#). Following the results from cloud-crushing simulations, that suggest that cool gas clouds with efficient cooling can survive in a hot medium (e.g., [Gronke & Oh 2018, 2020](#)), [Farber & Gronke \(2022\)](#) extends these previous studies by considering cooling down to  $T \sim 400$  K. They found that when these cold clouds surpass a certain critical size, they can fragment resulting in a mist of tiny droplets of molecular gas. Notably, the critical size for this fragmentation differs significantly from what has been previously found for cool clouds. These results are promising for understanding the survival of the molecular and dense atomic phase within the CGM, and also for establishing what are the physical scales important for resolving the molecular phase of the CGM in cosmological simulation studies.

## 1.4 Structure and goals of this thesis

The primary focus of this thesis has been the study of cold and cool gaseous reservoirs in high- $z$  galaxies, with particular emphasis on the circumgalactic gas and its link with the galaxies' properties. My work has involved a combination of observational and simulation-based approaches. Through observations, I have constrained the molecular gas reservoirs in quasar host galaxies at  $z \sim 3$  and link it to the known cool CGM phase, while with simulations I tried to predict [CII] emission levels in galactic and halo scales of star-forming galaxies at  $z \sim 4-6$ .

This thesis is based on one published paper ([Muñoz-Elgueta et al., 2022](#)) and one in preparation. The structure of the thesis is organized as follows:

- In Chapter 2, I aim to constrain the molecular gas reservoirs and investigate any relation between the molecular gas content and the large-scale atomic phase ( $T \sim 10^4$  K) as traced by the Ly $\alpha$  emission, in a sample of  $z \sim 3$  quasars. In particular, the population of quasars at  $z \sim 3$  has been targeted for CO emission only by very few works. In this Chapter, I present a spectroscopic survey with SEPIA band 5 at APEX, targeting the CO(7-6), CO(6-5) and [CI](2-1) transitions for a sample of nine quasars already observed with MUSE/VLT for the large-scale Ly $\alpha$  emission.
- In Chapter 3, I aim to reproduce the [CII] emission observed in star-forming galaxies

at  $z \sim 4-6$ , to enhance our general understanding of [CII] halos. In particular, I aim to understand what is the main mechanism powering [CII] halos in these star-forming galaxies. To achieve this, I applied a post-processing model to estimate the [CII] luminosities in galaxies from the cosmological TNG50 simulations and compared the results with observations from the ALPINE survey.

- Finally, in Chapter 4, I provide a summary of the main findings of this thesis and discuss future perspectives related to the research presented here.

Overall, this thesis contributes to advancing our knowledge of the astrophysical conditions at high- $z$  and shed light on the mechanisms responsible for the observed properties of galaxies in the early universe.



## Chapter 2

# APEX at the QSO MUSEUM: Molecular gas reservoirs associated with $z \sim 3$ quasars and their link to the extended Ly $\alpha$ emission

This work has been published in the Monthly Notices of the Royal Astronomical Society, Volume 511, Issue 1, pp.1462-1483.

Cool gas ( $T \sim 10^4$  K) traced by hydrogen Ly $\alpha$  emission is now routinely detected around  $z \sim 3$  quasars, but little is known about their molecular gas reservoirs. Here, we present an APEX spectroscopic survey of the CO(6-5), CO(7-6) and [C $\text{I}$ ](2-1) emission lines for 9 quasars from the QSO MUSEUM survey which have similar UV luminosities, but very diverse Ly $\alpha$  nebulae. These observations ( $\langle \text{rms} \rangle = 2.6$  mJy in  $300 \text{ km s}^{-1}$ ) detected three CO(6-5) lines with  $3.4 \leq I_{\text{CO}(6-5)} \leq 5.1 \text{ Jy km s}^{-1}$ ,  $620 \leq \text{FWHM} \leq 707 \text{ km s}^{-1}$ , and three [C $\text{I}$ ](2-1) lines with  $2.3 \leq I_{[\text{C}\text{I}](2-1)} \leq 15.7 \text{ Jy km s}^{-1}$ ,  $329 \leq \text{FWHM} \leq 943 \text{ km s}^{-1}$ . For the CO and [C $\text{I}$ ] detected sources, we constrain the molecular gas reservoirs to be  $M_{\text{H}_2} = (0.4 - 6.9) \times 10^{11} M_{\odot}$ , while the non-detections imply  $M_{\text{H}_2} < 1.1 \times 10^{11} M_{\odot}$ . We compare our observations with the extended Ly $\alpha$  properties to understand the link between the cool and the molecular gas phases. We find large velocity shifts between the bulk of Ly $\alpha$  and the molecular gas systemic redshift in five sources (from  $\sim -400$  to  $\sim +1200 \text{ km s}^{-1}$ ). The sources with the largest shifts have the largest Ly $\alpha$  line widths in the sample, suggesting more turbulent gas conditions and/or large-scale inflows/outflows around these quasars. We also find that the brightest ( $I_{[\text{C}\text{I}](2-1)} = 15.7 \pm 3.7 \text{ Jy km s}^{-1}$ ) and the widest (FWHM  $\sim 900 \text{ km s}^{-1}$ ) lines are detected for the smallest and dimmest Ly $\alpha$  nebulae. From this, we speculate that host galaxy obscuration can play an important role in reducing the ionizing and Ly $\alpha$  photons able to escape to halo scales, and/or that these systems are hosted by more massive halos.

## 2.1 Introduction

Super-massive black holes are found at the centre of massive galaxies (e.g., Richstone et al. 1998; Ferrarese & Merritt 2000; Kauffmann & Haehnelt 2000; Kormendy & Ho 2013). They become visible as extremely luminous active galactic nuclei (AGN) through episodes of intense accretion across the history of the Universe (e.g., Schmidt 1963; Bañados et al. 2018; Lyke et al. 2020). Because of the large budget of rest-mass energy available, these objects can regulate their own growth and the evolution of their host galaxies, even if only a small fraction of their feedback couples efficiently to the surrounding material (e.g., Silk & Rees 1998; Di Matteo et al. 2005; Steinborn et al. 2015).

Within AGN, quasars are the most luminous sources where we can see the nuclear emission directly (e.g., Antonucci 1993; Elvis 2000). Clustering measurements suggest that quasars preferentially inhabit dark matter halos with masses  $M_{\text{DM}} \sim 10^{12} - 10^{13} M_{\odot}$  (e.g., Porciani et al. 2004; Shen et al. 2007; White et al. 2012; Timlin et al. 2018 and references therein). This mass range should guarantee that a non-negligible fraction of cool ( $T \sim 10^4$  K) gas, inflowing from large intergalactic scales at redshifts  $z \gtrsim 2$  does not shock heat at the halo boundary, but accretes in cold form (e.g., Dekel & Birnboim 2006). Quasars at such epochs are therefore expected to sit in halos with both a cool and a warm/hot gas ( $T \sim 10^5 - 10^7$  K) phase (e.g., Kereš et al. 2005). Because the quasar number density peaks between  $z \sim 2$  and  $z \sim 3$  (e.g., Richards et al. 2006; Shen et al. 2020), these epochs ( $\sim 10.4 - 11.6$  Gyr ago) are frequently targeted by observations to understand how quasars are triggered and which reservoirs, from halo to galaxy scales, sustain their central engines.

The halo gas, known as circumgalactic medium (CGM, e.g., Tumlinson et al. 2017), has been studied around quasars mostly targeting the cool phase both in absorption (e.g., Hennawi et al. 2006; Prochaska et al. 2013; Farina et al. 2013, 2014; Lau et al. 2018) and in emission (e.g., Heckman et al. 1991; Bunker et al. 2003; Hennawi & Prochaska 2013; Farina et al. 2019; Fossati et al. 2021). While the absorption technique usually relies on only one background sightline per foreground halo to provide statistical information on the physical properties of the CGM of quasars, studies of the CGM in emission are currently able to map the quasar CGM around individual systems. At  $z \sim 2 - 3$ , the study of projected quasar pairs has led to a number of new insights: (i) the measurement of the anisotropic clustering of HI systems around quasars (Hennawi & Prochaska 2007; Jalan et al. 2019) suggested that their ionizing radiation escapes anisotropically or intermittently, (ii) the discovery of large reservoirs ( $> 10^{10} M_{\odot}$ ) of cool and metal-enriched ( $Z \gtrsim 0.1 Z_{\odot}$ ) halo gas (Prochaska et al. 2013, 2014; Lau et al. 2016), and (iii) the study of the kinematics of the halo, which seems to suggest that the gas is in virial equilibrium with the dark matter halo, though there is some evidence for outflowing gas (Prochaska et al. 2014; Lau et al. 2018).

In recent years, sensitive integral field unit spectrographs like the Multi-Unit Spectroscopic Explorer (MUSE; Bacon et al. 2010), the Keck Cosmic Web Imager (KCWI; Morrissey et al. 2012) and the Palomar Cosmic Web Imager (PCWI; Matuszewski et al. 2010) revolutionized the study of CGM gas through emission lines by allowing deeper ob-

servations in reasonable amount of time. The seminal papers by Rees (1988) and Haiman & Rees (2001), predicted that gas surrounding quasars reprocesses the impinging strong UV radiation as Ly $\alpha$  emission. Current studies routinely report extended Ly $\alpha$  emission, with  $\sim 200$  quasars surveyed to date at  $2 < z < 4$  (e.g., Husband et al. 2015; Fumagalli et al. 2016; Borisova et al. 2016; Arrigoni Battaia et al. 2019a; Cai et al. 2019; O’Sullivan et al. 2020; Mackenzie et al. 2021; Fossati et al. 2021). The bulk of the extended emission traces gas on a few tens of kpc near the quasars, while large-scale structures extending to  $R \gtrsim 80$  kpc are seen at lower surface brightness ( $SB_{Ly\alpha} \sim 10^{-18}$  erg s $^{-1}$  cm $^{-2}$  arcsec $^{-2}$ ).

These studies reveal few extended structures over hundreds of kpc with  $SB_{Ly\alpha} \gtrsim 10^{-17}$  erg s $^{-1}$  cm $^{-2}$  arcsec $^{-2}$  (Arrigoni Battaia et al. 2018b), likely pinpointing very dense environments (Hennawi et al. 2015, Nowotka et. al. subm.). These rare and bright large-scale nebulae are also known as enormous Ly $\alpha$  nebulae (ELAN; Cai et al. 2017). The Ly $\alpha$  kinematics in the extended nebulae is consistent with gravitational motions in halos with masses consistent with typical quasar hosts (e.g., Arrigoni Battaia et al. 2019a; O’Sullivan et al. 2020), with a few exceptions with possible quasar winds extending over tens of kpc (Travascio et al. 2020).

While a large fraction of these Ly $\alpha$  nebulae are likely powered by the quasars, the balance between different plausible mechanisms is still debated. Most previous work assumes that the Ly $\alpha$  emission is due to recombination radiation following quasar photoionization (e.g., Heckman et al. 1991; Cantalupo et al. 2014). Resonant scattering of quasar Ly $\alpha$  photons and active companions can, however, provide a non-negligible contribution on scales of tens of kpc near compact sources (e.g., Cantalupo et al. 2014; Husemann et al. 2018; Arrigoni Battaia et al. 2019b). On top of this, there are large uncertainties on the ionizing radiation that impinges on the surrounding gas, because quasars are expected to be anisotropic, intermittent sources with different degrees of obscuration. These uncertainties hamper the physical interpretation of properties of the emitting gas (e.g., density  $n_H$ , metallicity; Fossati et al. 2021).

There is, however, evidence in few systems that gas at large projected distances ( $\gtrsim 50$  kpc) is not affected by resonant scattering effects, namely: (i) non-resonant lines follow the kinematics of the Ly $\alpha$  emission (e.g., Heii, Cai et al. 2017), and (ii) there is no evidence for double-peaked line profiles at the current resolution of the observational data (e.g., Arrigoni Battaia et al. 2018b). Neglecting resonant scattering, photoionization models match the observed Ly $\alpha$  and low HeII emission only if interstellar-medium-like densities ( $n_h \gtrsim 1$  cm $^{-3}$ ) in small-scale structures ( $< 20$  pc) are invoked (Cantalupo et al. 2014; Hennawi et al. 2015; Arrigoni Battaia et al. 2015; Borisova et al. 2016). This finding suggests the presence of dense CGM gas whose survival and entrainment in the warm/hot halo seem plausible from current high resolution “cloud-crushing” simulations (e.g., Gronke & Oh 2018, 2020; Kanjilal et al. 2021). Note that such processes are still largely unresolved by current cosmological simulations, even when attempts are made to resolve the CGM (e.g., Hummels et al. 2019; Peoples et al. 2019).

It is therefore of interest to ascertain observationally the maximum density the cool CGM gas is able to reach, and whether a fraction of the gas is able to transform into a molecular phase. The molecular gas around quasars can be best probed through different

tracers depending on its physical properties (e.g., density, temperature) and those of the surrounding environment (e.g., radiation field) (e.g., Carilli & Walter 2013). Most previous works have focused on the rotational ( $J$ ) transitions of carbon monoxide  $^{12}\text{C}^{16}\text{O}$  (hereafter CO), which is the most abundant molecule after  $\text{H}_2$ . Low- $J$  CO transitions are good tracers of the total cold molecular gas due to their low excitation temperatures. The CO( $J=1-0$ ) ground transition requires an excitation temperature of only  $\sim 5.5\text{K}$  (e.g., Bolatto et al. 2013). Using observations of different CO transitions and radiative transfer models (e.g., large velocity gradient, LVG; e.g., van der Tak et al. 2007), it is possible to constrain the CO spectral line energy distribution (SLED), and probe the excitation conditions and physical properties of the gas, as the density and kinetic temperature (e.g., Weiß et al. 2007; Riechers et al. 2009).

The detection of CO emission in high- $z$  quasars greatly advanced thanks to the Atacama Large Millimeter/Submillimeter Array (ALMA, Wootten & Thompson 2009), and the Karl J. Jansky Very Large Array (JVLA, Perley et al. 2011). It is now possible to probe quasars at very high redshifts ( $z \sim 6 - 7.5$ , e.g., Wang et al. 2016; Venemans et al. 2017; Decarli et al. 2018; Novak et al. 2019). The population of  $z \sim 3$  quasars has also been studied in a number of previous studies (e.g., Weiß et al., 2007; Schumacher et al., 2012; Carilli & Walter, 2013; Bischetti et al., 2021). From these works, we know that at these redshifts, the low CO transitions (i.e.,  $J \leq 3$ ) are expected to be faint, and that the redshifted CO lines lie at challenging frequencies for current and past instruments. Past CO observations of  $z \sim 2 - 3$  quasars found molecular gas masses in the range of  $\sim 10^9 - 10^{11} M_\odot$ , similar to those found for quasars at higher redshift (e.g., Barvainis et al. 2002; Weiß et al. 2003; Beelen et al. 2004; Walter et al. 2011; Schumacher et al. 2012; Hill et al. 2019; Bischetti et al. 2021). These molecular reservoirs are characterized by densities of  $\sim 10^{3.0} - 10^{4.4} \text{cm}^{-3}$  and kinetic temperatures of  $\sim 30 - 90 \text{K}$  (Weiß et al., 2003; Weiss et al., 2007; Schumacher et al., 2012), and, when resolved, have an effective radius of  $\sim 0.5 - 2.5 \text{kpc}$  (e.g., Riechers et al., 2009; Schumacher et al., 2012; Stacey et al., 2021).

Currently, there is only tentative evidence for extended molecular gas reservoirs around individual quasars, but only few studies attempted long integrations. Riechers et al. (2006) presented CO(1-0) detections in three quasars at  $z \sim 4$ . Using single component LVG models, they found that all the flux detected in CO(1-0) was associated with the molecular gas traced by higher CO transitions. An extended component up to 30% of the total CO(1-0) luminosity was allowed by the observations. The extended component could have larger mass if the CO conversion factor was taken to be higher on larger scales. Emonts et al. 2019 targeted the CO(1-0) transition from the MAMMOTH-I ELAN located at  $z \sim 2$  (Cai et al. 2017; Arrigoni Battaia et al. 2018a) and reported emission extended over tens of kpc, with roughly 50% of the CO(1-0) emission outside of galaxies. Finally, Decarli et al. (2021) targeted the CO(3-2) transition for two  $z \sim 2$  ELANe, the Slug (Cantalupo et al. 2014) and the Jackpot (Hennawi et al. 2015). Their NOEMA observations did not unveil any extended molecular reservoir in these objects down to molecular gas surface densities typical of starbursting systems ( $\Sigma_{\text{H}_2} < 12 - 68 M_\odot \text{pc}^{-2}$ ).

Fine structure lines of atomic carbon, for example [C I], are an additional tracer to probe the cold molecular phase (e.g., Papadopoulos & Greve 2004; Valentino et al. 2018).

Observational studies in the local Universe have shown that CO and [C<sub>I</sub>] can coexist, suggesting that both transitions arise from the same regions (e.g., White et al. 1994; Ikeda et al. 2002; Israel & Baas 2002), though spatial variations could be present (e.g., Salak et al. 2019). Analysis using simultaneously [C<sub>I</sub>] and multi-transition CO observations at high redshifts found agreement between the H<sub>2</sub> masses determined through the two different tracers (e.g., Weiß et al. 2003; Alaghband-Zadeh et al. 2013), corroborating the assumption that [C<sub>I</sub>] and CO usually coexist (Carilli & Walter 2013). The carbon masses found in the literature for quasars at  $z \sim 2-3$  are typically of the order of  $10^6 - 10^7 M_{\odot}$  and do not differ significantly from those found for quasars at  $z > 3$  (e.g., Weiß et al., 2003; Walter et al., 2011; Schumacher et al., 2012; Venemans et al., 2017; Banerji et al., 2018; Yang et al., 2019). Molecular masses are then usually obtained by assuming the same abundance of [C<sub>I</sub>] relative to H<sub>2</sub> as found in high- $z$  quasars (e.g., Weiß et al. 2005).

In this framework, we targeted the CO(6-5) ( $\nu_{\text{rest}} = 691.4731$  GHz), CO(7-6) ( $\nu_{\text{rest}} = 806.6518$  GHz) and [C<sub>I</sub>]  $^3\text{P}_2-^3\text{P}_1$  (hereafter [C<sub>I</sub>](2-1),  $\nu_{\text{rest}} = 809.3420$  GHz) transitions with the SEPIA180 receiver (Belitsky et al. 2018b,a) on the Atacama Pathfinder Experiment (APEX) for a sample of nine  $z \sim 3$  quasars, whose halo gas has been studied in the QSO MUSEUM survey (Arrigoni Battaia et al., 2019a). With these observations, we aim to (i) constrain the molecular phase around these massive systems and thus start characterizing the multiphase nature of the halo gas, and (ii) investigate the relation between the molecular gas content and the large-scale cool phase. The molecular line detections also pin down the systemic redshift of the quasar very accurately, allowing us to probe the kinematics of the halo gas.

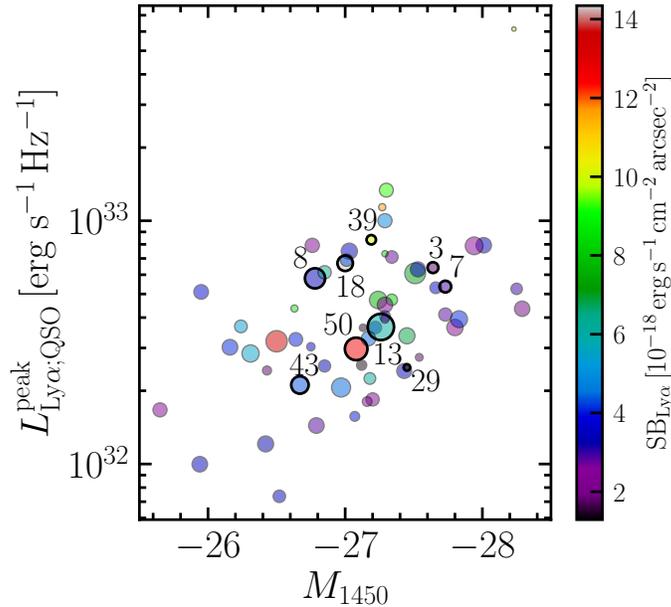
This chapter is structured as follows. In Section 2.2, we describe our sample, observations and data reduction. In Section 2.3 we present the observed line properties and refine the systemic redshift when possible. In Section 2.4, we describe the estimation of the molecular gas masses using different methods, and present results for these masses. In Section 2.5, we compare the derived molecular masses with the Ly $\alpha$  properties of our sources. In Section 2.6 we discuss our main results, and explore the link between the molecular gas content and the large scale Ly $\alpha$  emission. Finally, Section 2.7 summarizes our findings.

Throughout this work, we adopt the cosmological parameter  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_{\text{M}} = 0.3$  and  $\Omega_{\Lambda} = 0.7$ .

## 2.2 Sample and observational data

Our sample is composed of nine quasars at  $z \sim 3$  selected from the QSO MUSEUM survey (Arrigoni Battaia et al., 2019a), which targeted 61 quasars with MUSE for the study of their CGM rest-frame UV line emission. The nine quasars were observed with the SEPIA180<sup>1</sup> receiver mounted on the APEX antenna, located in Llano de Chajnantor, Chile. The targets were selected using the following constraints:

<sup>1</sup>The APEX/SEPIA180 dual polarisation 2SB receiver is a pre-production version of the ALMA Band 5 receiver, and covers the frequency range 159-211 GHz.



**Figure 2.1:** The peak specific luminosity at the Ly $\alpha$  line of the quasar versus  $M_{1450\text{\AA}}$  for the QSO MUSEUM sample (Arrigoni Battaia et al. 2019a). The symbols sizes are proportional to the area of the Ly $\alpha$  nebulae, while the colours indicate their average Ly $\alpha$  surface brightness. We highlight the nine sources observed with APEX with black circles and their ID numbers.

- Visibility from the telescope site and the presence of CO rotational transitions, CO(6-5) and CO(7-6), and the [C $\text{I}$ ](2-1) transition within the frequency range covered by the SEPIA180 instrument.
- The expected frequency of the targeted emission lines was required to be located far from the atmospheric 183 GHz water-absorption feature to best exploit the sensitivity of the SEPIA180 instrument, even under high water vapor conditions.
- Similar absolute magnitudes at rest frame 1450  $\text{\AA}$ , ranging between -27.64 and -26.67 mag, with a median of -27.20 (Figure 2.1).
- Coverage of a large portion of the physical parameter space of the QSO MUSEUM survey, namely Ly $\alpha$  nebulae with sizes spanning the range  $\sim 29 - 467 \text{ arcsec}^2$  (or  $\sim 1600 - 27000 \text{ kpc}^2$ ) and surface brightnesses  $\sim 1.25 \times 10^{-18} - 1.43 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-2}$  (Figure 2.1).
- One of the targets was selected to be radio-loud, reflecting a similar fraction of such objects in the parent quasar sample ( $\sim 10\%$ , Ivezić et al. 2002).

The CO(6-5) ( $\nu_{\text{obs}}^2 = 167.3 \text{ GHz}$ ) observations were carried out between October and December of 2018 under the ESO programme 0102.A-0394A (PI: F. Arrigoni Battaia), with

<sup>2</sup>Observed frequency of the line transition at  $z=3.133$ , the median redshift of the sample studied in this paper.

a total of  $\sim 133$  hours of telescope time. The [C<sub>I</sub>](2-1) ( $\nu_{\text{obs}} = 195.8$  GHz) and CO(7-6) ( $\nu_{\text{obs}} = 195.2$  GHz) observations were performed between May and December of 2019 under the ESO programme 0103.A-0306A (PI: F. Arrigoni Battaia), with a total of  $\sim 140$  hours of telescope time. The main beam full width half maximum (FWHM) of the SEPIA180 receiver is about  $\sim 32''$  ( $\sim 249$  kpc) for the CO(6-5) observations and  $\sim 30''$  ( $\sim 234$  kpc) for the [C<sub>I</sub>](2-1) - CO(7-6) observations. Figure 2.2 shows Ly $\alpha$  images of the nine targets analysed in this work, with superimposed APEX beams shown as white circles. The acquired data will therefore provide an integrated spectrum of the emission within such beams. The median value of precipitable water vapor (PWV) was 1.4 and 1.5 mm for CO(6-5) and [C<sub>I</sub>](2-1) - CO(7-6), respectively. The full histograms of the PWV values for the observations are shown in Fig. 2.3. A summary of the sample and observational setup is shown in Table 2.1. Due to source visibility and weather constraints, we obtained CO(6-5) data for 7 sources and [C<sub>I</sub>](2-1) - CO(7-6) for 8 sources.

The data reduction was performed using the GILDAS/CLASS<sup>3</sup> package version 1.1. For each source the data corresponding to a different date were processed separately, before combining them. In this procedure, for every target the noisy edges ( $\sim 3\%$ ) of the spectra were trimmed. Then, a velocity window<sup>4</sup> was chosen to encompass the expected location of the emission line, according to the redshift of the source. First-degree polynomial baselines were computed neglecting the data within that window, and subtracted from the individual scans. All data for each source were then combined into one final spectrum, after visual inspection of individual scans. These spectra cover an average spectral window of  $\sim 6000$  km s<sup>-1</sup>.

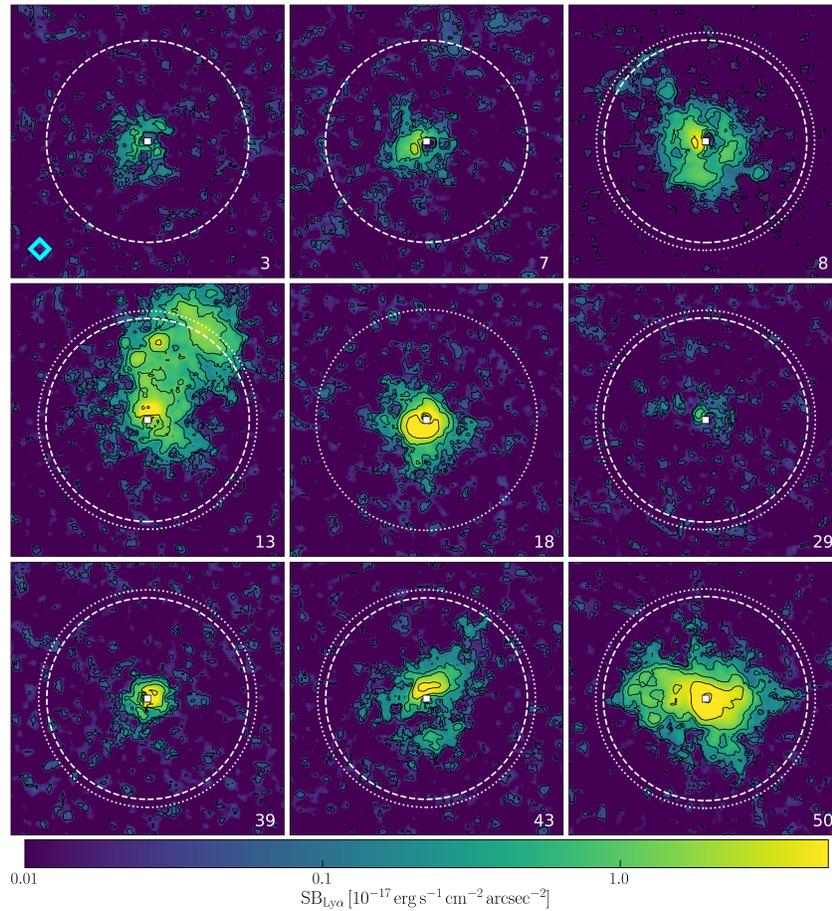
To further improve the root mean square (rms) of the final combined spectra, we applied the following procedure. For each target, we computed the rms for each used subscan in order to reject the noisiest data. We computed the median rms of the whole dataset (i.e., all dates for each source) and removed the data farthest away from this median. New reductions ignoring these subsamples were performed following the steps described above, checking if the final rms of the dataset improved. We found that the removal of the noisiest subsamples did not improve the final rms, because the decrease in exposure time compensates the improvement in rms, so we decided to keep all the data for the final reduction. The final rms for each tuning is reported in Table 2.1.

To illustrate the stability of the SEPIA180 instrument at the targeted frequencies, in Fig. 2.16 of Appendix 2.8 we show the rms as a function of the bin size (in km s<sup>-1</sup>), for the final combined spectra of each source, starting from the original resolution and up to 600 km s<sup>-1</sup>. At the bottom of each panel is shown by how much the observed rms deviates from the expected value. At a bin size of 300 km s<sup>-1</sup>, we found a mild median deviation of 12% and 14% for the CO(6-5) and [C<sub>I</sub>](2-1) - CO(7-6) observations, respectively.

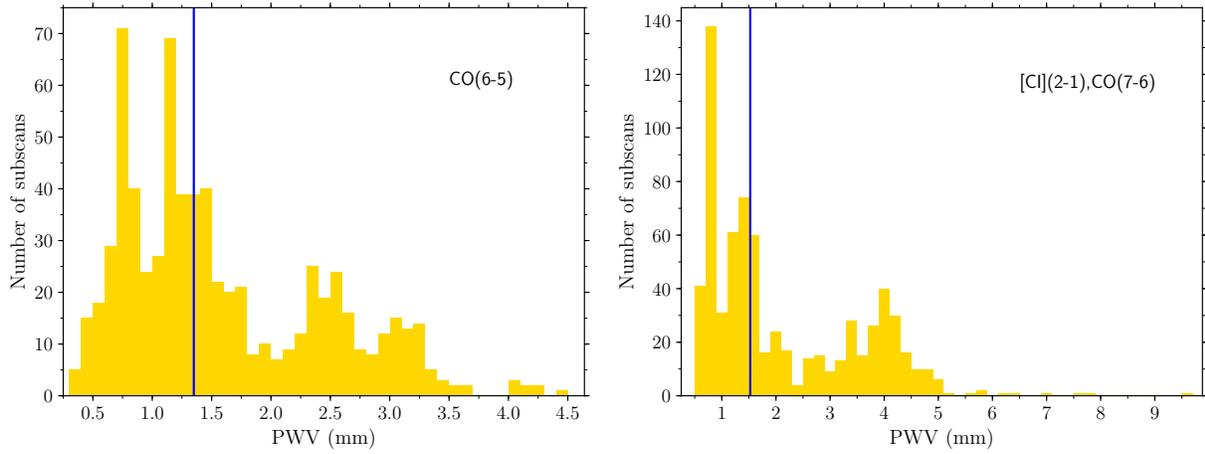
As last step, we transformed the intensity units of the spectra, originally in temperature (K), to physical flux units. For this purpose, we assumed the conversion factor  $38.4 \pm$

<sup>3</sup><http://www.iram.fr/IRAMFR/GILDAS/>

<sup>4</sup>The total width of the velocity window was in the range of  $\sim 1000 - 1500$  km s<sup>-1</sup> for the cases in which the CO(6-5) emission line was expected, and in the range of  $\sim 2000 - 2500$  km s<sup>-1</sup> for the cases in which the [C<sub>I</sub>](2-1) and CO(7-6) emission lines were expected.



**Figure 2.2:**  $40 \times 40$  (or  $\sim 312 \text{ kpc} \times 312 \text{ kpc}$ ) MUSE Ly $\alpha$  surface brightness maps of the nebulae surrounding the 9 targets of this study (adapted from Arrigoni Battaia et al. 2019a). The numbers in the bottom right corner are the IDs from the QSO MUSEUM survey (see Table 2.1, first column). The white dotted circles represent the APEX/SEPIA180 beam sizes for the CO(6-5) observations ( $\sim 32''$ ), while the white dashed circles are the beam sizes for the [C I](2-1) - CO(7-6) observations ( $\sim 30''$ ). The cyan diamond indicates that the source is radio-loud.



**Figure 2.3:** Histograms of the PWV values during the observations. Left panel: Values for the ESO programme 0102.A-0394A, the CO(6-5) observations. Right panel: Values for the ESO programme 0103.A-0306A, the [C I](2-1) - CO(7-6) observations. The blue vertical lines indicate the median values.

2.8 Jy K<sup>-1</sup>, calculated for the SEPIA180 receiver (Belitsky et al., 2018b)<sup>5</sup>.

<sup>5</sup>This value is consistent within uncertainties with the SEPIA180 efficiencies computed for the year of our observations (see listed values at <http://www.apex-telescope.org/telescope/efficiency/index.php>).

ID <sup>c</sup>	Quasar	RA (J2000)	DEC (J2000)	$z_{\text{sys}}^d$	0102.A-0394A <sup>a</sup>			0103.A-0306A <sup>b</sup>				
					Frequency <sup>e</sup> (GHz)	Exp. time <sup>f</sup> (hr)	RMS <sup>g</sup> (mK)	PWV <sup>h</sup> (mm)	Frequency <sup>e</sup> (GHz)	Exp. time <sup>f</sup> (hr)	RMS <sup>g</sup> (mK)	PWV <sup>h</sup> (mm)
3	J 0525-233	05:25:06.500	-23:38:10.00	3.110	-	-	-	-	196.000	4.8	0.131	3.6
7	SDSS J1209+1138	12:09:18.000	+11:38:31.00	3.117	-	-	-	-	195.932	5.4	0.086	1.8
8	UM683	03:36:26.900	-20:19:39.00	3.132	167.346	6.1	0.034	1.9	195.221	5.8	0.079	1.4
13	PKS-1017+109	10:20:10.000	+10:40:02.00	3.164	166.060	6.2	0.075	1.0	193.720	4.7	0.087	3.6
18	SDSS J1557+1540	15:57:43.300	+15:40:20.00	3.265	162.127	11.6	0.047	1.8	-	-	-	-
29	Q-0115-30	01:17:34.000	-29:46:29.00	3.180	164.600	11.5	0.060	1.5	192.000	13.6	0.065	1.7
39	SDSS J0100+2105	01:00:27.661	+21:05:41.57	3.100	168.000	8.1	0.071	1.7	196.888	6.0	0.047	2.1
43	CTSH22.05	01:48:18.130	-53:27:02.00	3.087	168.000	12.7	0.049	1.8	196.800	15.8	0.045	2.3
50	SDSS J0819+0823	08:19:40.580	+08:23:57.98	3.197	164.754	10.5	0.057	1.8	192.197	8.9	0.073	2.6

**Table 2.1:** Sample, observations setup and log. **Notes:** [a] ESO programme corresponding to the CO(6-5) observations; [b] ESO programme corresponding to the [C1](2-1) CO(7-6) observations; [c] Identification number taken from the QSO MUSEUM survey (Arrighi Battaia et al., 2019a); [d] Quasar systemic redshift taken from the QSO MUSEUM survey (Arrighi Battaia et al., 2019a), which has an intrinsic uncertainties of  $415 \text{ km s}^{-1}$ . See Table 2.3 for updated systemic redshifts ( $z_{\text{mol,sys}}$ ) calculated from the lines detected in this work; [e] Tuning frequency used for the observations; [f] Total ON-OFF exposure time per source. The total telescope time is roughly double this integration time; [g] RMS per bin of  $300 \text{ km s}^{-1}$  of the final combined spectrum, in antenna temperature  $T_A^*$  units; [h] Median PWV between all the observed dates for each source.

## 2.3 Observational results

The resulting APEX spectra, reduced and converted to flux density units (mJy), are presented in Figs. 2.4 and 2.5 for the CO(6-5) and [C<sub>I</sub>](2-1)-CO(7-6) observations, respectively. For all sources, the left panel spectrum has a bin size between 150 and 200 km s<sup>-1</sup> (depending on the depth of the data), and the right panel spectrum has a bin size of 300 km s<sup>-1</sup>. These two different bin sizes are shown to highlight the reliability of the detections. In this work we report as detections the lines that fulfill the following conditions: i) have a peak emission at S/N > 2σ, at bin sizes of 300 km s<sup>-1</sup>, ii) are also present at the resolution of 150 km s<sup>-1</sup> but with lower significance than at 300 km s<sup>-1</sup>, and consistent integrated fluxes, iii) have an integrated S/N > 3σ. In Appendix 2.9 we show that this detection algorithm is reliable, giving basically a zero-rate of false positive identifications in a negative and jack-knife tests.

Importantly, throughout this work we assume that the detected emission is due to the central quasar hosts, unless specified. We checked this assumption by computing the number of expected line-emitter companions for each transition within the APEX observations, down to the 3σ limiting luminosities of  $L' = 3.13 \times 10^{10}$ ,  $3.02 \times 10^{10}$  and  $2.65 \times 10^{10}$  K km s<sup>-1</sup> pc<sup>2</sup>, respectively for CO(6-5), CO(7-6), and [C<sub>I</sub>](2-1). Specifically, we assumed (i) a cylindrical volume defined by the APEX primary beam and the covered velocity range of  $\sim 6000$  km s<sup>-1</sup>, (ii) the luminosity function of CO(6-5), CO(7-6) and [C<sub>I</sub>](2-1) emission measured for similar redshifts ( $z \sim 2$ ; Decarli et al. 2020), and (iii) a deterministic bias model for the clustering of sources around quasars (e.g., García-Vergara et al. 2017, 2019). In this model, we assume a power-law shape for the clustering, with a fixed slope of  $\gamma = 1.8$ , and we use the  $z \sim 3$  quasar clustering (Shen et al. 2007) and the clustering of Lyman-break galaxies (LBGs) at  $z = 4$  (Ouchi et al., 2004), which are assumed to have similar clustering as CO and [C<sub>I</sub>] sources.

Following these assumptions, we found that the expected number of companions for the total number of observed fields per line (7, 8, 8) are 0.07, 0.01, 1.39, respectively for CO(6-5), CO(7-6), and [C<sub>I</sub>](2-1). We caution that the luminosity functions in Decarli et al. (2020) are still associated with large uncertainties, and only upper limits are reported at the bright-end sampled by our observations. We have thus extrapolated their measurements up to brighter fluxes, and therefore our estimations represent upper limits for the number of expected companions. Specifically, the worst case is the current [C<sub>I</sub>](2-1) luminosity function, which has a relatively flat shape (see Fig. 7 of Decarli et al. 2020). In this case, our extrapolation is flat and therefore represents a clear upper limit considering that the luminosity function is expected to steeply decrease for high luminosities. In summary, the only transition for which we may find a companion is [C<sub>I</sub>](2-1), with a conservative probability < 17% for each field. Follow-up high-resolution observations with interferometers (e.g., ALMA, NOEMA) are required to verify this assumption and assess whether any of the detected emission comes from companions and/or larger scales.

We also computed the probability that one detected line in one field is actually any CO or [C<sub>I</sub>] transition from an interloper galaxy at possible lower redshifts given our tunings, by assuming the luminosity functions in Decarli et al. (2019, 2020), and the comoving

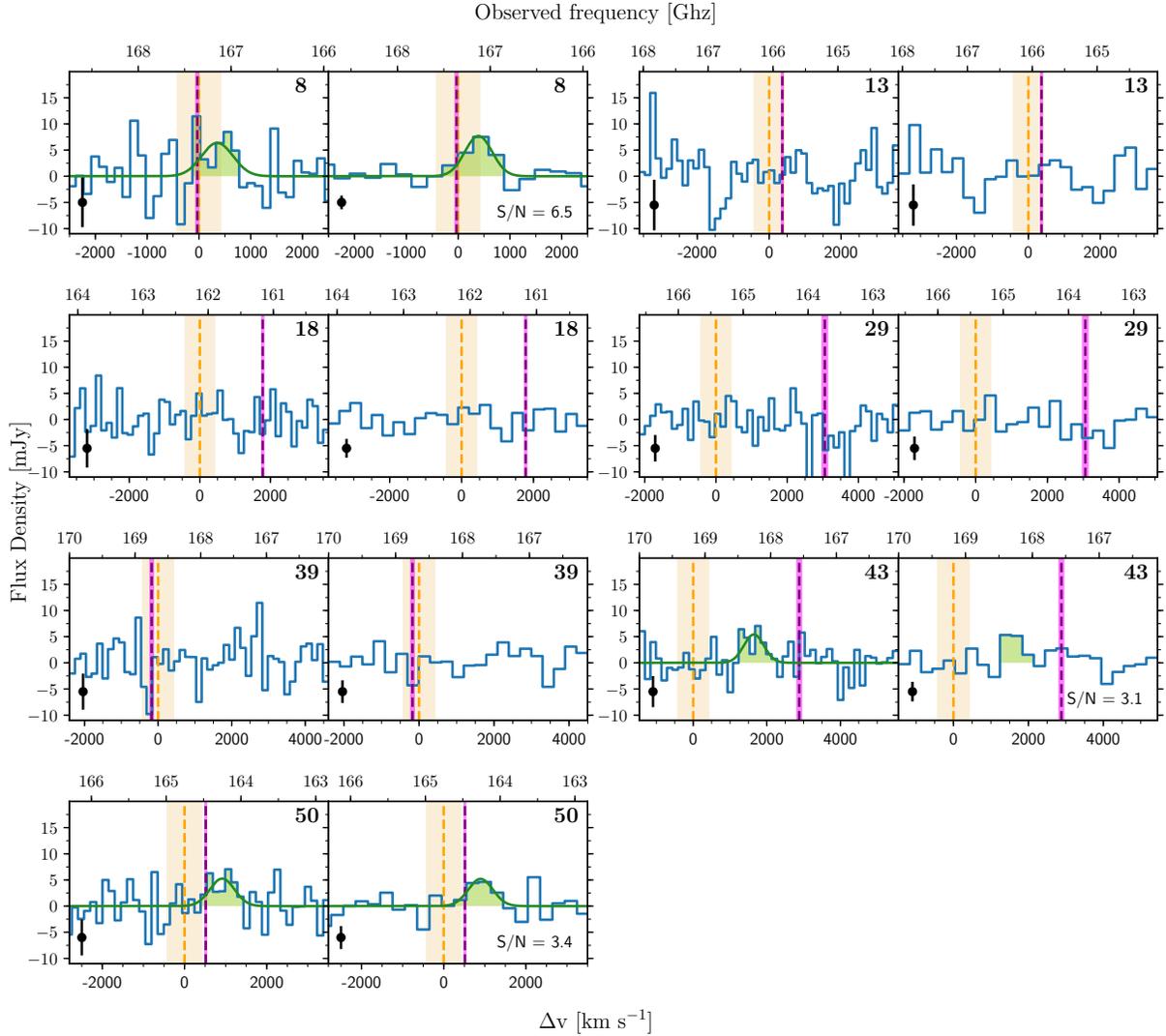
volume spanned by our observations. Once again, these estimates represent upper limits as our observations sample the bright-end of the luminosity functions. We found that the probability of observing a contaminant is  $<0.2\%$  for any CO or [C I] line. Therefore, it is very likely that any detected emission in our observations is associated with the quasar or its environment.

### 2.3.1 Emission line measurements

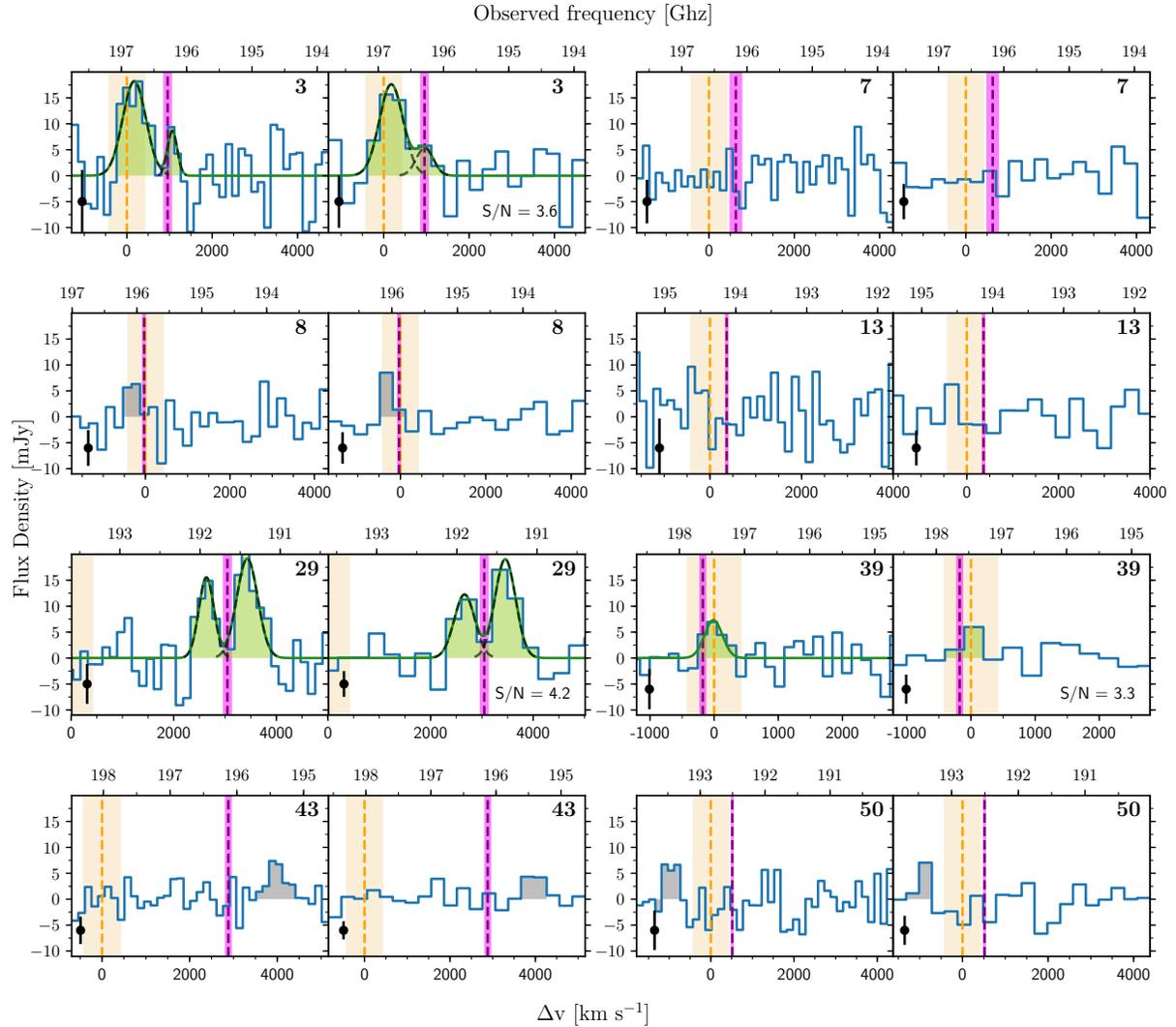
We measured the molecular velocity-integrated emission line fluxes  $I_{\text{CO}(6-5)}$ ,  $I_{\text{[C I]}(2-1)}$  and  $I_{\text{CO}(7-6)}$ , by fitting Gaussian to the detected lines. For those spectra that presented two peaks, a double Gaussian fit was applied. The uncertainties of the measured fluxes include the aforementioned error on the flux conversion factor. The Gaussian fits also provided an estimate of the full width at half maximum (FWHM) of the emission lines and their respective uncertainties.

When an emission line was not detected, we derived  $3\sigma$  upper limits using the rms noise within the same velocity range for the emission lines that were detected for that target, and within a velocity width of  $300 \text{ km s}^{-1}$  (expected average line width for quasar hosts, e.g., [Weiß et al. 2003](#); [Weiss et al. 2007](#); [Walter et al. 2011](#)) when no line was detected. In Table 2.2 we tabulate these fluxes (or upper limits), FWHM and their respective uncertainties for each source.

For the quasars with IDs 39 ([C I](2-1) line) and 43 (CO(6-5) line), it was not possible to obtain a good Gaussian fit at bin sizes of  $300 \text{ km s}^{-1}$  (see Figs. 2.4 and 2.5). Therefore, their integrated fluxes were first estimated by adding the area covered by each bin of  $300 \text{ km s}^{-1}$  contained within the emission line and verified against a Gaussian fit at bin sizes between  $150$  and  $200 \text{ km s}^{-1}$ . Their integrated fluxes computed from both bin sizes are consistent, as per our detection criteria. The FWHMs and fluxes given in Table 2.2 are those estimated with the Gaussian fit. For the quasar with ID 3 (or J 0525-233), all line properties are listed using bin sizes of  $150 \text{ km s}^{-1}$ , because at this resolution we obtained a better fit (see Fig. 2.5) than at bin sizes of  $300 \text{ km s}^{-1}$ .



**Figure 2.4:** The seven APEX CO(6-5) spectral line observations. Each observed spectrum is shown with bin sizes between  $150$  and  $200 \text{ km s}^{-1}$  (left column) and  $300 \text{ km s}^{-1}$  right (column). In each panel, the vertical black error bar represents the uncertainty (rms) per bin, while the ID number of each quasar is shown in the top right corner. The orange vertical lines at the velocity zero point represent the systemic redshift of the sources before our APEX observations (from CIV, see section 2.3.2), and the shaded regions correspond to its uncertainty ( $\sim 415 \text{ km s}^{-1}$ , Arrigoni Battaia et al. 2019a). The purple vertical lines represent the nebulae Ly $\alpha$  redshift from MUSE extracted within the APEX beam (see Fig. 2.2), with their corresponding uncertainty (magenta shaded regions). The green shaded area indicates the detected CO(6-5) emission line, while the dark green curve represents the Gaussian fit applied to this line, when possible.



**Figure 2.5:** Same as Figure 2.4, but for the eight APEX CO(7-6)-[C I](2-1) spectral line observations. Here, the zero velocities are placed at the expected [C I](2-1) transition, considering the systemic redshift from C IV (orange vertical line) with its uncertainty (shaded region) (see section 2.3.2). The grey shaded areas indicate a likely spurious source (ID8, see section 2.3.1 and Appendix 2.9) or tentative features (IDs 43 and 50) that do not belong to the quasar hosts. For sources with IDs 3 and 29, we applied a double Gaussian fit, where the black dashed lines represent the single Gaussian components. For ID 3, a second component was added not to overestimate the integrated flux of the [C I](2-1) line. The emission lines of quasar with ID 29 are both due to [C I](2-1) (see section 2.3.1), and its systemic redshift (orange line) lies outside the plotted window.

ID	Quasar	$I_{\text{CO}(6-5)}$ (Jy km s <sup>-1</sup> )	$S/N_{\text{CO}(6-5)}$	$\text{FWHM}_{\text{CO}(6-5)}$ (km s <sup>-1</sup> )	$I_{\text{Cl}(2-1)}$ (Jy km s <sup>-1</sup> )	$S/N_{\text{Cl}(2-1)}$	$\text{FWHM}_{\text{Cl}(2-1)}$ (km s <sup>-1</sup> )	$I_{\text{CO}(7-6)}$ (Jy km s <sup>-1</sup> )	$S/N_{\text{CO}(7-6)}$	$\text{FWHM}_{\text{CO}(7-6)}$ (km s <sup>-1</sup> )
3	J 0525-233	-	-	-	12.4±3.4 <sup>†</sup>	3.6	640±168 <sup>†</sup>	<6.6 (2.4±1.9) <sup>†*</sup>	- (1.3 <sup>†*</sup> )	- (265±217 <sup>†*</sup> )
7	SDSS J1209+1138	-	-	-	<3.1	-	-	<3.1	-	-
8	UM683	5.1±0.8	6.5	620±175	<3.9	-	-	<3.9	-	-
13	PKS-1017+109	<2.6	-	-	<3.0	-	-	<3.0	-	-
18	SDSS J1557+1540	<1.6	-	-	-	-	-	-	-	-
29	Q-0115-30	<3.6	-	-	15.7±3.7 <sup>†</sup>	4.2	943±249 <sup>†</sup>	<4.0	-	-
39	SDSS J0100+2105	<2.0	-	-	2.3±0.7 <sup>†</sup>	3.3	329±109 <sup>†</sup>	<1.7	-	-
43	CTSH22.05	3.4±1.1 <sup>†</sup>	3.1	643±279 <sup>†</sup>	<2.3	-	-	<2.3	-	-
50	SDSS J0819+0823	3.9±1.1	3.4	707±100	<3.9	-	-	<3.9	-	-

**Table 2.2:** CO(6-5), [Cl](2-1), CO(7-6) fluxes, FWHM, and integrated S/N ratios. Note that the integrated fluxes, S/N and FWHM have been calculated using a bin size of 300 km s<sup>-1</sup>. For the cases with non-detections, 3 $\sigma$  upper limits on the fluxes are provided. **Notes:** [\*] Feature added in the Gaussian fit of quasar with ID 3 (or J 0525-233), to not overestimate the integrated flux of the [Cl](2-1) line at bin size of 300 km s<sup>-1</sup> (see Fig. 2.5); [†] These values have been calculated using a bin size between 150 and 200 km s<sup>-1</sup>; [‡] These values have been calculated as the sum of the two Gaussian components (see Fig. 2.5). For each component, the individual values of  $I_{\text{Cl}(2-1)}$  are 6.1±2.9 km s<sup>-1</sup> and 9.5±2.3 Jy km s<sup>-1</sup>, with  $\text{FWHM}_{\text{Cl}(2-1)}$  values of 471±224 km s<sup>-1</sup> and 471±108 km s<sup>-1</sup>, respectively.

We started the analysis from the CO(6-5) tuning. From the spectra shown in Fig. 2.4, we found CO(6-5) emission line detections (represented by a green shaded area) for three sources: quasars with IDs 8, 43, and 50. Two of these detections are closer to the redshift of the Ly $\alpha$  nebula (IDs 43 and 50; purple vertical line) than the uncertain systemic redshift computed from CIV (orange vertical line). They present fluxes of  $3.4 \leq I_{\text{CO}(6-5)} < 5.1 \text{ Jy km s}^{-1}$ . The other targeted sources, quasars with IDs 13, 18, 29 and 39, do not present any CO(6-5) emission down to the current rms. We therefore report  $3\sigma$  upper limits for these sources (see Table 2.2).

From the spectra in Fig. 2.5, we found [C $\text{I}$ ](2-1) emission line detections in quasars with IDs 3, 29, and 39. They have fluxes of  $2.3 \leq I_{[\text{C}\text{I}](2-1)} \leq 15.7 \text{ Jy km s}^{-1}$ . The quasar with ID 8 (or UM683) has a feature at 196.1 GHz (or  $-328 \text{ km s}^{-1}$ , integrated S/N = 3.1,  $I_{\text{line}} \sim 3 \text{ Jy km s}^{-1}$ , which could be a spurious line given the results of the jack-knife test for this spectrum (see Appendix 2.9). For completeness, we also report that the quasars with ID 43 (or CTS22.05) and ID 50 (or SDSS J0819+0823) have tentative features at 195.4 GHz (or  $+3966 \text{ km s}^{-1}$ , integrated S/N = 2.7,  $I_{\text{line}} \sim 2.8 \text{ Jy km s}^{-1}$ ) and 193.4 GHz (or  $-957 \text{ km s}^{-1}$ , integrated S/N = 2.1,  $I_{\text{line}} \sim 2.5 \text{ Jy km s}^{-1}$ ), respectively. The redshift of these lines is not consistent with the CO(6-5) detections for these sources, but they might be an associated object. We note that sub-millimeter galaxies (SMG) have often been found in the surroundings of quasars (e.g., [Silva et al. 2015](#)), however the CO(6-5) transition (detected for IDs 43 and 50) is expected to be stronger in quasars than in SMG ([Carilli & Walter, 2013](#)). Also, we stress that the detections in the CO(6-5) tuning must come from the quasar, due to the low probability of finding a companion in this transition in our observations (see Section 2.3). The rest of the sources do not have any detected emission lines at the rms of the current observations.

It is important to stress that we considered our detected lines as [C $\text{I}$ ](2-1) and not CO(7-6) because they are closer to the expected position of [C $\text{I}$ ](2-1) based on the systemic redshift of the sources (within  $1\sigma$ , IDs 3 and 39), or because if they were CO(7-6) emission, CO(6-5) detections would be also expected unless unphysical line ratios are assumed (quasar with ID 29, see next paragraph). We acknowledge that the used systemic redshifts (based on CIV) are uncertain (see Section 2.3.2) and further data are needed to verify our line identifications.

The spectrum of the quasar with ID 3 (or J 0525-233) shows the [C $\text{I}$ ](2-1) emission line (at 196.8 GHz) and a S/N = 1.4 bump (binning of  $300 \text{ km s}^{-1}$ ) at the expected location of CO(7-6) (196.2 GHz), which we estimate would contribute  $\sim 20\%$  of the [C $\text{I}$ ](2-1) flux when using the  $300 \text{ km s}^{-1}$  binning. In contrast, the quasar with ID 29 (or Q-0115-30)<sup>6</sup> presents two emission lines (at 191.9 and 191.4 GHz) consistent with the redshift of its Ly $\alpha$  nebula. Both these lines likely correspond to [C $\text{I}$ ](2-1). To confirm this, we first assumed that the line at frequency 191.4 GHz was CO(7-6) and estimated the observed  $I_{\text{CO}(7-6)}/I_{\text{CO}(6-5)}$  ratio ( $> 3.73$ ). We then compared this observed ratio to theoretical predictions from large velocity gradient (LVG) models ([Goldreich & Kwan, 1974](#)) with different physically plausible kinetic temperatures and densities (see Section 2.4.2 for a detailed explanation

<sup>6</sup>See Figure 2.14 for a higher resolution spectrum for this source.

of these models and input parameters). This test did not allowed us to find any modelled  $I_{\text{CO}(7-6)}/I_{\text{CO}(6-5)}$  ratio as high as that observed, indicating that the line peak at lower frequencies also corresponds to [C<sub>I</sub>](2-1)<sup>7</sup>. We stress that this identification needs further confirmation, e.g. by targeting additional transitions from this source. Quasar ID 29, with a total  $I_{[\text{C}\text{I}](2-1)} = 15.7 \pm 3.7 \text{ Jy km s}^{-1}$ , is therefore the brightest [C<sub>I</sub>](2-1) detection in this sample. Furthermore, it is  $\sim 3$  times brighter than the lensed Cloverleaf quasar ( $I_{[\text{C}\text{I}](2-1)} \sim 5.2 \text{ Jy km s}^{-1}$ , Weiß et al. 2003), which is as far as we know the brightest [C<sub>I</sub>](2-1)  $z \sim 3$  quasar detected in the literature. We further discuss the nature of this double-peak emission of quasar with ID 29 (or Q-0115-30) in Section 2.6.1.

Overall, our observations suggest that [C<sub>I</sub>](2-1) is stronger than CO(7-6) in these  $z \sim 3$  quasars. This is in contrast with other observations of the same transitions in different  $z \sim 4 - 7$  quasars (Riechers et al. 2009; Venemans et al. 2017; Lu et al. 2018; Wang et al. 2019; Yang et al. 2019) and  $z \sim 2.5-3$  quasars (Weiß et al., 2003; Walter et al., 2011; Schumacher et al., 2012), likely indicating that the excitation conditions in the molecular gas are different. Also, Banerji et al. (2018) observed these transitions in two quasars and their companions at  $z \sim 2.5$ , finding a [C<sub>I</sub>](2-1) line slightly stronger than CO(7-6) in one quasar and one companion. To our knowledge, the studies mentioned above are the only available observations of these lines in the literature. Therefore, if our observations are confirmed in deeper datasets, current literature could be affected by low number statistics. We further note that [C<sub>I</sub>](2-1) is stronger than CO(7-6) in high-redshift radio galaxies ( $\sim 5$  times stronger, e.g., Gullberg et al. 2016), which can be explained as enhancement of atomic carbon in cosmic ray dominated regions (e.g., Bisbas et al., 2017).

### 2.3.2 Redshifts from molecular lines

A precise estimate of the systemic redshifts of quasars plays a fundamental role in understanding the physical processes and kinematics of each system. For instance, an accurate systemic redshift would allow us to better constrain the cool gas kinematics mapped on large scales by the Ly $\alpha$  emission, and to compare it to cosmological simulations (Arrigoni Battaia et al., 2019a).

The uncertain systemic redshifts for our sample, estimated by Arrigoni Battaia et al. (2019a), are shown in the fifth column of Table 2.1. These redshifts were determined from the peak of the C<sub>IV</sub> line, after correcting from the expected luminosity-dependent blue-shift (Shen et al., 2016), and have an intrinsic uncertainty of  $415 \text{ km s}^{-1}$ . This large uncertainty, comparable to outflow/inflow velocities expected in quasar halos, hampers any kinematical study of these systems. However, the molecular emission lines should provide a more robust measure of the quasar’s systemic redshift (e.g., Banerji et al. 2017;

<sup>7</sup>All the LVG models explored in this work have  $I_{\text{CO}(7-6)}/I_{\text{CO}(6-5)} < 1.34$  when considering kinetic temperatures of  $20 - 200 \text{ K}$ , and densities of  $10^3 - 10^5 \text{ cm}^{-3}$ , as frequently used in the literature (e.g., Riechers et al. 2006; Weiss et al. 2007). These LVG models allow a maximum  $I_{\text{CO}(7-6)} = 3.48 \text{ Jy km s}^{-1}$  at a kinetic temperature of  $200 \text{ K}$ , which represent the most extreme SLED found in the literature (Weiss et al. 2007). Even in this extreme case, a fraction of 63% of the observed second line peak has to correspond to [C<sub>I</sub>](2-1).

[Bischetti et al. 2021](#)). We therefore derived new systemic redshifts for the objects with detected molecular lines. For the case of ID 29, the redshift was estimated as an average between the centroids of the two peaks. These new redshifts and their uncertainties are listed in Table 2.3. They have on average an uncertainty of  $74.8 \text{ km s}^{-1}$  and a difference of  $+1045 \text{ km s}^{-1}$  (or  $\Delta z = +0.014$ ) with respect to the systemic redshifts from CIV (see also Section 2.5.2). Hereafter we will assume these new values as systemic redshifts.

ID	Quasar	$z_{\text{CO}(6-5)}^a$	$\nu_{\text{CO}(6-5)}$ (GHz)	$z_{\text{C}i(2-1)}^b$	$\nu_{\text{C}i(2-1)}$ (GHz)	$z_{\text{CO}(7-6)}^c$	$\nu_{\text{CO}(7-6)}$ (GHz)	$z_{\text{mol,sys}}^d$	$\Delta \nu_{\text{Ly}\alpha}^e$ (km s <sup>-1</sup> )	$\text{FWHM}_{\text{Ly}\alpha}$ (km s <sup>-1</sup> )
3	J 0525-233	-	-	3.112±0.001	196.81±0.05	-(3.111±0.001)	-(196.17±0.06)	3.112±0.001	786±109	1542±196
7	SDSS J1209+1138	-	-	-	-	-	-	-	-	-
8	UM683	3.137±0.001	167.13±0.04	-	-	-	-	3.137±0.001	-423±76	915±46
13	PKS-1017+109	-	-	-	-	-	-	-	-	-
18	SDSS J1557+1540	-	-	-	-	-	-	-	-	-
29	Q-0115-30	-	-	3.223±0.001 <sup>‡</sup>	191.65±0.06 <sup>‡</sup>	-	-	3.223±0.001	-11±120	508±172
39	SDSS J0100+2105	-	-	3.100±0.001	197.42±0.03	-	-	3.100±0.001	-152±64	781±100
43	CTSH22.05	3.110±0.002	168.26±0.07	-	-	-	-	3.110±0.002	1236±133	1068±142
50	SDSS J0819+0823	3.210±0.001	164.26±0.02	-	-	-	-	3.210±0.001	-377±44	967±38

**Table 2.3:** Redshifts and central frequencies obtained from the targeted molecular lines, and velocity shifts from the large-scale Ly $\alpha$  emission within the same APEX aperture. **Notes:** [a] Redshift obtained from the observed CO(6-5) emission line; [b] Redshift obtained from the observed [C]i(2-1) emission line; [c] Redshift obtained from the observed CO(7-6) emission line; [d] Quasar systemic redshift estimated using the observed molecular lines; [e] Velocity shift between the Ly $\alpha$  line peak of the nebularities (extracted within the APEX beam, see Fig. 2.2) and the systemic redshift  $z_{\text{mol,sys}}$ ; [‡] These values have been calculated as the average between the two Gaussian components (see Fig. 2.5). For each component, the individual values of  $\nu_{\text{C}i(2-1)}$  are 191.91±0.11 GHz and 191.39±0.06 GHz.

## 2.4 Molecular mass estimates

In this section we estimate molecular gas masses  $M_{\text{H}_2}$  by using different methods. Specifically, we compute i) carbon masses  $M_{\text{C I}}$  and derive the respective  $M_{\text{H}_2}$  by assuming a carbon abundance relative to  $\text{H}_2$ , ii)  $M_{\text{H}_2}$  for sources that have a measured constraint (upper limit) on the  $\text{CO}(7-6)/\text{CO}(6-5)$  ratio, iii)  $M_{\text{H}_2}$  by combining the two previously obtained mass ranges from [C I] and the CO ratio, and iv)  $M_{\text{H}_2}$  for sources with no clear constraint on the CO ratio (i.e. sources with non-detections).

### 2.4.1 Atomic carbon mass

The mass of the atomic carbon can be estimated from the [C I](2-1) line luminosity through the formulation presented in Weiß et al. (2003, 2005), under the assumption that this [C I] transition is optically thin:

$$\frac{M_{\text{C I}}}{M_{\odot}} = 4.566 \times 10^{-4} Q(T_{\text{ex}}) \frac{1}{5} e^{T_2/T_{\text{ex}}} L'_{[\text{C I}]} \quad (2.1)$$

where  $Q(T_{\text{ex}}) = 1 + 3e^{-T_1/T_{\text{ex}}} + 5e^{-T_2/T_{\text{ex}}}$  corresponds to the C I partition function,  $T_{\text{ex}}$  is the excitation temperature, and  $T_1 = 23.6$  K and  $T_2 = 62.5$  K correspond to the energies above the ground state. We used  $T_{\text{ex}} = 30$  K as frequently found in high-redshift quasars (e.g., Weiß et al. 2003; Walter et al. 2011). The [C I] line luminosity,  $L'_{[\text{C I}]}$ , can be estimated via (Solomon et al., 1992)

$$L'_{\text{line}} = 3.25 \times 10^7 I_{\text{line}} \nu_{\text{obs}}^{-2} D_{\text{L}}^2 (1+z)^{-3} \quad (2.2)$$

where  $I_{\text{line}}$  is the velocity-integrated line flux in units of  $\text{Jy km s}^{-1}$ ,  $\nu_{\text{obs}}$  is the observed frequency in units of GHz, and  $D_{\text{L}}$  corresponds to the luminosity distance in Mpc. The final units of  $L'_{\text{line}}$  are  $\text{K km s}^{-1} \text{pc}^2$ .

As noted in Section 2.3.1, the [C I](2-1) emission line is detected for three sources of our sample: quasars with IDs 3, 29 and 39. Using equation 2.1, we found atomic carbon masses in the range of  $\sim 4.0 \times 10^7 M_{\odot} - 3.1 \times 10^8 M_{\odot}$ . For the five sources with upper limits on the [C I](2-1) transition (quasars with IDs 7, 8, 13, 43 and 50), we found atomic carbon masses  $< 8.0 \times 10^7 M_{\odot}$ . The detected sources show higher values than usually reported for quasars in the literature. Indeed, it is common to find values of the order  $10^6 - 10^7 M_{\odot}$ , when the same excitation temperature of 30 K is assumed (e.g., Walter et al., 2011; Venemans et al., 2017; Yang et al., 2019). The value for each source is listed in the third column of Table 2.4. It is important to note that the assumption of a higher  $T_{\text{ex}}$  in equation 2.1, for instance  $T_{\text{ex}} = 50$  K, implies a  $M_{\text{C I}} \sim 38$  % lower.

The atomic carbon mass can be used to determine the molecular gas mass using the atomic carbon abundance relative to  $\text{H}_2$ :

$$\frac{X[\text{C I}]}{X[\text{H}_2]} = \frac{M_{\text{C I}}}{6M_{\text{H}_2}}. \quad (2.3)$$

ID	Quasar	$M_{\text{C I}}^a$ ( $10^8 M_{\odot}$ )	$M_{\text{H}_2}^b$ ( $10^{11} M_{\odot}$ )	$M_{\text{H}_2}^c$ ( $10^{11} M_{\odot}$ )	$M_{\text{H}_2}^d$ ( $10^{11} M_{\odot}$ )	$M_{\text{dyn}} \sin^2(i)^e$ ( $10^{11} M_{\odot}$ )
3	J 0525-233	$2.3 \pm 0.6$	$4.6 \pm 2.3$	$6.9$	$2.3-6.9^*$	$1.6 \pm 0.8$
7	SDSS J1209+1138	$< 0.6$	$< 1.1$	$< 3.2$	$< 1.1$	-
8	UM683	$< 0.7$	$< 1.5$	$0.7-64.1$	$0.7-1.5$	$1.5 \pm 0.8$
13	PKS-1017+109	$< 0.6$	$< 1.1$	$< 3.2$	$< 1.1$	-
18	SDSS J1557+1540	-	-	$< 0.6$	$< 0.6$	-
29	Q-0115-30	$3.1 \pm 0.7$	$6.1 \pm 2.9$	$< 4.4$	$3.2-4.4$	$4.4 \pm 1.0^{\ddagger}$ $12.4 \pm 6.6^{\diamond}$
39	SDSS J0100+2105	$0.4 \pm 0.1$	$0.8 \pm 0.4$	$< 1.8$	$0.4-1.8^{\dagger}$	$0.4 \pm 0.3$
43	CTSH22.05	$< 0.4$	$< 0.9$	$0.5-42.7$	$0.5-0.9$	$1.6 \pm 1.4$
50	SDSS J0819+0823	$< 0.8$	$< 1.5$	$0.5-51.3$	$0.5-1.5$	$1.9 \pm 0.6$

**Table 2.4:** Molecular, carbon and dynamical mass estimations. Note that upper limits of masses provided in each column are due to non-detections of [C I] and/or CO (see Section 2.4). **Notes:** [a] Atomic carbon mass assuming an excitation temperature of 30 K (see Section 2.4.1), [b] Molecular gas mass derived from the atomic carbon mass (see Section 2.4.1, equation 2.3), [c] Molecular gas mass derived from CO (see Section 2.4.2), [d] Molecular gas mass derived from applying both CO and [C I] constraints (see Section 2.4.3), [e] Inclination-dependent dynamical masses (see Section 2.6.1). The reported uncertainties include only the errors on the FWHM values, [\*] In the case the detected line for ID 3 is CO(7-6), we estimate  $1.2 < M_{\text{H}_2}/[10^{11} M_{\odot}] < 4.5$  from the final CO SLEDs (see Fig. 2.8), [†] In the case the detected line for ID 39 is CO(7-6), we estimate  $0.2 < M_{\text{H}_2}/[10^{11} M_{\odot}] < 0.6$  from the final CO SLEDs (see Fig. 2.8), [‡] This value correspond to the dynamical mass for the disk scenario (see Section 2.6.1), [◊] This value correspond to the dynamical mass for the merger scenario (see Section 2.6.1).

In this work we assumed  $X[\text{C I}]/X[\text{H}_2] = (8.4 \pm 3.5) \times 10^{-5}$  as usually done in the literature for high-redshift quasars (e.g., Walter et al. 2011; Venemans et al. 2017). We note that this value is higher than what has been recently found for  $z \sim 1-3$  main sequence galaxies ( $\sim 1.7 \times 10^{-5}$ , Valentino et al. 2018; Boogaard et al. 2020).

Using this method, we constrained molecular gas masses for quasars with IDs 3, 29 and 39, which span from  $\sim 8.0 \times 10^{10} M_{\odot}$  to  $\sim 6.1 \times 10^{11} M_{\odot}$ . For quasars with IDs 7, 8, 13, 43 and 50, we found molecular gas masses  $< 1.5 \times 10^{11} M_{\odot}$ . Most of these molecular gas masses are higher than the typical range ( $\sim 10^9-10^{11} M_{\odot}$ ) found in the literature for other high- $z$  quasars (e.g., Weiß et al., 2003; Walter et al., 2011; Anh et al., 2013; Hill et al., 2019). We list all these values in the fourth column of Table 2.4.

Importantly, we are assuming that all the [C I](2-1) emission comes from the quasar hosts. If part of this emission comes from larger areas, the  $T_{\text{ex}}$  might break into different values (e.g., lower/higher on larger/smaller scales), affecting the molecular masses estimated in this work. To constraint  $T_{\text{ex}}$  around these quasars, resolved [C I](2-1) and [C I](1-0) observations are needed (e.g., ALMA, NOEMA).

### 2.4.2 LVG Models and CO constraint

The molecular gas mass can also be derived from CO following the equation

$$M_{\text{H}_2} = \alpha L'_{\text{CO}(1-0)} \quad (2.4)$$

with  $\alpha$  being the CO luminosity-to-gas mass conversion factor, and  $L'_{\text{CO}(1-0)}$  is the luminosity of the CO(1-0) that can be estimated from equation 3.5. In this work, we assume a value of  $\alpha = 0.8 \text{ M}_\odot (\text{K km s}^{-1} \text{ pc}^2)^{-1}$  (Downes & Solomon, 1998), which has been estimated for local ultra-luminous infrared galaxies and has been typically adopted to calculate molecular masses in high-redshift quasars (e.g., Wang et al., 2010; Venemans et al., 2017).

To estimate the molecular gas masses from the CO(6-5) and CO(7-6) transitions, we need to assume a CO spectral line energy distribution (hereafter SLED) to find the CO(1-0) line intensity. We did not find in the literature any quasar characterized by a SLED<sup>8</sup> that agrees with the observed constraints on the CO ratios. For this reason, we modelled the CO SLED using the large velocity gradient (hereafter LVG) method, which has been widely applied to high- $z$  quasars by several authors (e.g., Weiß et al., 2007; Riechers et al., 2009; Schumacher et al., 2012). These models include a velocity gradient (that indicates the change in the line of sight velocity in the turbulent medium) that is considerably larger than local thermal velocities of the gas, leading to photons being able to escape due to the different velocities along the cloud, following a photon escape probability. This method finds the populations of the molecular energy levels excited by collisions with H<sub>2</sub> (main collision partner for CO) for certain given parameters as CO abundance, kinetic temperature ( $T_{\text{kin}}$ ), H<sub>2</sub> density ( $n_{\text{H}_2}$ ) and velocity gradient ( $dv/dr$ ). It is then possible to identify the physical parameters that best describe the conditions of the gas through the comparison of the model predictions to the observed line ratios (e.g., Carilli & Walter, 2013).

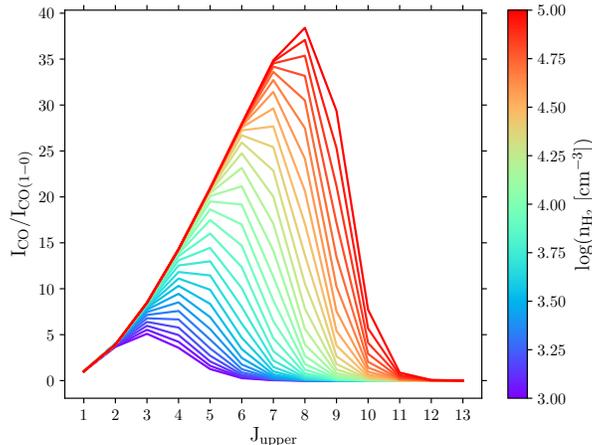
In this work, we have used the radiative transfer code RADEX<sup>9</sup> (van der Tak et al., 2007), considering a spherical and single component LVG model. For the calculations, we adopted an H<sub>2</sub> ortho-to-para ratio of 3 and collision rates from Yang et al. (2010). We set the following input parameters and explore a grid of different models in which we vary some of these parameters:  $T_{\text{kin}}$  equal to 30 K,  $n_{\text{H}_2}$  range of  $10^3$ - $10^5 \text{ cm}^{-3}$  (typical for quasar host galaxies, Carilli & Walter 2013), background temperature of  $\sim 11$  K (cosmic microwave background at  $z \sim 3$ ), and a column density of the molecular gas given by:

$$N_{\text{H}_2} = 3.086 \times 10^{18} n_{\text{H}_2} \frac{\Delta V_{\text{turb}}}{dv/dr} \quad (2.5)$$

where  $\Delta V_{\text{turb}}$  corresponds to the turbulence line width fixed here at  $100 \text{ km s}^{-1}$ , and  $[\text{CO}]/dv/dr = 1 \times 10^5 \text{ pc (km s}^{-1})^{-1}$ , following procedures commonly adopted in the literature (e.g., Weiß et al., 2007; Schumacher et al., 2012). LVG models are intrinsically

<sup>8</sup>There are only a handful of  $z \sim 3$  quasars with a well-characterized SLED (see Fig. 2.8).

<sup>9</sup><https://home.strw.leidenuniv.nl/~moldata/radex.html>



**Figure 2.6:** CO SLEDs obtained from LVG modelling, for  $T_{\text{kin}} = 30$  K, using RADEX (van der Tak et al., 2007). The CO line flux normalized to the CO(1-0) line is plotted as a function of the upper rotational quantum number  $J$ . Different values in  $\text{H}_2$  density are shown by the colour bar.

degenerate in the parameters  $T_{\text{kin}}$  and  $n_{\text{H}_2}$ , meaning that different combinations of these parameters can give the same SLED. We focused only on models with  $T_{\text{kin}} = 30$  K because the value of the kinetic temperature is expected to be comparable to the excitation temperature of neutral carbon (Israel & Baas, 2002), which we assumed to be  $\sim 30$  K (see Section 2.4.1). In total, 31 different CO ladders were modelled, which are shown in Fig. 2.6. In this figure, the CO line flux normalized to the CO(1-0) line is shown as a function of rotational quantum number  $J$ , and the colour bar represents the different values of  $n_{\text{H}_2}$ . The peak of the modelled ladders varies from  $J \sim 3$  to  $J \sim 8$ , for  $n_{\text{H}_2} = 10^3 \text{ cm}^{-3}$  and  $10^5 \text{ cm}^{-3}$ , respectively. We note that according to current observations, quasars are expected to show the SLED peak between  $J \sim 6$  and  $J \sim 8$  (e.g., Riechers et al., 2009, 2011; Carilli & Walter, 2013; Banerji et al., 2018; Wang et al., 2019; Bischetti et al., 2021).

Importantly, the emitted CO flux is proportional to the source solid angle (e.g., Weiß et al., 2007). As our observations are not able to constrain the size of the emitting source, we focus on line ratios in this paper. These LVG models will therefore represent only solutions for the component of highly excited gas, which likely coexist with a less excited component emitting the [C I](2-1) emission that we observe in some of our targets. Indeed, it has been shown that emission from species with very different critical densities likely originate from gas at different densities (e.g., Harrington et al. 2021), and also on different scales (e.g., Emonts et al. 2014; Casey et al. 2018; Spingola et al. 2020).

We first find models that reproduce the observed upper limits on the CO(7-6)/CO(6-5) line ratios. In our dataset there are only three sources for which it was possible to estimate an observed  $I_{\text{CO}(7-6)}/I_{\text{CO}(6-5)}$  upper limit: the quasars with IDs 8, 43 and 50 (see Table 2.2). For these sources, we selected LVG models constrained by their observed  $I_{\text{CO}(7-6)}/I_{\text{CO}(6-5)}$  ratio and found densities in the range  $n_{\text{H}_2} \sim 10^3\text{-}10^{4.4} \text{ cm}^{-3}$ . These models constrained the peak of the SLED to be at  $J \sim 3\text{-}7$ , and the molecular gas masses to be (0.5

-  $64.1) \times 10^{11} M_{\odot}$  when considering the predicted  $I_{\text{CO}(1-0)}$  from the SLEDs. These results are shown in the fifth column of Table 2.4.

It is noteworthy that using a higher  $T_{\text{kin}}$  in our LVG models, for instance  $T_{\text{kin}} = 50$  K (found in some high- $z$  quasars, e.g., Weiss et al. 2007; Riechers et al. 2011), implies an upper limit in the molecular masses  $\sim 32\%$  lower, i.e. this difference does not alter our results significantly. In contrast, using a lower  $T_{\text{kin}}$ , for example  $T_{\text{kin}} = 20$  K, implies non-physical values for the upper limit in the molecular masses of up to  $\sim 10^{13} M_{\odot}$ .

### 2.4.3 Molecular mass constraints from a joint CO and [CI] analysis

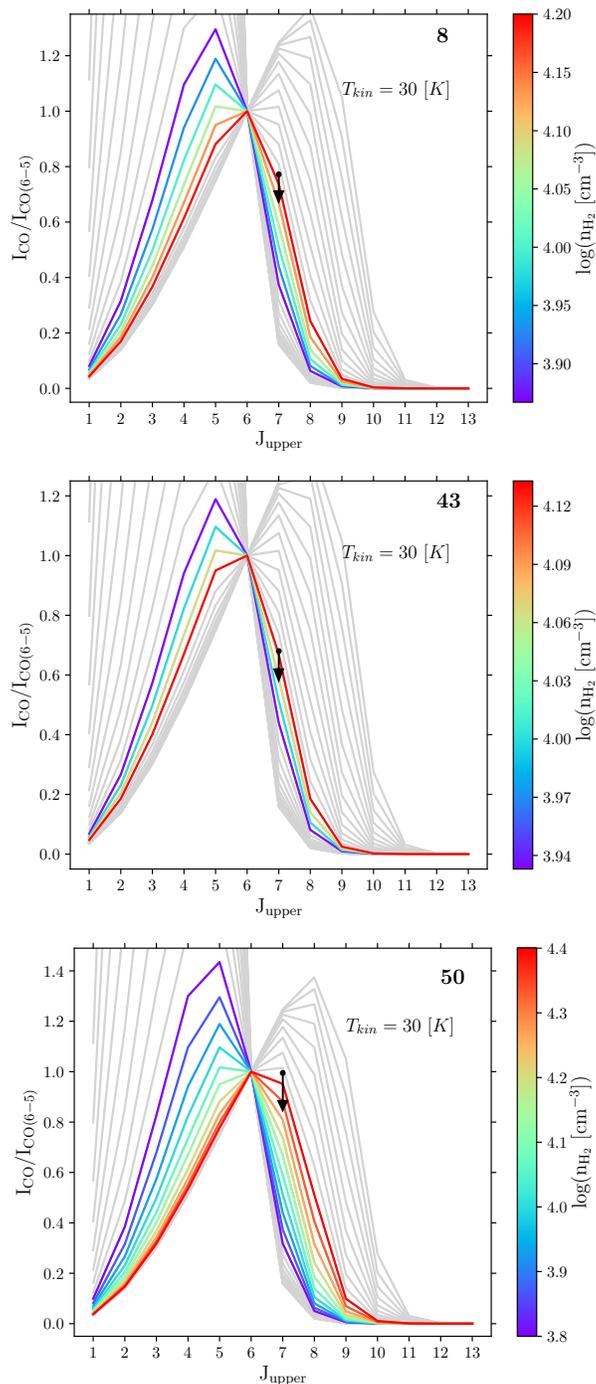
After applying the CO constraint explained above, we set another condition based on the molecular gas masses already estimated from the [CI](2-1) transition (see Section 2.4.1). As  $\alpha$  is fixed, this condition is identical to impose a constraint on the CO(1-0) luminosity. We caution that this approach may introduce a bias in our calculation, which depends on the goodness of the assumed parameters to model the molecular mass from [CI], and on  $\alpha$  itself with respect to the physical conditions in each individual source. Observations of the CO(1-0) transition for these sources are definitely needed to confirm this methodology.

For each source, this step excluded some of the LVG models at the lowest densities selected in Section 2.4.2. In this way, we obtained the final CO SLEDs from the union of the CO and [CI] constraints. We show these final constrained CO SLEDs in Fig. 2.7 for quasars with IDs 8, 43 and 50. Their ranges of molecular masses using these joint constraints are  $(0.7-1.5) \times 10^{11} M_{\odot}$ ,  $(0.5-0.9) \times 10^{11} M_{\odot}$  and  $(0.5-1.5) \times 10^{11} M_{\odot}$ , respectively. These values are also listed in the second to last column of Table 2.4.

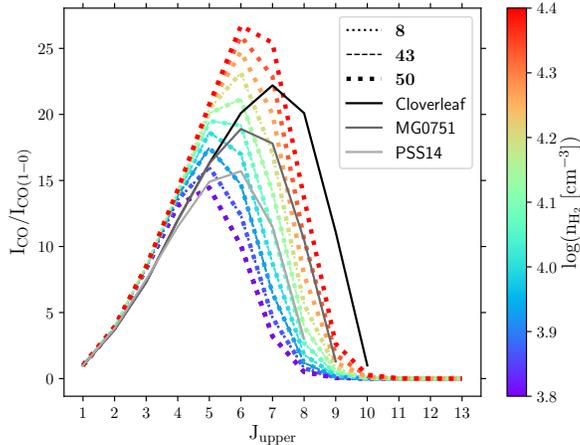
Fig. 2.8 shows the CO SLEDs obtained in this work (discontinuous lines) in comparison to three  $z \sim 3$  quasars with well sampled SLEDs (grey continuous lines), Cloverleaf ( $z = 2.6$ , Weiss et al. 2007; Bradford et al. 2009), MG 0751+2716 ( $z = 3.2$ , Weiss et al. 2007) and PSS1409 ( $z = 2.6$ , Weiss et al. 2007). From this figure, we see that none of the SLEDs obtained in this work are similar in shape to those three obtained previously, indicating that our quasars have different physical properties. Our results yield SLEDs with J peak somewhat lower than the expected range for quasars (J $\sim$ 6-8), varying from J $\sim$ 5 to J $\sim$ 7.

The six sources with IDs 3, 7, 13, 18, 29 and 39, have only upper limits for the targeted CO transitions. Our observations are therefore not able to put any constraint on the SLED of these sources. To compute the upper limits for the CO(1-0) luminosity, we assumed that the models selected for the previous three sources with IDs 8, 43 and 50, apply to the full observed sample.

We first selected the CO SLEDs found for the quasars with IDs 8, 43 and 50 (shown in Fig. 2.8). Depending on which emission line upper limit flux was measured for each of the six sources mentioned above, we used the CO(1-0)/CO(6-5) and/or CO(1-0)/CO(7-6) ratios. Then, we multiplied these CO ratios by the observed upper limit flux for each of them. After this, we estimated upper limits for the molecular gas masses of the six quasars with IDs 3, 7, 13, 18, 29 and 39. For the sources that have flux upper limits for both CO



**Figure 2.7:** CO SLEDs (at  $T_{\text{kin}} = 30$  K) obtained with the joint CO and [C I] constraints, for the sources with IDs 8, 43 and 50 (top, middle and bottom panels, respectively), the SLEDs that can fit the joint constraints are represented by the coloured curves. The ladders show the CO line flux normalized to the CO(6-5) line as a function of rotational quantum number  $J$ , and the vertical black arrow represents the observed upper limit on the CO(7-6)/CO(6-5) ratio for each source. The different values in  $\text{H}_2$  density are shown by the colour bar. The grey curves in the background show the discarded CO SLEDs (from Fig. 2.6).



**Figure 2.8:** CO SLEDs for the sources with IDs 8, 43 and 50 (discontinuous lines; see legend) compared to the SLEDs for the quasars Cloverleaf, MG0751 and PSS14, obtained by Weiss et al. (2007) (continuous grey lines). The ladders correspond to the CO line flux normalized to the CO(1-0) line as a function of rotational quantum number  $J$ .

lines (quasars with IDs 13, 29 and 39), two upper limits in molecular gas mass were derived and the higher value was selected.

To summarize: the resulting molecular masses of our sample, obtained from applying only the CO constraints, are in the range of  $(0.5 - 64.1) \times 10^{11} M_{\odot}$  for the sources with IDs 8, 43 and 50, and are  $< 6.9 \times 10^{11} M_{\odot}$  for the objects with IDs 3, 7, 13, 18, 29 and 39<sup>10</sup>. We note that the range mentioned includes masses of the order of  $10^{12} M_{\odot}$ , which are implausible given the expected dark-matter halo mass for these quasars ( $M_{\text{DM}} \sim 10^{12.5} M_{\odot}$ , e.g., White et al. 2012). These values are also much higher than the values found in the literature for other high- $z$  quasars ( $\sim 10^9 - 10^{11} M_{\odot}$ , e.g., Weiß et al. 2003; Walter et al. 2011; Anh et al. 2013; Hill et al. 2019). These high masses are caused by models with densities of  $n_{\text{H}_2} \sim 10^3 - 10^{3.8} \text{ cm}^{-3}$ . This disagreement with the literature values suggests that such low densities are not plausible in explaining our current constraints on the emission from the high- $J$  CO transitions. We note that the molecular gas masses obtained from applying the CO and [C I] constraints jointly, remove such models.

We emphasize that the final selected densities ( $n_{\text{H}_2} \sim 10^{3.8} - 10^{4.4} \text{ cm}^{-3}$ , see Fig. 2.8) represent those values needed by the high- $J$  CO lines to match the molecular masses estimated from [C I](2-1). These densities should not be regarded as representative of those emitting the [C I](2-1) line, unless this emission is more extended than the high- $J$  CO transitions.<sup>11</sup>

<sup>10</sup>We note that the latter upper limit is based on the final CO SLEDs obtained for the sources with IDs 8, 43, and 50 (Figure 2.8).

<sup>11</sup>For completeness, we report here the predictions of the [C I](2-1)/CO(6-5) ratio using our 10 finally selected LVG models. These models with  $T_{\text{kin}} = 30$  K are characterized by [C I](2-1)/CO(6-5)  $< 1.33$ . In the case of ID 29, this value is lower than the observed [C I](2-1)/CO(6-5)  $> 4.36$ . This apparent discrepancy could be solved by invoking a [C I](2-1) solid angle  $> 3 \times$  the high- $J$  CO solid angle (which corresponds

The molecular masses that are finally derived after applying a more restricted set of models are in the range of  $0.4 - 6.9 \times 10^{11} M_{\odot}$  for quasars with IDs 3, 8, 29, 39, 43 and 50, and  $< 1.1 \times 10^{11} M_{\odot}$  for quasars with IDs 7, 13 and 18. The brightest of our detected sources still has an estimated molecular gas mass that is higher than typically found in the literature. We will discuss this further in the next sections.

We finally check the molecular gas mass estimate for ID 3 and ID 39, for which we identify [C $\text{I}$ ](2-1) based on the uncertain C $\text{IV}$  systemic redshift (section 2.3.1). In the case the detected lines are CO(7-6) instead of [C $\text{I}$ ](2-1), the final models shown in Figure 2.8 would imply molecular gas masses of  $(1.2 - 4.5) \times 10^{11} M_{\odot}$  and  $(0.2 - 0.6) \times 10^{11} M_{\odot}$ , respectively. These values would still confirm our findings, with ID 3 having a large molecular mass.

## 2.5 Comparison with the large-scale Ly $\alpha$ emission

To investigate any relation between the molecular gas phase and the cool halo gas, in this section we compare the APEX observations with the Ly $\alpha$  line properties of the associated large-scale nebulae discovered with MUSE (Arrigoni Battaia et al., 2019a). First, we focus on finding any trend in the properties of each phase when compared with the other. Secondly, we study the line emission locations and profiles.

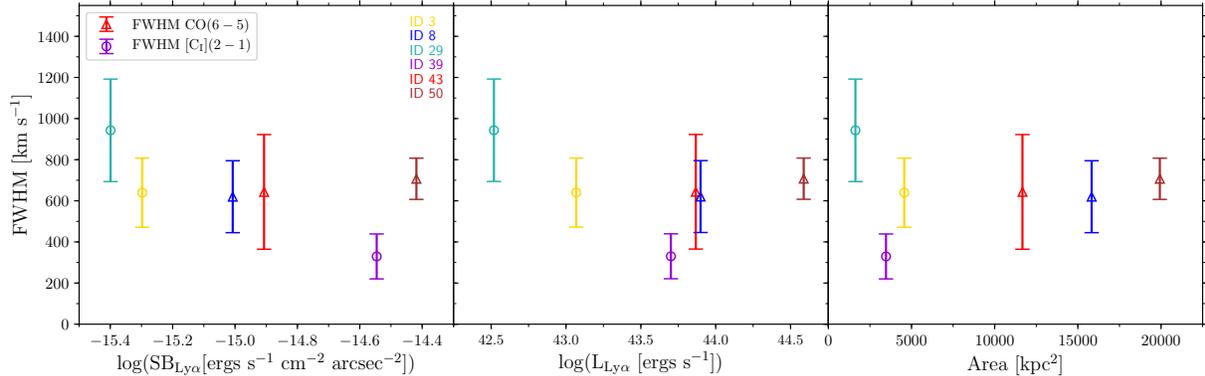
### 2.5.1 Molecular lines versus Ly $\alpha$ line properties

The FWHM of molecular lines from high-redshift galaxies are frequently used as a dynamical mass estimator (assuming a rotating disk geometry, e.g., Walter et al., 2003; Narayanan et al., 2009; Wang et al., 2013), and hence their stellar mass and therefore their halo mass can be determined (see Section 2.6.1 for further discussion). As halos of different masses are expected to be characterised by different fractions of cool and hot gas (e.g., Dekel & Birnboim, 2006), the FWHM of molecular lines could show important trends with the Ly $\alpha$  properties. For instance, more massive halos should in principle show larger FWHM of the molecular lines and smaller fractions of cool gas, with consequently smaller Ly $\alpha$  luminosities ( $L_{\text{Ly}\alpha}$ ) and Ly $\alpha$  areas compared to less massive halos. We start by comparing the FWHM of the observed molecular lines with the total  $L_{\text{Ly}\alpha}$ , the average Ly $\alpha$  surface brightness  $\text{SB}_{\text{Ly}\alpha}$ , and the area encompassed by the Ly $\alpha$  emission by the nebulae surrounding the quasars in our sample.

Figure 2.9 shows this comparison for the different sources in our sample (represented by different colours). Specifically, the left panel shows the FWHM of the molecular lines as a function of  $\text{SB}_{\text{Ly}\alpha}$  (corrected for the cosmological dimming), the central panel shows the FWHM versus  $L_{\text{Ly}\alpha}$ , and the right panel shows the FWHM versus the area of the Ly $\alpha$

---

to  $> 1.7 \times$  the high-J CO source radius) and/or with multi-component models (e.g., Weiß et al. 2003; Harrington et al. 2021).



**Figure 2.9:** FWHM of the detected CO(6-5) (triangles) and [C I](2-1) (circles) emission lines versus Ly $\alpha$  surface brightness corrected for cosmological dimming (first panel), the total Ly $\alpha$  luminosity (second panel), and physical area (third panel) for the large-scale Ly $\alpha$  nebulae discovered around the quasars in our sample (Arrigoni Battaia et al., 2019a).

nebulae. The legend in each panel indicates the different markers used for the different molecular lines (CO(6-5) and [C I](2-1)).

From this figure, we find that there is no clear correlation between FWHM of the molecular lines and Ly $\alpha$  properties. Also, we note that all the quasars show similar values of FWHM (in the range of 329 - 943 km s $^{-1}$ , average FWHM  $647 \pm 129$  km s $^{-1}$ ), considering the uncertainties. In the literature, the molecular linewidths (CO and [C I] lines) found for high- $z$  quasars have values between  $\sim 150$  - 450 km s $^{-1}$  (e.g., Weiß et al., 2003, 2007; Walter et al., 2011; Schumacher et al., 2012; Venemans et al., 2017). The values found for our quasars are larger on average. For example, quasars with IDs 29 and ID 50 have a FWHM of  $943 \pm 249$  km s $^{-1}$  ([C I](2-1) emission) and  $707 \pm 100$  km s $^{-1}$  (CO(6-5) emission) respectively, suggesting that they have a different kinematics and/or physical properties compared to the quasars studied in the literature.

As molecular gas is expected to form from the cooling of  $10^4$ K gas (e.g., Dalgarno & McCray, 1972), one may naively expect to observe the largest molecular reservoirs in sources with the largest and brightest Ly $\alpha$  nebulae. As a next step, we focus on the molecular gas masses obtained from the joint constraints from CO and [C I] (see Section 2.4.3), and compare them to  $L_{\text{Ly}\alpha}$ ,  $\text{SB}_{\text{Ly}\alpha}$  and the area of the large-scale nebulae.

The results of this analysis are presented in Figure 2.10, where the colours represent different sources, the vertical dashed lines are ranges and the arrows upper limits of the estimated molecular gas masses. The horizontal black lines represent the expected molecular mass for the quasar host galaxies on the main sequence (MS) of star-forming galaxies ( $\log(M_{\text{H}_2}[M_{\odot}]) = 10.98^{+0.32}_{-0.51}$ ), on  $4 \times \text{MS}$  ( $\log(M_{\text{H}_2}[M_{\odot}]) = 11.29$ ) and  $10 \times \text{MS}$  ( $\log(M_{\text{H}_2}[M_{\odot}]) = 11.51$ ). These values were computed under the following assumptions: i) a quasar halo mass of  $\log(M_{\text{halo}}[M_{\odot}]) = 12.68^{0.81}_{-0.67}$  (for  $z \sim 3$  quasars), estimated in the study of Kim & Croft (2008) from Ly $\alpha$  forest statistics (this value encompasses estimates from quasar clustering, e.g., Shen et al. 2007; White et al. 2012; Timlin et al. 2018), ii) the  $M_{\text{halo}} - M_*$

relation from [Moster et al. \(2018\)](#) to estimate the stellar mass of the objects in our sample, and iii) a molecular gas fraction expected for objects on MS, 4 $\times$ MS and 10 $\times$ MS of star-forming galaxies, as defined in the empirical relation of [Tacconi et al. \(2018\)](#). The grey shaded region represents the large uncertainties in the calculation for 1 $\times$ MS. We note that quasars hosts are estimated to have star formation rates higher or similar to MS galaxies at the same redshift (e.g., [Zhang et al. 2016](#); [Shangguan et al. 2020](#); [Jarvis et al. 2020](#); [Circosta et al. 2021](#)). The histogram shows the  $M_{\text{H}_2}$  distribution for our sample (red) and for a sample of  $z \sim 2.5 - 3$  quasars (blue) extracted from the literature as explained in the figure caption. The molecular masses reported for these quasars are overall below the computed predictions (see histogram in Fig. 2.10), possibly indicating some gas depletion with respect to similarly massive star-forming galaxies.

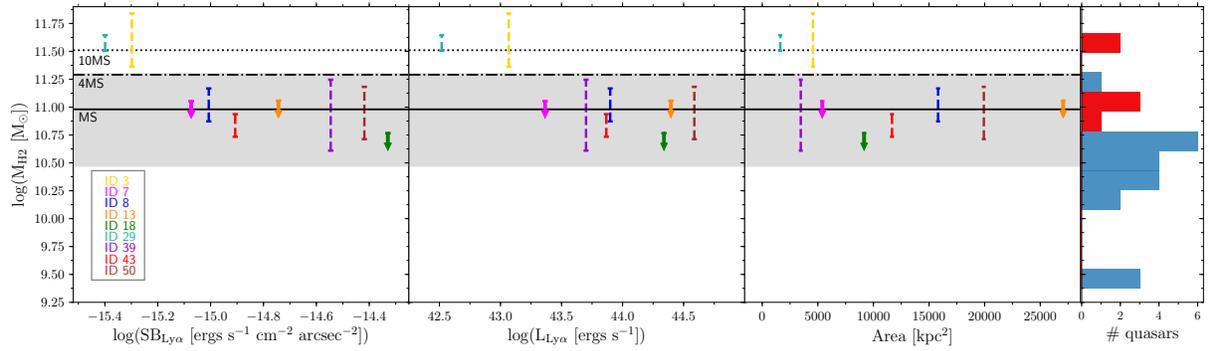
It is clear that the molecular gas masses of the quasars with APEX detections are in agreement with the expected values for MS galaxies, with the exception of quasars with IDs 3 and 29, which have masses well above the MS. Intriguingly, these two objects characterised by the highest molecular gas masses are associated with the Ly $\alpha$  nebulae with the lowest values of  $L_{\text{Ly}\alpha}$  and  $\text{SB}_{\text{Ly}\alpha}$ . The third panel of Fig. 2.10 shows that the most massive molecular reservoirs are associated with some of the smallest nebulae. They also have the highest FWHM [C I](2-1) emission lines (see Fig. 2.9). We will discuss this result in more detail in Section 2.6.3.

### 2.5.2 Molecular tracers vs Ly $\alpha$ line profiles

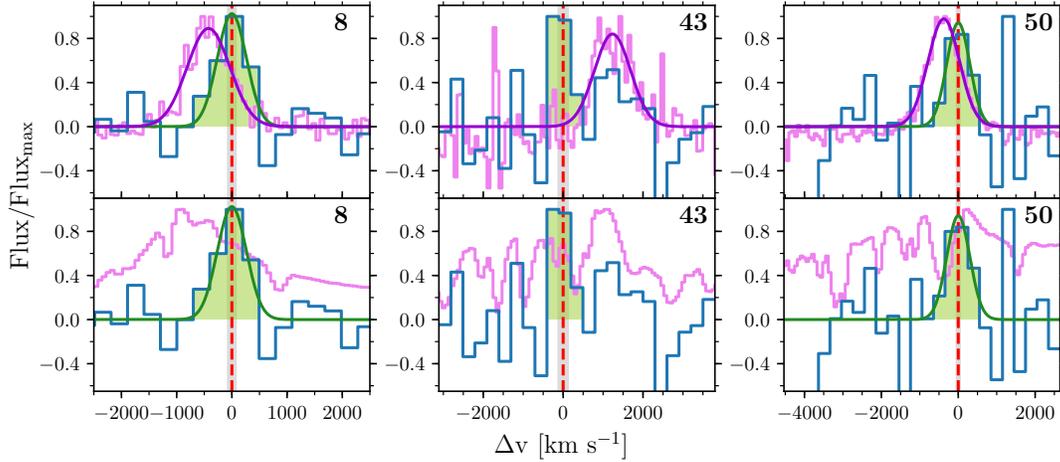
Now that we have a more robust estimate of the systemic redshift  $z_{\text{mol,sys}}$  for some of our sources, we can compare its location with the redshift of the discovered Ly $\alpha$  nebulae, assessing if the Ly $\alpha$  emission spans similar velocity ranges with respect to the molecular phase. For this purpose, we compare the molecular and Ly $\alpha$  line profiles extracted within the same aperture (the APEX beam; see Fig. 2.2), and also compare those with the quasar spectra.

Figures 2.11 and 2.12 show the normalized APEX CO(6-5) and [C I](2-1) detections (blue) in comparison to their normalized Ly $\alpha$  nebula spectrum (upper row) and quasar spectrum (bottom row). The quasar spectra have been extracted within circular apertures of 1.5 arcsec radius. The vertical red dashed line represents the current systemic redshift estimated from the molecular lines and the grey shaded area corresponds to its uncertainty (see Section 2.3.2).

From both figures, we note that some of the peaks of the Ly $\alpha$  nebulae present a significant shift from the current systemic redshift. For each object, we listed these velocity shifts  $\Delta v_{\text{Ly}\alpha\text{Neb-mol}}$  between the quasar systemics  $z_{\text{mol,sys}}$  and the Ly $\alpha$  of the nebularities in the last column of Table 2.3. The values of  $\Delta v_{\text{Ly}\alpha\text{Neb-mol}}$  are in the range of -423 to 1236 km s $^{-1}$ , with the quasars with ID 8 (or UM683) and ID 43 (or CTSH22.05) having the bluest and the reddest shift, respectively. The velocity shifts do not show any trend with respect to the Ly $\alpha$  physical properties explored in this paper. Considering that the Ly $\alpha$  photons experiment changes in frequency due to the scattering processes, these shifts can be an indication of bulk inflows (blueshift) or outflows (redshift) (e.g., [Prescott et al.](#),



**Figure 2.10:** Molecular gas mass  $M_{\text{H}_2}$  derived from the joint CO and [C I] constraints versus  $\text{Ly}\alpha$  surface brightness corrected for cosmological dimming (first panel), the total  $\text{Ly}\alpha$  luminosity (second panel), and physical area (third panel) for the large-scale  $\text{Ly}\alpha$  nebulae discovered around the quasars in our sample (Arrigoni Battaia et al. 2019a). The ID numbers of the sources are shown in the first panel. The vertical dashed lines represent ranges of  $M_{\text{H}_2}$ , and the arrows indicate upper limits (see Table 2.4). The horizontal lines show the expected location of massive star-forming main sequence MS,  $4\times\text{MS}$  and  $10\times\text{MS}$  galaxies using the gas fractions in Tacconi et al. (2018) and the halo mass-stellar mass relation in Moster et al. (2018). The grey area is the uncertainty on this expectation, for  $1\times\text{MS}$ . The histogram shows the  $M_{\text{H}_2}$  distribution for our sample considering the average values of the mass ranges (red), and for a sample of  $z \sim 2.5 - 3$  quasars (blue) extracted from the literature (Barvainis et al. 2002; Weiß et al. 2003; Beelen et al. 2004; Walter et al. 2011; Schumacher et al. 2012; Hill et al. 2019; Bischetti et al. 2021), whose  $M_{\text{H}_2}$  have been estimated with different molecular gas tracers and have a typical uncertainty of  $\sim 0.1$  dex.

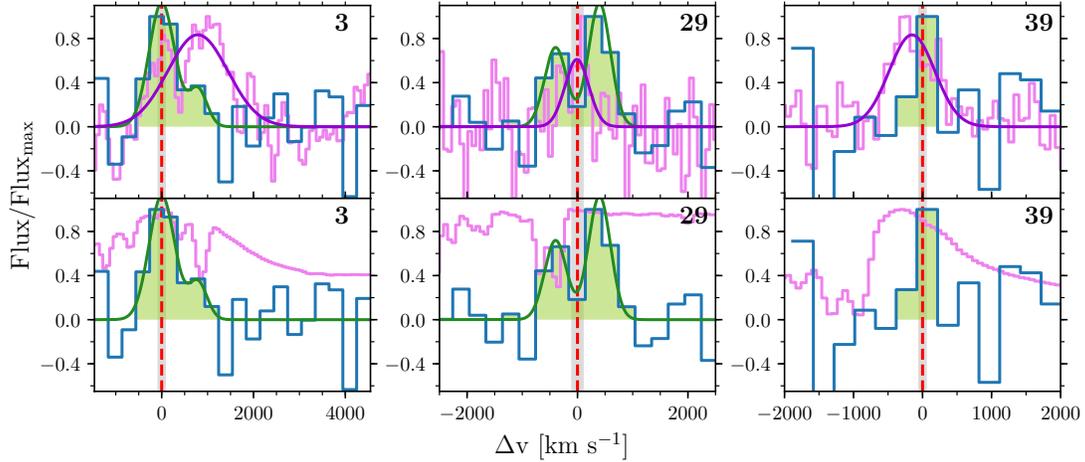


**Figure 2.11:** *Top panels:* Normalized APEX CO(6-5) emission (blue) of the three detected sources, ID 8, ID 43, and ID 50 compared with the nebulae Ly $\alpha$  emission from MUSE (magenta) extracted within the APEX beam (see Figure 2.2). The APEX spectra are shown with a bin size of  $300 \text{ km s}^{-1}$ . Each quasar ID number is shown in the top right corner of each panel. The red vertical lines with grey shaded area correspond to the systemic redshift estimated in this work with its uncertainty. The dark green curves and green shaded areas follow the same notation as in Figure 2.4. To guide the eye, the purple curves represent a Gaussian fit applied to the Ly $\alpha$  emission lines. *Bottom panels:* same as above, but comparing with the quasar spectra extracted within a circular aperture of  $1.5 \text{ arcsec}$  radius.

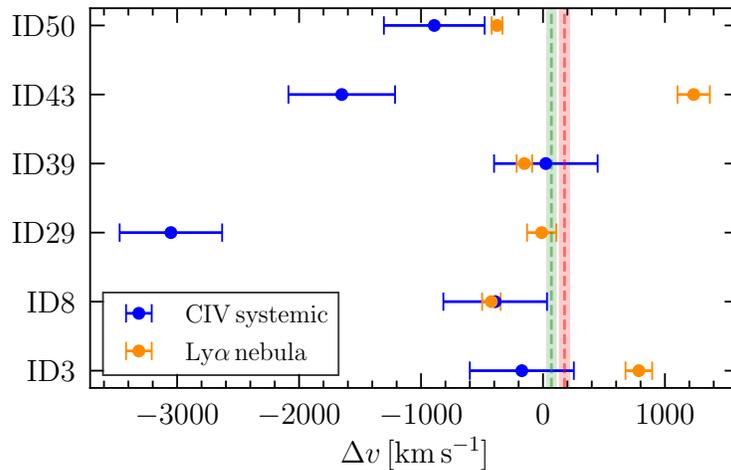
2015; Dijkstra, 2017).

Figure 2.13 shows the velocity shifts between  $z_{\text{sys}}$  and  $z_{\text{mol,sys}}$  (blue), and redshift of the Ly $\alpha$  nebulae and  $z_{\text{mol,sys}}$  (orange). The now obtained systemic redshifts from molecular tracers are found to be, in most of the cases, more consistent with the Ly $\alpha$  nebular redshifts, showing an average shift of  $176 \pm 39 \text{ km s}^{-1}$  (red vertical line). We note that this shift is larger than the average shift found at  $z \sim 6$ ,  $69 \pm 36 \text{ km s}^{-1}$  (Farina et al., 2019), obtained by comparing [CII] redshifts and the nebular redshifts for nine sources (green vertical line). The small number statistics hampers any conclusion from this comparison. For example, if we remove from our sample the radio-loud object, ID 3, we would get a smaller average shift of  $46 \pm 35 \text{ km s}^{-1}$ , which is consistent with the work by Farina et al. (2019).

This analysis shows that the systemic redshifts obtained from CIV are not reliable in 4 out of 6 cases ( $\sim 70\%$ ). This poor reliability was already indicated by the large velocity shifts of the extended nebular Ly $\alpha$  emission and the CIV redshifts in surveys targeting  $z \sim 2 - 3$  quasars, reporting values as high as  $\sim 6000 \text{ km s}^{-1}$  (Arrigoni Battaia et al., 2019a; Cai et al., 2019). The same works found that the peak of the nebular emission has smaller velocity shifts with respect to the observed peak of the Ly $\alpha$  emission of the quasar. Figures 2.11 and 2.12 indicate that for 3 out of 6 quasars (ID 3, 29, and 50), this was the case because of that peak being at the real systemic. We discuss the most significant shifts between  $z_{\text{mol,sys}}$  and the redshift of the Ly $\alpha$  nebulae in Section 2.6.2.



**Figure 2.12:** Same as Figure 2.11, but for the APEX [Cl](2-1) detections (blue).



**Figure 2.13:** Velocity shifts between  $z_{\text{sys}}$  (from CIV) and  $z_{\text{mol,sys}}$  (blue), and redshift of Ly $\alpha$  nebulae and  $z_{\text{mol,sys}}$  (orange). The green and red vertical dashed lines with errors are the average shift for Ly $\alpha$  nebulae for the sample in Farina et al. (2019) ( $69 \pm 36 \text{ km s}^{-1}$ ; for 9  $z \sim 6$  quasars with [CII] systemic redshifts), and for our work ( $176 \pm 39 \text{ km s}^{-1}$ ), respectively.

## 2.6 Discussion

In this section, we discuss different scenarios regarding the quasar host properties found in this work, and we try to interpret the link between the molecular gas content and the large scale Ly $\alpha$  emission. Future studies are essential to confirm or reject what we propose in the following sections.

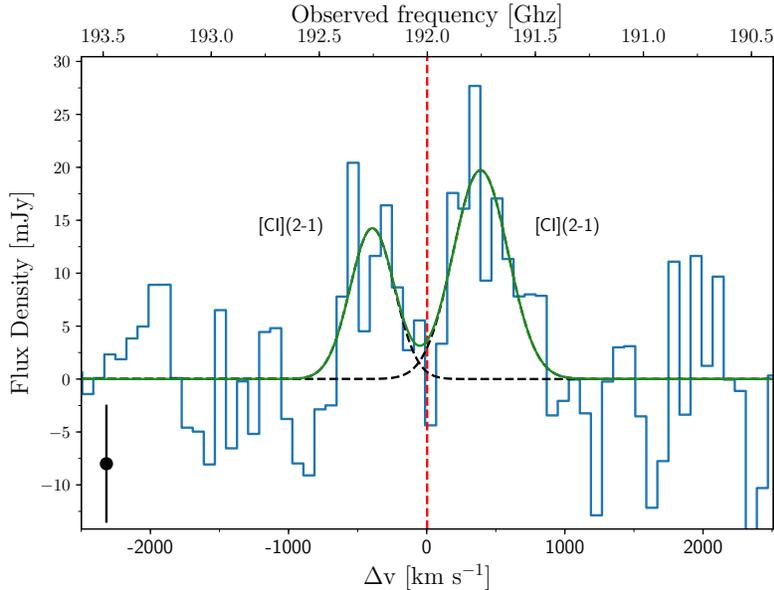
### 2.6.1 Quasar host properties

#### Dynamical masses

For high- $z$  galaxies, it is possible to derive the dynamical mass of the system using the formula commonly used in the literature (e.g., [Walter et al., 2003](#); [Wang et al., 2013](#); [Venemans et al., 2016](#)), assuming rotational support:  $M_{\text{dyn}} = 2.33 \times 10^5 v_{\text{circ}}^2 R [M_{\odot}]$ , where  $R$  is the disk radius from the observed line and  $v_{\text{circ}}$  is the maximum circular velocity of the gas disk, which is calculated as  $v_{\text{circ}} = 0.75 \times \text{FWHM} / \sin(i)$ , i.e. it depends on the FWHM of the observed molecular line and inclination  $i$  of the galaxy (with respect to the plane of the sky). From our molecular data it is not possible to estimate the parameters  $R$  and  $i$ , for this reason we proceed with a set of assumptions. [Arrigoni Battaia et al.](#) in preparation studied the quasar with ID 13 (or PK-1017+109) through ALMA observations, and found a radius of  $\sim 3$  kpc for this object. This is on the high side of the typical disk radii range found in literature (from molecular gas) for high- $z$  quasar hosts ( $\sim 0.5 - 2.5$  kpc, e.g., [Carilli et al. 2003](#); [Walter et al. 2004](#); [Aravena et al. 2008](#); [Riechers et al. 2009](#); [Polletta et al. 2011](#); [Schumacher et al. 2012](#); [Molina et al. 2021](#); [Stacey et al. 2021](#)). Here we adopt  $R = 3$  kpc, which is in agreement with our ALMA observations of one of these systems.

Applying the above to five sources in our sample with FWHM estimations, excluding quasar ID 29 that is discussed in Section 2.6.1), we find inclination-dependent dynamical masses in the range of  $M_{\text{dyn}} \sin^2(i) \sim (0.4 - 1.9) \times 10^{11} M_{\odot}$  for quasars with IDs 3, 8, 39, 43 and 50. The quasar with ID 39 (or SDSS J0100+2105) is the least massive and the quasar with ID 50 (or SDSS J0819+0823) the most massive. The specific values for these dynamical masses are tabulated in the last column of Table 2.4, with uncertainties considering only the errors from their FWHMs. Even though these dynamical masses are highly uncertain, we can get a rough idea on the inclination. Indeed, assuming that the amount of dark matter is negligible on host scales,  $M_{\text{dyn}} \approx M_* + M_{\text{H}_2}$ , where  $M_*$  is the stellar mass. Assuming that the stellar mass is at least as high as the highest allowed molecular mass from our calculations, we find that two sources, IDs 3 and 39, require relatively small angles ( $\sim 20^\circ$ ), while the other four sources, IDs 8, 29, 43 and 50, require an inclination ( $\sim 50^\circ - 60^\circ$ ) close to the mean expected value of  $\sin(i) = 0.79$  ( $i = 57.3^\circ$ , [Law et al. 2009](#)).

One caveat is that the estimates are subject to the different molecular lines used. We computed that the CO lines reported in the literature for  $z \sim 3$  quasars are on average larger than [CI] lines by  $112 \pm 43$  km s $^{-1}$  (e.g., [Weiß et al. 2003](#); [Schumacher et al. 2012](#); [Banerji et al. 2018](#); [Yang et al. 2019](#)). As these transitions have different critical densities,



**Figure 2.14:** APEX [C I](2-1) spectral line observation for Q-0115-30 (quasar with ID 29). The spectrum is shown with a bin size of  $80 \text{ km s}^{-1}$ , the highest resolution possible. The red vertical line represent the systemic redshift estimated from the molecular lines. The curves and symbol follow the same notation as Figure 2.5.

this could be an indication that the lines trace different gas components, which can be the case if [C I] and the high- $J$  CO transitions (as CO(6-5)) traces gas also on different scales, with [C I] extending to outer scales (Schumacher et al., 2012).

### The case of Q-0115-30 (ID 29): disk or merger?

In Section 2.3.1, we showed that the quasar with ID 29 (or Q-0115-30) has a [C I](2-1) line emission consisting of two Gaussian components separated by  $\sim 800 \text{ km s}^{-1}$ . While this line identification needs further confirmation, in this section we discuss their possible origins. This double-peaked feature is unique within our sample and could be produced by a rotating disk or a merger between two molecular gas reservoirs (e.g., Neri et al., 2003; Narayanan et al., 2006; Greve & Sommer-Larsen, 2008; Polletta et al., 2011). In Figure 2.14 we show this double-peaked line emission at the highest resolution possible (bin size of  $80 \text{ km s}^{-1}$ ) for which there is still a decent peak S/N of  $\sim 2.5$  and  $\sim 3.5$ .

We now calculate the dynamical masses, under the assumption that the system is virialized. For the case of the gas distributed in a rotating disk, we estimate the dynamical mass  $M_{\text{dyn}}^{\text{disk}}$  as explained in Section 2.6.1, but considering the velocity difference between the two [C I](2-1) peaks  $\Delta v$ , following the formalism of Neri et al. (2003). For the disk radius, we assume a value of  $R = 3 \text{ kpc}$ . We obtain  $M_{\text{dyn}}^{\text{disk}} \sin^2(i) \sim 4.4 \times 10^{11} M_{\odot}$ . This quasar has the highest dynamical mass of the sample, which is somewhat expected given its larger molecular mass. A larger molecular mass and dynamical mass with respect to

the other quasars could also imply a larger dark-matter halo mass.

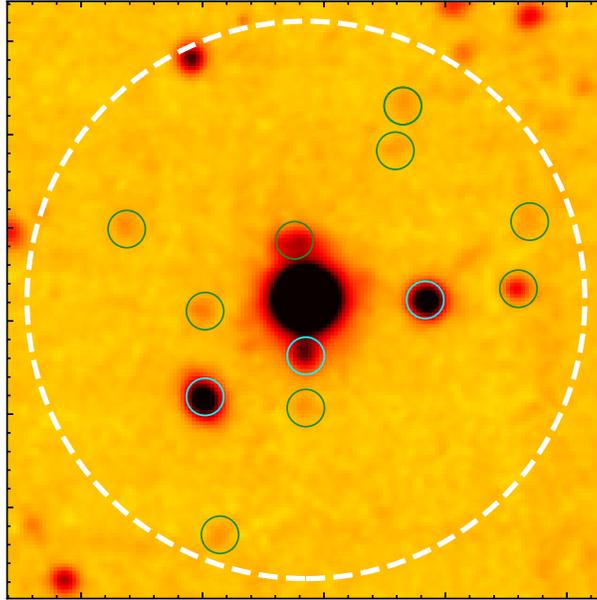
The combination of an almost edge-on disk and a larger halo mass, could explain a low level of extended Ly $\alpha$  emission as observed for this system. In a unification scenario for AGN (e.g., Urry & Padovani 1995) radiation from the quasar escapes through two ionization cones determined by obscuration from the dust distribution on small scales. Since we observe this object as a quasar, our line-of-sight is inside its ionizing cones. A large fraction of the quasar emission will pass through its massive host molecular disk, and will be absorbed before reaching the CGM. In other words, the obscuration from such a misaligned host galaxy decreases the budget of ionizing photons able to boost the Ly $\alpha$  emission on CGM scales, resulting in a dimmer and less extended Ly $\alpha$  nebula. Any Ly $\alpha$  emission boosted on the scales of the host galaxy likely experience a large number of scatterings and final absorption due to dust, expected to be present in large quantities in such a large molecular reservoir (e.g., Venemans et al. 2017; see Section 2.6.3 for a rough indication of the dust mass).

On the other hand, a large halo mass implies a smaller fraction of cool gas able to survive the shock-heating process and penetrate the halo down to the host galaxy (e.g., Dekel & Birnboim 2006). In turn, warmer halo and a smaller cool CGM reservoir would result in a lower level of Ly $\alpha$  emission. We therefore conclude that in this scenario, the mass of the halo together with the geometry of the system and Ly $\alpha$  radiative transfer effects could determine the small extent and low SB level of the Ly $\alpha$  nebula surrounding this quasar. Radiative transfer calculations possibly coupled to cosmological hydrodynamical simulations with dust implementations are needed to confirm this interpretation.

For the merger scenario, we can estimate the dynamical mass  $M_{\text{dyn}}^{\text{merger}}$  following Greve & Sommer-Larsen (2008):  $M_{\text{dyn}}^{\text{merger}} = 2 \times (2.33 \times 10^5) \text{FWHM}^2 R [M_{\odot}]$ , where FWHM is the full width half maximum of the observed [C I](2-1) emission lines, and  $R$  is the half projected distance between the two merging objects. We used the same  $R$  value assumed in the disk scenario (to obtain a lower limit in  $M_{\text{dyn}}^{\text{merger}}$ )<sup>12</sup>, and  $\text{FWHM} = 943 \pm 249 \text{ km s}^{-1}$  (see Table 2.2), which was calculated as the sum of each single Gaussian fit corresponding to each [C I](2-1) observed line. We obtain a mass of  $M_{\text{dyn}}^{\text{merger}} = 1.2 \times 10^{12} M_{\odot}$ . Assuming that the amount of dark matter is negligible on host scales, the stellar mass is given by  $M_* \approx M_{\text{dyn}}^{\text{merger}} - M_{\text{H}_2}$ . Then, using the molecular gas mass range derived from applying the CO and [C I] constraints, we estimate a total stellar mass in the range  $M_* = (7.6 - 8.8) \times 10^{11} M_{\odot}$ . Following Moster et al. (2018), this stellar mass implies an extremely large dark matter halo mass of  $M_{\text{halo}} \sim 10^{15.0} - 10^{15.1} M_{\odot}$ .

We also estimated the individual stellar masses of each component of the merging system. The molecular gas masses for the two different objects computed from the atomic carbon masses are  $M_{\text{H}_2}^{\text{obj1}} \sim 2.4 \times 10^{11} M_{\odot}$  and  $M_{\text{H}_2}^{\text{obj2}} \sim 3.7 \times 10^{11} M_{\odot}$ , respectively. Then, assuming  $M_{\text{H}_2}^{\text{obj1}}/M_{\text{H}_2}^{\text{obj2}} = M_*^{\text{obj1}}/M_*^{\text{obj2}}$ , we obtain  $M_*^{\text{obj1}} \sim 2.3 \times 10^{11} M_{\odot}$  and  $M_*^{\text{obj2}} \sim$

<sup>12</sup>In this scenario the companion object could be located anywhere inside the APEX beam, with the  $800 \text{ km s}^{-1}$  velocity shift representing a large peculiar velocity or a distance of at maximum  $\sim 1.6 \text{ Mpc}$  within the Hubble flow. Given the expected virial radius of such massive objects, the latter scenario will represent a merger in its very early phases or two objects still separated in the Hubble flow.



**Figure 2.15:** MUSE white-light image for the field around ID 29 (or Q-0115-30, at the center) highlighting the location of continuum sources within the APEX beam for the [C<sub>I</sub>](2-1) observations (white dashed circle). Low- $z$  interlopers (cyan circles) and sources with unknown redshift (green circles) are indicated (see Section 2.6.1 for details).

$3.6 \times 10^{11} M_{\odot}$ . The current MUSE data are not able to verify whether such a massive companion object exists, but this field is characterized by nine faint continuum objects with unknown redshifts within the beam of our APEX observations (Figure 2.15). These objects do not have any emission lines or absorption features that allow us to verify their redshifts. In particular, we note the presence of one continuum source at  $\sim 3$  (or  $\sim 23$  projected kpc) from the quasar. In summary, given the brightness of the quasar and the seeing-limited MUSE observations, we are not able to assess whether there is an ongoing merger on very small scales. A strongly dust-obscured and massive companion could be missed by our optical observations (e.g., Omont et al. 1996). Finally, we note that within the APEX beam we also find three low- $z$  interlopers (cyan circles in Figure 2.15), whose redshifts can be clearly determined ( $z \sim 0.52, 1.05, 0.07$ ). These redshifts ensure that the molecular emission from our targeted object is not contaminated by these sources.

Our consideration of ID 29 shows that its dynamical mass is very large, independent of the origin of the double-peaked [C<sub>I</sub>](2-1) line emission. Observing this system with an interferometer (e.g., ALMA, NOEMA) would allow us to map the [C<sub>I</sub>](2-1) emission and ascertain whether ID 29 is associated with an exceptionally massive molecular disk or is merging with a similarly massive companion. Such observations, by probing the  $\sim 1-2$  mm continuum, would be in turn able to verify the reason why the Ly $\alpha$  nebula around ID 29 is dimmer and less extended than similarly bright quasars.

### 2.6.2 Ly $\alpha$ nebulae kinematics with respect to molecular gas

The use of Ly $\alpha$  emission as a tracer of gas kinematics is a complex task due to its resonant nature (e.g., Neufeld 1990). Ly $\alpha$  photons are expected to interact several times with Hydrogen gas before escaping most astrophysical systems (e.g., Dijkstra 2017). During this process, scattered photons can experience both large changes in their frequency and a large displacement in space, possibly washing out any information on gas kinematics. Also, the larger the number of scatterings in a medium, the higher the probability for a Ly $\alpha$  photon to be absorbed by dust. If the dust distribution is not homogeneous, the Ly $\alpha$  line shape could be affected, possibly hiding information on the kinematics of the system.

In Section 2.5.2 we compared the molecular emission lines observed with APEX, with the Ly $\alpha$  emission obtained with MUSE and integrated within the APEX beam. The main result is that we find velocity shifts between the Ly $\alpha$  emission and the systemic redshift  $z_{\text{sys,mol}}$  in the range  $-423$  to  $1236$  km s $^{-1}$ . The presence of a velocity shift between this Ly $\alpha$  emission line integrated on halos scales and the systemic redshift of the quasar host galaxy, could be associated with a variety of physical processes occurring in the halo, e.g. substructures or gas infalling onto the central quasar, large-scale outflows, rotating structures or projection effects along the line-of-sight (see Arrigoni Battaia et al., 2018b, and references therein). In the following discussion we focus on the possibility that these shifts represent bulk inflows or outflows.

In particular, as indicated by Ly $\alpha$  radiative transfer modelling (e.g, Verhamme et al., 2006; Laursen et al., 2009), photons scattering through outflowing or inflowing gas should appear redshifted or blueshifted, respectively, from the systemic redshift. The quasar with ID 29 has a shift consistent with  $z_{\text{sys,mol}}$ . We cannot therefore draw any firm conclusion for this source with the current dataset. It could be that Ly $\alpha$  radiative transfer effects and a balanced interplay of outflows and inflows result in the absence of a strong line shift or that very fast outflows or inflows bring the gas out of resonance, allowing the observation of Ly $\alpha$  at  $z_{\text{mol,sys}}$ . For the remaining four sources, IDs 3, 8, 39, 43 and 50, we find significant offsets. In particular, IDs 3 and 43 present the largest shifts and also show the largest values of FWHM of the Ly $\alpha$  emission lines in the sample (with  $\text{FWHM}_{\text{Ly}\alpha} > 1000$  km s $^{-1}$ ), which is typically characterized by relatively quiescent kinematics (median  $\text{FWHM}_{\text{Ly}\alpha} = 915$  km s $^{-1}$ )<sup>13</sup>. In the following, we review these four sources.

The quasars with IDs 8, 39 and 50 show negative shifts of  $-423 \pm 76$ ,  $-152 \pm 64$  and  $-377 \pm 44$  km s $^{-1}$ , respectively, which could indicate an overall inflow signature from CGM scales onto the central quasar. The quasar with ID 39 (or SDSS J0100+2105) is surrounded by one of the less extended Ly $\alpha$  nebulae in our sample (see Fig. 2.10; Ly $\alpha$  extends out to  $\sim 50$  projected kpc). The observed blueshift for this source is lower than the inflow velocities ( $v_{\text{in}} \approx 0.9v_{\text{vir}} \sim 320 - 460$  km s $^{-1}$ ; Goerdt & Ceverino 2015) expected for the cool gas in quasar host halos ( $10^{12} - 10^{13}$  M $_{\odot}$ , e.g., White et al. 2012; Timlin et al. 2018

<sup>13</sup>The  $\text{FWHM}_{\text{Ly}\alpha}$  values reported in this work are from Gaussian fits of the Ly $\alpha$  emission integrated within the APEX beam. They therefore differ from the 2D first moment analysis on resolved maps presented in Arrigoni Battaia et al. (2019a). The  $\text{FWHM}_{\text{Ly}\alpha}$  values are not corrected for the instrument resolution.

and references therein). This could be due to projection effects or to the presence of other violent motions (e.g. winds, turbulences) that could wash out the inflow signature.

The observed blueshift for the quasars with ID 8 (or UM683) and ID 50 (or SDSS J0819+0823) is comparable with the inflow velocities expected for the cool gas in quasar host halos. The quasar with ID 8 has an average molecular reservoir and a Ly $\alpha$  nebula with intermediate surface brightness, extending out to  $\sim 145$  kpc (Arrigoni Battaia et al. 2019a.) Interestingly, ID 50 is surrounded by one of the brightest and more extended (in area) Ly $\alpha$  nebulae in the QSO museum sample (see Fig. 2.10; Ly $\alpha$  extends out to  $\sim 130$  projected kpc), with a clearly asymmetric morphology. The widths of their integrated Ly $\alpha$  lines,  $\text{FWHM}_{\text{Ly}\alpha} = 915 \pm 46 \text{ km s}^{-1}$  and  $\text{FWHM}_{\text{Ly}\alpha} = 967 \pm 38 \text{ km s}^{-1}$  (for IDs 8 and 50, respectively), may therefore be explained by gravitational motions in such massive halos ( $\text{FWHM}_{\text{ID}}^{\text{DM}} \sim 580 - 870 \text{ km s}^{-1}$ ; Arrigoni Battaia et al. 2019a) and by the presence of large turbulence within the cool gas reservoir.

In contrast, the quasars with IDs 3 and 43 show positive shifts of  $786 \pm 109 \text{ km s}^{-1}$  and  $1236 \pm 133 \text{ km s}^{-1}$ , respectively, which are suggestive of bulk large-scale outflows. These velocity shifts are overall higher than what has been found in Ly $\alpha$ -emitting galaxies (average of  $\Delta v \sim 200 \text{ km s}^{-1}$ , e.g., Trainor et al. 2015; Prescott et al. 2015; Verhamme et al. 2018), Lyman-break galaxies (average of  $\Delta v \sim 460 \text{ km s}^{-1}$ , see Dijkstra 2017 and references therein), and are also higher than the values found for other Ly $\alpha$  nebulae, the so called Ly $\alpha$  blobs ( $\sim 0 - 230 \text{ km s}^{-1}$ , e.g., Yang et al. 2011; McLinden et al. 2013; Prescott et al. 2015).

The quasar with ID 3 is radio-loud, characterized by a large reservoir of molecular gas and has one of the Ly $\alpha$  nebulae with the lowest surface brightness, extending out to a maximum of  $\sim 70$  projected kpc (Arrigoni Battaia et al. 2019a). Its integrated Ly $\alpha$  emission shows the most active kinematics of the sample ( $\text{FWHM}_{\text{Ly}\alpha} = 1542 \pm 196 \text{ km s}^{-1}$ ), which together with the observed large Ly $\alpha$  shift likely indicates that small-scale radio feedback is affecting the surrounding gas distribution in this system. The quasar with ID 43 has an average molecular reservoir and a Ly $\alpha$  nebula with intermediate surface brightness, extending out to  $\sim 118$  kpc (Arrigoni Battaia et al. 2019a). Its radio nature is not clear because there are no radio observations of this target in the literature. Its relatively active Ly $\alpha$  kinematics ( $\text{FWHM}_{\text{Ly}\alpha} = 1068 \pm 142 \text{ km s}^{-1}$ ), together with the large observed positive shift of the Ly $\alpha$  line with respect to CO(6-5) and the asymmetric morphology of the nebula with a northern bright region, might all be hints of outflowing gas in this system.

### 2.6.3 Relation between large molecular reservoirs and Ly $\alpha$ nebulae

In section 2.5.1, we found a tentative trend indicating that the two quasars with the lowest surface brightness for their Ly $\alpha$  nebulae are associated with the highest molecular gas

masses. As the molecular gas content is also closely linked to the amount of dust<sup>14</sup>, this could imply that the radiation of these quasars is more obscured by their host galaxies, therefore reducing the reprocessed Ly $\alpha$  emission from the surrounding gas distribution. Two effects could then be in play in the radiative transfer of these systems: first, scattered Ly $\alpha$  photons on small scales (tens of kpc) encounter a very high column density and dusty medium resulting in absorption of the emission, and second, most ionizing photons are not able to escape the massive host galaxy towards the CGM, limiting the detectability of cool halo gas in emission.

On top of this, the halo mass could play an important role in determining the reservoir of cool gas on CGM scales, with more massive halos having a smaller fraction of their gas in this phase. Taking also this aspect into consideration, there are at least two possible interpretations for this tentative trend: 1) IDs 3 and 29 (which present the most massive molecular reservoirs together with dimmer nebulae) occupy similar massive dark matter halos as the rest of quasars in the sample, but their molecular gas fraction (and consequently dust) are higher, resulting in smaller escape of ionizing and Ly $\alpha$  photons, 2) IDs 3 and 29 live in more massive dark matter halos, which are warmer and characterized by a smaller cool gas fraction, hence lower levels of Ly $\alpha$ .

Unfortunately, we cannot unambiguously infer the halo masses with the available observations, nor we can firmly quantify obscuration in these systems. Future analysis based on high-resolution interferometric observations and spectra from other CO transitions (e.g., with ALMA, NOEMA) are needed to better constrain the nature of these sources. Such follow-up studies together with constraints on the far-infrared continuum will be able to accurately determine the dust and molecular content of these systems, allowing a better description of the radiative transfer around these quasars.

## 2.7 Summary

With the purpose of gaining a new leverage on the gas cycle around  $z \sim 3$  quasars, we exploit the frequency window opened-up by SEPIA180 at APEX, targeting the CO(6-5), CO(7-6) and [C $\text{I}$ ](2-1) transitions for a sample of 9  $z \sim 3$  quasars already observed with MUSE/VLT. Specifically, these 9 targets (quasars with IDs 3, 8, 13, 18, 29, 39, 43 and 50) cover the most of the parameter space of Ly $\alpha$  nebulae properties in the QSO MUSEUM survey (Arrigoni Battaia et al., 2019a). These data (average rms of 2.6 mJy) allowed us to investigate any relation between the molecular gas content and the large-scale atomic phase ( $\sim 10^4\text{K}$ ) as traced by the Ly $\alpha$  emission. The main results of this work are summarized as follows.

---

<sup>14</sup>To have an idea of the dust masses of our sample, one could assume a gas-to-dust ratio of 70 (e.g., Riechers et al., 2013; Venemans et al., 2017) and considered that the  $\sim 75\%$  of the total gas mass corresponds to molecular gas (e.g., Wang et al., 2016). Under these very simple assumptions, our quasars should have a dust mass in the range of  $M_{\text{dust}} = 0.8 - 12.9 \times 10^9 M_{\odot}$  for the sources with IDs 3, 8, 29, 39, 43 and 50, and are  $< 2.1 \times 10^9 M_{\odot}$  for the objects with IDs 7 and 18. Under these assumptions objects with ID 3 and 29 with the most massive molecular reservoirs also present the highest dust masses.

- CO(6-5) emission is detected in three sources (quasars with IDs 8, 43 and 50) with fluxes of  $3.4 \leq I_{\text{CO}(6-5)} \leq 5.1 \text{ Jy km s}^{-1}$  and  $620 \leq \text{FWHM} \leq 707 \text{ km s}^{-1}$ , while the [C I](2-1) emission is detected in other three sources (quasars with IDs 3, 29, and 39) with fluxes of  $2.3 \leq I_{[\text{C I}](2-1)} \leq 15.7 \text{ Jy km s}^{-1}$  and  $329 \leq \text{FWHM} \leq 943 \text{ km s}^{-1}$ .
- The brightest and widest detection in the sample is the [C I](2-1) line ( $I_{[\text{C I}](2-1)} \sim 15.7 \text{ Jy km s}^{-1}$ ,  $\text{FWHM} \sim 943 \text{ km s}^{-1}$ ) associated to ID 29, which also presents the dimmest and smallest Ly $\alpha$  nebula in the sample.
- We obtained molecular gas masses  $M_{\text{H}_2}$  from the CO and [C I](2-1) constraints (Sections 2.4.1 and 2.4.2). Applying the two estimates together, the molecular masses for sources with detections are in the range  $M_{\text{H}_2} = (0.4 - 6.9) \times 10^{11} M_{\odot}$ , while for quasars with non-detections, the upper limits are  $M_{\text{H}_2} < 1.1 \times 10^{11} M_{\odot}$ . These masses are on the high side of the typical ranges found in the literature ( $\sim 10^9 - 10^{11} M_{\odot}$ ) for other  $z \sim 3$  quasars, with ID 29 and ID 3 being outliers.
- The systemic redshift of 6 quasars is refined using the CO(6-5) and [C I](2-1) detections. These new systemic redshifts have an average uncertainty of  $74.8 \text{ km s}^{-1}$ .
- We found significant offsets between the extended Ly $\alpha$  nebulae (Ly $\alpha$  emission extracted within the APEX beam) and the newly estimated systemic redshifts in five sources (quasars with IDs 3, 8, 39, 43 and 50). Two of these sources also show the largest values of  $\text{FWHM}_{\text{Ly}\alpha}$  in our sample ( $> 1000 \text{ km s}^{-1}$ ), which is otherwise characterized by relatively quiescent kinematics (median  $\text{FWHM}_{\text{Ly}\alpha} = 915 \text{ km s}^{-1}$ ). In Section 2.6.2 we discuss how these velocity shifts could be signatures of bulk inflows/outflows. For the quasars with IDs 8, 39 and 50, the nebular Ly $\alpha$  line shows a blueshift of  $-423 \pm 76$ ,  $-152 \pm 64$  and  $-377 \pm 44 \text{ km s}^{-1}$ , respectively, which could indicate a large-scale inflow. In contrast, quasars with IDs 3 and 43 have the bulk of the Ly $\alpha$  line redshifted by  $786 \pm 109 \text{ km s}^{-1}$  and  $1236 \pm 133 \text{ km s}^{-1}$ , respectively, which can indicate that the bulk of the large-scale gas is outflowing in both cases. For ID 3, this scenario is further strengthened by the fact that this quasar is radio-loud.
- We found that the two most massive molecular reservoirs in our sample (quasars with IDs 3 and 29,  $M_{\text{H}_2} = (2.3 - 6.9) \times 10^{11} M_{\odot}$  and  $M_{\text{H}_2} = (3.2 - 4.4) \times 10^{11} M_{\odot}$ , respectively) are associated with the dimmest and smallest Ly $\alpha$  nebulae. This suggests that the quasar host galaxy properties are key in understanding the powering and physics of surrounding Ly $\alpha$  nebulae. Obscuration from the host galaxy, due to physical properties (e.g., higher molecular gas fraction, higher dust fraction) or geometry, could reduce the escape of ionizing and Ly $\alpha$  photons emitted by the quasar, ultimately reducing the emission from the cool CGM. This interpretation is further strengthened by the detection of a double-peaked [C I](2-1) line emission for ID 29, which could be due to an inclined massive molecular disk, likely misaligned with the ionization cones of ID 29 (see Section 2.6.1). Another possibility is that these quasars

with more massive molecular reservoirs could be hosted by more massive dark-matter halos. Such halos are expected to be warmer and to show a smaller fraction of cool CGM gas, which in turn would result in lower values of Ly $\alpha$  emission.

These APEX observations therefore suggest that the quasar host-galaxy properties could affect the Ly $\alpha$  radiative transfer, thus regulating the amount of detectable emission on large scales around quasars. Overall, this work stresses the importance of simultaneously studying the physical properties of the multiphase gas reservoirs surrounding quasars. By characterizing the quasar host galaxy properties and system geometry (e.g., systemic redshift, inflow/outflow), future studies will be able to assess the radiative transfer effects affecting the quasar radiation and ultimately better constrain the powering mechanisms for extended gas reservoirs. Deep high-resolution interferometric observations (e.g., with ALMA, NOEMA) are indispensable to map the extent and geometry of the molecular reservoirs, ultimately linking its physics to the other CGM phases.

## 2.8 Appendix A: Stability of data

We demonstrated the stability of the SEPIA180 receiver by checking if the noise scales down properly when using larger velocity bins. In Figure 2.16 we show the rms of the final reduced spectra as a function of the bin size, for each of our sources in the different observed tunings. The smallest bin size corresponds to the original resolution of the spectra (i.e., the unbinned data). The solid lines represent the expected variation of the rms ( $\text{rms}_{\text{expected}}$ ) with respect to the bin size, which was estimated as:

$$\text{rms}_{\text{expected}} = \text{rms}_{\text{original}} \times \sqrt{\text{binsize}_{\text{original}}/\text{binsize}_{\text{observed}}} \quad (2.6)$$

where  $\text{rms}_{\text{original}}(\text{binsize}_{\text{original}})$  is the rms (bin size) at the original resolution, and  $\text{binsize}_{\text{observed}}$  is the bin size used in the final combined spectrum. Each panel has a lower subpanel, which shows the deviation of the observed rms from the predicted value, estimated as  $(\text{rms}_{\text{observed}} - \text{rms}_{\text{expected}})/\text{rms}_{\text{expected}}$ , where  $\text{rms}_{\text{observed}}$  is the final rms obtained in the final binned spectrum.

At a bin size of 300 km s $^{-1}$ , we obtained a median deviation of only 12% and 14% for CO(6-5) and [C $\text{I}$ ](2-1), respectively. Such a small deviation ensures that we can preform the analysis discussed in the main text.

Further, as a proof of the stability over time of the calibration of the receiver, in Fig. 2.17 we show the rms for each scan versus its PWV for the source with ID 39 (J 0100+2105). Each colour indicates a different observation date, spanning several months of APEX operations. These values correspond to both the CO(6-5) and [C $\text{I}$ ](2-1) - CO(7-6) observations. A linear fit shown by the grey solid line has been applied to the points, with a resulting squared correlation coefficient of  $r^2 = 0.93$  and a residual standard error of  $\sigma = 0.009$  K. The dashed lines enclose 95 per cent of the measurements (i.e.  $\pm 2\sigma$ ). Abrupt (vertical)

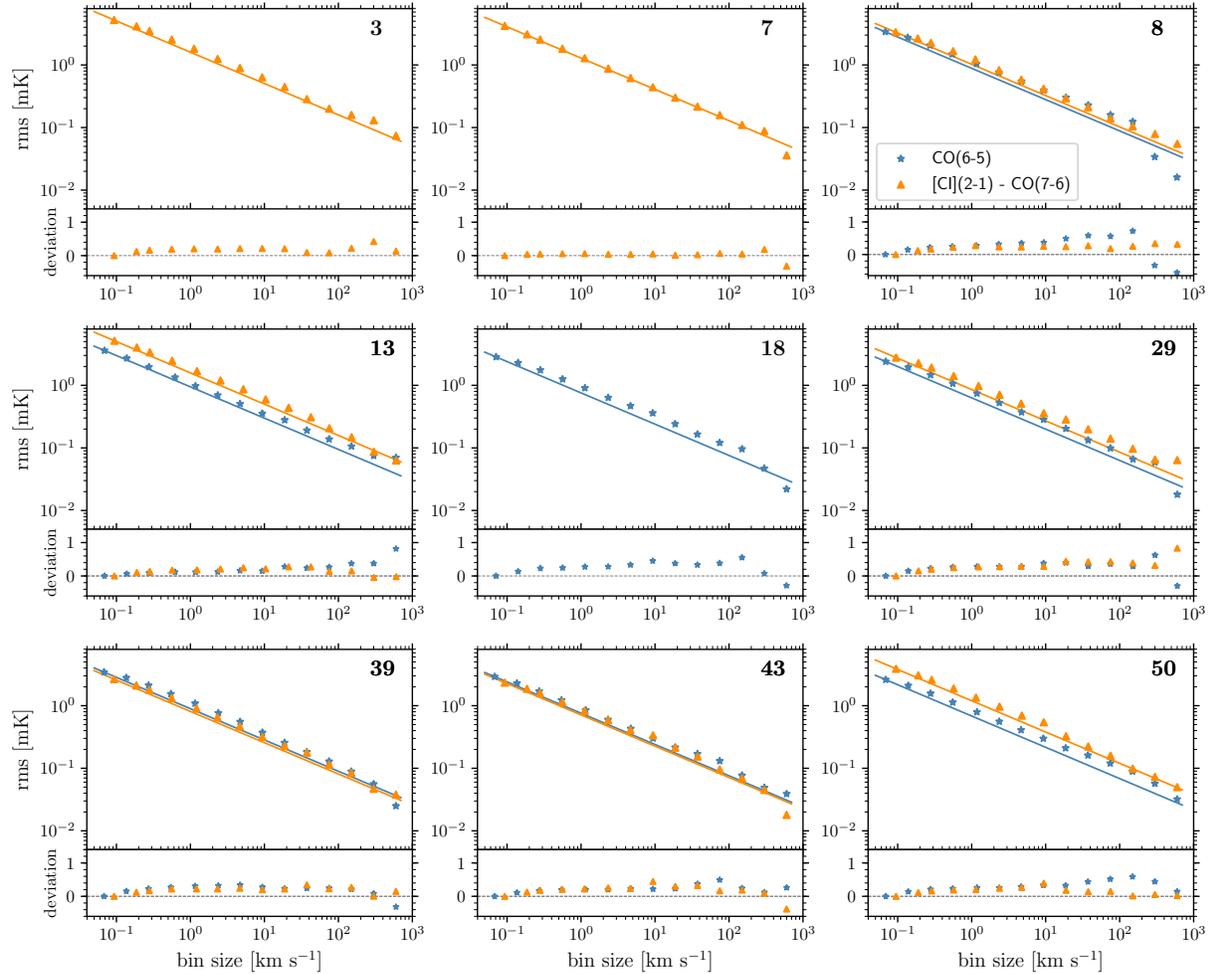
changes of the rms during one observing date (e.g., blue points) indicate strong changes in the airmass of the source. From this figure, we note that data taken under similar PWV and elevation, but in different dates, are characterized by the same rms.

## 2.9 Appendix B: Reliability of line detections

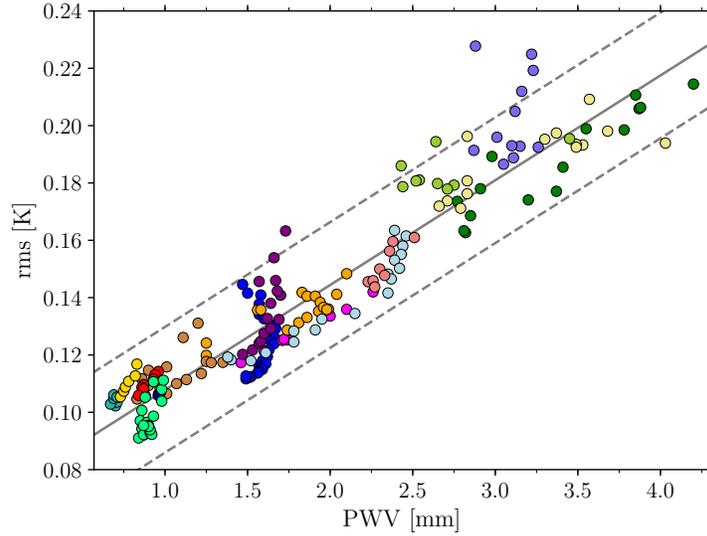
We performed two tests to confirm the reliability of the reported lines. First, we performed a negative test that consists in inverting the final combined spectra (multiplying the fluxes by -1) and applying our detection algorithm. Second, we produced jack-knife spectra by inverting each second scan during data reduction. We then apply the detection algorithm to these spectra which should contain only noise.

We stress that in both tests we applied the same detection algorithm used to identify line detections in our spectra (see Section 2.3). The results of these tests are shown in Table 2.5. For completeness, we list in footnotes the presence of features that fulfill only partially our selection criteria (i.e. with peak  $S/N > 2$ ).

In summary, we did not find any detections in the negative test. While, in the jack-knife test we found only one line detection (integrated  $S/N=3.3$ ) for ID8, [C<sub>I</sub>](2-1)-CO(7-6) observations. We note that the final [C<sub>I</sub>](2-1)-CO(7-6) combined spectrum for this source has a feature at 196.1 GHz (or  $-328 \text{ km s}^{-1}$ , integrated  $S/N = 3.1$ ,  $I_{\text{line}} \sim 3 \text{ Jy km s}^{-1}$ ) that according to our jack-knife test, could be a spurious line.



**Figure 2.16:** RMS of the final combined spectra as a function of the bin size for each of the nine sources, for CO(6-5) (blue stars) and [CII](2-1) - CO(7-6) (orange triangles) observations. The ID number of each quasar is shown in the top right corner of each panel. The solid lines represent the expected scaling of the rms with increasing bin size (see section 2.8). The lower subpanels show the deviation of the observed rms with respect to the predicted value,  $(\text{rms}_{\text{observed}} - \text{rms}_{\text{expected}}) / \text{rms}_{\text{expected}}$ .



**Figure 2.17:** RMS values (for each scan) versus PWV relation for the source SDSS J0100+2105 (ID 39), corresponding to the CO(6-5) and [C I](2-1) - CO(7-6) observations. Each colour indicates a different observation date. The grey solid line represents a linear fit to the points, and the two dashed lines represent the  $\pm 2\sigma$  ( $\pm 0.018$  K) deviation from the fit.

ID <sup>a</sup>	Quasar	Negative test		Jackknife test	
		CO(6-5) tuning	[C I],CO(7-6) tuning	CO(6-5) tuning	[C I],CO(7-6) tuning
3	J 0525-233	-	0	-	0 <sup>e</sup>
7	SDSS J1209+1138	-	0	-	0 <sup>f</sup>
8	UM683	0	0	0 <sup>g</sup>	1 <sup>h</sup>
13	PKS-1017+109	0	0	0	0 <sup>i</sup>
18	SDSS J1557+1540	0 <sup>b</sup>	-	0	-
29	Q-0115-30	0	0	0	0 <sup>j</sup>
39	SDSS J0100+2105	0 <sup>c</sup>	0	0	0 <sup>k</sup>
43	CTSH22.05	0	0	0	0
50	SDSS J0819+0823	0	0 <sup>d</sup>	0	0

**Table 2.5:** Numbers of spurious lines in the negative and jack-knife tests. **Notes:** [a] Identification number taken from the QSO MUSEUM survey (Arrigoni Battaia et al., 2019a), [b] The negative spectrum has one feature at  $+1200$  km s<sup>-1</sup> with peak S/N > 2, but its integrated S/N is only S/N<sub>int</sub> = 2.3, [c] The negative spectrum has one feature at  $+3500$  km s<sup>-1</sup> with peak S/N > 2, but its integrated S/N is only S/N<sub>int</sub> = 1.9, [d] The negative spectrum has one feature at  $+2000$  km s<sup>-1</sup> with peak S/N > 2, but its integrated S/N is only S/N<sub>int</sub> = 2.5, [e] This jackknife spectrum has one feature with peak S/N > 2, but its integrated S/N is only S/N<sub>int</sub> = 2.8, [f] This jackknife spectrum has one feature with peak S/N > 2, but its integrated S/N is only S/N<sub>int</sub> = 1.9, [g] This jackknife spectrum has one feature with peak S/N > 2, but its integrated S/N is only S/N<sub>int</sub> = 2.3, [h] This jackknife spectrum has one feature with peak S/N = 2.8, and its integrated S/N is S/N<sub>int</sub> = 3.3, [i] This jackknife spectrum has one feature with peak S/N > 2, but its integrated S/N is only S/N<sub>int</sub> = 2.1, [j] This jackknife spectrum has two features with peak S/N > 2, but their integrated S/N is only S/N<sub>int</sub> = 2.1 and 2.5, [k] This jackknife spectrum has one feature with peak S/N > 2, but its integrated S/N is only S/N<sub>int</sub> = 1.9.

# Chapter 3

## Modelling [CII] emission in high- $z$ star-forming galaxies

The content of this chapter is based on *Muñoz-Elgueta et al. 2023, in prep.*, which will be submitted soon.

### 3.1 Introduction

The gas bound to a galaxy dark-matter halo, known as the circumgalactic medium (CGM), plays an essential role in understanding galaxy formation and evolution. It is a reservoir of gas that interacts with the galactic environment, regulating key processes such as star formation, accretion, and feedback. Investigating the CGM provides crucial insights into the complex interplay between galaxies and their surrounding gas reservoirs (e.g., [Tumlinson et al., 2017](#)). Studying it is challenging as it is predominantly composed of diffuse and tenuous gas, making direct observations difficult. Recent advancements in observational capabilities (e.g., Atacama Large Millimeter/Submillimeter Array (ALMA), [Wootten & Thompson, 2009](#)), combined with state-of-the-art simulations (e.g., IllustrisTNG, [Nelson et al. 2019](#); Pillepich et al. 2019, EAGLE, [Crain et al. 2015](#); [Schaye et al. 2015](#)), have opened new avenues for exploring the CGM and unveiling its properties. However, the limited resolution of the simulations and the requirement of deep observations are obstacles that remain.

The [CII]  $158\mu\text{m}$  fine-structure line has emerged as a valuable tracer of the CGM. This line is one of the brightest emission lines in the infrared spectra of star-forming galaxies (e.g., [Stacey et al., 1991](#); [Braucher et al., 2008](#)) and is an important tracer of cold gas in both the interstellar medium (ISM) and CGM. It is emitted from the upper fine-structure level  $J=3/2$  to the lower level  $J=1/2$  of the  $\text{C}^+$  ion. The higher energy level exhibits an equivalent temperature of 91 K, while its critical densities are 16,  $2.4 \times 10^3$  and  $4.8 \times 10^3 \text{ cm}^3$  for collisions with electrons, hydrogen atoms and  $\text{H}_2$  molecules, respectively ([Goldsmith et al., 2012](#)). Since the ionization potential of atomic carbon is only 11.3 eV (versus 13.6 eV of neutral hydrogen), [CII] can originate from atomic ISM, diffuse molecular and ionized

gas. At high redshifts, both simulations (e.g., Olsen et al. 2015; Vallini et al. 2015) and observations (e.g., Stacey et al. 2010; Gullberg et al. 2015) have found that the [CII] line primarily emanates from photodissociation regions (PDRs).

Observations of [CII] emission in extragalactic sources notably increased with the advent of the *Herschel Space Observatory* (Pilbratt et al., 2010). Most of the studies have focused on local starburst and luminous infrared galaxies, investigating the properties of the ISM and its link with the [CII] properties (e.g., Sargsyan et al. 2012; Díaz-Santos et al. 2014; Herrera-Camus et al. 2015; Samsonyan 2022; Díaz-Santos et al. 2017). In particular, there has been progress in investigating [CII] as a SFR tracer. While different works found a tight correlation between [CII] luminosity and SFR (e.g., De Looze et al. 2014; Herrera-Camus et al. 2015), other *Herschel* observations showed that  $L_{[\text{CII}]} / L_{\text{IR}}$  (or equivalently,  $L_{[\text{CII}]} / \text{SFR}$ ) ratio decreases continuously with increasing  $L_{\text{IR}}$  (at  $L_{\text{IR}} \gtrsim 10^{11} L_{\odot}$ , e.g., Sargsyan et al. 2012; Díaz-Santos et al. 2017; Contursi et al. 2017). This could be an indication that the  $L_{[\text{CII}]} / \text{SFR}$  ratio is influenced by other galaxy properties, such as metallicity and dust temperature (e.g., Díaz-Santos et al. 2014).

While locally the [CII] line can only be observed from space due to atmospheric limitations, at high- $z$  the frequency of the [CII] line significantly shifts, making it accessible from ground-based observatories like ALMA. In particular, the Earth’s atmosphere allows us for optimal access to the [CII] transition in the frequency range 345.5-211.2 GHz, which corresponds to the redshift range  $4.5 < z < 8.5$ . Therefore, [CII] has become a key target for observational studies of high-redshift galaxies, where it can provide insights into the early stages of galaxy formation and evolution. Extended [CII] emission (known as [CII] halos) has now been regularly observed in several high-redshift objects such as dusty starburst and normal star-forming galaxies (e.g., Fujimoto et al. 2019, 2020; Ginolfi et al. 2020b; Carniani et al. 2020; Rybak et al. 2020; Herrera-Camus et al. 2015, 2021; Meyer et al. 2022; Akins et al. 2022). These halos can extend up to tens of kiloparsecs from the galaxies and their physical properties and origins are still poorly understood. As reference, the typical effective radius of star-forming galaxies at  $4 < z < 6$  is expected to be  $\lesssim 1.5$  kpc (e.g., Shibuya et al. 2015). In the context of star-forming galaxies, Fujimoto et al. (2019) reported the first identification of [CII] halos ( $\sim 10$  kpc scales) around  $z \sim 5-7$  star-forming galaxies using deep ALMA data, through a  $uv$ -visibility plane stacking. Subsequently, with development of the ALPINE survey (Le Fèvre et al., 2020), further observations of [CII] halos around high- $z$  star-forming galaxies have been reported.

Ginolfi et al. (2020b) studied a sample of 50 [CII] emitting galaxies at  $z \sim 4-6$  from ALPINE. By stacking the [CII] spectra and cubes, they find outflows signatures together with extended [CII] emission ( $\sim 15$  kpc) in their galaxy sample with the higher SFR ( $\text{SFR} > 25 M_{\odot} \text{year}^{-1}$ ). In particular, they obtained an average [CII] surface-brighthness radial profile for the high-SFR galaxies, which is characterized by a compact inner component and an extended component (radial scales  $> 10$  kpc). Later, Fujimoto et al. (2020) performed individual measurements for the ALPINE galaxies, finding that a  $\sim 30$  per cent of them present [CII] halos ( $\sim 10$  kpc scales). Recently, an additional ALMA effort, the REBELS survey (Bouwens et al., 2022) has been conducted. This study consists of ALMA [CII] observations of a sample of 40 star-forming galaxies at  $z \sim 6.5-9$ . Based on this survey,

Fudamoto et al. (2022) stacked 28 [CII] emitting galaxies at  $z \sim 7$ . Their results indicate that the average [CII] surface-brightness profile is well fit with a single component (i.e., no extended component), with an effective radius of  $\sim 2$  kpc. Possible reasons of such absence of an extended component could be related to the fact that they stacked moment maps rather than the data cubes, and that they included both high and low-SFR galaxies in the analysis. Despite the aforementioned result, this [CII] emission is spatially more extended compared to the dust continuum and the rest-frame UV emission, in agreement with previous works (e.g., Ginolfi et al. 2020b; Fujimoto et al. 2020).

To fully understand these [CII] halos, it is essential to perform theoretical investigations. Currently, cosmological simulations are not able to reproduce extended [CII] emission at high- $z$  (e.g., Pallottini et al. 2017, 2019). For instance, Fujimoto et al. (2019) compared their observational results with two sets of numerical simulations of massive galaxies ( $M_{\text{halo}} \sim 10^{11}$ - $10^{12} M_{\odot}$ ) at  $z \sim 6$  (Pallottini et al., 2017; Arata et al., 2019). By using mock data from both simulations, they find that the [CII] emission is not as extended as the observational data. Specifically, they obtained surface-brightness profiles which decline rapidly beyond  $r \sim 3$  kpc, which is significantly smaller than the reported [CII] extension of 10 kpc observed by Fujimoto et al. (2019). These profiles are at least one order of magnitude lower than the observed flux at inner-CGM scales. This disagreement suggests that certain crucial physical processes for the production of [CII], such as feedback or metal enrichment, are not well reproduced by the simulations. It is also possible that there are limitations or uncertainties in the modelling of [CII] emissivities. This discrepancy between simulations and observations highlights the need for additional investigation and theoretical refinement.

The origin of this extended [CII] emission remains unclear. Different scenarios have been proposed, including the presence of satellite galaxies, extended PDRs, extended HII regions, inflows of cold streams, outflows produced by stellar feedback or AGN feedback (see Fujimoto et al. 2019 for a detailed description). Currently, the scenario in which star-formation driven outflows are responsible of the extended [CII] emission seems to be the most plausible according to observations (e.g., Ginolfi et al. 2020b; Fujimoto et al. 2020). Indeed the observed [CII] halos are only evident around more star-forming galaxies and the line-emission stacking of these sources seem to highlight the presence of broad wings (Ginolfi et al., 2020b). This scenario has also been supported by semi-analytical models (e.g., Pizzati et al. 2020, 2023). These models primarily focus on simulating the cooling outflows and their resulting [CII] emission, suggesting that the observed [CII] halos could be attributed to gas that is cooling after being heated and expelled in supernova-driven outflows. The effectiveness of these models is largely attributed to their ability to accurately capture the catastrophic cooling<sup>1</sup> process within central kpc scales. Specifically, they find that supernova-driven outflows expand into the CGM at velocities of 200-500  $\text{km s}^{-1}$ . The outflowing gas undergoes rapid cooling ( $\sim$ few hundred K) within the first

---

<sup>1</sup>Catastrophic cooling refers to the rapid cooling of gas under certain conditions, leading to the formation of cold ( $T \sim 10^2$ - $4$ K) winds. These winds propagate outward, characterized by lower velocities and higher densities (Pizzati et al., 2023).

kpc, then it is slowly heated ( $T \sim 10^3$  K) by the cosmic UV background. These conditions favor the formation and survival of  $C^+$  ions.

In this framework, we post-processed a sample of Illustris-TNG50 halos hosting  $z \sim 4$ -6 star-forming galaxies to model their [CII] emission. In particular, we produce mock observations and compare them with the ALPINE observations (Ginolfi et al., 2020b), ultimately discussing the possible origin for the observed extended [CII] emission in the inner CGM of high- $z$  galaxies. This chapter is structured as follows. In section 3.2 we present a summary of the observations of Ginolfi et al. (2020b). In section 3.3 we described the methods used in this work. In section 3.4 we present our main results. In section 3.5 we discuss our main results and explore the different factors involved in the production of [CII] halos. In section 3.6 we present the caveats. Finally, section 3.7 summarizes our findings.

## 3.2 Observations

The ALPINE survey is an ALMA large program that targeted [CII] and FIR-continuum emission of a sample of 118 normal (i.e., main sequence) star-forming galaxies. These galaxies are at  $4.4 < z < 5.8$ , and have  $M_* \sim 10^{8.5} - 10^{11} M_\odot$  and  $SFR \sim 3-650 M_\odot \text{year}^{-1}$ . The average resolution of these observations is 0.9 arcsec, i.e., ranging from  $\sim 5.2$  to 6 kpc for the different redshifts considered. This coarse resolution implies that the galaxies are not resolved in these ALMA observations. All the details regarding the observational setup and data reduction are described in Béthermin et al. (2020). In total, 75 out of 118 galaxies were robustly detected in [CII], with signal-to-noise ratio (S/N)  $> 3.5$ . Ginolfi et al. (2020b) studied these [CII]-emitting galaxies, excluding  $\sim 30\%$  of those objects that present signs of ongoing mergers (major or minor). Thus, their final sample<sup>2</sup> consists of 50 normal star-forming galaxies at  $4.4 < z < 5.8$ , with  $M_* \sim 10^{8.5} - 10^{11} M_\odot$  and  $SFR \sim 5-650 M_\odot \text{year}^{-1}$ .

In order to investigate the efficiency of galactic feedback in these galaxies and the spatial distribution of the [CII] emission, Ginolfi et al. (2020b) investigated the [CII] emission data through a stacking analysis. For the purposes of this paper, in the following paragraphs we summarize only their stacking of the cubes. This procedure consists in a combination of the [CII] cubes of each galaxy  $C_i$  with  $i$  slices (previously aligned), based on a vector variance stacking technique, where the stacked cube  $C_i^{\text{stack}}$  is estimated as:

$$C_i^{\text{stack}} = \frac{\sum_{k=1}^N C_{i,k} \cdot w_{i,k}}{\sum_{k=1}^N w_{i,k}}$$

where  $C_{i,k}$  corresponds to the cube of the  $k$ th galaxy. The weighting factor  $w_{i,k}$  is the inverse variance, equal to  $1/\sigma_{i,k}^2$ , where  $\sigma_{i,k}$  corresponds to the spatial rms associated with each slice of each galaxy. In this way, noisier cubes are down-weighted.

<sup>2</sup>Some of the physical properties of Ginolfi's sample listed here are different than the values reported in Ginolfi et al. (2020b). These discrepancies are due to the fact that Ginolfi et al. (2020b) used the values from a preliminary analysis of the ALPINE sample. Therefore, we use the values reported in the catalogs of the ALPINE survey (Le Fèvre et al., 2020), which was published after Ginolfi et al. (2020b).

This stacking was performed separately for two sub-samples separated by SFR: high-SFR ( $\text{SFR} > 25 M_{\odot} \text{ yr}^{-1}$ ) galaxies and low-SFR ( $\text{SFR} \leq 25 M_{\odot} \text{ yr}^{-1}$ ) galaxies. The SFRs were computed through spectral energy distribution (SED) fitting, assuming a constant star-formation history over 100 Myr (Faisst et al., 2020). In order to obtain velocity-integrated flux maps of the core of the [CII] emission, the spectral slices of the stacked cubes were collapsed in the velocity range of  $[-200: +200] \text{ km s}^{-1}$ . The final stacked synthesized beam (or point spread function, PSF) was estimated by stacking the PSF of each input galaxy cube. Then, the channels of this stacked PSF cube are collapsed over the same velocity range mentioned above. This final PSF has a major axis full-width-half-maximum (FWHM) of 0.98 arcsec, a minor axis FWHM of 0.89 arcsec, and a position angle of  $-30 \text{ deg}$ . From the resulting [CII] flux maps of the subsamples, Ginolfi et al. (2020b) finally computed circularly averaged radial surface-brightness (SB) profiles, considering radial bins of 0.15 arcsec.

For most of our comparative analysis, we present a single observed SB profile, derived by averaging the two SB profiles obtained from Ginolfi et al. (2020b) for the high and low-SFR galaxy subsamples.

## 3.3 Methods

### 3.3.1 IllustrisTNG simulations

In this work, we explore different high- $z$  halos from the IllustrisTNG cosmological magnetohydrodynamical simulations (Nelson et al., 2019; Pillepich et al., 2019). Using the moving-mesh code AREPO (Springel, 2010), these simulations follow the coupled dynamics of the dark-matter and gas. This method is based on a spatial discretization consisting in a Voronoi tessellation of the simulation box. The physical processes of baryonic TNG runs, among which are star-formation, stellar evolution and feedback, supermassive black hole formation and evolution, follow the ‘‘TNG galaxy formation model’’ (methods described in Weinberger et al., 2017; Pillepich et al., 2018). Due to limitations in the resolution, the processes of star formation and pressurization in the multiphase ISM are modeled based on the approach of Springel & Hernquist (2003). In this model, gas with densities above a threshold of  $\simeq 0.1 \text{ cm}^{-3}$  undergoes stochastic star formation according to the Kennicutt-Schmidt relation<sup>3</sup>, assuming a Chabrier initial mass function (Chabrier, 2003). To account for the effects of unresolved supernovae, a two-phase effective equation of state model is employed to incorporate the pressurization of star-forming gas.

This TNG model has been calibrated in order to reproduce several of the observed galaxy properties mainly at  $z=0$ . The cosmology assumed has parameters  $\Omega_{\Lambda,0} = 0.6911$ ,  $\Omega_{m,0} = 0.3089$ ,  $\Omega_{b,0} = 0.0486$ ,  $\sigma_8 = 0.8159$ ,  $n_s = 0.9667$  and  $h = 0.6774$  ((Planck Collaboration et al., 2016)).

---

<sup>3</sup>The Kennicutt-Schmidt law is an empirical relation between the SFR surface density and the gas surface density in galaxies.

Specifically, in this work we used the TNG50-1 box, which presents the smallest volume ( $51.7^3 \text{ cMpc}^3$ ) but the highest resolution ( $m_{\text{baryon}} = 5.6 \times 10^4 M_{\odot}/h$ ) of the TNG runs. The latter is a key factor to perform our [CII] emission predictions, especially considering that high densities are needed for [CII] production (see critical densities in Section 3.1). Compared to previous works, the TNG50 simulations offer key advantages. They strike a balance between computational efficiency and resolution, allowing for a larger sample size and a more comprehensive representation of the high- $z$  halo population. Additionally, these simulations incorporate a self-consistent treatment of various physical processes, which are uniformly implemented across the simulation volume. This ensures a consistent modeling approach for the entire sample of halos.

### 3.3.2 Selection of simulated galaxies

We selected a sample of TNG50-1 star-forming galaxies<sup>4</sup> with similar physical properties to the galaxies of the full ALPINE survey (Le Fèvre et al., 2020). Specifically, they were chosen to match the ranges in stellar mass ( $\log(M_{*}) \sim 8.5 - 11.0 M_{\odot}$ ) and star-formation rate ( $\text{SFR} \sim 3 - 623 M_{\odot}\text{year}^{-1}$ )<sup>5</sup>. In total, we found 357 TNG50 galaxies matching these properties. We then optimized our selection by constraining our distribution of SFR and  $M_{*}$  to conform to those exhibited by the ALPINE [CII] emitting galaxies (Ginolfi et al., 2020b). Throughout this work, we adopt for the simulated galaxies: i) stellar masses within twice the stellar half mass radius, and ii) SFRs averaged across the last 100 Myrs within twice the stellar half mass radius (see Donnari et al., 2019; Pillepich et al., 2019), comparable to the observationally derived SFRs (see Section 3.2). These galaxies were also selected to have similar redshifts of those of the observations ( $z \sim 4.4-5.7$ ). However, it was not possible to precisely match them because snapshots with intermediate redshifts (the so-called mini snapshots) do not contain information on the metal abundances which are needed to estimate the [CII] emission. Thus, the final selected halos were extracted from the three full snapshots 13, 17 and 21, corresponding to the redshifts 6, 5 and 4, respectively.

The individual halos were cut from the full snapshot (at  $\sim 3 R_{200}$ ) in order to facilitate their analysis. In total, we selected 75 halos. Analogous to the observations, we excluded from our analysis those halos that would have resulted in observable merging systems. Specifically, following the observational limitations, we did not consider systems with companion galaxies with  $\text{SFR} > 1.5 M_{\odot}\text{year}^{-1}$ , within a 3D radial distance of  $1.1R_{200}$ . We also note that our selected galaxies contain AGNs, while the observed sample excludes type I AGNs. However, these AGNs of our simulated sample have very low bolometric luminosities<sup>6</sup> ( $L_{\text{BOL}} \lesssim 10^{45.5} \text{ erg s}^{-1}$ ), which would not be detected in observations. Figure 3.1 shows the distribution of the selected simulated galaxies (red) in the SFR and  $M_{*}$

<sup>4</sup>In this work, the selected TNG50-1 galaxies consist in the central subhalos of their corresponding FoF halos.

<sup>5</sup>A  $\sim 37$  per cent of the TNG50-1 galaxies with  $M_{*}$  in the range  $\sim 8.5 - 11.0 M_{\odot}$  have SFRs in the range  $\sim 3 - 623 M_{\odot}\text{year}^{-1}$ .

<sup>6</sup>We estimated the bolometric luminosity of the central AGNs following Nelson et al. (2019).

planes, in comparison to the sample of [Ginolfi et al. \(2020b\)](#) (blue), using density contours. Importantly, this selection represents one possible realization of Ginolfi’s sample extracted from the TNG50 galaxies in the mass and SFR ranges of the ALPINE survey. In this figure, the spacing between the blue contours (observations) appears larger than the spacing between the red contours (simulations). This suggests that the observed sample has a more spread-out distribution of galaxies in the joint space of  $M_*$  and SFR compared to the simulated sample. Areas where the blue and red contours overlap suggest regions of similar  $M_*$  and SFR values between both samples.

To gain insight on the physical properties of the selected galaxies, Fig. 3.2 presents a temperature-density ( $T - n_{\text{H}}$ ) phase diagram within  $R_{200}$ , of one  $z=5$  system. The color bar indicates the gas mass distribution, where the yellow (blue) regions correspond to areas of high (low) gas mass concentration within the halo. The yellowest region ( $\sim T < 10^{4.5}$  K,  $10^{-2.5} < n_{\text{H}} < 10^{-1} \text{cm}^{-3}$ ) is dominated by ISM gas. Note that all dense gas cells with  $n_{\text{H}} > 0.1 \text{cm}^{-3}$  are by construction star-forming in these simulations. The temperatures of this star-forming gas are determined by the effective equation of state ([Springel & Hernquist, 2003](#)) and they do not represent the physical temperatures of the gas ([Pillepich et al., 2019](#)).

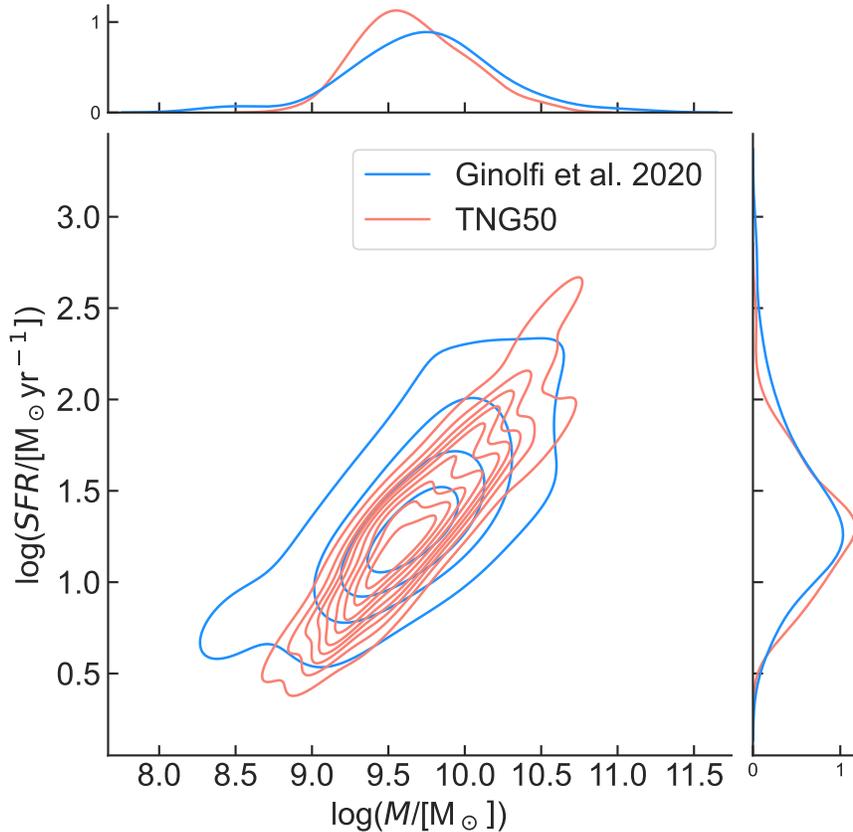
### 3.3.3 [CII] cooling rate and flux estimation

In order to obtain an estimate of the [CII] emission for each selected halo, we used the analytical [CII] cooling rate ( $\Lambda_{\text{CII}}$ ) from [Bisbas et al. \(2022\)](#). In their study, [Bisbas et al. \(2022\)](#) performed [CII] synthetic observations using the GRIFFIN smooth particle hydrodynamics simulations ([Lahén et al., 2020](#)). These simulations have a resolution of  $4M_{\odot}$  per gas particle and use a modified version of GADGET-3 ([Springel, 2005](#)) described in [Hu et al. \(2014\)](#). They assume a non-equilibrium model of cooling and chemistry that traces  $\text{H}_2$ ,  $\text{H}^+$ ,  $\text{CO}$ ,  $\text{H}$ ,  $\text{C}^+$ ,  $\text{O}$  and free electrons ([Nelson & Langer 1997](#); [Glover & Mac Low 2007](#)), only when the gas temperature is  $< 3 \times 10^4$  K. For temperatures above this threshold, equilibrium cooling tables based on 12 different metal species ( $\text{H}$ ,  $\text{He}$ ,  $\text{N}$ ,  $\text{C}$ ,  $\text{O}$ ,  $\text{Si}$ ,  $\text{Mg}$ ,  $\text{Fe}$ ,  $\text{S}$ ,  $\text{Ca}$ ,  $\text{Ne}$ , and  $\text{Zn}$ ) are used (from [Wiersma et al. 2009](#)). Furthermore, the simulations also include routines to account for the interstellar UV radiation field and mechanisms for stellar feedback, such as photoionization, photo-electric heating, and supernovae.

Based on the direct  $\Lambda_{\text{CII}}$  estimations from the simulations, [Bisbas et al. \(2022\)](#) proposed a theoretical approach to get  $\Lambda_{\text{CII}}$ , based on analytical expectations. They demonstrated that the  $\Lambda_{\text{CII}}$  values obtained from the analytical expressions are in agreement with the majority of the  $\Lambda_{\text{CII}}$  values obtained directly from their simulations. This approach assumes that the gas is optically thin and is applicable for  $T > 150$  K. The analytical cooling rate can be expressed as:

$$\Lambda_{\text{CII}} = \frac{R_{\text{c,dex}} \times 2e^{-91.25/T}}{R_{\text{s}} + R_{\text{c,dex}}(1 + 2e^{-91.25/T})} E_{\text{CII}} n_{\text{C}^+} (\text{erg s}^{-1} \text{cm}^{-3}) \quad (3.1)$$

where  $T$  is the temperature of each gas cell and the excitation temperature of CII transition

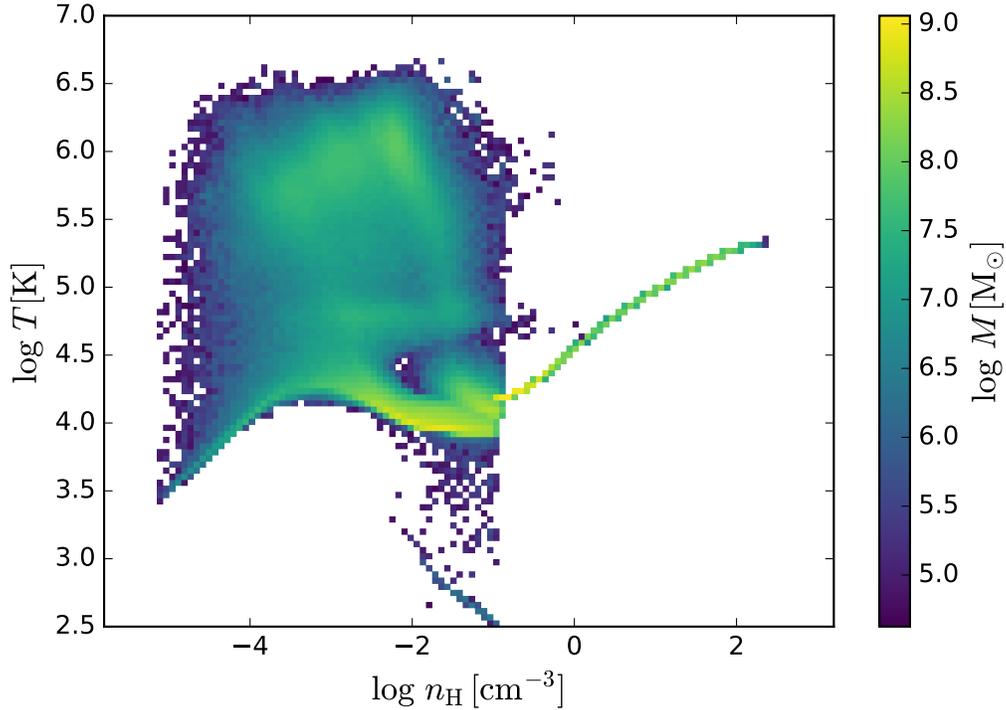


**Figure 3.1:** Comparison of the distribution of the SFRs and  $M_*$  values between the simulated (red) and observed (blue) samples. The joint distribution is visualized using density contours. The contour spacing represents the density of data points, with closer contours indicating regions of higher data density. The histograms along the x and y axes provide marginal distributions of the  $M_*$  and SFR, respectively, for both observed and simulated samples.

is 91.25 K,  $E_{\text{CII}}$  is the energy  $h\nu_{\text{CII}} = 1.25988 \times 10^{-14}$  erg,  $n_{\text{C}^+}$  is the number density of  $\text{C}^+$  particles,  $R_s$  is the spontaneous emission equal to the corresponding Einstein  $A$ -coefficient  $R_s = 2.291 \times 10^{-6} \text{s}^{-1}$ , and  $R_{\text{c,dex}}$  is the total de-excitation rate given by:

$$R_{\text{c,dex}} = R_{\text{c},10}(\text{H})n_{\text{H}} + R_{\text{c},10}(\text{H}_2)n_{\text{H}_2} + R_{\text{c},10}(\text{e}^-)n_{\text{e}} \quad (3.2)$$

where  $R_{\text{c},10}(\text{e}^-)$ ,  $R_{\text{c},10}(\text{H}_2)$  and  $R_{10}(\text{H})$  are collisional de-excitation rates with  $\text{e}^-$ ,  $\text{H}$ , and  $\text{H}_2$  as colliding partners, respectively. We consider the contribution of collisions with  $\text{H}_2$  negligible. This assumption is based on the fact that TNG50 does not follow the formation and evolution of molecular gas, consistent with the simulation setup. Additionally, previous works argue that the contribution of molecular gas to the total [CII] luminosity is minor (e.g., Franeck et al. 2018; Tarantino et al. 2021; Bisbas et al. 2022). Therefore, we focus on [CII] emission originating from PDRs where sufficient amounts of  $\text{HI}$  and  $\text{e}^-$  are present as collision partners (e.g., Olsen et al. 2015).



**Figure 3.2:** Phase-diagram for one of the selected TNG50 galaxies at  $z=5$ , weighted and colour coded by gas mass. The diagram encompasses all gas cells within  $1R_{200}$ . For densities above the star-formation threshold  $n_{\text{H}}=0.1$ , the temperatures are determined by the effective equation of state (Springel & Hernquist, 2003).

Then,  $R_{\text{c},10}(\text{e}^-)$  and  $R_{\text{c},10}(\text{H})$  are:

$$R_{\text{c},10}(\text{e}^-) = 2.426206 \times 10^{-7} \left( \frac{T}{100} \right)^{-0.345} (\text{cm}^{-3} \text{ s}^{-1}) \quad (3.3)$$

$$R_{\text{c},10}(\text{H}) = 3.113619 \times 10^{-10} \left( \frac{T}{100} \right)^{0.385} (\text{cm}^{-3} \text{ s}^{-1}) \quad (3.4)$$

To estimate  $\Lambda_{\text{CII}}$  (equation 3.1), all the needed parameters can be directly estimated from the outputs of TNG50-1, except from  $n_{\text{C}^+}$ . The star-forming gas cells ( $n_{\text{H}} > 0.13 \text{ cm}^{-3}$ ) are assumed to be at their cold-phase temperature of 1000 K (Springel & Hernquist, 2003), as this cold-phase is dominant in mass with respect to the warmer component in these cells. This approach is common in modelling emission lines arising from the cold-phase in TNG simulations (e.g., Nelson et al. 2021).

In order to estimate the values of  $n_{\text{C}^+}$  we need the  $\text{C}^+$  fractions for each gas cell. As an initial simplified approach, we assume a scenario where all neutral carbon is converted into  $\text{C}^+$ . Subsequently, we explore other cases in which the  $\text{C}^+$  fractions are obtained using the spectral synthesis code CLOUDY version 17.03 (Ferland et al., 2017). Our CLOUDY calculations were run in single-zone mode and iterate until line and continuum optical

depths become stable. Importantly, our models tested two different ionizing conditions: i) ultraviolet background (UVB) only, and ii) UVB together with the local ionization field due to young stars (ages  $<10$  Myr). For the first case, we use the UVB model of [Khaire & Srianand \(2019\)](#) at the redshift of each simulated galaxy. This UVB spectrum includes updated inputs based on observations, such as type-1 quasar emissivities and distribution of HI in the intergalactic medium. For the second case, we use the star-forming regions spectra model of [Cerviño et al. \(2002\)](#). As the young stars are being formed embedded in molecular clouds, we assume that only a 5 per cent of the stellar radiation can escape the birth cloud (i.e., escape fraction  $f_{\text{esc}} = 0.05$ ). We assume that all the young stars present the same SED. This allows us to compute the normalization for the radiation field ( $\phi$ ) at the position of a gas cell using an optically thin approximation. We sum the contributions from all the young stellar particles in the simulation, taking into account the distance between each source and the gas cells ([Obreja et al., 2019](#)). This normalization process helps us quantify the radiation received by the gas cells from the surrounding star-forming regions. Figure 3.3 shows examples of the input spectra used in our CLOUDY calculations.

Once the  $\text{C}^+$  fractions are computed, all the quantities in equation 3.1 are available. We can then obtain the [CII] luminosities per cell and estimate their respective flux densities following [Solomon et al. \(1992\)](#):

$$I_{\text{line}} = \frac{L_{\text{line}}}{1.04 \times 10^{-3} \nu_{\text{obs}} D_{\text{L}}^2} (L_{\odot}) (\text{Jy km s}^{-1}) \quad (3.5)$$

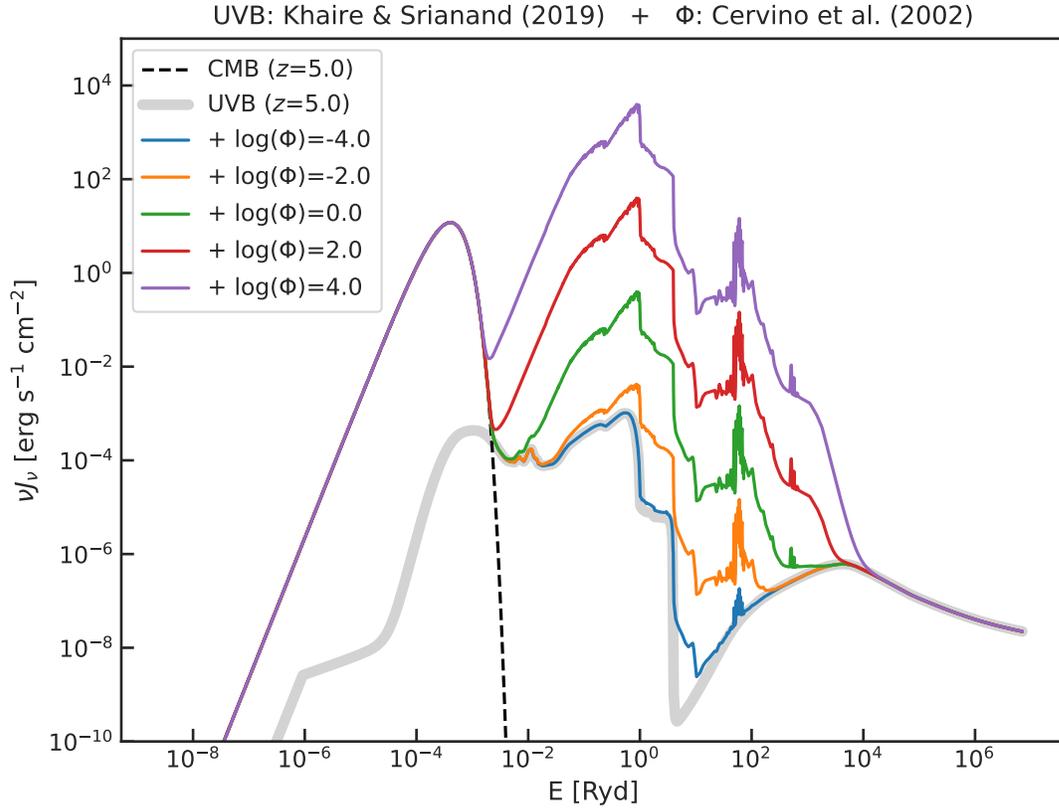
where  $L_{\text{[CII]}}$  is the [CII] line luminosity in  $L_{\odot}$  units,  $\nu_{\text{obs}}$  is the observed frequency in units of GHz, and  $D_{\text{L}}$  corresponds to the luminosity distance in Mpc.

### 3.3.4 [CII] surface brightness maps and profiles

After getting  $L_{\text{[CII]}}$  for each gas cell, we constructed [CII] surface brightness maps. For this, we obtained 2D projections by integrating the emission within a slice encompassing a range of  $[-R_{200}, +R_{200}]$  (or  $\sim[-40, +40]$  kpc at the median  $z=5$ ) from the galaxy's center<sup>7</sup>. We note that the velocity range used in the observed maps ( $[-200, +200]$  km s<sup>-1</sup>) corresponds to a larger range of  $\sim[-350, +350]$  physical kpc, assuming that these velocities trace the Hubble flow at  $z=5$ . However, it is reasonable to consider a smaller velocity range in the simulations compared to the observations. This is because structures at larger distances from the galaxy, beyond the chosen velocity range, are not expected to contribute to the [CII] signal in a monotonically decreasing manner as a function of radius, as seen in the observations.

In order to make these 2D maps comparable with the observations, we applied a 2D Gaussian smoothing with a FWHM equal to the stacked synthesized beam obtained in the ALMA observations (0.9 arcsec or  $\sim 6$  kpc at  $z=5$ , see Section 3.2). Figure 3.4 shows an example of [CII] 2D surface brightness maps for one of the selected galaxies at  $z=5$ , at full

<sup>7</sup>The center of the galaxies considered here is the position of the particle with the minimum gravitational potential energy.

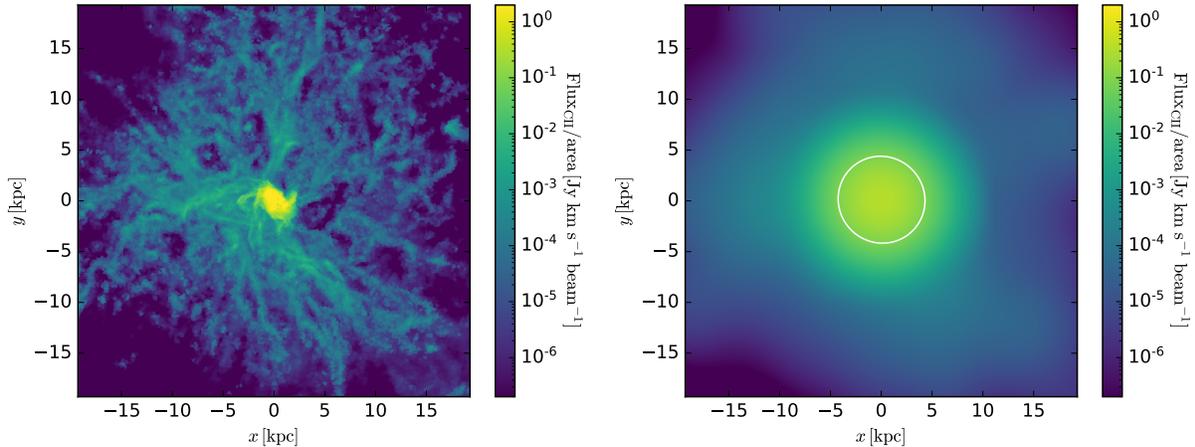


**Figure 3.3:** Example of different input spectra used in the CLOUDY calculations. The solid grey line represents the UVB SED of Khaire & Srianand (2019), and the dashed black line the CMB, both at  $z=5$ . The coloured lines represent the contribution to the SED due to young stars (Cerviño et al., 2002) at different normalizations, plus the contribution of the UVB. The units of the radiation field  $\phi$  are  $M_{\odot} \text{ year}^{-1} \text{ kpc}^{-2}$ .

resolution (left panel) and after applied the smoothing (right panel). We have not added noise to our modeled smoothed SB maps because our primary focus is to assess whether we can achieve similar flux levels as observed. Future simulations of ALMA observations with SIMALMA<sup>8</sup> are needed to incorporate noise levels and achieve a more realistic representation.

For each galaxy, we then extracted circularly averaged radial profiles from their 2D smoothed SB maps. The profiles were obtained by considering radial bins of 0.15 arcsec, as done in the observations. Subsequently, total averaged profiles were obtained by stacking the individual profiles. To accomplish this, we utilized a total of 20 orientations per galaxy and performed 1000 iterations of the experiment, randomly selecting one orientation per

<sup>8</sup>SIMALMA is a task from CASA used to simulate observations with ALMA and the Atacama Compact Array (ACA). This task takes into account instrumental and atmospheric effects that can affect the observed data, including the antenna layout, atmospheric conditions, and noise sources.



**Figure 3.4:** [CII] emission surface-brightness map for one of the selected TNG50 galaxies at  $z=5$ . Left: the map is presented at the full-resolution of the simulation. Right: the map has been convolved with the ALMA beam of the [Ginolfi et al. \(2020b\)](#) observations. The white contour represents the average  $2\sigma$  level of the observations ( $0.08 \text{ Jy km s}^{-1} \text{ beam}^{-1}$ ).

galaxy and calculating the final average based on these realizations.

## 3.4 Results

In this section, we present the results obtained from our post-processing model, which generates [CII] emission in high-redshift galaxies using the TNG50 simulations. By comparing our model predictions with observational data from the ALPINE survey, we assess the validity and performance of our approach, and try to provide some intuition on the aforementioned [CII] emission scenarios.

### 3.4.1 Relation between [CII] luminosity and star formation rate

Several studies have found that the [CII] luminosity is tightly correlated with the SFR, with a nearly linear relation (e.g. [De Looze et al., 2014](#); [Olsen et al., 2017](#); [Lagache et al., 2018](#); [Schaerer et al., 2020](#)). However, at high- $z$  the scatter around the local relation tends to increase. This fact has been primarily attributed to variations in galaxy ISM physical properties, such as gas density and metallicity (e.g., [Lagache et al. 2018](#)). In this context, we explore the [CII]-SFR relation obtained from our post-processing model, and compare the obtained results with the observational data from the ALPINE survey. Therefore, the subsequent analysis checks the consistency between our model predictions and the observations.

The results are shown in Fig. 3.5, where both the simulated (left panel) and observational (right panel) data are color-coded by stellar mass. For the simulated galaxies, the x-axis shows the SFR averaged over the last 100 Myrs, and the y-axis the [CII] luminosity

obtained by assuming UVB and young stars as photoionization fields, and by integrating inside the area within the  $2\sigma$  level of the observations (stacked flux map,  $r \sim 1.3$  arcsec). Comparing the two panels, it is evident that our results are in agreement with the observations, when considering the simulated galaxies that can be detected by the observations ( $L_{[\text{CII}]} > 10^{7.8} L_{\odot}$ ). In addition, for both sets of data we can appreciate a trend in the stellar masses along the relation. However, when comparing the distributions of normalized distances<sup>9</sup> between  $M_*$ ,  $L_{[\text{CII}]}$  and SFR through a Kolmogorov-Smirnov<sup>10</sup> test, it is found that the two samples are not drawn from the same distribution (p-value = 0.002). A possible explanation for this disparity is the significant scatter in  $M_*$  and SFR values in the observed data compared to TNG50. Our results are also roughly in agreement with the expected relationship for local starburst galaxies (blue line, De Looze et al. 2014) and for high- $z$  galaxies (red dashed line, Lagache et al. 2018). We note that three of our simulated galaxies (datapoints inside black box) lie below the expected relationships, exhibiting luminosities below the observational limit. Since these galaxies represent non-detections in the Ginolfi’s observations, we exclude them from any future analysis.

Overall, the results of this analysis provide validation for the predictions made by our theoretical model on galactic scales. In the next sections we will explore what these models predict for the [CII] emission in the CGM.

### 3.4.2 [CII] Surface brightness profiles

#### Averaged profiles of all gas

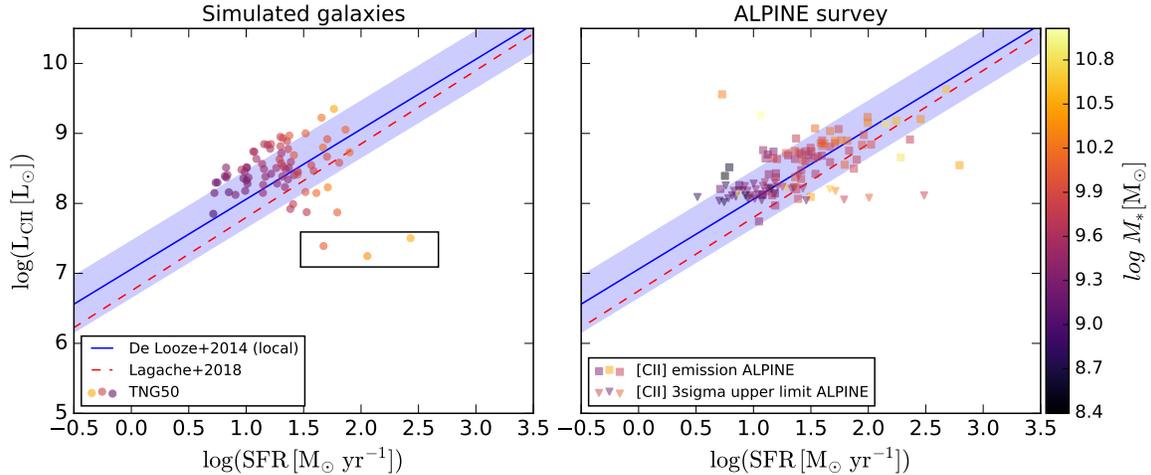
The study of radial [CII] surface brightness profiles in high-redshift galaxies enables a deeper understanding of the spatial distribution of [CII] emission, the properties of the cold ISM, the interplay between star formation and feedback, and the properties of the CGM (e.g., Fujimoto et al., 2019, 2020; Ginolfi et al., 2020b). In figure 3.6 we present the resulting [CII] surface brightness profiles of our simulated galaxies, for the three different cases of  $\text{C}^+$  fractions assumed (represented by different colors, see legend), as explained in Section 3.3.3. The black squares show the averaged profile from Ginolfi et al. (2020b), the solid lines show the average (including  $z=4, 5$  and  $6$ ) profiles from TNG50, and the shaded areas encompass 95 per cent of the distribution of averaged profiles through multiple realizations of the experiment, accounting for different orientations of the galaxies (see section 3.3.4).

<sup>9</sup>The normalized Euclidean distances for the ALPINE and TNG50 datasets are estimated as:

$$r_{\text{data}} = \left( \frac{(M_{*\text{data}} - M_{*\text{min}})^2}{(M_{*\text{max}} - M_{*\text{min}})^2} \right) + \left( \frac{(\text{SFR}_{\text{data}} - \text{SFR}_{\text{min}})^2}{(\text{SFR}_{\text{max}} - \text{SFR}_{\text{min}})^2} \right) + \left( \frac{(L_{[\text{CII}]\text{data}} - L_{[\text{CII}]\text{min}})^2}{(L_{[\text{CII}]\text{max}} - L_{[\text{CII}]\text{min}})^2} \right)$$

where the subscript *data* refers to the specific data set being considered (ALPINE or TNG50), *min* and *max* represent the minimum and maximum values obtained for the variables  $M_*$ , SFR and  $L_{[\text{CII}]}$  between the two samples, respectively.

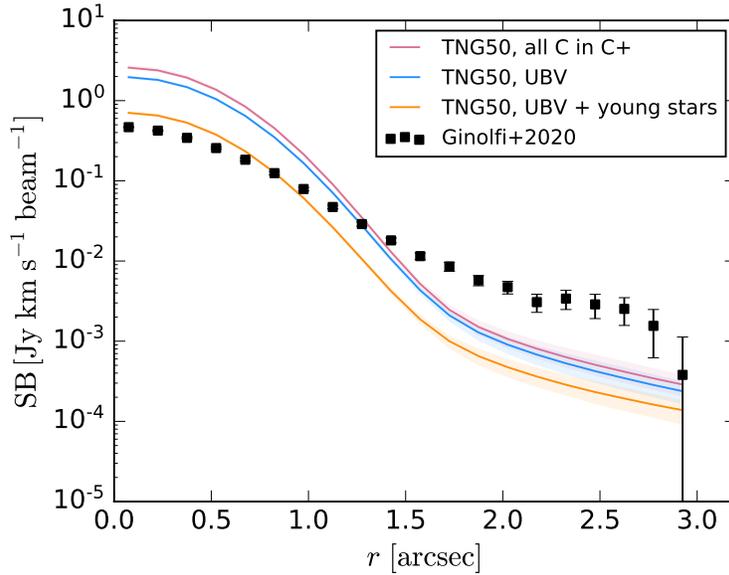
<sup>10</sup>The Kolmogorov-Smirnov test (Chakravarti et al., 1967) is a non-parametric test used to evaluate whether two samples or datasets are drawn from the same underlying distribution, or if one sample conforms to a particular theoretical distribution.



**Figure 3.5:** [CII] luminosity versus SFR for the TNG50 galaxies (left panel) and the ALPINE survey (right panel). The shaded blue area represents the  $1\sigma$  relationship of De Looze et al. (2014) for starburst local galaxies. The dashed red line shows the relationship of Lagache et al. (2018) for  $z=5$  star-forming galaxies. The black box in the left panel contains three simulated galaxies with  $L_{[CII]}$  below the observational limit.

The radial distances are shown up to  $\sim 3$  arcsec (or  $\sim 19$  kpc at  $z=5$ ), which corresponds to the maximum radius analyzed in the observational profiles.

For all cases, the most central part of the profile is dominated by the PSF of the observations (up to a radius of  $\sim 1.5$  arcsec), while the flatter part at larger scales correspond to the inner CGM of these galaxies. In the scenario where all carbon (C) is assumed to be converted into [CII] (dark pink line), comparing with the observations we get an overestimation of the emission at central scales ( $\sim 8$  times higher) and an underestimation at scales larger than  $\sim 1.5$  arcsec ( $\sim 6$  times lower). When considering UVB as ionizing source (light blue line), we can see that our predicted [CII] flux is also higher ( $\sim 7$  times) than the observations at central scales, and lower ( $\sim 7$  times) at large scales. In the last case which assumes UVB and young stars as ionizing sources (orange line), the simulated flux is comparable to the observed one at central scales (in agreement with our analysis in Section 3.4.1), but it is even lower at larger scales ( $\sim 10$  times lower). Hereafter we focus on this last case which provides more realistic predictions for each galaxy emission, and discuss what could cause the discrepancy on CGM scales. Indeed, the results of this model suggest that there is extended [CII] emission within the inner CGM. However, the flux normalization of these extended [CII] halos compared to the observations is not matched. In what follows we will analyze in detail the predicted SB profiles and explore the contributions from different gas phases.

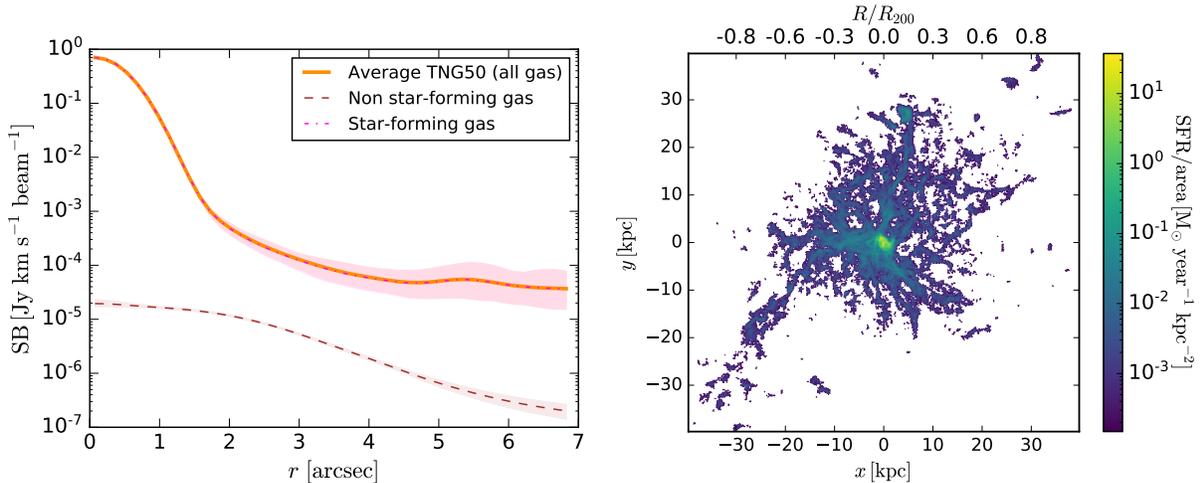


**Figure 3.6:** Averaged [CII] SB profiles obtained for our sample of 72 TNG50 halos at  $z=4, 5$  and  $6$  (solid lines), with shaded areas enclosing a 95 percent of different results obtained by different galaxy orientations. The different line colors represent the TNG50 profiles obtained assuming  $C^+$  fractions equal to the total C fraction (dark pink),  $C^+$  fraction obtained considering the UVB as incident field (light blue) and  $C^+$  fraction considering UVB and local young stars as ionization sources (orange). As a comparison, the black squares represent the results from the observations.

### Contribution of star-forming gas

By examining the contribution of star-forming gas to the total averaged [CII] surface brightness profile, we can refine our understanding of the underlying physical mechanisms responsible for the observed emission. Figure 3.7 shows the total contribution of the dense star-forming gas (magenta dot-dashed line) to the total [CII] SB profile (orange line, including UVB and local young stars). The non star-forming gas contribution is shown by the brown dashed line. We can clearly see that the total [CII] emission is almost completely dominated by the star-forming gas. Since the [CII] emission in the ISM is primarily emitted from PDRs regions where ionized carbon atoms are heated by UV radiation from nearby young stars (e.g., [Stacey et al. 2010](#)), this result is expected at central galactic scales (up to  $\sim 1.5$  arcsec), where most of the star-formation processes take place. On the other hand, understanding the origin of star-forming gas dominating in the inner CGM ( $0.1 \leq R/R_{200} < 0.6$ , or  $\sim 1.5 \leq r < 4$  arcsec at  $z=5$ ) is more complex.

Indeed, our selected TNG50 galaxies present star-forming gas extended up to  $R_{200}$ . This can be appreciated in Fig. 3.7 (right panel), that shows a projection of star-forming gas for one galaxy at  $z = 5$  (same galaxy shown in Fig. 3.4). In section 3.5.1 we discuss the possible origin of this large-scale star-forming gas.

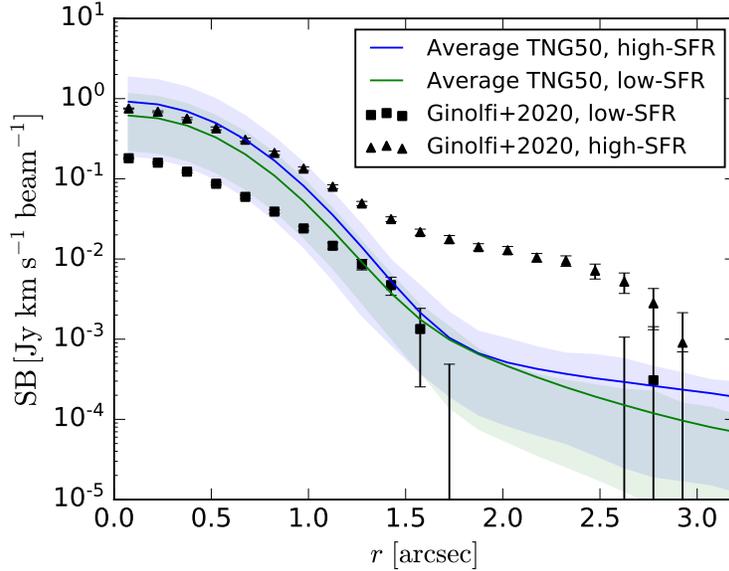


**Figure 3.7:** Left: Averaged contribution of the star-forming gas (magenta dot-dashed line) and non star-forming gas (brown dashed line) to the total [CII] emitting gas (orange line). The shaded areas enclose a 95 percent of different results obtained by different galaxy orientations. Right: SFR density map for a TNG50 galaxy at  $z=5$ . The box size radius is  $\sim 40$  kpc or  $\sim 1R_{200}$ .

### High and low-SFR subsamples

In order to understand the relationship between the [CII] emission and SFR, [Ginolfi et al. \(2020b\)](#) divided their sample in two subsamples: high and low-SFR galaxies, using as threshold  $25 M_{\odot} \text{ year}^{-1}$ . Then, to determine the typical size of the stacked [CII] line core, they computed SB radial profiles for both subsamples (see Fig. 3.8). They find that the subsample of high-SFR galaxies show a more extended profile (radial scales  $>10$  kpc at  $z=5$ ), while the profile of low-SFR galaxies drops around  $r \sim 1.5$  arcsec (or  $r \sim 10$  kpc). They claim that this result supports the scenario where star-formation feedback is responsible for the presence of [CII] halos.

In order to further compare with the observations, we did the same analysis dividing our sample in high and low-SFR using the same threshold of  $25 M_{\odot} \text{ year}^{-1}$ . This comparison is shown in Fig. 3.8, where the triangles (squares) represent the high (low) SFR subsample of [Ginolfi et al. \(2020b\)](#), and the blue (green) solid curve represents the high (low) SFR subsample of the TNG50-1 galaxies. The shaded areas encompass the percentiles 2 and 98 of the whole distribution of averaged profiles per galaxy per rotation. The black symbols are the averaged profiles obtained by the observations. At central galactic scales ( $r < 1.5$  arcsec), the profile of the high-SFR galaxies is higher ( $\sim 2$  times) than the one of low-SFR galaxies, which is expected. Between  $1.5 < r < 2$  arcsec both profiles are in agreement. Then, at larger scales (i.e., CGM scales), the profile of the high-SFR subsample overcomes the one of low-SFR galaxies. However this is a slight difference compared to what is found in the observations. Hence, according to our model, [CII] halos are expected with similar intensities in both high and low-SFR galaxies. A possible reason for this discrepancy is that the simulations may not accurately capture the complex physical processes responsible



**Figure 3.8:** Averaged [CII] SB profiles (solid lines) assuming  $C^+$  fractions from the UVB and local young stars case. The blue line is the averaged profile for the high-SFR subsample, and the green one for the low-SFR subsample. The black triangles and squares represent the results for high and low-SFR galaxies of the observations, respectively. The shaded areas encompass the 2.5 and 97.5 percentiles of the distribution of all profiles (for each galaxy and each rotation).

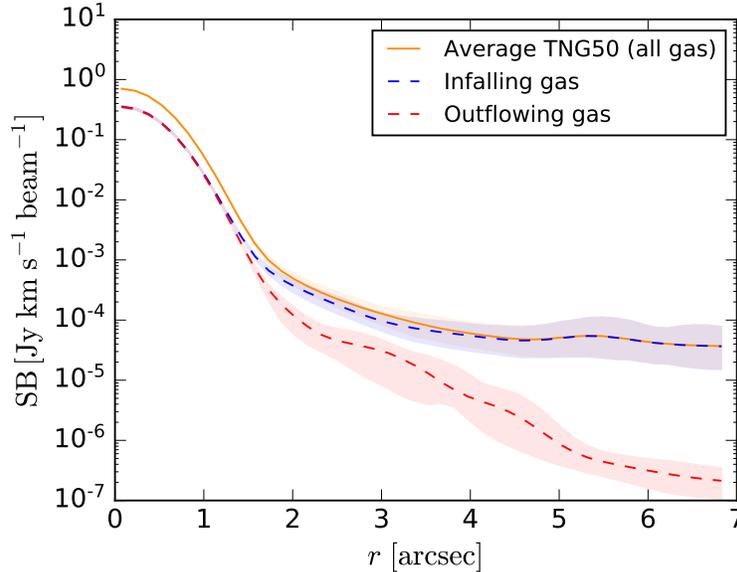
for producing [CII] halos in galaxies, such as the effects of feedback from star formation or active galactic nuclei. A detailed discussion is presented in section 3.5.1.

### Contribution of inflows and outflows

Understanding the role of inflows and outflows in galaxy formation and evolution is essential to fully comprehend the physical processes that govern the growth and properties of galaxies. In this context, the identification of the [CII] contribution from each of these components allow us to disentangle the impact of inflows and outflows on the overall [CII] emission properties of the studied galaxies. In particular, cold ( $T \sim 10^4 K$ ) inflows and outflows have been proposed as a possible origin of [CII] halos (Fujimoto et al., 2019). In the case of cold gas inflows, the infalling streams can produce shock heating, resulting in the production of [CII]. Conversely, extended [CII] emission may be generated by ionized carbon, powered by outflows originating from AGN or star formation processes. In this context, we estimated the gas radial velocities  $v_{\text{rad}}$  with respect to the center of the halos, specifically the position at the minimum gravitational potential. We denote as outflow the gas with  $v_{\text{rad}} > 0$ , and inflow as  $v_{\text{rad}} < 0 \text{ km s}^{-1}$ .

The results of this analysis are shown in Fig. 3.9, where we separate the contribution to the [CII] radial profile from inflowing and outflowing gas. The blue dashed line indicates the contribution of inflowing gas, while the red dashed line indicates outflows. We can appreciate that on galactic scales (up to  $\sim 1.5 \text{ arcsec}$ ) the contribution of inflows and

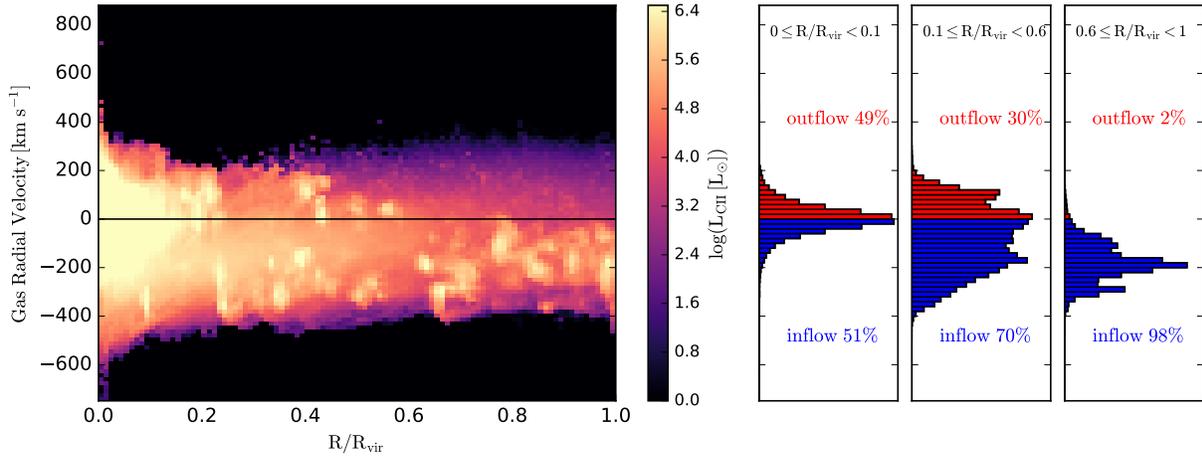
outflows is balanced. On larger scales the contribution of the infalling gas dominates, increasing with radial distance.



**Figure 3.9:** Averaged contribution of the infalling gas (blue dashed line) and outflowing gas (red dashed line) to the total [CII] emitting gas (orange line). The shaded areas represent the  $2\sigma$  uncertainty obtained using 1000 samples of galaxies with randomly picked orientations, as explained in Section 3.3.4

In order to further investigate [CII] as a kinematical tracer, the left panel of figure 3.10 shows the gas radial velocity versus radial distance normalized by  $R_{200}$ , weighted by [CII] luminosity, for a stacking of all the selected halos at redshift 4, 5, 6. The histograms at the right panels show the total distribution of inflows and outflows at different radial ranges:  $0 \leq R/R_{200} < 0.1$ ,  $0.1 \leq R/R_{200} < 0.6$ , and  $0.6 \leq R/R_{200} < 1$ . As a reference, the average  $R_{200}$  for the TNG50 selected galaxies at  $z=5$  is  $\sim 40$  kpc (or  $\sim 6$  arcsec). For the first radial bin (inner galactic scales), our predictions indicate a balance between inflows and outflows. For the second radial bin (inner CGM), we note that the contribution of inflows exceeds the contribution of outflows by a 40 per cent. At larger scales (third radial bin), the contribution of inflows (98 %) dominates the [CII] emission.

These results suggest that infalling gas may play a more significant role in the [CII] emission at larger scales, potentially indicating the presence of gas accretion processes that contribute to the buildup of gas reservoirs in high-redshift galaxies. These results are intriguing due to the expectation that inflowing gas from the intergalactic medium typically possesses low metallicity (e.g., Pallottini et al. 2014). Consequently, it is anticipated that this gas would have a minimal impact on [CII] halos. Therefore, these results raise interesting implications for our understanding of metal-enriched gas inflows. We further explore these results in section 3.5.



**Figure 3.10:** Left: Gas radial velocity versus radial distance (normalized to the virial radius), weighted by [CII] luminosity, for a stacking of all the selected TNG50 galaxies (i.e. at redshifts 4, 5 and 6). Velocities greater than 0 denote outflows, and less than 0 inflows. Right: the histograms shows the total distribution of inflowing and outflowing gas, at different radial bins (see legend).

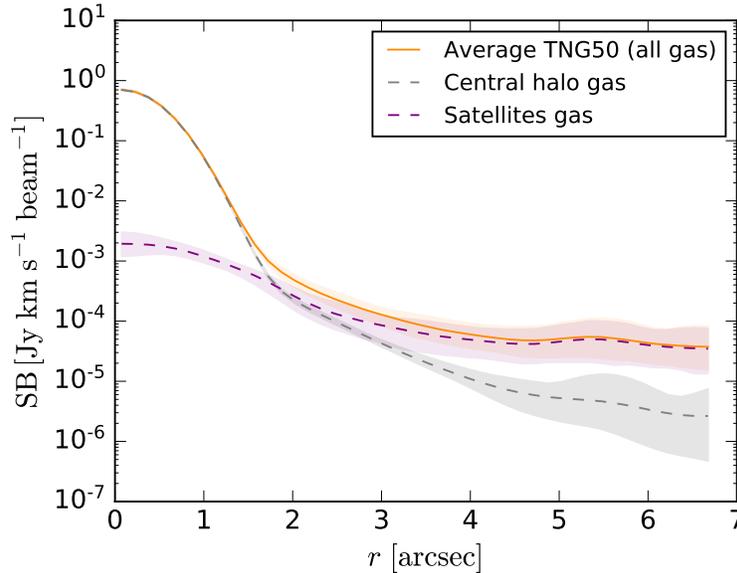
### Contribution of satellites

Another potential origin of [CII] halos is related to satellite galaxies (e.g., Fujimoto et al., 2019). This scenario suggests that the presence of satellite galaxies surrounding central star-forming galaxies gives rise to observed extended [CII] emission. In this context, we analyzed the contribution of gas that is not part of the central galaxy, which consists in gas from satellites, to the total [CII] emission. Our sample excludes systems with the presence of observable companion galaxies, i.e. with  $\text{SFR} > 1.5 M_{\odot}$  and at a projected distance  $< 3$  arcsec from the central galaxy, to be in agreement with Ginolfi et al. (2020b). However, we analysed the contribution of low-SFR satellite galaxies (that would not be detectable by observations) that still might be present at small and large scales.

In this context, Figure 3.11 shows the contribution of satellites (purple dashed line) and the contribution of the central galaxies (grey dashed line) to the total [CII] profile (orange). As expected, we note that for the inner part of the galaxies, i.e between 0 and  $\sim 2$  arcsec, the contribution from the central galaxy dominates. However, there is also a very low contribution from the satellites because of projection effects. We also note that after  $\sim 2$  arcsec the contribution of satellites starts to dominate. This contribution increases up to a  $\sim 100$  per cent at  $R_{200}$  ( $\sim 7$  arcsec).

To further explore the contribution of satellites, Fig. 3.12 shows the same analysis as that in Fig. 3.10, but only for the gas from satellites, neglecting the contribution of the central galaxies. From the 2D histogram (left panel), it is evident that companion galaxies (clumpy structures) are already present on small scales,  $\sim 0.1 R/R_{200}$  (or 4 kpc/0.7 arcsec at  $z=5$ ). From the histograms at the right panel, we note that after  $0.6 R/R_{200}$  almost all (99 per cent) the gas from satellites is inflowing towards the central halo.

Interestingly, both these results suggest that gas that is part of satellites galaxies may



**Figure 3.11:** Averaged contribution of the satellites gas (purple dashed line) and central galaxies gas (grey dashed line) to the total [CII] emitting gas (orange line). The shaded areas enclose a 95 percent of different results obtained by different galaxy orientations.

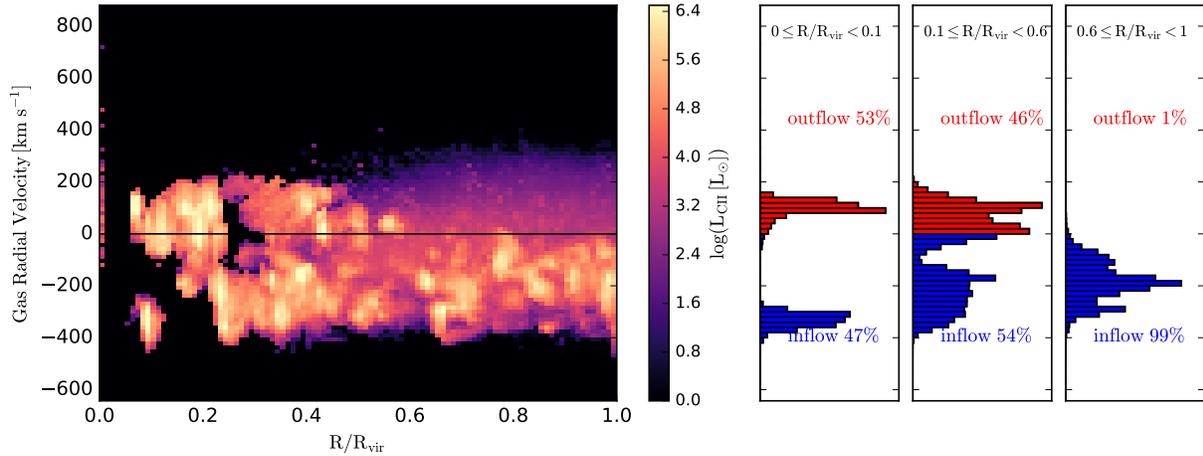
play an important role in observations of [CII] halos, especially when analyzing stacked data, in contrast to what has been argued in observational works (e.g., Fujimoto et al., 2019, 2020). These results will be further elaborated and discussed in the next section.

## 3.5 Discussion

In this section we explore different properties associated with extended [CII] emission based on the results found in this work. Subsequently, we aim to interpret these properties and discuss potential scenarios concerning the origin of [CII] halos.

### 3.5.1 Origin of large-scale star-forming gas

The CGM is a complex and dynamic environment influenced by different physical processes operating across various scales. These processes include gas accretion, feedback from star formation and black holes, and interactions with the surrounding environment (see Tumlinson et al. 2017 and references therein). In this context, at high- $z$  the presence of star-forming gas extending beyond the body of the central galaxies is not unexpected in the TNG model (e.g., Donnari et al., 2019). In particular, Pillepich et al. (2019) studied the evolution of stellar and gaseous components in star-forming TNG50 galaxies at  $0 < z < 6$ . They found that, for all masses and redshifts, the star-forming gas (as traced by  $H\alpha$ ) settles into disky or elongated morphologies. At high- $z$ , they also find that the star-forming gas often exhibits a complex structure, characterized by significant asymmetries on large



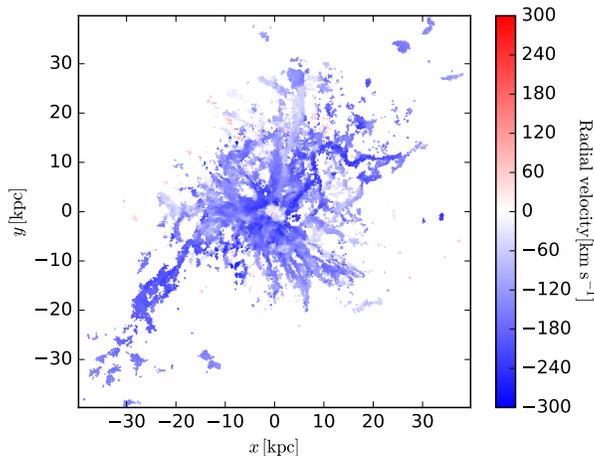
**Figure 3.12:** Same as in Figure 3.10, but for a stacking of all satellites (rejecting the contribution of the central galaxies) of all the selected TNG50 halos.

scales (beyond the stellar body), which can be attributed to the prevalence of inflows and outflows during this epoch.

In Section 3.4.2, it was shown that the star-forming gas in the TNG50 galaxies is extended up to  $\sim R_{200}$ , dominating the [CII] SB profiles from inner galactic to CGM scales. One option is that this extended star-forming gas comes from the central galaxy. Galaxies undergoing active star formation can generate strong outflows that expel gas into their surroundings. This gas could then cool down and lead to star formation in the halo. Another option is that this star-forming gas comes from satellites galaxies. Interactions can trigger star formation and redistribute gas within and around galaxies, or gas can be stripped or ejected from satellites galaxies. Furthermore, the extended star-forming gas may be fueled by accretion of cold gas from the intergalactic medium along filamentary structures.

Considering the kinematic analysis performed in Section 3.4.2 together with the contribution of satellites in Section 3.4.2, we conclude that the large-scale ( $r > 1.5$  arcsec) star-forming gas is mostly infalling and belongs to the satellites galaxies. Specifically, from Fig. 3.12 we obtain that at  $R/R_{200} > 0.6$  almost all the gas belonging to satellites is infalling (right third histogram). Another illustrative example is presented in Fig. 3.13, that shows an averaged SFR-weighted radial velocity map of the star-forming gas for a galaxy at  $z=5$ . Inflows are represented in blue and outflows in red. It is evident from the figure that the majority of the large-scale star-forming gas primarily consists of infalling gas.

As a consistency check, we also analyzed the metallicity distribution of the star-forming gas. The chemical signatures within the CGM are influenced by different mechanisms such as galactic outflows from supernovae and black holes, and the ejection or stripping of metals from satellite galaxies (e.g., Tegmark et al. 1993; Shen et al. 2013; Nelson et al. 2019; Torrey et al. 2019; Péroux et al. 2020). Figure 3.14 shows an averaged SFR-weighted metallicity map (right panel) of the star-forming gas for a galaxy at  $z=5$ . Its corresponding averaged

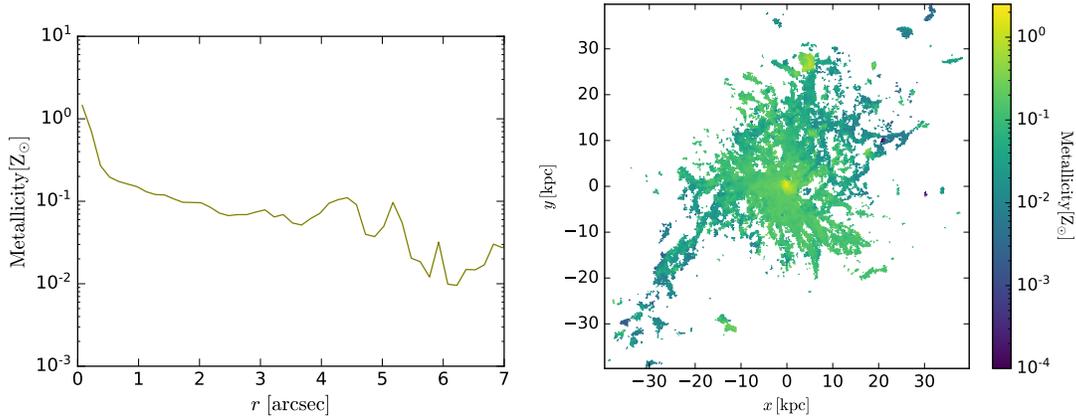


**Figure 3.13:** Average SFR-weighted radial velocity map for a galaxy at  $z=5$ . Only the star-forming gas is shown. The box size radius is  $\sim 40$  kpc or  $\sim 1R_{200}$ .

radial profile is shown in the left panel. As expected, the gas that in the central body of the galaxy exhibits higher metal enrichment (by almost 2 orders of magnitude) compared to the large-scale accreting gas. Additionally, we note that the infalling gas streams in the CGM are already metal polluted at this redshift, with typical values in the range of  $10^{-1}$ - $10^{-2}Z_{\odot}$ . As a reference, the dominant metallicity of outflowing gas for  $z \sim 4$  galaxies ( $M_{*} \sim 10^{9.5}M_{\odot}$ ) has been found to be  $\sim 10^{-0.7}Z_{\odot}$  (Nelson et al., 2019), while gas from the intergalactic medium is expected to have a low-metallicity of  $\sim 10^{-3}Z_{\odot}$  at  $z \sim 4-6$  (Pallottini et al., 2014). Overall, this result is expected since metal-enriched gas is necessary for the presence of  $C^{+}$  ions and consequently [CII] emission.

### 3.5.2 Role of satellite galaxies in producing [CII] halos

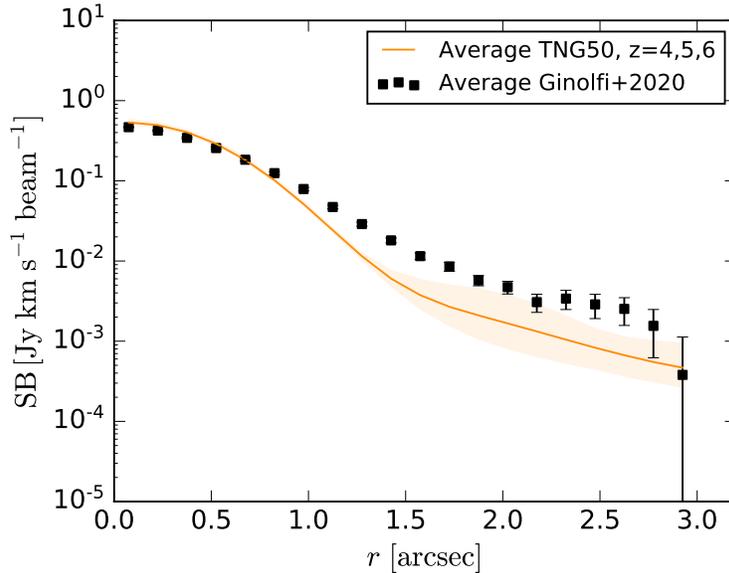
Most observational works studying [CII] halos have focused on isolated galaxies (i.e., with no observable minor or major mergers, e.g., Fujimoto et al. 2019; Ginolfi et al. 2020b; Fujimoto et al. 2020; Herrera-Camus et al. 2021; Fudamoto et al. 2022). However, contamination from possible faint undetected satellites cannot be discarded. The presence of satellites may affect the gas dynamics, heating processes, or feedback mechanisms, influencing the overall [CII] emission properties. In this context, some attempts from observations have been made to estimate the contribution of such undetected satellites to the observed [CII] halos. For instance, Fujimoto et al. (2019) investigated the possible contribution of satellites to the extended [CII] emission they found in a stacking of data of star-forming galaxies at  $z = 5-7$ . They estimated radial values of the  $L_{\text{CII}}/\text{SFR}$  ratio, finding the highest ratios in the outer regions. These values are in disagreement with those expected for local dwarf galaxies. The authors interpret this as an indication that satellites are likely not originating the observed [CII] halo. On the other hand, Ginolfi et al. (2020a) studied one merging system at  $z \sim 4.6$ , from the ALPINE survey. They find a [CII] halo extending



**Figure 3.14:** Left: Average metallicity (weighted by SFR) radial profile of the star-forming gas for one galaxy at  $z=5$ . Right: Average SFR-weighted metallicity map for a galaxy at  $z=5$ . Only the star-forming gas is shown. Its corresponding radial profile is shown in the left panel. The box size radius is  $\sim 40$  kpc or  $\sim R_{200}$ .

up to  $r \sim 15$  kpc, and argue that its origin is mainly interstellar gas stripped by strong gravitational interactions, with a possible less significant contribution from galactic outflows and star formation from small faint satellites. Recently, Schimek et al. (2023) presented a theoretical effort of simulating [CII] emission in a zoom-in simulation of a system at  $z=6.5$  undergoing a major merger. Their results indicate that the [CII] emission in the CGM is higher in an accreting filament and tidal features from the merging galaxies.

From the previous section and the results presented in Section 3.4.2, we find that star-forming gas from satellites or companion galaxies have a substantial contribution to the extended [CII] emission after  $\sim 2$  arcsec. This suggests that faint satellites, which may be undetected in observations, could be important contributors to the [CII] halos observed around galaxies. However, as we are not obtaining the emission level of the observed [CII] profile on CGM scales, we cannot confirm that. A possibility for this missing flux is that there might be other physical processes affecting the extended [CII] emission that are not well reproduced by TNG50 (see next Section). To test what would be the maximum contribution of satellites in our model, we derived the average [CII] SB profile for a larger sample of TNG50 galaxies (169) that covers the physical properties range of Ginolfi’s sample (see Section 3.3.2), but without imposing a similar distribution in SFR and stellar mass. Importantly, we select these galaxies without removing the systems with mergers. The result is presented in Fig. 3.15, where the  $C^+$  fractions were estimated by assuming UVB and young stars as ionizing sources. Interestingly, the level of [CII] emission at large scales is  $\sim 8$  times higher than what is found in our cleaned sample (see Fig. 3.6, bottom panel), being closer to the observations. This supports the scenario in which the presence of satellites and their interactions with the central galaxies could have a significant impact on the [CII] emission in the inner CGM. This effect could be especially important in observational results based on stacking techniques.

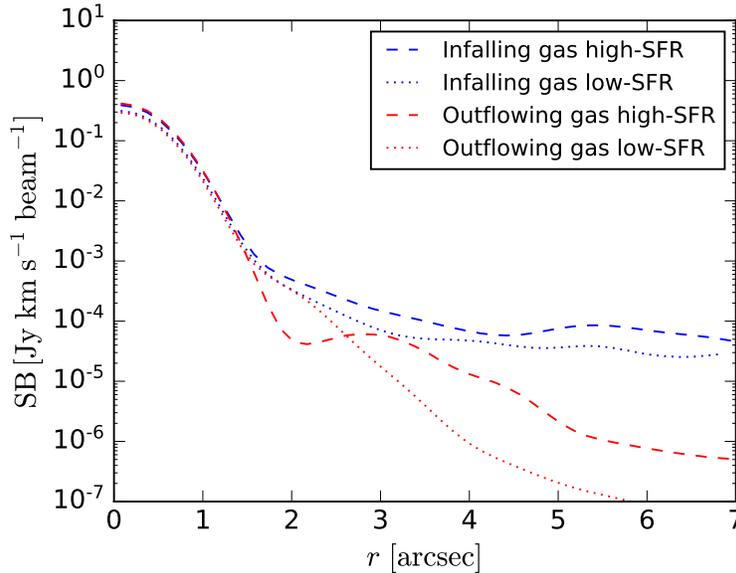


**Figure 3.15:** Averaged [CII] SB profiles obtained for a sample of 169 TNG50 galaxies at  $z=4, 5$  and  $6$  (orange line) which include close satellites. The shaded area enclosing a 95 percent of different results obtained by different galaxy orientations. The  $C^+$  fractions have been estimated considering UVB and local young stars as ionization sources.

In summary, our model, which shows that [CII] traces star-forming gas, suggest that satellites (including faint and undetected ones) can play an important role in shaping the extended [CII] emission in the inner CGM of high- $z$  galaxies. However, further investigation is needed to fully understand the specific role of satellites in reproducing the exact levels of the observed extended [CII] emission. Even when including galaxies with close satellites, we are not able to fully match the observations. Future theoretical works exploring the detailed properties of satellites, their gas content, and interactions with the central galaxies could provide insights into the mechanisms responsible for the observed [CII] halos of high- $z$  galaxies.

### 3.5.3 Role of outflows in [CII] emission of high and low-SFR galaxies

Based on a spectral and cube stacking analysis, [Ginolfi et al. \(2020b\)](#) argue that [CII] halos trace circumgalactic gas that has been previously enriched by past outflows driven by star formation activity. They found extended [CII] emission only for the subsample of galaxies with high-SFR ( $SFR > 25 M_{\odot} \text{ yr}^{-1}$ ), implying that galaxies with more active star formation should present more extended [CII] emission. This interpretation is further supported by semi-analytical models ([Pizzati et al. 2020, 2023](#)). For instance, [Pizzati et al. \(2023\)](#) implemented a model to predict [CII] halos originated from outflowing gas, which



**Figure 3.16:** Averaged [CII] SB profiles for the outflowing (inflowing) gas only, represented by the red (blue) color. Our original TNG50 sample has been divided in two subsamples: high-SFR (dashed lines) and low-SFR (dotted lines). The  $C^+$  fractions have been estimated considering UVB and local young stars as ionization sources.

incorporates key factors such as the outflow mass loading factor<sup>11</sup> ( $\eta$ ), the parent galaxy SFR, and the circular velocity of the dark-matter halo. In particular, the value of  $\eta$  is related to the radiative cooling of the wind and its capability to reach low temperatures ( $T \sim 10^2 - 10^4$  K). Conversely, the gravitational potential of the dark-matter halo can substantially decelerate the gas expansion. These authors compare their results with ALPINE individual galaxies (from Fujimoto et al. 2020) finding that detected [CII] haloes are a natural outcome of starburst-driven outflows. Their model also suggests that low-mass systems also present [CII] halos, however their emission levels are too faint to be detected with the current observational capabilities.

In Section 3.4.2, we analyzed the [CII] SB profiles of our high-SFR and low-SFR subsamples, finding no major difference between both profiles at CGM scales, in contrast with the observational results of Ginolfi et al. (2020b). Both subsamples present extended [CII] emission, however the emission of the high-SFR subsample should be higher (up to  $\sim 20$  times) to match the observations. To have more insights regarding the reason of this discrepancy, we constructed the [CII] radial SB profile of the outflowing gas only and look for differences in both SFR subsamples (for completeness, the same is done for the inflowing gas). This exercise is shown in Fig. 3.16, where the dashed line indicates the average for high-SFR galaxies ( $SFR > 25 M_{\odot} \text{ yr}^{-1}$ ), and the dotted line for low-SFR galaxies ( $SFR < 25 M_{\odot} \text{ yr}^{-1}$ ). The radial axis goes until  $\sim 7$  arcsec ( $\sim R_{200}$ ). We note that for galactic scales ( $r < 2$  arcsec), the SB of the outflowing gas is  $\sim 2$  times higher for the

<sup>11</sup> $\eta = \dot{M}/SFR$ , with  $\dot{M}$  the outflow rate.

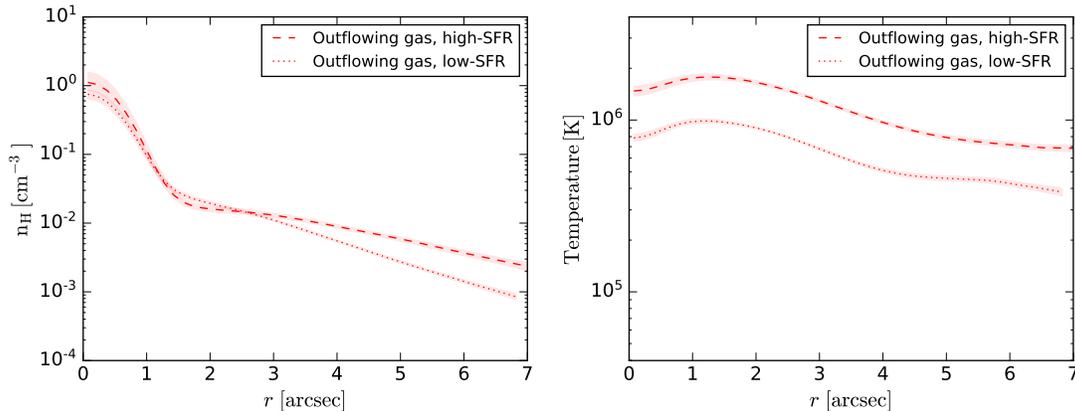
high-SFR galaxies. Then, at  $r \sim 2$  arcsec there is a drop in the profile of the high-SFR subsample, which then raises again, overcoming the outflowing gas profile of low-SFR galaxies at  $r \sim 2.8$  arcsec, until  $\sim R_{200}$  (by a factor of  $\sim 10$ ). This indicates that outflows in the high-SFR galaxies are more energetic and hence more extended, contributing to a larger observable [CII] SB for scales larger than  $r \sim 2.8$  arcsec. However, this increased outflow contribution is still much smaller than the contribution from infalling star-forming gas in our model. If outflows are the main contributors to the [CII] halos, their phase diagram in the TNG50 simulations are not accurate enough to produce observable levels of emission.

In particular, the TNG model assumes that star formation feedback drives galactic outflows, which are launched from star-forming gas with a wind velocity ( $v_w$ ) that depends on the local dark matter velocity dispersion ( $\sigma_{\text{DM}}$ , e.g., [Vogelsberger et al. 2013](#)). The outflow mass loading is determined by the supernova energy and the assumed wind speed, while the metal content is a fraction of ISM metallicity. The model uses a kinetic wind scheme where wind particles are created and decoupled from the gas until they leave the dense ISM. Outside the dense medium, the wind particles recouple with the gas, transferring their mass, momentum, metals, and thermal energy to the surrounding medium. Wind particles re-couple to the gas cell when: 1) the density drops below a threshold ( $n_{\text{H}} = 0.05 \times$  density threshold for star formation  $= 0.006 \text{ cm}^{-3}$ ), or 2) after a maximum travel time has been reached ( $0.025 \times$  Hubble time at respective  $z$ ) ([Pillepich et al., 2018](#)).

For the TNG50 galaxies at  $z=5$  (median of the sample) studied here, the average distance<sup>12</sup> at which the wind particles would re-couple with the gas from the condition 1), is  $\sim 5$  arcsec (or  $\sim 32$  kpc) and  $\sim 4.6$  arcsec (or  $\sim 30$  kpc) for the high and low-SFR subsamples, respectively. On the other hand, following condition 2), the maximum distance<sup>13</sup> after the maximum travel time is  $\sim 2.8$  arcsec (or  $\sim 18$  kpc) and  $\sim 2.4$  arcsec (or  $\sim 15$  kpc) for the high and low-SFR subsamples, respectively. These results indicate that for the case of our selected halos, the wind particles re-couple with the gas cells according to the second criteria, which is first met. The different radial distances estimated for both subsamples confirm that the outflows of the high-SFR galaxies propagate at larger distances compared to the low-SFR galaxies (Fig. 3.16). On the other hand, the distance of  $\sim 18$  kpc ( $\sim 2.8$  arcsec) obtained for the high-SFR subsample is comparable to the distance that reach the outflows in the observations of [Ginolfi et al. \(2020b\)](#) ( $r \sim 15$  kpc), and is also close to the radius where there is a drop in the outflowing gas SB profile ( $\sim 2$  arcsec, Fig. 3.16). This suggests that as the wind particles re-couple with the gas and transfer their properties, it can lead to a change in the gas properties and potentially a decrease in the [CII] emission. Therefore, if all of the observed extended [CII] emission is attributed to stellar winds, it would be necessary to incorporate a new subgrid prescription able to enhance the wind

<sup>12</sup>To estimate this distance, an average  $n_{\text{H}}$  profile has been computed, considering 3D radial bins of  $\sim 0.4$  arcsec. This profile then provided the average distance at which the density threshold is reached.

<sup>13</sup>This distance is given by  $v_w \times$  current Hubble time  $\times 0.025$ . We first estimated  $v_w$  that depends on  $\sigma_{\text{DM}}$  (see equation 1 in [Pillepich et al. 2018](#)). To estimate  $\sigma_{\text{DM}}$ , we assumed that  $V_{\text{max}} \sim 1.45\sigma_{\text{DM}}$  ([Okamoto et al., 2010](#)), with  $V_{\text{max}}$  the maximum circular velocity of the host dark matter halo. By assuming a NFW profile,  $V_{\text{max}}$  was estimated as  $V_{\text{max}} = \sqrt{0.2162c/f(c)}V_{\text{vir}}$ , with  $V_{\text{vir}}$  the virial halo velocity,  $c$  the concentration parameter, and  $f(c) = \ln(1+c) - c/(1+c)$  (e.g., [Dutton & Macciò 2014](#)).



**Figure 3.17:** Left: Averaged density profile, weighted by temperature (colder gas has higher weight), obtained for the outflowing gas only. Right: Averaged temperature profile, weighted by density (denser gas has higher weight), obtained for the outflowing gas only.

contribution significantly. This increase could be achieved by including a prescription for an entrained clumpy, cold dense phase that is currently missing (e.g., Hamann et al. 2013; Liang et al. 2016; Faucher-Giguère et al. 2016; McCourt et al. 2018; Gronke & Oh 2018, 2020).

Interestingly, the drop at  $r \sim 2$  arcsec could be directly related to the results shown in Fig. 3.8: around  $r \sim 2$  arcsec is when the profiles of both subsamples coincide. To better understand the physics behind this drop, we analyzed the radial density and temperature distributions (smoothed by the ALMA beamsize) of the outflowing gas, for both subsamples (Fig. 3.17). While the temperature profiles show a similar behavior for both subsamples (having the high-SFR subsample higher temperatures along all radii), the density profiles of both subsamples behave differently. The density profile for the high-SFR galaxies is higher than the one for the low-SFR galaxies at central scales, and then decreases becoming lower than the low-SFR profile at  $r \sim 2$  arcsec. Such a decrease in density would result in a smaller cooling rate for [CII].

Overall, this analysis suggests that the general difference in the [CII] profiles between the low and high-SFR samples in the outflows case can be attributed to the distance that wind particles can travel in each scenario. On the other hand, the drop in the outflow radial profile of the high-SFR subsample around 2 arcsec is closely linked to the decrease in density at the same radius, and potentially to the average radius at which wind particles re-couple to the gas. In turn, this different behavior of the outflows in the high and low-SFR subsamples would be influencing their total [CII] SB profiles.

On the other hand, it is important to discuss the effect of AGN feedback. While the observed samples do not report the evidence for AGN activity in the targeted galaxies down to current observational limits, all our selected TNG50 galaxies have AGNs ( $L_{\text{BOL}} \lesssim 10^{45.5}$  erg  $\text{s}^{-1}$ ) affecting the surrounding gas with thermal mode feedback<sup>14</sup> (i.e.,

<sup>14</sup>In the TNG model, the feedback from supermassive black holes is implemented based on the Bondi-

high accretion rates, average  $\dot{M}/\dot{M}_{\text{Edd}} = 0.5$ ). The average AGN feedback energy is higher for the high star-forming sample compared to the low star-forming sample ( $\Delta E = 10^{43.9} \text{erg s}^{-1}$  and  $10^{43.3} \text{erg s}^{-1}$ , respectively). This stronger AGN feedback in the most star-forming galaxies may lead to higher gas temperatures (right panel of Figure 3.17), resulting in similar or lower [CII] emission compared to the low star-forming galaxies. Importantly, an AGN would also affect the  $C^+$  fractions with its hard radiation. This effect is not taken into account in our calculations, but would increase the tension with observations as we expect Carbon to be more ionized when accounting for an AGN contribution.

Overall the discrepancy between the simulated [CII] profile for the high-SFR subsample and the observed one could be attributed to different factors: 1) low statistic or absence of very high SFR galaxies in the simulations. Cosmological simulations have difficulties in reproducing the violent star-formation rates invoked by observations (e.g., Furlong et al. 2015; Donnari et al. 2019; Katsianis et al. 2020). The number of high-SFR galaxies studied here is smaller than the number of low-SFR galaxies. Our high-SFR subsample consists of 21 galaxies (versus 50 for the low-SFR subsample), and are in the range of 25-270  $M_{\odot} \text{ year}^{-1}$ , while in the observations they analyze 20 high-SFR galaxies in the range of 25-623  $M_{\odot} \text{ year}^{-1}$  (versus 20 low-SFR). The most star-forming systems in the observations, those at  $\sim 600 M_{\odot} \text{ year}^{-1}$ , are dominating the averaged [CII] halo emission. However, in TNG50 such systems are simply not available. 2) Bias in the observations. The results of the observations may be affected by a selection bias favoring more luminous [CII] halos at a given SFR. 3) Subgrid prescription regulating the feedback processes (from star-formation and AGN) in TNG50 may not accurately capture the complex physical processes responsible for producing extended [CII] emission in galaxies. 4) Other resolution limitations and missing physical processes in our model, as the caveats explained in Section 3.5.4.

In summary, our predicted [CII] emission in TNG50 galaxies suggests that [CII] halos originate through a combination of contribution of star-forming gas (associated with satellite galaxies on CGM scales) and outflows from galactic scales. The contribution of inflowing gas from satellites dominates after  $\sim 2$  arcsec. However, compared to the observations we do not match the level of [CII] emission on CGM scales. This discrepancy could be related to the limited resolution of the simulations, issues in the feedback model of TNG, and/or other physical processes not implemented in our model. The next section details the caveats of our model.

### 3.5.4 Caveats

There are certain limitations and considerations related to the simulations and model assumptions that need to be taken into account when interpreting the results and conclusions presented in the previous sections:

---

to-Eddington mass growth ratio ( $\dot{M}_{\text{Bondi}}/\dot{M}_{\text{Edd}}$ ). When this ratio is larger (smaller) than a threshold of  $\dot{M}_{\text{Bondi}}/\dot{M}_{\text{Edd}} = \min[0.02(M_{\text{BH}}/10^8 M_{\odot})^2, 0.1]$ , the feedback energy is injected in “thermal mode” (“kinetic mode”) (Weinberger et al., 2017).

- **Cold dense gas clumps in the CGM:** The TNG50 simulations may not fully resolve cold high-density gas within the CGM, potentially resulting in an underestimation of the [CII] emission. It is possible that the presence of cold clumps in the CGM would contribute to higher [CII] emission levels than those captured in this work. The presence and survival of dense CGM gas is supported by observations (e.g., [Cantalupo et al. 2014](#); [Arrigoni Battaia et al. 2015](#); [Borisova et al. 2016](#)) and high-resolution "cloud-crushing" simulations (e.g., [Gronke & Oh, 2018, 2020](#); [Kanji-lal et al., 2021](#)). In particular, [Nelson et al. \(2020\)](#) attempted to resolve small-scale cold CGM gas in TNG50, finding the presence of cold clouds whose number and mass increase with resolution. Despite these efforts, such processes are still largely unresolved. In light of the critical densities associated with  $C^+$ , it becomes necessary to incorporate higher densities, potentially through the application of density distributions on sub-grid scales (e.g., [Olsen et al. 2021](#); [Schimek et al. 2023](#)).
- **Coldest temperature:** The temperature of the star-forming gas cells is assumed to be fixed at 1000 K (cold mode, [Springel & Hernquist 2003](#)). While this specific value is motivated by the subgrid model used in the simulations, in reality star-forming gas is characterized by a distribution of temperatures which spans also lower values (e.g., [Kainulainen et al. 2009](#); [Draine 2011](#); [Olsen et al. 2021](#)). Colder dense gas (currently not modelled in the TNG50 suite) could enhance the production of [CII] emission.
- **Turbulence dissipation in the CGM:** The CGM is expected to be a turbulent environment, influenced by different perturbations associated with galaxy interactions, accretion, and outflows. Turbulence dissipation affects the dynamics, thermal properties, and chemical enrichment of the CGM (see [Faucher-Giguere & Oh 2023](#) and references therein). In the cold CGM, turbulent widths have typical values of  $\sim 10\text{-}30 \text{ km s}^{-1}$ , which indicates random motions with velocity dispersions of  $\sigma \sim 1000 \text{ km s}^{-1}$  (e.g., [Borisova et al. 2016](#); [McCourt et al. 2018](#); [Rudie et al. 2019](#)). Considering this, the inclusion of sub-grid scale high-velocity turbulences could be essential to enhance the frequency of collisions between  $C^+$  and its colliding partners, leading to an increased [CII] emission in the CGM (e.g., [Godard et al. 2014](#)).
- **$H_2$  as collisional partner:** In our post-processing model, we have neglected  $H_2$  as colliding partner because of the reasons explained in Section 3.3.3, and especially because its physics is not modelled by the simulations. [CII] emission due to collisions with  $H_2$  is characterized by the highest critical density when compared to collisions with  $e^-$  and  $H$  (see Section 3.1), and such densities are not present on CGM scales in the TNG50 simulations. However,  $H_2$  might be an important collisional partner out to the inner CGM. There is growing observational evidence of the presence of molecular gas outside the main body of high- $z$  galaxies, i.e. reaching the inner CGM (e.g., [Ginolfi et al. 2017](#); [Falgarone et al. 2017](#); [Emonts et al. 2019](#); [Li et al. 2020](#); [Vidal-García et al. 2021](#); [Emonts et al. 2023](#)). This suggests that  $H_2$  may indeed be an important contributor to the [CII] emission in those regions.

- **Assumptions in the CLOUDY calculations:** To obtain the  $C^+$  fractions in the simulated systems, we used CLOUDY tables by assuming a fixed escape fraction of ionising photons of 5 per cent. However, it is important to note that this escape fraction might be variable at different radial scales, as well as the optical thickness of the medium. Some works suggest that  $f_{esc}$  should be very small ( $<1$  per cent) to produce the observed [CII] halos (e.g., Pizzati et al. 2020, 2023), and it may also vary for different galaxy mass ranges (e.g., Xu et al. 2016.)

### 3.6 Summary

To investigate the nature of [CII] halos detected in observations around galaxies at  $z > 4$ , we conducted a post-processing analysis on the Illustris-TNG50 simulation. We aimed to generate [CII] emission in a sample of simulated galaxies at  $z \sim 4-6$ . Specifically, by incorporating  $C^+$  fractions derived by assuming UV background and young stars as radiation sources, we estimate [CII] emissivities and generate mock observations, comparing our results with the ALPINE [CII] emitting galaxies investigated by Ginolfi et al. (2020b). This approach allowed us to study the possible origin of extended [CII] emission in high- $z$  galaxies, and discuss the contribution of different factors to this emission. The main results of this work are summarized as follows:

i) Our model reproduces the  $L_{\text{CII}}$  versus SFR relationship found by previous works in local and high- $z$  galaxies. This result provides validation for the predictions from our model on galactic scales.

ii) We derived an averaged [CII] SB profile and compared it to the observed data. Our predicted [CII] emission levels on central scales ( $r < 1.5$  arcsec or  $\sim 10$  kpc at  $z=5$ ) are in agreement with the observations. However, at CGM scales our results underestimate the [CII] emission levels by a factor of  $\sim 10$ .

iii) The [CII] emission in our systems is dominated by the dense star-forming gas, as expected. This gas extends up to the virial radius.

iv) From a kinematic analysis, we find that on inner galactic scales ( $r < 0.7$  arcsec or 5 kpc at  $z=5$ ) the [CII] emission traces a balance between inflows and outflows (contribution of 51 and 49 per cent, respectively)

v) The contribution of infalling gas from satellites dominates the [CII] emission in the CGM, specifically after  $r \sim 2$  arcsec (or  $\sim 13$  kpc at  $z=5$ ). This suggests that faint and undetected satellites may play a significant role in shaping the observed extended [CII] emission.

vi) We find further tension with the observations when comparing the SB profiles for the low and high-SFR galaxies, finding no significant difference in the extended [CII] emission levels (up to  $r \sim 3$  arcsec or  $\sim 20$  kpc at  $z=5$ ) between the two subsamples. We propose that this could be an effect of the feedback model and the limited resolution of the TNG simulations.

vii) According to our results, the origin of [CII] halos is a combination of contribution

of gas from satellites and outflows from the central galaxies. However, the contribution of infalling gas by galaxies completely dominates at  $4 < r < 7$  arcsec ( $0.6 < r/R_{200} < 1$ ).

Therefore, this theoretical effort highlights the complex nature of extended [CII] emission in the inner CGM of high-redshift galaxies. Further investigations based on the state-of-art simulations and observations are necessary to fully understand the specific role of satellites and other physical processes (e.g., galactic winds, turbulences) in reproducing the observed [CII] halos. In particular, cosmological simulations with higher resolution in the ISM and CGM are needed. While running cosmological simulations with higher refinements is currently computationally expensive and requires tuning, one could achieve improved results with the use of more detailed subgrid prescriptions based on radiative transfer calculations run on high-resolution turbulent boxes (Buhlmann et al. in prep). On the other hand, future observations from large sub-mm single-dish telescopes (e.g., AtLAST, 50 m telescope) in combination with high-resolution interferometric observations (e.g., CRISTAL, large ALMA project) will be able to improve our general knowledge of the CGM and specifically of [CII] halos in high- $z$  objects.



# Chapter 4

## Summary and Future outlook

In this chapter, I provide a summary of the key findings and contributions of this thesis to the field of galaxy evolution. By placing my work in context, I emphasize its significance and impact on the field. Additionally, I outline the future prospects and potential extensions of this thesis, focusing on the projects presented herein.

### 4.1 Summary and implications of the main results

#### 4.1.1 Link between molecular gas reservoirs and extended Ly $\alpha$ emission in $z \sim 3$ quasars

To unveil the multi-phase nature of the gas reservoirs around  $z \sim 3$  quasars, I performed an observational work that was presented in Chapter 2. I conducted a deep spectroscopic survey with APEX/SEPIA180, targeting the CO(7-6), CO(6-5) and [CI](2-1) transitions for a sample of nine  $z \sim 3$  quasars, which were previously observed with VLT/MUSE, to study the molecular gas phase and its link with the Ly $\alpha$  properties on larger scales. The findings from this pilot sample revealed intriguing insights. Our results suggest that the most massive molecular reservoirs are associated with the dimmest and some of the smallest Ly $\alpha$  nebulae (see Fig. 2.10), possibly indicating that a large fraction of the quasar photons are blocked on galaxy scales. Additionally, indications of potential large-scale inflows and outflows were detected. We found large velocity shifts (ranging from  $\sim -400$  to  $\sim +1200$  km s $^{-1}$ ) between the bulk of Ly $\alpha$  and the systemic redshift in five systems, with the broadest Ly $\alpha$  lines associated to the sources with the largest velocity shifts.

Overall, to our knowledge this work represents one of the first observations providing constraints and correlations between the cool gas traced by Ly $\alpha$  emission and the molecular gas content in high- $z$  quasars. By doing so, it enhances our understanding of the gas properties, kinematics, and the interplay between different gas phases in the context of galaxy formation and evolution. Specifically at  $z \sim 3$ , that represents the epoch at which the quasar density peaks, and the lowest redshift accessible by MUSE. The findings presented in this study provide important insights into the underlying physical processes that shape

these systems. Recent works are now attempting similar analysis, aiming to link both gas phases in high- $z$  quasars using ALMA observations. For instance, Drake et al. (2022) compared the Ly $\alpha$  and [CII] emission observed with ALMA, in eight  $z \sim 6$  quasars. Their results suggest that the kinematics of the halo gas (as traced by Ly $\alpha$ ) is decoupled from that of the host galaxies' disks (as traced by [CII]). On the other hand, Li et al. (2023) studied a sample of ten quasars with enormous Ly $\alpha$  nebulae at  $z \sim 2$ , with ALMA+ACA CO(4-3) observations. They found extended (15-100 kpc) reservoirs of molecular gas in six sources and no trend between  $M_{\text{H}_2}$  and brightness/size of the Ly $\alpha$  nebulae. However, their sample only included extremely large and bright Ly $\alpha$  nebulae, and they explored less physical properties between molecular gas and Ly $\alpha$  nebulae compared to our work.

Finally, our results provides a foundation for further investigations and follow-up projects (see Section 4.2), ensuring advancements in our knowledge of these type of systems.

#### 4.1.2 [CII] emission in simulated TNG50 galaxies at $z \sim 4-6$

In Chapter 3, I continued investigating cold gas reservoirs at high redshift, but now from a theoretical perspective. I implemented a post-processing model to predict the [CII] emission levels for TNG50 star-forming galaxies at  $4 < z < 6$ , and compared them with ALPINE survey observations. Our model predicted similar [CII] emission on galactic scales, but underestimated the emission at larger scales by  $\sim 10x$ . It is important to note that several caveats in our models may contribute to this discrepancy (Section 3.5.4). Also, there was no significant difference in [CII] emission between low and high-SFR galaxies, possibly due to resolution limitations and feedback modeling of the simulation. Finally, our findings suggest that the extended [CII] emission is likely a result of both gas from satellite galaxies and outflows from central galaxies, with satellites playing a dominant role within  $0.6 < r/R_{200} < 1$ .

This work represents the first attempt to simulate [CII] emission in galaxies within the context of the IllustrisTNG simulation suite. In summary, this work advances our understanding of [CII] emission in high-redshift galaxies and its connection to galactic-scale and extended structures. Interestingly, the dominance of infalling gas from satellites in the large-scale [CII] emission suggests their important role in enriching the CGM and shaping the overall gas dynamics. Understanding the relative contributions of different gas sources provides valuable insights into the mechanisms driving the assembly and evolution of galaxies, as well as the influence of their surrounding environment.

Other theoretical works have used zoom-in simulations to predict [CII] emission, failing in reproducing [CII] halos (e.g., Pallottini et al. 2017; Arata et al. 2019). It is suggested that this discrepancy is due to that certain physical processes such as feedback or metal enrichment, are not well reproduced by the simulations, or due to limitations in the modelling of [CII] emissivities. In a recent work, Schimek et al. (2023) applied different post-processing models, based on CLOUDY and radiative transfer codes, to produce [CII] emission in a zoom-in simulation of a star-forming galaxy at  $z \sim 6.5$ . Their calculations were not able to match the  $L_{[\text{CII}]}$ -SFR relationship found in observations (e.g., De Looze et al. 2014; Herrera-Camus et al. 2015). Instead, they significantly overestimate the total

[CII] emission, obtaining luminosities up to 1.5 order of magnitudes higher. They propose that higher resolution able to resolve more substructures in the cold gas, as well as incorporating variations in the stellar initial mass function and accounting for turbulent mixing, could potentially improve their results.

Finally, the study of this thesis emphasizes the need of cosmological simulations with higher resolution in the ISM and CGM to accurately capture the underlying physical processes. It also highlights the necessity of more detailed subgrid prescriptions, possibly incorporating feedback mechanisms that can effectively entrain compact and cold structures in outflows. This would allow to bridge the gap between simulations and observations, leading to a more comprehensive understanding of galaxy formation and evolution. A future outlook of this project is discussed in Section 4.2.2.

## 4.2 Future directions and projects extensions

### 4.2.1 Multi-phase gas reservoirs in $z \sim 3$ quasars

The results from my first project (summarized in Section 4.1.1) raised several important scientific questions, leading to the submission of observing proposals to acquire new data and address key inquiries:

- What are the physical properties characterizing the ISM of the uncovered most massive systems, such as star formation activity, gas metallicities, and gas densities? (Section 4.2.1)
- Is the tentative trend between Ly $\alpha$  nebula surface brightness/size and molecular mass confirmed or refuted when examining other  $z \sim 3$  quasars with dim and small Ly $\alpha$  nebulae? (Section 4.2.1)
- Does the molecular gas in the most massive systems extend over CGM scales or is it primarily contained within galactic scales? (Section 4.2.1)

The acquisition of additional data through the proposed observations is expected to shed light on these questions and offer further insights into the complex nature of gas reservoirs surrounding  $z \sim 3$  quasars.

In this Section, I outline my future projects, which are based on the observing proposals I submitted (and got approved through competitive calls) as Principal Investigator (PI) during my PhD studies. These projects represent the next steps in my research and build upon the knowledge and experience gained from my previous work.

#### Exploring the most massive $z \sim 3$ quasar host galaxy in [CII]

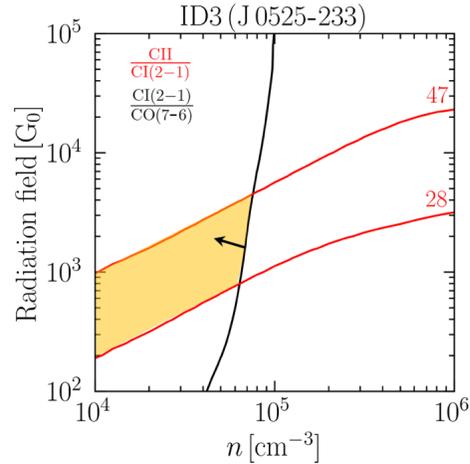
In order to further investigate the physical properties of the most massive system found in Muñoz-Elgueta et al. (2022), I submitted an observing proposal to get data from the APEX/nFLASH460<sup>1</sup> receiver, which covers the frequency window between 378 and 508 GHz. This proposal was successfully accepted and the data are currently available (15 hours

<sup>1</sup><https://www.apex-telescope.org/ns/nflash/>

of telescope time). The target of this proposal is J 0525-233 (or ID 3), which exhibits one of the brightest [CI](2-1) line and has the most massive molecular gas reservoir in our sample ( $2.3 \times 10^{11} \leq M_{\text{H}_2} \leq 6.8 \times 10^{11} M_{\odot}$ ; Section 2.4.3). If our predictions are correct, this dataset should detect at  $8\sigma$  the [CII] line emission from this system. This will be, at the best of our knowledge, the first  $3 < z < 4$  quasar with a constrain at their [CII] line frequency.

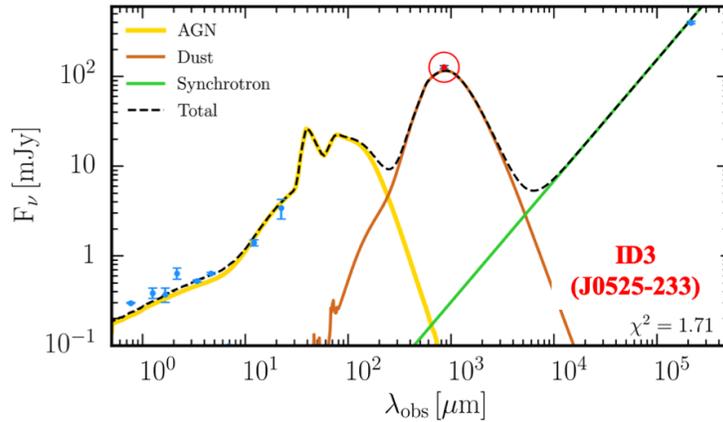
Considering the previous results suggesting that quasar’s host galaxy properties possibly affect the visible surrounding cool large-scale structure, with the acquired data we aim to further characterize the physical properties of our target. Several studies have detected [CII] in high- $z$  quasars, mostly at  $z > 4$ , reporting that PDR models can reproduce observed line ratios (e.g., Venemans et al. 2017; Hailey-Dunsheath et al. 2010; Maiolino et al. 2005; Iono et al. 2006). However, to our knowledge there are no works targeting this line in the  $3 < z < 4$  population. This lack of studies is due to the absence of instruments covering [CII] at these  $z$  ( $380.1 - 475.1$  GHz) in the past (only recently in ALMA Band 8 and APEX/nFLASH460), and due to the strong water atmosphere-absorption line at  $\sim 448$  GHz. This in turn will open up the possibility of future statistical studies of resolved molecular reservoirs (i.e. with ALMA) and their surrounding halo gas (i.e. VLT/MUSE).

**Figure 4.1:** PDR models of Meijerink et al. (2007), as a function of  $n$  and  $G$ , for ID 3. The red lines indicate the extrema of the predicted [CII]/[CI](2-1) luminosity ratios. The black continuous line shows the observed [CI](2-1)/CO(7-6) luminosity ratio. The arrow of the same color represents lower limit. The yellow area indicates the parameter space allowed by the current data.



Based on the finding that the [CII] emission in QSO hosts is consistent with models for PDRs (e.g., Kaufman et al. 2006; Meijerink et al. 2007), we use the PDR grids of models as a function of density  $n$  and radiation field  $G$  in Meijerink et al. (2007) (see Fig. 4.1), and constrain the expected luminosity of [CII] for our source given the known quantities (or upper limits) from previous observations:  $L_{\text{FIR}}$ ,  $L_{[\text{CI}](2-1)}$ ,  $L_{\text{CO}(7-6)}$ , and the current inferred molecular mass (or  $L_{\text{CO}(1-0)}$ ). The FIR luminosities ( $L_{\text{FIR}}$ ) were derived through a spectral energy distribution fit (Fig. 4.2) and strongly rely on LABOCA data at  $870\mu\text{m}$  (Project: M9519C\_105, PI: F. Arrigoni Battaia). In particular, the luminosity ratios  $L_{[\text{CII}]} / L_{\text{FIR}}$  and  $L_{[\text{CII}]} / L_{\text{CO}(1-0)}$  are sensitive to the star formation activity (e.g., Stacey et al. 1991), and have been commonly used in the literature to investigate the nature of the emission and some physical properties in high- $z$  galaxies (i.e., metallicities, gas densities, UV radiation field, e.g., Stacey et al. 2010; Hailey-Dunsheath et al. 2010; Ferkinhoff et al. 2014; Gullberg

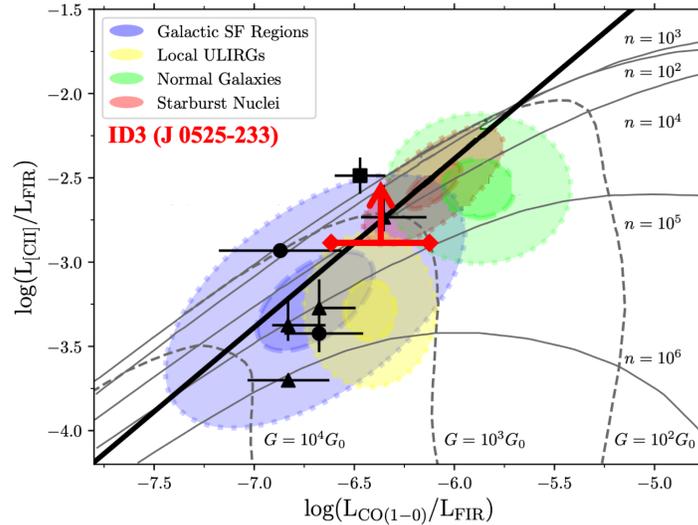
et al. 2015; Venemans et al. 2017). Figure 4.3 shows a diagnostic plot  $L_{[\text{CII}]} / L_{\text{FIR}}$  versus  $L_{\text{CO}(1-0)} / L_{\text{FIR}}$ , with the expected position of our targeted quasar (red symbol, the arrow indicates lower limit). The CO(1-0) luminosity  $L_{\text{CO}(1-0)}$  for our source was computed from the previously derived  $M_{\text{H}_2}$ . Regions occupied by different type of galaxies are indicated by ellipses. The solid black line indicates the typical ratio  $L_{[\text{CII}]} / L_{\text{CO}(1-0)} = 4100$ , which has been estimated in local starburst galaxies and star-forming regions in the Milky Way (e.g., Stacey et al. 1991) and in dusty star-forming galaxies at  $z > 2$  (e.g., Gullberg et al. 2015). In principle, J 0525-233 falls in a region shared between local ULIRGs, galactic star-forming regions, normal and starburst galaxies, and it is close to other two quasars taken from the literature (black symbols), located at  $z \sim 2.5$  and  $z > 5$ . These predictions, along with other estimations, will be validated or refuted through future analysis of these APEX/nFLASH460 data.



**Figure 4.2:** SEDs for J 0525-233 (ID 3). The LABOCA data recently obtained are key in constraining  $L_{\text{FIR}}$ . ID3 is a radio-loud quasars and presents very strong  $870 \mu\text{m}$  emission.

### Large reservoirs in $z \sim 3$ quasars with small and dim $\text{Ly}\alpha$ nebulae

To provide further confirmation of the association between the most massive molecular reservoirs and the dimmest and smallest  $\text{Ly}\alpha$  nebulae (Chapter 2), I submitted an observing proposal to get data from the APEX/SEPIA180 receiver. This proposal was successfully accepted and the data are currently available (75 hours of telescope time). The aim is to target the  $[\text{CII}](2-1)$  transition for all the five visible/feasible  $z \sim 3$  quasars known to have small and dim  $\text{Ly}\alpha$  nebulae (see Fig. 4.4). These five objects are the only currently published  $z \sim 3$  quasars visible from APEX and with a redshift that places their expected  $[\text{CII}](2-1)$  transition at a feasible frequency, far away from the low transmission band at 183 GHz (atmospheric water vapour). All these targets have been found in the QSO MUSEUM survey (Arrigoni Battaia et al., 2019a). This study will allow us to (i) constrain the molecular gas mass of the QSO hosts, (ii) confirm the link between small and dim  $\text{Ly}\alpha$  nebulae and large molecular reservoirs, and (iii) simultaneously constrain the CO(7-6)

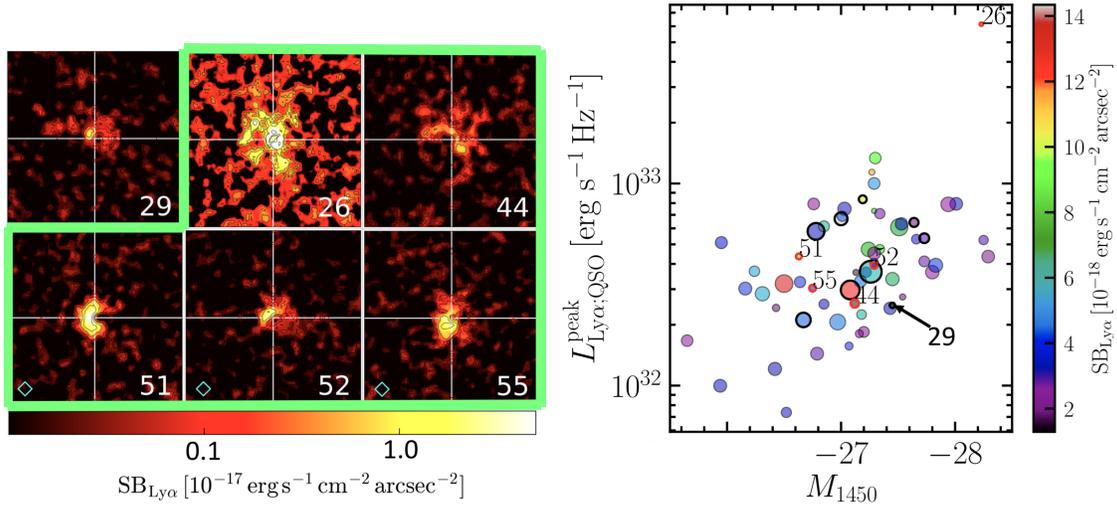


**Figure 4.3:**  $L_{\text{CII}}/L_{\text{FIR}}$  versus  $L_{\text{CO}(1-0)}/L_{\text{FIR}}$ , adapted from Venemans et al. (2017). Overlaid are the UV radiation fields (dashed grey lines) and gas densities contours (continuous grey lines) from Kaufman et al. (2006). The zones occupied by different objects are shown in coloured ellipses. A ratio of  $L_{\text{CII}}/L_{\text{CO}(1-0)} = 4100$ , found in local starburst galaxies and star-forming regions in the Milky Way (e.g., Stacey et al. 1991) and in dusty star-forming galaxies at  $z > 2$  (e.g., Gullberg et al. 2015), is indicated by the continuous black line. Quasars from the literature (Hailey-Dunsheath et al. 2010; Ferkinhoff et al. 2015; Uzgil et al. 2016; Venemans et al. 2017) at different redshifts are indicated by black symbols:  $z \sim 2.5$  (square),  $4 < z < 5$  (circles),  $z > 5$  (triangles). The targeted source J 0525-233 is represented by the red symbols.

transition, which is usually expected to be at the peak of the CO ladder for high- $z$  quasars (e.g., Carilli & Walter 2013). This work will act as benchmark for an intense campaign of follow-up observations (also for low- $J$  CO transitions) with the aim of extending the study of the interplay between the molecular reservoir and the large-scale cool gas to a statistical sample of quasars with known  $\text{Ly}\alpha$  nebulae (i.e., with ALMA, NOEMA).

### Uncovering the extended molecular gas reservoir around two promising $z \sim 3$ quasars

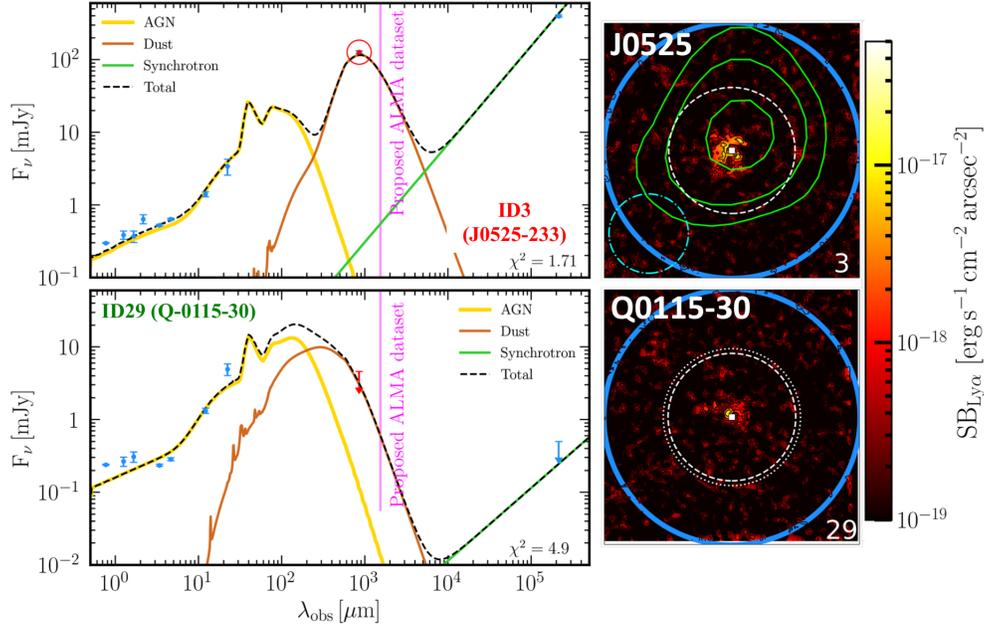
I proposed deep ALMA band 5 observations of the  $[\text{CI}](2-1)$  line for the two most massive objects (in molecular mass) found in (Muñoz-Elgueta et al., 2022) (ID 3 and ID 29, see Chapter 2), which also present the strongest  $[\text{CI}](2-1)$  emission. The molecular gas masses of these sources have been constrained inside a large beam size of  $\sim 30''$  ( $\sim 230$  kpc). The aim is to assess whether these large molecular gas reservoirs extend out to the inner CGM scales ( $\sim 30$  kpc scales). This proposal was accepted (6.3 hours) with a rank C designation, but unfortunately, the allocated telescope time was not used to observe the data. Currently, I have resubmitted this proposal which is still under review process. One of the proposed targets is ID 29, which is an extremely interesting source because presents a bright double-peaked  $[\text{CI}](2-1)$  emission (see Fig. 2.5, left panel), which could be due to a



**Figure 4.4:** Left:  $30'' \times 30''$  Ly $\alpha$  SB maps of the five targets proposed (green) compared to the one of ID 29 (adapted from Arrigoni Battaia et al. 2019a). Right: Quasars' peak specific luminosity at the Ly $\alpha$  line versus  $M_{1450}$  for the QSO MUSEUM sample. The markers sizes are proportional to the area of the Ly $\alpha$  nebulae, while the colors indicate the SB. The targets of the pilot SEPIA180 sample are indicated by black circles, with ID 29 clearly marked. The targets of this proposal are highlighted with their IDs and with red circles.

massive molecular disk or a merger of two massive objects. On the other hand, ID 3 is a radio-loud quasar and presents an impressive 127 mJy 870  $\mu\text{m}$  emission (see Fig. 4.5, left panel, circled data point). This remarkable flux density is comparable to that of protocluster cores (e.g., Miller et al. 2018), and would translate to a dust mass of  $8 \pm 4 \times 10^{10} M_{\odot}$  and a star-formation rate of  $1825 \pm 400 M_{\odot} \text{ yr}^{-1}$ , if coming from a single source (SED fit in Fig. 4.5). The proposed dataset will allow us to:

- (i) Determine the extent and geometry of the molecular reservoirs around the two proposed targets. By contrasting the [CI](2-1) emission from ALMA to that detected by the APEX pilot programme, we can assess if ALMA recovers less flux than APEX, resolving out large-scale emission ( $>30$  kpc). This would be a clear indication of molecular gas on halo scales. On the contrary, if all the molecular mass is on small scales, our targets would represent the most massive objects of the  $z \sim 3$  quasar population (see Fig. 2.10).
- (ii) Refine the molecular gas masses estimates, and study the gas kinematics out to halo scales. Specifically, from [CI](2-1) it is possible to compute the CI mass, and then derive the respective  $M_{\text{H}_2}$  (e.g., Weiß et al. 2003; Venemans et al. 2017).
- (iii) Unveil the nature of the double-peaked [CI](2-1) emission in Q-0115-30, constituted of two Gaussian components separated by  $\sim 800 \text{ km s}^{-1}$  (Fig. 2.5). By mapping the [CI](2-1) emission we will be able to ascertain whether ID 29 is associated with an exceptionally massive molecular disk or is merging with a similarly massive companion (e.g., Polletta et al. 2011).
- (iv) Detect at very high significance the underlying continuum at  $\sim 1.5 \text{ mm}$  allowing us to refine the SEDs (therefore dust masses) of the sources presented in Fig. 4.5 (left panel)



**Figure 4.5:** Left: SEDs for the two targets of this proposal. ID 3 is a radio-loud quasars and presents very strong  $870 \mu\text{m}$  emission. ID 29 presents a  $2\sigma$  upper limit at  $870 \mu\text{m}$ . The magenta line at  $1.5\text{mm}$  is drawn down to the  $5\sigma$  flux limit for the proposed observations ( $\sim 60 \mu\text{Jy}$ ). Right:  $60'' \times 60''$  MUSE Ly $\alpha$  SB maps of the nebulae surrounding the 2 proposed targets (from Arrigoni Battaia et al. 2019a). We indicate the beam sizes of APEX/SEPIA180 (white circles), the primary beam of the proposed ALMA observations (blue), and the APEX/LABOCA  $870 \mu\text{m}$  detection for ID3 (with S/N=5, 10, 20 contours; green) with the respective beam (cyan).

and assess the presence of both extended continuum emission or multiple components (e.g., Miller et al. 2018).

#### 4.2.2 Extension of [CII] predictions

As outlined in Section 4.1.2, the results from my second project revealed a discrepancy between the predicted [CII] emission from star-forming galaxies at large scales ( $\sim r > 10 \text{ kpc}$ ) and observational data. This tension highlights the necessity of employing higher resolution cosmological simulations capable of accurately resolving the ISM and CGM in high- $z$  galaxies. While zoom-in simulations have been attempted to address this issue, they still get discrepancies with the observations (e.g., Pallottini et al. 2017; Arata et al. 2019; Schimek et al. 2023). Therefore, an alternative approach could be to use more detailed subgrid prescriptions and applying them to cosmological simulations such as IllustrisTNG, that has the advantage of providing a large universe with a statistically significant sample of high- $z$  galaxies.

Importantly, it would be fundamental to apply similar subgrid prescriptions to different

cosmological simulations in order to assess the extent to which different prescriptions for AGN and SFR affect the [CII] emission on CGM scales. By comparing the results of different simulations with consistent subgrid models, we can get more insights about the real impact of these physical processes on the [CII] emission and its distribution over the CGM. Matching the observations from [Ginolfi et al. \(2020b\)](#) would likely require the presence of cold, small-scale clouds entrained in the outflows of high- $z$  galaxies.

In this context, a future continuation of this work would be to incorporate subgrid prescriptions based on radiative transfer calculations performed on high-resolution turbulent boxes. These calculations have been started with an effort led by Stefanie Walch and Lennart Buhmann at the University of Cologne, in which I have been involved. Additionally, these subgrid models can be refined by calculating [CII] emission from the SILCC simulations ([Walch et al., 2015](#)), which account for models of the small-scale structure of the multiphase ISM ([Rathjen et al., 2021](#)), possibly able to resolve the clumpiness of the CGM. These refinements have the potential to improve our modelling and possibly reduce the discrepancies with observations, particularly regarding the extended [CII] emission observed in high- $z$  galaxies.

Once the modeling of [CII] emission around star-forming galaxies would be acceptable, an additional problem to be tackled is how the extended [CII] emission level change in the presence of a bright quasar. In contrast to star-forming galaxies at  $4 < z < 6$ , the presence of [CII] halos in quasars within the same redshift range has not been clearly established thus far. Several works targeted the [CII] emission around  $z \sim 6$  quasars reporting contradictory results, with the latest work suggesting the [CII] emission is not extended and is not tracing any outflow (i.e., there is no presence of a broad line component in the stacked spectrum; [Novak et al. 2020](#)). In this context, it would be interesting to investigate the predicted [CII] emission in quasars host galaxies selected from cosmological simulations and compare the results with observations (e.g., [Gallerani et al. 2012](#); [Decarli et al. 2018](#); [Venemans et al. 2020](#)). However, this analysis is currently not feasible due to certain limitations. In general, most of the observational studies have targeted the [CII] line high- $z$  in bright type I quasars, with  $L_{\text{BOL}} > 10^{46} \text{ erg s}^{-1}$ , residing in massive halos of  $10^{12}$ - $10^{13} M_{\odot}$ . Such types of halos are not found in cosmological boxes due to their rarity (typical numerical density of  $\lesssim 10^{-6.2} \text{ cMpc}^{-3}$ , [Shen et al. 2020](#)). For example, any of the IllustrisTNG volumes are too small. While TNG300 does contain 75 massive halos with masses  $> 10^{12} M_{\odot}$ , their AGNs have much lower luminosities ( $L_{\text{BOL}} < 10^{46} \text{ erg s}^{-1}$ ) in comparison to observed type I quasars. Nevertheless, recent JWST observations have started to unveil obscured quasars at high- $z$  with lower masses and luminosities (e.g., [Kocevski et al. 2023](#); [Barro et al. 2023](#)). These results are promising and suggest that future JWST observations will likely identify obscured quasars at  $4 < z < 6$ . Subsequent follow-up observations with ALMA for [CII] can be performed, allowing for a comparison with IllustrisTNG halos and, consequently, a future [CII] modeling.

Another future follow-up project would be to study the [CII] predictions for a sample of quasar host halos at  $z \sim 3$ . The IllustrisTNG simulations indeed have some halos available for such a study (e.g., 2 (4) halos with  $M_{\text{halo}} > 10^{12} M_{\odot}$  and AGNs with  $L_{\text{BOL}} > 10^{46} \text{ erg s}^{-1}$  in TNG50 (TNG100)). However, currently there are no [CII] observations of  $z \sim 3$  quasars

for direct comparison. By using the in-hand APEX data (see Section 4.2.1), we will be able to get first observational constraints on [CII] in  $z \sim 3$  quasars, which can therefore be used to compare with the TNG simulations post-processed with our future models. These predictions will be very useful to test the physics implemented in the simulations (e.g., AGN feedback) and will help obtaining further observational data aiming at [CII] emission in and around  $z \sim 3$  quasars.

# Bibliography

- Akins, H. B., Fujimoto, S., Finlator, K., et al. (2022), *ALMA Reveals Extended Cool Gas and Hot Ionized Outflows in a Typical Star-forming Galaxy at  $Z = 7.13$* , ApJ, 934(1), 64
- Alaghband-Zadeh, S., Chapman, S. C., Swinbank, A. M., et al. (2013), *Using [C I] to probe the interstellar medium in  $z \sim 2.5$  sub-millimeter galaxies*, MNRAS, 435(2), 1493
- Anderson, M. E., Churazov, E., and Bregman, J. N. (2016), *A deep XMM-Newton study of the hot gaseous halo around NGC 1961*, MNRAS, 455(1), 227
- Anglés-Alcázar, D., Faucher-Giguère, C.-A., Kereš, D., et al. (2017), *The cosmic baryon cycle and galaxy mass assembly in the FIRE simulations*, MNRAS, 470(4), 4698
- Anh, P. T., Boone, F., Hoai, D. T., et al. (2013), *Resolving the molecular gas around the lensed quasar RXJ0911.4+0551*, A&A, 552, L12
- Antonucci, R. (1993), *Unified models for active galactic nuclei and quasars.*, ARA&A, 31, 473
- Arata, S., Yajima, H., Nagamine, K., et al. (2019), *Radiative properties of the first galaxies: rapid transition between UV and infrared bright phases*, MNRAS, 488(2), 2629
- Aravena, M., Bertoldi, F., Schinnerer, E., et al. (2008), *Properties of the molecular gas in a starbursting QSO at  $z = 1.83$  in the COSMOS field*, A&A, 491(1), 173
- Arrigoni Battaia, F., Chen, C.-C., Fumagalli, M., et al. (2018a), *Overdensity of submillimeter galaxies around the  $z \sim 2.3$  MAMMOTH-1 nebula. The environment and powering of an enormous Lyman- $\alpha$  nebula*, A&A, 620, A202
- Arrigoni Battaia, F., Hennawi, J. F., Prochaska, J. X., and Cantalupo, S. (2015), *Deep He II and C IV Spectroscopy of a Giant Ly $\alpha$  Nebula: Dense Compact Gas Clumps in the Circumgalactic Medium of a  $z \sim 2$  Quasar*, ApJ, 809(2), 163
- Arrigoni Battaia, F., Hennawi, J. F., Prochaska, J. X., et al. (2019a), *QSO MUSEUM I: a sample of 61 extended Ly  $\alpha$ -emission nebulae surrounding  $z \sim 3$  quasars*, MNRAS, 482(3), 3162

- Arrigoni Battaia, F., Obreja, A., Prochaska, J. X., et al. (2019b), *Discovery of intergalactic bridges connecting two faint  $z \sim 3$  quasars*, A&A, 631, A18
- Arrigoni Battaia, F., Prochaska, J. X., Hennawi, J. F., et al. (2018b), *Inspiraling halo accretion mapped in Ly  $\alpha$  emission around a  $z \sim 3$  quasar*, MNRAS, 473(3), 3907
- Augustin, R., Quiret, S., Milliard, B., et al. (2019), *Emission from the circumgalactic medium: from cosmological zoom-in simulations to multiwavelength observables*, MNRAS, 489(2), 2417
- Bañados, E., Venemans, B. P., Mazzucchelli, C., et al. (2018), *An 800-million-solar-mass black hole in a significantly neutral Universe at a redshift of 7.5*, Nature, 553(7689), 473
- Bacon, R., Accardo, M., Adjali, L., et al. (2010), *The MUSE second-generation VLT instrument*, in *Ground-based and Airborne Instrumentation for Astronomy III*, edited by I. S. McLean, S. K. Ramsay, H. Takami, volume 7735 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 773508
- Baker, J. G. and Menzel, D. H. (1938), *Physical Processes in Gaseous Nebulae. III. The Balmer Decrement.*, ApJ, 88, 52
- Banerji, M., Carilli, C. L., Jones, G., et al. (2017), *The discovery of gas-rich, dusty starbursts in luminous reddened quasars at  $z \sim 2.5$  with ALMA*, MNRAS, 465(4), 4390
- Banerji, M., Jones, G. C., Wagg, J., et al. (2018), *The interstellar medium properties of heavily reddened quasars and companions at  $z \sim 2.5$  with ALMA and JVL A*, MNRAS, 479(1), 1154
- Barro, G., Perez-Gonzalez, P. G., Kocevski, D. D., et al. (2023), *Extremely red galaxies at  $z = 5 - 9$  with MIRI and NIRSpec: dusty galaxies or obscured AGNs?*, arXiv e-prints, arXiv:2305.14418
- Barvainis, R., Alloin, D., and Bremer, M. (2002), *A CO survey of gravitationally lensed quasars with the IRAM interferometer*, A&A, 385, 399
- Beelen, A., Cox, P., Pety, J., et al. (2004), *Starburst activity in the host galaxy of the  $z = 2.58$  quasar J1409+5628*, A&A, 423, 441
- Belitsky, V., Bylund, M., Desmaris, V., et al. (2018a), *ALMA Band 5 receiver cartridge. Design, performance, and commissioning*, A&A, 611, A98
- Belitsky, V., Lapkin, I., Fredrixon, M., et al. (2018b), *SEPIA - a new single pixel receiver at the APEX telescope*, A&A, 612, A23
- Bergeron, J., Aracil, B., Petitjean, P., and Pichon, C. (2002), *The warm-hot intergalactic medium at  $z \sim 2.2$ : Metal enrichment and ionization source*, A&A, 396, L11

- Bertone, S. and Schaye, J. (2012), *Rest-frame ultraviolet line emission from the intergalactic medium at  $2 \leq z \leq 5$* , MNRAS, 419(1), 780
- B  thermin, M., Fudamoto, Y., Ginolfi, M., et al. (2020), *The ALPINE-ALMA [CII] survey: Data processing, catalogs, and statistical source properties*, A&A, 643, A2
- Binney, J. (1977), *The physics of dissipational galaxy formation.*, ApJ, 215, 483
- Binney, J. and Tabor, G. (1995), *Evolving cooling flows.*, MNRAS, 276, 663
- Birnboim, Y. and Dekel, A. (2003), *Virial shocks in galactic haloes?*, MNRAS, 345(1), 349
- Bisbas, T. G., Tan, J. C., and Tanaka, K. E. I. (2021), *Photodissociation region diagnostics across galactic environments*, MNRAS, 502(2), 2701
- Bisbas, T. G., van Dishoeck, E. F., Papadopoulos, P. P., et al. (2017), *Cosmic-ray Induced Destruction of CO in Star-forming Galaxies*, ApJ, 839(2), 90
- Bisbas, T. G., Walch, S., Naab, T., et al. (2022), *The Origin of the [C II] Deficit in a Simulated Dwarf Galaxy Merger-driven Starburst*, ApJ, 934(2), 115
- Bischetti, M., Feruglio, C., Piconcelli, E., et al. (2021), *The WISSH quasars project. IX. Cold gas content and environment of luminous QSOs at  $z \sim 2.4-4.7$* , A&A, 645, A33
- Blitz, L. and Rosolowsky, E. (2004), *The Role of Pressure in Giant Molecular Cloud Formation*, ApJ, 612(1), L29
- Blitz, L. and Rosolowsky, E. (2006), *The Role of Pressure in GMC Formation II: The H<sub>2</sub>-Pressure Relation*, ApJ, 650(2), 933
- Bolatto, A. D., Wolfire, M., and Leroy, A. K. (2013), *The CO-to-H<sub>2</sub> Conversion Factor*, ARA&A, 51(1), 207
- Bond, J. R., Kofman, L., and Pogosyan, D. (1996), *How filaments of galaxies are woven into the cosmic web*, Nature, 380(6575), 603
- Boogaard, L. A., van der Werf, P., Weiss, A., et al. (2020), *The ALMA Spectroscopic Survey in the Hubble Ultra Deep Field: CO Excitation and Atomic Carbon in Star-forming Galaxies at  $z = 1-3$* , ApJ, 902(2), 109
- Borisova, E., Cantalupo, S., Lilly, S. J., et al. (2016), *Ubiquitous Giant Ly $\alpha$  Nebulae around the Brightest Quasars at  $z \sim 3.5$  Revealed with MUSE*, ApJ, 831(1), 39
- Boselli, A., Gavazzi, G., Lequeux, J., and Pierini, D. (2002), *[CII] at 158  $\mu$  m as a star formation tracer in late-type galaxies*, A&A, 385, 454
- Bouwens, R. J., Smit, R., Schouws, S., et al. (2022), *Reionization Era Bright Emission Line Survey: Selection and Characterization of Luminous Interstellar Medium Reservoirs in the  $z \lesssim 6.5$  Universe*, ApJ, 931(2), 160

- Bower, R. G., Benson, A. J., Malbon, R., et al. (2006), *Breaking the hierarchy of galaxy formation*, MNRAS, 370(2), 645
- Bradford, C. M., Aguirre, J. E., Aikin, R., et al. (2009), *The Warm Molecular Gas around the Cloverleaf Quasar*, ApJ, 705(1), 112
- Brauher, J. R., Dale, D. A., and Helou, G. (2008), *A Compendium of Far-Infrared Line and Continuum Emission for 227 Galaxies Observed by the Infrared Space Observatory*, ApJS, 178(2), 280
- Bryan, G. L. and Norman, M. L. (1998), *Statistical Properties of X-Ray Clusters: Analytic and Numerical Comparisons*, ApJ, 495(1), 80
- Bunker, A., Smith, J., Spinrad, H., et al. (2003), *Illuminating protogalaxies? The discovery of extended Lyman- $\alpha$  emission around a QSO at  $z=4.5$* , Ap&SS, 284(2), 357
- Burchett, J. N., Tripp, T. M., Wang, Q. D., et al. (2018), *Warm-hot gas in X-ray bright galaxy clusters and the H I-deficient circumgalactic medium in dense environments*, MNRAS, 475(2), 2067
- Byrohl, C., Nelson, D., Behrens, C., et al. (2021), *The physical origins and dominant emission mechanisms of Lyman alpha haloes: results from the TNG50 simulation in comparison to MUSE observations*, MNRAS, 506(4), 5129
- Cai, Z., Cantalupo, S., Prochaska, J. X., et al. (2019), *Evolution of the Cool Gas in the Circumgalactic Medium of Massive Halos: A Keck Cosmic Web Imager Survey of Ly $\alpha$  Emission around QSOs at  $z \approx 2$* , ApJS, 245(2), 23
- Cai, Z., Fan, X., Yang, Y., et al. (2017), *Discovery of an Enormous Ly $\alpha$  Nebula in a Massive Galaxy Overdensity at  $z = 2.3$* , ApJ, 837(1), 71
- Cantalupo, S., Arrigoni-Battaia, F., Prochaska, J. X., et al. (2014), *A cosmic web filament revealed in Lyman- $\alpha$  emission around a luminous high-redshift quasar*, Nature, 506(7486), 63
- Carilli, C. L., Lewis, G. F., Djorgovski, S. G., et al. (2003), *A Molecular Einstein Ring: Imaging a Starburst Disk Surrounding a Quasi-Stellar Object*, Science, 300(5620), 773
- Carilli, C. L. and Walter, F. (2013), *Cool Gas in High-Redshift Galaxies*, ARA&A, 51(1), 105
- Carniani, S., Ferrara, A., Maiolino, R., et al. (2020), *Missing [C II] emission from early galaxies*, MNRAS, 499(4), 5136
- Casey, C. M., Narayanan, D., Carilli, C., et al. (2018), *Imaging Cold Gas to 1 kpc Scales in High-Redshift Galaxies with the ngVLA*, in *Science with a Next Generation Very Large Array*, edited by E. Murphy, volume 517 of *Astronomical Society of the Pacific Conference Series*, 629

- Cattaneo, A., Faber, S. M., Binney, J., et al. (2009), *The role of black holes in galaxy formation and evolution*, Nature, 460(7252), 213
- Cen, R. and Zheng, Z. (2013), *The Nature of Ly $\alpha$  Blobs: Powered by Extreme Starbursts*, ApJ, 775(2), 112
- Cerviño, M., Mas-Hesse, J. M., and Kunth, D. (2002), *Evolutionary synthesis models of starbursts. IV. Soft X-ray emission*, A&A, 392, 19
- Chabrier, G. (2003), *Galactic Stellar and Substellar Initial Mass Function*, PASP, 115(809), 763
- Chakravarti, I., Roy, J., and Laha, R. (1967), *Handbook of Methods of Applied Statistics*, number Bd. 2 in Wiley series in probability and mathematical statistics, Wiley
- Cicone, C., Maiolino, R., Gallerani, S., et al. (2015), *Very extended cold gas, star formation and outflows in the halo of a bright quasar at  $z \approx 6$* , A&A, 574, A14
- Cicone, C., Maiolino, R., Sturm, E., et al. (2014), *Massive molecular outflows and evidence for AGN feedback from CO observations*, A&A, 562, A21
- Ciotti, L. and Ostriker, J. P. (2001), *Cooling Flows and Quasars. II. Detailed Models of Feedback-modulated Accretion Flows*, ApJ, 551(1), 131
- Circosta, C., Mainieri, V., Lamperti, I., et al. (2021), *SUPER. IV. CO( $J = 3-2$ ) properties of active galactic nucleus hosts at cosmic noon revealed by ALMA*, A&A, 646, A96
- Cole, S., Percival, W. J., Peacock, J. A., et al. (2005), *The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications*, MNRAS, 362(2), 505
- Contursi, A., Baker, A. J., Berta, S., et al. (2017), *Interstellar medium conditions in  $z \approx 0.2$  Lyman-break analogs*, A&A, 606, A86
- Corlies, L., Peeples, M. S., Tumlinson, J., et al. (2020), *Figuring Out Gas & Galaxies in Enzo (FOGGIE). II. Emission from the  $z = 3$  Circumgalactic Medium*, ApJ, 896(2), 125
- Costa, T., Arrigoni Battaia, F., Farina, E. P., et al. (2022), *AGN-driven outflows and the formation of Ly $\alpha$  nebulae around high- $z$  quasars*, MNRAS, 517(2), 1767
- Cowie, L. L. and McKee, C. F. (1977), *The evaporation of spherical clouds in a hot gas. I. Classical and saturated mass loss rates.*, ApJ, 211, 135
- Crain, R. A., McCarthy, I. G., Frenk, C. S., et al. (2010), *X-ray coronae in simulations of disc galaxy formation*, MNRAS, 407(3), 1403
- Crain, R. A., Schaye, J., Bower, R. G., et al. (2015), *The EAGLE simulations of galaxy formation: calibration of subgrid physics and model variations*, MNRAS, 450(2), 1937

- Cresci, G., Marconi, A., Zibetti, S., et al. (2015), *The MAGNUM survey: positive feedback in the nuclear region of NGC 5643 suggested by MUSE*, A&A, 582, A63
- Croton, D. J., Springel, V., White, S. D. M., et al. (2006), *The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies*, MNRAS, 365(1), 11
- Croxall, K. V., Smith, J. D., Pellegrini, E., et al. (2017), *The Origins of [C II] Emission in Local Star-forming Galaxies*, ApJ, 845(2), 96
- Dalgarno, A. and McCray, R. A. (1972), *Heating and Ionization of HI Regions*, ARA&A, 10, 375
- Das, S., Mathur, S., Gupta, A., and Krongold, Y. (2021), *The Hot Circumgalactic Medium of the Milky Way: Evidence for Supervirial, Virial, and Subvirial Temperatures; Non-solar Chemical Composition; and Nonthermal Line Broadening*, ApJ, 918(2), 83
- De Looze, I., Cormier, D., Lebouteiller, V., et al. (2014), *The applicability of far-infrared fine-structure lines as star formation rate tracers over wide ranges of metallicities and galaxy types*, A&A, 568, A62
- Decarli, R., Aravena, M., Boogaard, L., et al. (2020), *The ALMA Spectroscopic Survey in the Hubble Ultra Deep Field: Multiband Constraints on Line-luminosity Functions and the Cosmic Density of Molecular Gas*, ApJ, 902(2), 110
- Decarli, R., Arrigoni-Battaia, F., Hennawi, J. F., et al. (2021), *A search for dust and molecular gas in enormous Ly $\alpha$  nebulae at  $z \approx 2$* , A&A, 645, L3
- Decarli, R., Walter, F., González-López, J., et al. (2019), *The ALMA Spectroscopic Survey in the HUDF: CO Luminosity Functions and the Molecular Gas Content of Galaxies through Cosmic History*, ApJ, 882(2), 138
- Decarli, R., Walter, F., Venemans, B. P., et al. (2018), *An ALMA [C II] Survey of 27 Quasars at  $z \lesssim 5.94$* , ApJ, 854(2), 97
- Dekel, A. and Birnboim, Y. (2006), *Galaxy bimodality due to cold flows and shock heating*, MNRAS, 368(1), 2
- Dekel, A., Sari, R., and Ceverino, D. (2009), *Formation of Massive Galaxies at High Redshift: Cold Streams, Clumpy Disks, and Compact Spheroids*, ApJ, 703(1), 785
- Di Matteo, T., Springel, V., and Hernquist, L. (2005), *Energy input from quasars regulates the growth and activity of black holes and their host galaxies*, Nature, 433(7026), 604
- Díaz-Santos, T., Armus, L., Charmandaris, V., et al. (2017), *A Herschel/PACS Far-infrared Line Emission Survey of Local Luminous Infrared Galaxies*, ApJ, 846(1), 32

- Díaz-Santos, T., Armus, L., Charmandaris, V., et al. (2014), *Extended [C II] Emission in Local Luminous Infrared Galaxies*, ApJ, 788(1), L17
- Diemer, B., Stevens, A. R. H., Forbes, J. C., et al. (2018), *Modeling the Atomic-to-molecular Transition in Cosmological Simulations of Galaxy Formation*, ApJS, 238(2), 33
- Dijkstra, M. (2017), *Saas-Fee Lecture Notes: Physics of Lyman Alpha Radiative Transfer*, arXiv e-prints, arXiv:1704.03416
- Dijkstra, M., Haiman, Z., and Spaans, M. (2006), *Ly $\alpha$  Radiation from Collapsing Protogalaxies. I. Characteristics of the Emergent Spectrum*, ApJ, 649(1), 14
- Dijkstra, M. and Loeb, A. (2008), *Ly $\alpha$ -driven outflows around star-forming galaxies*, MNRAS, 391(1), 457
- Donnari, M., Pillepich, A., Nelson, D., et al. (2019), *The star formation activity of IllustrisTNG galaxies: main sequence, UVJ diagram, quenched fractions, and systematics*, MNRAS, 485(4), 4817
- Downes, D. and Solomon, P. M. (1998), *Rotating Nuclear Rings and Extreme Starbursts in Ultraluminous Galaxies*, ApJ, 507(2), 615
- Draine, B. (2011), *Physics of the Interstellar and Intergalactic Medium*, Princeton Series in Astrophysics, Princeton University Press
- Drake, A. B., Farina, E. P., Neeleman, M., et al. (2019), *Ly $\alpha$  Halos around  $z \sim 6$  Quasars*, ApJ, 881(2), 131
- Drake, A. B., Neeleman, M., Venemans, B. P., et al. (2022), *The Decoupled Kinematics of High- $z$  QSO Host Galaxies and Their Ly $\alpha$  Halos*, ApJ, 929(1), 86
- Dutta, R., Fumagalli, M., Fossati, M., et al. (2021), *Metal-enriched halo gas across galaxy overdensities over the last 10 billion years*, MNRAS, 508(3), 4573
- Dutta, R., Fumagalli, M., Fossati, M., et al. (2020), *MUSE Analysis of Gas around Galaxies (MAGG) - II: metal-enriched halo gas around  $z \sim 1$  galaxies*, MNRAS, 499(4), 5022
- Dutton, A. A. and Macciò, A. V. (2014), *Cold dark matter haloes in the Planck era: evolution of structural parameters for Einasto and NFW profiles*, MNRAS, 441(4), 3359
- Elvis, M. (2000), *A Structure for Quasars*, ApJ, 545(1), 63
- Emonts, B. H. C., Cai, Z., Prochaska, J. X., et al. (2019), *The Cold Circumgalactic Environment of MAMMOTH-I: Dynamically Cold Gas in the Core of an Enormous Ly $\alpha$  Nebula*, ApJ, 887(1), 86

- Emonts, B. H. C., Lehnert, M. D., Dannerbauer, H., et al. (2018), *Giant galaxy growing from recycled gas: ALMA maps the circumgalactic molecular medium of the Spiderweb in [C I]*, MNRAS, 477(1), L60
- Emonts, B. H. C., Lehnert, M. D., Villar-Martín, M., et al. (2016), *Molecular gas in the halo fuels the growth of a massive cluster galaxy at high redshift*, Science, 354(6316), 1128
- Emonts, B. H. C., Lehnert, M. D., Yoon, I., et al. (2023), *A cosmic stream of atomic carbon gas connected to a massive radio galaxy at redshift 3.8*, Science, 379(6639), 1323
- Emonts, B. H. C., Norris, R. P., Feain, I., et al. (2014), *CO(1-0) survey of high- $z$  radio galaxies: alignment of molecular halo gas with distant radio sources*, MNRAS, 438(4), 2898
- Faisst, A. L., Schaerer, D., Lemaux, B. C., et al. (2020), *The ALPINE-ALMA [C II] Survey: Multiwavelength Ancillary Data and Basic Physical Measurements*, ApJS, 247(2), 61
- Falgarone, E., Zwaan, M. A., Godard, B., et al. (2017), *Large turbulent reservoirs of cold molecular gas around high-redshift starburst galaxies*, Nature, 548(7668), 430
- Farber, R. J. and Gronke, M. (2022), *Molecular Shattering*, arXiv e-prints, arXiv:2209.13622
- Farina, E. P., Arrigoni-Battaia, F., Costa, T., et al. (2019), *The REQUIEM Survey. I. A Search for Extended Ly $\alpha$  Nebular Emission Around 31  $z \gtrsim 5.7$  Quasars*, ApJ, 887(2), 196
- Farina, E. P., Falomo, R., Decarli, R., et al. (2013), *On the cool gaseous haloes of quasars*, MNRAS, 429(2), 1267
- Farina, E. P., Falomo, R., Scarpa, R., et al. (2014), *The extent of the Mg II absorbing circumgalactic medium of quasars*, MNRAS, 441(1), 886
- Faucher-Giguère, C.-A., Feldmann, R., Quataert, E., et al. (2016), *A stellar feedback origin for neutral hydrogen in high-redshift quasar-mass haloes*, MNRAS, 461(1), L32
- Faucher-Giguère, C.-A., Kereš, D., Dijkstra, M., et al. (2010), *Ly $\alpha$  Cooling Emission from Galaxy Formation*, ApJ, 725(1), 633
- Faucher-Giguère, C.-A., Kereš, D., and Ma, C.-P. (2011), *The baryonic assembly of dark matter haloes*, MNRAS, 417(4), 2982
- Faucher-Giguere, C.-A. and Oh, S. P. (2023), *Key Physical Processes in the Circumgalactic Medium*, arXiv e-prints, arXiv:2301.10253
- Ferkinhoff, C., Brisbin, D., Nikola, T., et al. (2015), *Band-9 ALMA Observations of the [N II] 122  $\mu$ m Line and FIR Continuum in Two High- $z$  galaxies.*, ApJ, 806(2), 260

- Ferkinhoff, C., Brisbin, D., Parshley, S., et al. (2014), *The Second-generation  $z$  (Redshift) and Early Universe Spectrometer. I. First-light Observation of a Highly Lensed Local- $ulirg$  Analog at High- $z$* , ApJ, 780(2), 142
- Ferland, G. J., Chatzikos, M., Guzmán, F., et al. (2017), *The 2017 Release Cloudy*, Rev. Mex. Astron. Astrofis., 53, 385
- Ferrarese, L. and Merritt, D. (2000), *A Fundamental Relation between Supermassive Black Holes and Their Host Galaxies*, ApJ, 539(1), L9
- Fielding, D., Quataert, E., McCourt, M., and Thompson, T. A. (2017), *The impact of star formation feedback on the circumgalactic medium*, MNRAS, 466(4), 3810
- Fossati, M., Fumagalli, M., Lofthouse, E. K., et al. (2021), *MUSE analysis of gas around galaxies (MAGG) - III. The gas and galaxy environment of  $z = 3-4.5$  quasars*, MNRAS, 503(2), 3044
- Franeck, A., Walch, S., Seifried, D., et al. (2018), *Synthetic [C II] emission maps of a simulated molecular cloud in formation*, MNRAS, 481(4), 4277
- Fudamoto, Y., Smit, R., Bowler, R. A. A., et al. (2022), *The ALMA REBELS Survey: Average [C II] 158  $\mu\text{m}$  Sizes of Star-forming Galaxies from  $z = 7$  to  $z = 4$* , ApJ, 934(2), 144
- Fujimoto, S., Ouchi, M., Ferrara, A., et al. (2019), *First Identification of 10 kpc [C II] 158  $\mu\text{m}$  Halos around Star-forming Galaxies at  $z = 5-7$* , ApJ, 887(2), 107
- Fujimoto, S., Silverman, J. D., Bethermin, M., et al. (2020), *The ALPINE-ALMA [C II] Survey: Size of Individual Star-forming Galaxies at  $z = 4-6$  and Their Extended Halo Structure*, ApJ, 900(1), 1
- Fumagalli, M., Cantalupo, S., Dekel, A., et al. (2016), *MUSE searches for galaxies near very metal-poor gas clouds at  $z \sim 3$ : new constraints for cold accretion models*, MNRAS, 462(2), 1978
- Furlong, M., Bower, R. G., Theuns, T., et al. (2015), *Evolution of galaxy stellar masses and star formation rates in the EAGLE simulations*, MNRAS, 450(4), 4486
- Gallerani, S., Neri, R., Maiolino, R., et al. (2012), *Resolved [CII] emission in a lensed quasar at  $z = 4.4$* , A&A, 543, A114
- García-Vergara, C., Hennawi, J. F., Barrientos, L. F., and Arrigoni Battaia, F. (2019), *Clustering of Ly $\alpha$  Emitters around Quasars at  $z \sim 4$* , ApJ, 886(2), 79
- García-Vergara, C., Hennawi, J. F., Barrientos, L. F., and Rix, H.-W. (2017), *Strong Clustering of Lyman Break Galaxies around Luminous Quasars at  $Z \sim 4$* , ApJ, 848(1), 7

- Geach, J. E., Alexander, D. M., Lehmer, B. D., et al. (2009), *The Chandra Deep Proto-cluster Survey: Ly $\alpha$  Blobs are Powered by Heating, Not Cooling*, ApJ, 700(1), 1
- Genzel, R., Tacconi, L. J., Lutz, D., et al. (2015), *Combined CO and Dust Scaling Relations of Depletion Time and Molecular Gas Fractions with Cosmic Time, Specific Star-formation Rate, and Stellar Mass*, ApJ, 800(1), 20
- Ginolfi, M., Jones, G. C., Béthermin, M., et al. (2020a), *The ALPINE-ALMA [CII] survey. Circumgalactic medium pollution and gas mixing by tidal stripping in a merging system at  $z \sim 4.57$* , A&A, 643, A7
- Ginolfi, M., Jones, G. C., Béthermin, M., et al. (2020b), *The ALPINE-ALMA [C II] survey: Star-formation-driven outflows and circumgalactic enrichment in the early Universe*, A&A, 633, A90
- Ginolfi, M., Maiolino, R., Nagao, T., et al. (2017), *Molecular gas on large circumgalactic scales at  $z = 3.47$* , MNRAS, 468(3), 3468
- Glover, S. C. O. and Mac Low, M.-M. (2007), *Simulating the Formation of Molecular Clouds. I. Slow Formation by Gravitational Collapse from Static Initial Conditions*, ApJS, 169(2), 239
- Gnedin, N. Y. and Kravtsov, A. V. (2011), *Environmental Dependence of the Kennicutt-Schmidt Relation in Galaxies*, ApJ, 728(2), 88
- Godard, B., Falgarone, E., and Pineau des Forêts, G. (2014), *Chemical probes of turbulence in the diffuse medium: the TDR model*, A&A, 570, A27
- Goerdt, T. and Ceverino, D. (2015), *Inflow velocities of cold flows streaming into massive galaxies at high redshifts*, MNRAS, 450(4), 3359
- Goldreich, P. and Kwan, J. (1974), *Molecular Clouds*, ApJ, 189, 441
- Goldsmith, P. F. and Langer, W. D. (1978), *Molecular cooling and thermal balance of dense interstellar clouds.*, ApJ, 222, 881
- Goldsmith, P. F., Langer, W. D., Pineda, J. L., and Velusamy, T. (2012), *Collisional Excitation of the [C II] Fine Structure Transition in Interstellar Clouds*, ApJS, 203(1), 13
- Gould, A. and Weinberg, D. H. (1996), *Imaging the Forest of Lyman Limit Systems*, ApJ, 468, 462
- Greve, T. R. and Sommer-Larsen, J. (2008), *CO line emission from Lyman break galaxies. Cosmological simulations and predictions for ALMA*, A&A, 480(2), 335
- Gronke, M., Dijkstra, M., McCourt, M., and Oh, S. P. (2017), *Resonant line transfer in a fog: using Lyman-alpha to probe tiny structures in atomic gas*, A&A, 607, A71

- Gronke, M. and Oh, S. P. (2018), *The growth and entrainment of cold gas in a hot wind*, MNRAS, 480(1), L111
- Gronke, M. and Oh, S. P. (2020), *How cold gas continuously entrains mass and momentum from a hot wind*, MNRAS, 492(2), 1970
- Gullberg, B., De Breuck, C., Vieira, J. D., et al. (2015), *The nature of the [C II] emission in dusty star-forming galaxies from the SPT survey*, MNRAS, 449(3), 2883
- Gullberg, B., Lehnert, M. D., De Breuck, C., et al. (2016), *ALMA finds dew drops in the dusty spider's web*, A&A, 591, A73
- Hailey-Dunsheath, S., Nikola, T., Stacey, G. J., et al. (2010), *Detection of the 158  $\mu\text{m}$  [C II] Transition at  $z = 1.3$ : Evidence for a Galaxy-wide Starburst*, ApJ, 714(1), L162
- Haiman, Z. and Rees, M. J. (2001), *Extended Ly $\alpha$  Emission around Young Quasars: A Constraint on Galaxy Formation*, ApJ, 556(1), 87
- Haiman, Z., Spaans, M., and Quataert, E. (2000), *Ly $\alpha$  Cooling Radiation from High-Redshift Halos*, ApJ, 537(1), L5
- Hamann, F., Chartas, G., McGraw, S., et al. (2013), *Extreme-velocity quasar outflows and the role of X-ray shielding*, MNRAS, 435(1), 133
- Harrington, K. C., Weiss, A., Yun, M. S., et al. (2021), *Turbulent Gas in Lensed Planck-selected Starbursts at  $z \sim 1-3.5$* , ApJ, 908(1), 95
- Hayes, M., Scarlata, C., and Siana, B. (2011), *Central powering of the largest Lyman- $\alpha$  nebula is revealed by polarized radiation*, Nature, 476(7360), 304
- Heckman, T. M. and Best, P. N. (2014), *The Coevolution of Galaxies and Supermassive Black Holes: Insights from Surveys of the Contemporary Universe*, ARA&A, 52, 589
- Heckman, T. M., Lehnert, M. D., van Breugel, W., and Miley, G. K. (1991), *Spatially Resolved Optical Images of High-Redshift Quasi-stellar Objects*, ApJ, 370, 78
- Hennawi, J. F. and Prochaska, J. X. (2007), *Quasars Probing Quasars. II. The Anisotropic Clustering of Optically Thick Absorbers around Quasars*, ApJ, 655(2), 735
- Hennawi, J. F. and Prochaska, J. X. (2013), *Quasars Probing Quasars. IV. Joint Constraints on the Circumgalactic Medium from Absorption and Emission*, ApJ, 766(1), 58
- Hennawi, J. F., Prochaska, J. X., Burles, S., et al. (2006), *Quasars Probing Quasars. I. Optically Thick Absorbers near Luminous Quasars*, ApJ, 651(1), 61

- Hennawi, J. F., Prochaska, J. X., Cantalupo, S., and Arrigoni-Battaia, F. (2015), *Quasar quartet embedded in giant nebula reveals rare massive structure in distant universe*, *Science*, 348(6236), 779
- Herrera-Camus, R., Bolatto, A. D., Wolfire, M. G., et al. (2015), *[C II] 158  $\mu\text{m}$  Emission as a Star Formation Tracer*, *ApJ*, 800(1), 1
- Herrera-Camus, R., Förster Schreiber, N., Genzel, R., et al. (2021), *Kiloparsec view of a typical star-forming galaxy when the Universe was  $\sim 1$  Gyr old. I. Properties of outflow, halo, and interstellar medium*, *A&A*, 649, A31
- Hill, R., Chapman, S. C., Scott, D., et al. (2019), *The SCUBA-2 web survey: I. Observations of CO(3-2) in hyper-luminous QSO field*, *MNRAS*, 485(1), 753
- Hollas, J. (2002), *Basic Atomic and Molecular Spectroscopy*, Tutorial chemistry texts, Royal Society of Chemistry
- Hollenbach, D. J. and Tielens, A. G. G. M. (1999), *Photodissociation regions in the interstellar medium of galaxies*, *Reviews of Modern Physics*, 71(1), 173
- Hu, C.-Y., Naab, T., Walch, S., et al. (2014), *SPHGal: smoothed particle hydrodynamics with improved accuracy for galaxy simulations*, *MNRAS*, 443(2), 1173
- Hummels, C. B., Smith, B. D., Hopkins, P. F., et al. (2019), *The Impact of Enhanced Halo Resolution on the Simulated Circumgalactic Medium*, *ApJ*, 882(2), 156
- Husband, K., Bremer, M. N., Stanway, E. R., and Lehnert, M. D. (2015), *Dissecting the complex environment of a distant quasar with MUSE*, *MNRAS*, 452(3), 2388
- Husemann, B., Worseck, G., Arrigoni Battaia, F., and Shanks, T. (2018), *Discovery of a dual AGN at  $z = 3.3$  with 20 kpc separation*, *A&A*, 610, L7
- Ikeda, M., Oka, T., Tatematsu, K., et al. (2002), *The Distribution of Atomic Carbon in the Orion Giant Molecular Cloud 1*, *ApJS*, 139(2), 467
- Iono, D., Yun, M. S., Elvis, M., et al. (2006), *A Detection of [C II] Line Emission in the  $z = 4.7$  QSO BR 1202-0725*, *ApJ*, 645(2), L97
- Ishibashi, W. and Fabian, A. C. (2012), *Active galactic nucleus feedback and triggering of star formation in galaxies*, *MNRAS*, 427(4), 2998
- Israel, F. P. and Baas, F. (2002), *Neutral atomic carbon in centers of galaxies*, *A&A*, 383, 82
- Ivezić, Ž., Menou, K., Knapp, G. R., et al. (2002), *Optical and Radio Properties of Extragalactic Sources Observed by the FIRST Survey and the Sloan Digital Sky Survey*, *AJ*, 124(5), 2364

- Jalan, P., Chand, H., and Srianand, R. (2019), *Probing the Environment of High- $z$  Quasars Using the Proximity Effect in Projected Quasar Pairs*, ApJ, 884(2), 151
- Jarvis, M. E., Harrison, C. M., Mainieri, V., et al. (2020), *High molecular gas content and star formation rates in local galaxies that host quasars, outflows, and jets*, MNRAS, 498(2), 1560
- Jeans, J. H. (1902), *The Stability of a Spherical Nebula*, Philosophical Transactions of the Royal Society of London Series A, 199, 1
- Jones, G. C., Maiolino, R., Circosta, C., et al. (2023), *Evidence for extended gaseous reservoirs around AGN at cosmic noon from ALMA CO(3-2) observations*, MNRAS, 518(1), 691
- Kaaret, P., Koutroumpa, D., Kuntz, K. D., et al. (2020), *A disk-dominated and clumpy circumgalactic medium of the Milky Way seen in X-ray emission*, Nature Astronomy, 4, 1072
- Kainulainen, J., Beuther, H., Henning, T., and Plume, R. (2009), *Probing the evolution of molecular cloud structure. From quiescence to birth*, A&A, 508(3), L35
- Kanjilal, V., Dutta, A., and Sharma, P. (2021), *Growth and structure of multiphase gas in the cloud-crushing problem with cooling*, MNRAS, 501(1), 1143
- Katsianis, A., Gonzalez, V., Barrientos, D., et al. (2020), *The high-redshift SFR- $M^*$  relation is sensitive to the employed star formation rate and stellar mass indicators: towards addressing the tension between observations and simulations*, MNRAS, 492(4), 5592
- Katz, N., Keres, D., Dave, R., and Weinberg, D. H. (2003), *How Do Galaxies Get Their Gas?*, in *The IGM/Galaxy Connection. The Distribution of Baryons at  $z=0$* , edited by J. L. Rosenberg, M. E. Putman, volume 281 of *Astrophysics and Space Science Library*, 185
- Kauffmann, G. and Haehnelt, M. (2000), *A unified model for the evolution of galaxies and quasars*, MNRAS, 311(3), 576
- Kaufman, M. J., Wolfire, M. G., and Hollenbach, D. J. (2006), *[Si II], [Fe II], [C II], and  $H_2$  Emission from Massive Star-forming Regions*, ApJ, 644(1), 283
- Kereš, D., Katz, N., Fardal, M., et al. (2009), *Galaxies in a simulated  $\Lambda$ CDM Universe - I. Cold mode and hot cores*, MNRAS, 395(1), 160
- Kereš, D., Katz, N., Weinberg, D. H., and Davé, R. (2005), *How do galaxies get their gas?*, MNRAS, 363(1), 2
- Khairé, V. and Srianand, R. (2019), *New synthesis models of consistent extragalactic background light over cosmic time*, MNRAS, 484(3), 4174

- Kim, Y.-R. and Croft, R. A. C. (2008), *Constraining quasar host halo masses with the strength of nearby Ly $\alpha$  forest absorption*, MNRAS, 387(1), 377
- Klein, R. I., McKee, C. F., and Colella, P. (1994), *On the Hydrodynamic Interaction of Shock Waves with Interstellar Clouds. I. Nonradiative Shocks in Small Clouds*, ApJ, 420, 213
- Kocevski, D. D., Onoue, M., Inayoshi, K., et al. (2023), *Hidden Little Monsters: Spectroscopic Identification of Low-Mass, Broad-Line AGN at  $z > 5$  with CEERS*, arXiv e-prints, arXiv:2302.00012
- Kormendy, J. and Ho, L. C. (2013), *Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies*, ARA&A, 51(1), 511
- Krumholz, M. R. (2013), *The star formation law in molecule-poor galaxies*, MNRAS, 436(3), 2747
- Krumholz, M. R., McKee, C. F., and Tumlinson, J. (2008), *The Atomic-to-Molecular Transition in Galaxies. I. An Analytic Approximation for Photodissociation Fronts in Finite Clouds*, ApJ, 689(2), 865
- Kusakabe, H., Verhamme, A., Blaizot, J., et al. (2022), *The MUSE eXtremely Deep Field: Individual detections of Ly $\alpha$  haloes around rest-frame UV-selected galaxies at  $z \sim 2.9$ -4.4*, A&A, 660, A44
- Lagache, G., Cousin, M., and Chatzikos, M. (2018), *The [CII] 158  $\mu$ m line emission in high-redshift galaxies*, A&A, 609, A130
- Lagos, C. d. P., Crain, R. A., Schaye, J., et al. (2015), *Molecular hydrogen abundances of galaxies in the EAGLE simulations*, MNRAS, 452(4), 3815
- Lahén, N., Naab, T., Johansson, P. H., et al. (2020), *The GRIFFIN Project—Formation of Star Clusters with Individual Massive Stars in a Simulated Dwarf Galaxy Starburst*, ApJ, 891(1), 2
- Lau, M. W., Prochaska, J. X., and Hennawi, J. F. (2016), *Quasars Probing Quasars. VIII. The Physical Properties of the Cool Circumgalactic Medium Surrounding  $z \sim 2$ -3 Massive Galaxies Hosting Quasars*, ApJS, 226(2), 25
- Lau, M. W., Prochaska, J. X., and Hennawi, J. F. (2018), *Quasars Probing Quasars. IX. The Kinematics of the Circumgalactic Medium Surrounding  $z \sim 2$  Quasars*, ApJ, 857(2), 126
- Laursen, P., Razoumov, A. O., and Sommer-Larsen, J. (2009), *Ly $\alpha$  Radiative Transfer in Cosmological Simulations Using Adaptive Mesh Refinement*, ApJ, 696(1), 853

- Law, D. R., Steidel, C. C., Erb, D. K., et al. (2009), *The Kiloparsec-scale Kinematics of High-redshift Star-forming Galaxies*, ApJ, 697(2), 2057
- Le Fèvre, O., Béthermin, M., Faisst, A., et al. (2020), *The ALPINE-ALMA [CII] survey. Survey strategy, observations, and sample properties of 118 star-forming galaxies at  $4 < z < 6$* , A&A, 643, A1
- Leclercq, F., Bacon, R., Verhamme, A., et al. (2020), *The MUSE Hubble Ultra Deep Field Survey. XIII. Spatially resolved spectral properties of Lyman  $\alpha$  haloes around star-forming galaxies at  $z \lesssim 3$* , A&A, 635, A82
- Leclercq, F., Bacon, R., Wisotzki, L., et al. (2017), *The MUSE Hubble Ultra Deep Field Survey. VIII. Extended Lyman- $\alpha$  haloes around high- $z$  star-forming galaxies*, A&A, 608, A8
- Leroy, A. K., Walter, F., Brinks, E., et al. (2008), *The Star Formation Efficiency in Nearby Galaxies: Measuring Where Gas Forms Stars Effectively*, AJ, 136(6), 2782
- Li, J., Emonts, B. H. C., Cai, Z., et al. (2023), *The SUPERCOLD-CGM survey: I. Probing the extended CO(4-3) Emission of the Circumgalactic medium in a sample of 10 Enormous Ly $\alpha$  Nebulae at  $z \sim 2$* , arXiv e-prints, arXiv:2304.02041
- Li, J., Emonts, B. H. C., Cai, Z., et al. (2021), *Massive Molecular Outflow and 100 kpc Extended Cold Halo Gas in the Enormous Ly $\alpha$  Nebula of QSO 1228+3128*, ApJ, 922(2), L29
- Li, J., Wang, R., Riechers, D., et al. (2020), *Probing the Full CO Spectral Line Energy Distribution (SLED) in the Nuclear Region of a Quasar-starburst System at  $z = 6.003$* , ApJ, 889(2), 162
- Liang, C. J., Kravtsov, A. V., and Agertz, O. (2016), *Column density profiles of multiphase gaseous haloes*, MNRAS, 458(2), 1164
- Lifshitz, E. M. (1946), *On the gravitational stability of the expanding universe*, Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki, 16, 587
- Lu, N., Cao, T., Díaz-Santos, T., et al. (2018), *CO (7-6), [C I] 370  $\mu$ m, and [N II] 205  $\mu$ m Line Emission of the QSO BRI 1335-0417 at Redshift 4.407*, ApJ, 864(1), 38
- Lyke, B. W., Higley, A. N., McLane, J. N., et al. (2020), *The Sloan Digital Sky Survey Quasar Catalog: Sixteenth Data Release*, ApJS, 250(1), 8
- Mackenzie, R., Pezzulli, G., Cantalupo, S., et al. (2021), *Revealing the impact of quasar luminosity on giant Ly  $\alpha$  nebulae*, MNRAS, 502(1), 494
- Mackey, J., Walch, S., Seifried, D., et al. (2019), *Non-equilibrium chemistry and destruction of CO by X-ray flares*, MNRAS, 486(1), 1094

- Madden, S. C., Cormier, D., Hony, S., et al. (2020), *Tracing the total molecular gas in galaxies: [CII] and the CO-dark gas*, A&A, 643, A141
- Magorrian, J., Tremaine, S., Richstone, D., et al. (1998), *The Demography of Massive Dark Objects in Galaxy Centers*, AJ, 115(6), 2285
- Maiolino, R., Caselli, P., Nagao, T., et al. (2009), *Strong [CII] emission at high redshift*, A&A, 500(2), L1
- Maiolino, R., Cox, P., Caselli, P., et al. (2005), *First detection of [CII]158  $\mu\text{m}$  at high redshift: vigorous star formation in the early universe*, A&A, 440(2), L51
- Mandelker, N., van den Bosch, F. C., Springel, V., and van de Voort, F. (2019), *Shattering of Cosmic Sheets due to Thermal Instabilities: A Formation Channel for Metal-free Lyman Limit Systems*, ApJ, 881(1), L20
- Marinacci, F., Grand, R. J. J., Pakmor, R., et al. (2017), *Properties of H I discs in the Auriga cosmological simulations*, MNRAS, 466(4), 3859
- Martin, D. C., O’Sullivan, D., Matuszewski, M., et al. (2019), *Multi-filament gas inflows fuelling young star-forming galaxies*, Nature Astronomy, 3, 822
- Matuszewski, M., Chang, D., Crabill, R. M., et al. (2010), *The Cosmic Web Imager: an integral field spectrograph for the Hale Telescope at Palomar Observatory: instrument design and first results*, in *Ground-based and Airborne Instrumentation for Astronomy III*, edited by I. S. McLean, S. K. Ramsay, H. Takami, volume 7735 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 77350P
- McCourt, M., Oh, S. P., O’Leary, R., and Madigan, A.-M. (2018), *A characteristic scale for cold gas*, MNRAS, 473(4), 5407
- McLinden, E. M., Malhotra, S., Rhoads, J. E., et al. (2013), *[O III] Emission and Gas Kinematics in a Lyman-alpha Blob at  $z \sim 3.1$* , ApJ, 767(1), 48
- McMullin, J. P., Waters, B., Schiebel, D., et al. (2007), *CASA Architecture and Applications*, in *Astronomical Data Analysis Software and Systems XVI*, edited by R. A. Shaw, F. Hill, D. J. Bell, volume 376 of *Astronomical Society of the Pacific Conference Series*, 127
- McNamara, B. R. and Nulsen, P. E. J. (2007), *Heating Hot Atmospheres with Active Galactic Nuclei*, ARA&A, 45(1), 117
- Meijerink, R., Spaans, M., and Israel, F. P. (2006), *Irradiated ISM: Discriminating between Cosmic Rays and X-Rays*, ApJ, 650(2), L103
- Meijerink, R., Spaans, M., and Israel, F. P. (2007), *Diagnostics of irradiated dense gas in galaxy nuclei. II. A grid of XDR and PDR models*, A&A, 461(3), 793

- Meiksin, A., Bolton, J. S., and Puchwein, E. (2017), *Gas around galaxy haloes - III: hydrogen absorption signatures around galaxies and QSOs in the Sherwood simulation suite*, MNRAS, 468(2), 1893
- Meyer, R. A., Walter, F., Cicone, C., et al. (2022), *Physical Constraints on the Extended Interstellar Medium of the  $z = 6.42$  Quasar J1148+5251:  $[C II]_{158\mu m}$ ,  $[N II]_{205\mu m}$ , and  $[O I]_{146\mu m}$  Observations*, ApJ, 927(2), 152
- Miller, T. B., Chapman, S. C., Aravena, M., et al. (2018), *A massive core for a cluster of galaxies at a redshift of 4.3*, Nature, 556(7702), 469
- Mo, H., van den Bosch, F., and White, S. (2010), *Galaxy Formation and Evolution*, Galaxy Formation and Evolution, Cambridge University Press
- Molina, J., Wang, et al. (2021), *Compact Molecular Gas Distribution in Quasar Host Galaxies*, arXiv e-prints, arXiv:2101.00764
- Mori, M. and Umemura, M. (2006), *Early metal enrichment and Lyman  $\alpha$  emission*, New Astron. Rev., 50(1-3), 199
- Morrissey, P., Matuszewski, M., Martin, C., et al. (2012), *The Keck Cosmic Web Imager: a capable new integral field spectrograph for the W. M. Keck Observatory*, in *Ground-based and Airborne Instrumentation for Astronomy IV*, edited by I. S. McLean, S. K. Ramsay, H. Takami, volume 8446 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 844613
- Moster, B. P., Naab, T., and White, S. D. M. (2018), *EMERGE - an empirical model for the formation of galaxies since  $z \sim 10$* , MNRAS, 477(2), 1822
- Muñoz-Elgueta, N., Arrigoni Battaia, F., Kauffmann, G., et al. (2022), *APEX at the QSO MUSEUM: molecular gas reservoirs associated with  $z \sim 3$  quasars and their link to the extended Ly  $\alpha$  emission*, MNRAS, 511(1), 1462
- Narayanan, D., Cox, T. J., Hayward, C. C., et al. (2009), *The star-forming molecular gas in high-redshift Submillimetre Galaxies*, MNRAS, 400(4), 1919
- Narayanan, D., Cox, T. J., Robertson, B., et al. (2006), *Molecular Outflows in Galaxy Merger Simulations with Embedded Active Galactic Nuclei*, ApJ, 642(2), L107
- Nelson, D., Byrohl, C., Ogorzalek, A., et al. (2023), *Resonant scattering of the O VII X-ray emission line in the circumgalactic medium of TNG50 galaxies*, MNRAS, 522(3), 3665
- Nelson, D., Byrohl, C., Peroux, C., et al. (2021), *The cold circumgalactic medium in emission: Mg II haloes in TNG50*, MNRAS, 507(3), 4445
- Nelson, D., Pillepich, A., Springel, V., et al. (2019), *First results from the TNG50 simulation: galactic outflows driven by supernovae and black hole feedback*, MNRAS, 490(3), 3234

- Nelson, D., Sharma, P., Pillepich, A., et al. (2020), *Resolving small-scale cold circumgalactic gas in TNG50*, MNRAS, 498(2), 2391
- Nelson, R. P. and Langer, W. D. (1997), *The Dynamics of Low-Mass Molecular Clouds in External Radiation Fields*, ApJ, 482(2), 796
- Neri, R., Downes, D., Cox, P., and Walter, F. (2014), *High-resolution C<sup>+</sup> imaging of HDF850.1 reveals a merging galaxy at z = 5.185*, A&A, 562, A35
- Neri, R., Genzel, R., Ivison, R. J., et al. (2003), *Interferometric Observations of Powerful CO Emission from Three Submillimeter Galaxies at z=2.39, 2.51, and 3.35*, ApJ, 597(2), L113
- Nesvadba, N. P. H., Boulanger, F., Salomé, P., et al. (2010), *Energetics of the molecular gas in the H<sub>2</sub> luminous radio galaxy 3C 326: Evidence for negative AGN feedback*, A&A, 521, A65
- Neufeld, D. A. (1990), *The Transfer of Resonance-Line Radiation in Static Astrophysical Media*, ApJ, 350, 216
- Novak, M., Bañados, E., Decarli, R., et al. (2019), *An ALMA Multiline Survey of the Interstellar Medium of the Redshift 7.5 Quasar Host Galaxy J1342+0928*, ApJ, 881(1), 63
- Novak, M., Venemans, B. P., Walter, F., et al. (2020), *No Evidence for [C II] Halos or High-velocity Outflows in z ≳ 6 Quasar Host Galaxies*, ApJ, 904(2), 131
- Obreja, A., Macciò, A. V., Moster, B., et al. (2019), *Local photoionization feedback effects on galaxies*, MNRAS, 490(2), 1518
- Ohyama, Y., Taniguchi, Y., Kawabata, K. S., et al. (2003), *On the Origin of Ly $\alpha$  Blobs at High Redshift: Kinematic Evidence of a Hyperwind Galaxy at z = 3.1*, ApJ, 591(1), L9
- Okamoto, T., Frenk, C. S., Jenkins, A., and Theuns, T. (2010), *The properties of satellite galaxies in simulations of galaxy formation*, MNRAS, 406(1), 208
- Olsen, K., Greve, T. R., Narayanan, D., et al. (2017), *SÍGAME Simulations of the [CII], [OI], and [OIII] Line Emission from Star-forming Galaxies at z 6*, ApJ, 846(2), 105
- Olsen, K. P., Burkhart, B., Mac Low, M.-M., et al. (2021), *SÍGAME v3: Gas Fragmentation in Postprocessing of Cosmological Simulations for More Accurate Infrared Line Emission Modeling*, ApJ, 922(1), 88
- Olsen, K. P., Greve, T. R., Narayanan, D., et al. (2015), *Simulator of Galaxy Millimeter/Submillimeter Emission (SiGAME): The [C ii]-SFR Relationship of Massive z = 2 Main Sequence Galaxies*, ApJ, 814(1), 76

- Omont, A., Petitjean, P., Guilloteau, S., et al. (1996), *Molecular gas and dust around a radio-quiet quasar at redshift 4.69*, *Nature*, 382(6590), 428
- Oppenheimer, B. D., Crain, R. A., Schaye, J., et al. (2016), *Bimodality of low-redshift circumgalactic O VI in non-equilibrium EAGLE zoom simulations*, *MNRAS*, 460(2), 2157
- O’Sullivan, D. B., Martin, C., Matuszewski, M., et al. (2020), *The FLASHES Survey. I. Integral Field Spectroscopy of the CGM around 48 z 2.3-3.1 QSOs*, *ApJ*, 894(1), 3
- Ouchi, M., Shimasaku, K., Okamura, S., et al. (2004), *Subaru Deep Survey. VI. A Census of Lyman Break Galaxies at z~4 and 5 in the Subaru Deep Fields: Clustering Properties*, *ApJ*, 611(2), 685
- Pallottini, A., Ferrara, A., Bovino, S., et al. (2017), *The impact of chemistry on the structure of high-z galaxies*, *MNRAS*, 471(4), 4128
- Pallottini, A., Ferrara, A., Decataldo, D., et al. (2019), *Deep into the structure of the first galaxies: SERRA views*, *MNRAS*, 487(2), 1689
- Pallottini, A., Ferrara, A., Gallerani, S., et al. (2014), *Simulating cosmic metal enrichment by the first galaxies*, *MNRAS*, 440(3), 2498
- Pantoni, L., Massardi, M., Lapi, A., et al. (2021), *An ALMA view of 11 dusty star-forming galaxies at the peak of cosmic star formation history*, *MNRAS*, 507(3), 3998
- Papadopoulos, P. P. and Greve, T. R. (2004), *C I Emission in Ultraluminous Infrared Galaxies as a Molecular Gas Mass Tracer*, *ApJ*, 615(1), L29
- Peebles, P. J. E. (1980), *The large-scale structure of the universe*
- Peebles, M. S., Corlies, L., Tumlinson, J., et al. (2019), *Figuring Out Gas & Galaxies in Enzo (FOGGIE). I. Resolving Simulated Circumgalactic Absorption at  $2 \leq z \leq 2.5$* , *ApJ*, 873(2), 129
- Pensabene, A., Decarli, R., Bañados, E., et al. (2021), *ALMA multiline survey of the ISM in two quasar host-companion galaxy pairs at z < 6*, *A&A*, 652, A66
- Percival, W. J., Cole, S., Eisenstein, D. J., et al. (2007), *Measuring the Baryon Acoustic Oscillation scale using the Sloan Digital Sky Survey and 2dF Galaxy Redshift Survey*, *MNRAS*, 381(3), 1053
- Perley, R. A., Chandler, C. J., Butler, B. J., and Wrobel, J. M. (2011), *The Expanded Very Large Array: A New Telescope for New Science*, *ApJ*, 739(1), L1
- Péroux, C., Nelson, D., van de Voort, F., et al. (2020), *Predictions for the angular dependence of gas mass flow rate and metallicity in the circumgalactic medium*, *MNRAS*, 499(2), 2462

- Piacitelli, D. R., Solhaug, E., Faerman, Y., and McQuinn, M. (2022), *Absorption-based circumgalactic medium line emission estimates*, MNRAS, 516(2), 3049
- Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. (2010), *Herschel Space Observatory. An ESA facility for far-infrared and submillimetre astronomy*, A&A, 518, L1
- Pillepich, A., Nelson, D., Springel, V., et al. (2019), *First results from the TNG50 simulation: the evolution of stellar and gaseous discs across cosmic time*, MNRAS, 490(3), 3196
- Pillepich, A., Springel, V., Nelson, D., et al. (2018), *Simulating galaxy formation with the IllustrisTNG model*, MNRAS, 473(3), 4077
- Pineda, J. L., Langer, W. D., and Goldsmith, P. F. (2014), *A Herschel [C II] Galactic plane survey. III. [C II] as a tracer of star formation*, A&A, 570, A121
- Pizzati, E., Ferrara, A., Pallottini, A., et al. (2020), *Outflows and extended [C II] haloes in high-redshift galaxies*, MNRAS, 495(1), 160
- Pizzati, E., Ferrara, A., Pallottini, A., et al. (2023), *[C II] Haloes in ALPINE galaxies: smoking-gun of galactic outflows?*, MNRAS, 519(3), 4608
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. (2016), *Planck 2015 results. XIII. Cosmological parameters*, A&A, 594, A13
- Planck Collaboration, Aghanim, N., Akrami, Y., et al. (2020), *Planck 2018 results. VI. Cosmological parameters*, A&A, 641, A6
- Polletta, M., Nesvadba, N. P. H., Neri, R., et al. (2011), *Disk, merger, or outflow? Molecular gas kinematics in two powerful obscured QSOs at  $z \geq 3.4$* , A&A, 533, A20
- Popping, G., Pillepich, A., Somerville, R. S., et al. (2019), *The ALMA Spectroscopic Survey in the HUDF: the Molecular Gas Content of Galaxies and Tensions with IllustrisTNG and the Santa Cruz SAM*, ApJ, 882(2), 137
- Porciani, C., Magliocchetti, M., and Norberg, P. (2004), *Cosmic evolution of quasar clustering: implications for the host haloes*, MNRAS, 355(3), 1010
- Prescott, M. K. M., Martin, C. L., and Dey, A. (2015), *Spatially Resolved Gas Kinematics within a Ly $\alpha$  Nebula: Evidence for Large-scale Rotation*, ApJ, 799(1), 62
- Press, W. H. and Schechter, P. (1974), *Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation*, ApJ, 187, 425
- Prochaska, J. X., Hennawi, J. F., and Simcoe, R. A. (2013), *A Substantial Mass of Cool, Metal-enriched Gas Surrounding the Progenitors of Modern-day Ellipticals*, ApJ, 762(2), L19

- Prochaska, J. X., Lau, M. W., and Hennawi, J. F. (2014), *Quasars Probing Quasars. VII. The Pinnacle of the Cool Circumgalactic Medium Surrounds Massive  $z \sim 2$  Galaxies*, ApJ, 796(2), 140
- Rathjen, T.-E., Naab, T., Girichidis, P., et al. (2021), *SILCC VI - Multiphase ISM structure, stellar clustering, and outflows with supernovae, stellar winds, ionizing radiation, and cosmic rays*, MNRAS, 504(1), 1039
- Rees, M. J. (1988), *Quasars as probes of gas in extended protogalaxies.*, MNRAS, 231, 91P
- Rees, M. J. and Ostriker, J. P. (1977), *Cooling, dynamics and fragmentation of massive gas clouds: clues to the masses and radii of galaxies and clusters.*, MNRAS, 179, 541
- Reid, B. A., Percival, W. J., Eisenstein, D. J., et al. (2010), *Cosmological constraints from the clustering of the Sloan Digital Sky Survey DR7 luminous red galaxies*, MNRAS, 404(1), 60
- Richards, G. T., Strauss, M. A., Fan, X., et al. (2006), *The Sloan Digital Sky Survey Quasar Survey: Quasar Luminosity Function from Data Release 3*, AJ, 131(6), 2766
- Richstone, D., Ajhar, E. A., Bender, R., et al. (1998), *Supermassive black holes and the evolution of galaxies.*, Nature, 385(6701), A14
- Riechers, D. A., Bradford, C. M., Clements, D. L., et al. (2013), *A dust-obscured massive maximum-starburst galaxy at a redshift of 6.34*, Nature, 496(7445), 329
- Riechers, D. A., Carilli, C. L., Maddalena, R. J., et al. (2011), *CO( $J = 1 \rightarrow 0$ ) in  $z \sim 2$  Quasar Host Galaxies: No Evidence for Extended Molecular Gas Reservoirs*, ApJ, 739(1), L32
- Riechers, D. A., Walter, F., Carilli, C. L., et al. (2006), *CO( $1-0$ ) in  $z \sim 4$  Quasar Host Galaxies: No Evidence for Extended Molecular Gas Reservoirs*, ApJ, 650(2), 604
- Riechers, D. A., Walter, F., Carilli, C. L., and Lewis, G. F. (2009), *Imaging The Molecular Gas in a  $z = 3.9$  Quasar Host Galaxy at 0farcs3 Resolution: A Central, Sub-Kiloparsec Scale Star Formation Reservoir in APM 08279+5255*, ApJ, 690(1), 463
- Rodighiero, G., Daddi, E., Baronchelli, I., et al. (2011), *The Lesser Role of Starbursts in Star Formation at  $z = 2$* , ApJ, 739(2), L40
- Rosdahl, J. and Blaizot, J. (2012), *Extended Ly $\alpha$  emission from cold accretion streams*, MNRAS, 423(1), 344
- Rudie, G. C., Steidel, C. C., Pettini, M., et al. (2019), *Column Density, Kinematics, and Thermal State of Metal-bearing Gas within the Virial Radius of  $z \sim 2$  Star-forming Galaxies in the Keck Baryonic Structure Survey*, ApJ, 885(1), 61

- Rybak, M., Zavala, J. A., Hodge, J. A., et al. (2020), *First Detection of the [O I] 63  $\mu$ m Emission from a Redshift 6 Dusty Galaxy*, ApJ, 889(1), L11
- Salak, D., Nakai, N., Seta, M., and Miyamoto, Y. (2019), *ALMA Observations of Atomic Carbon [C I] ( $^3 P_1 \rightarrow ^3 P_0$ ) and Low-J CO Lines in the Starburst Galaxy NGC 1808*, ApJ, 887(2), 143
- Samsonyan, A. L. (2022), *Analysis of Emission Line Widths of [CII] 158 $\mu$ m*, Astrophysics, 65(2), 151
- Sanders, R. L., Shapley, A. E., Jones, T., et al. (2023), *CO Emission, Molecular Gas, and Metallicity in Main-sequence Star-forming Galaxies at  $z \sim 2.3$* , ApJ, 942(1), 24
- Sargsyan, L., Lebouteiller, V., Weedman, D., et al. (2012), *[C II] 158  $\mu$ m Luminosities and Star Formation Rate in Dusty Starbursts and Active Galactic Nuclei*, ApJ, 755(2), 171
- Scannapieco, E. and Brügger, M. (2015), *The Launching of Cold Clouds by Galaxy Outflows. I. Hydrodynamic Interactions with Radiative Cooling*, ApJ, 805(2), 158
- Schaerer, D., Ginolfi, M., Béthermin, M., et al. (2020), *The ALPINE-ALMA [C II] survey. Little to no evolution in the [C II]-SFR relation over the last 13 Gyr*, A&A, 643, A3
- Schaye, J., Crain, R. A., Bower, R. G., et al. (2015), *The EAGLE project: simulating the evolution and assembly of galaxies and their environments*, MNRAS, 446(1), 521
- Schimek, A., Decataldo, D., Shen, S., et al. (2023), *High resolution modeling of [CII], [CI], [OIII] and CO line emission from the ISM and CGM of a star forming galaxy at  $z \sim 6.5$* , arXiv e-prints, arXiv:2306.00583
- Schmidt, M. (1963), *3C 273 : A Star-Like Object with Large Red-Shift*, Nature, 197(4872), 1040
- Schneider, N., Simon, R., Kramer, C., et al. (2003), *A multiwavelength study of the S 106 region. II. Characteristics of the photon dominated region*, A&A, 406, 915
- Schöier, F. L., van der Tak, F. F. S., van Dishoeck, E. F., and Black, J. H. (2005), *An atomic and molecular database for analysis of submillimetre line observations*, A&A, 432(1), 369
- Scholtz, J., Maiolino, R., Jones, G. C., and Carniani, S. (2023), *Evidence of extended cold molecular gas and dust haloes around  $z \sim 2.3$  extremely red quasars with ALMA*, MNRAS, 519(4), 5246
- Schumacher, H., Martínez-Sansigre, A., Lacy, M., et al. (2012), *Gas and dust in a  $z = 2.8$  obscured quasar*, MNRAS, 423(3), 2132

- Scoville, N., Lee, N., Vanden Bout, P., et al. (2017), *Evolution of Interstellar Medium, Star Formation, and Accretion at High Redshift*, ApJ, 837(2), 150
- Shangguan, J., Ho, L. C., Bauer, F. E., et al. (2020), *AGN Feedback and Star Formation of Quasar Host Galaxies: Insights from the Molecular Gas*, ApJ, 899(2), 112
- Shen, S., Madau, P., Guedes, J., et al. (2013), *The Circumgalactic Medium of Massive Galaxies at  $z \sim 3$ : A Test for Stellar Feedback, Galactic Outflows, and Cold Streams*, ApJ, 765(2), 89
- Shen, X., Hopkins, P. F., Faucher-Giguère, C.-A., et al. (2020), *The bolometric quasar luminosity function at  $z = 0-7$* , MNRAS, 495(3), 3252
- Shen, Y., Brandt, W. N., Richards, G. T., et al. (2016), *The Sloan Digital Sky Survey Reverberation Mapping Project: Velocity Shifts of Quasar Emission Lines*, ApJ, 831(1), 7
- Shen, Y., Strauss, M. A., Oguri, M., et al. (2007), *Clustering of High-Redshift ( $z \approx 2.9$ ) Quasars from the Sloan Digital Sky Survey*, AJ, 133(5), 2222
- Shibuya, T., Ouchi, M., and Harikane, Y. (2015), *Morphologies of  $\sim 190,000$  Galaxies at  $z = 0-10$  Revealed with HST Legacy Data. I. Size Evolution*, ApJS, 219(2), 15
- Shin, J., Woo, J.-H., Chung, A., et al. (2019), *Positive and Negative Feedback of AGN Outflows in NGC 5728*, ApJ, 881(2), 147
- Silk, J. (1977), *On the fragmentation of cosmic gas clouds. I. The formation of galaxies and the first generation of stars.*, ApJ, 211, 638
- Silk, J. and Rees, M. J. (1998), *Quasars and galaxy formation*, A&A, 331, L1
- Silva, A., Sajina, A., Lonsdale, C., and Lacy, M. (2015), *ALMA Detected Overdensity of Sub-millimeter Sources Around WISE/NVSS-selected  $z \sim 2$  Dusty Quasars*, ApJ, 806(2), L25
- Smith, D. J. B., Jarvis, M. J., Simpson, C., and Martínez-Sansigre, A. (2009), *An 80-kpc Ly $\alpha$  halo around a high-redshift type-2 quasi-stellar object*, MNRAS, 393(1), 309
- Solomon, P. M., Downes, D., and Radford, S. J. E. (1992), *Warm Molecular Gas in the Primeval Galaxy IRAS 10214+4724*, ApJ, 398, L29
- Speagle, J. S., Steinhardt, C. L., Capak, P. L., and Silverman, J. D. (2014), *A Highly Consistent Framework for the Evolution of the Star-Forming “Main Sequence” from  $z \sim 0-6$* , ApJS, 214(2), 15
- Spergel, D. N., Bean, R., Doré, O., et al. (2007), *Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology*, ApJS, 170(2), 377

- Spingola, C., McKean, J. P., Vegetti, S., et al. (2020), *SHARP - VI. Evidence for CO (1-0) molecular gas extended on kpc-scales in AGN star-forming galaxies at high redshift*, MNRAS, 495(2), 2387
- Spitzer, L. (1978), *Physical processes in the interstellar medium*
- Springel, V. (2005), *The cosmological simulation code GADGET-2*, MNRAS, 364(4), 1105
- Springel, V. (2010), *E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh*, MNRAS, 401(2), 791
- Springel, V., Di Matteo, T., and Hernquist, L. (2005), *Modelling feedback from stars and black holes in galaxy mergers*, MNRAS, 361(3), 776
- Springel, V. and Hernquist, L. (2003), *Cosmological smoothed particle hydrodynamics simulations: a hybrid multiphase model for star formation*, MNRAS, 339(2), 289
- Sravan, N., Faucher-Giguère, C.-A., van de Voort, F., et al. (2016), *Strongly time-variable ultraviolet metal-line emission from the circum-galactic medium of high-redshift galaxies*, MNRAS, 463(1), 120
- Stacey, G. J., Geis, N., Genzel, R., et al. (1991), *The 158 Micron [C ii] Line: A Measure of Global Star Formation Activity in Galaxies*, ApJ, 373, 423
- Stacey, G. J., Hailey-Dunsheath, S., Ferkinhoff, C., et al. (2010), *A 158  $\mu\text{m}$  [C II] Line Survey of Galaxies at  $z \sim 1-2$ : An Indicator of Star Formation in the Early Universe*, ApJ, 724(2), 957
- Stacey, H. R., McKean, J. P., Powell, D. M., et al. (2021), *The rocky road to quiescence: compaction and quenching of quasar host galaxies at  $z \sim 2$* , MNRAS, 500(3), 3667
- Steinborn, L. K., Dolag, K., Hirschmann, M., et al. (2015), *A refined sub-grid model for black hole accretion and AGN feedback in large cosmological simulations*, MNRAS, 448(2), 1504
- Stevens, A. R. H., Lagos, C. d. P., Contreras, S., et al. (2017), *How to get cool in the heat: comparing analytic models of hot, cold, and cooling gas in haloes and galaxies with EAGLE*, MNRAS, 467(2), 2066
- Stevens, A. R. H., Lagos, C. d. P., Cortese, L., et al. (2021), *Molecular hydrogen in IllustrisTNG galaxies: carefully comparing signatures of environment with local CO and SFR data*, MNRAS, 502(3), 3158
- Stutzki, J., Graf, U. U., Haas, S., et al. (1997), *Atomic Carbon in M82: Physical Conditions Derived from Simultaneous Observations of the [C I] Fine-Structure Submillimeter-Wave Transitions*, ApJ, 477(1), L33

- Sunyaev, R. A. and Zeldovich, Y. B. (1972), *Formation of Clusters of Galaxies; Protocluster Fragmentation and Intergalactic Gas Heating*, A&A, 20, 189
- Tacconi, L. J., Genzel, R., Saintonge, A., et al. (2018), *PHIBSS: Unified Scaling Relations of Gas Depletion Time and Molecular Gas Fractions*, ApJ, 853(2), 179
- Tacconi, L. J., Genzel, R., and Sternberg, A. (2020), *The Evolution of the Star-Forming Interstellar Medium Across Cosmic Time*, ARA&A, 58, 157
- Taniguchi, Y. and Shioya, Y. (2000), *Superwind Model of Extended Ly $\alpha$  Emitters at High Redshift*, ApJ, 532(1), L13
- Tarantino, E., Bolatto, A. D., Herrera-Camus, R., et al. (2021), *Characterizing the Multiphase Origin of [C II] Emission in M101 and NGC 6946 with Velocity-resolved Spectroscopy*, ApJ, 915(2), 92
- Tegmark, M., Silk, J., and Evrard, A. (1993), *Late Reionization by Supernova-driven Winds*, ApJ, 417, 54
- Tegmark, M., Strauss, M. A., Blanton, M. R., et al. (2004), *Cosmological parameters from SDSS and WMAP*, Phys. Rev. D, 69(10), 103501
- Timlin, J. D., Ross, N. P., Richards, G. T., et al. (2018), *The Clustering of High-redshift ( $2.9 \leq z \leq 5.1$ ) Quasars in SDSS Stripe 82*, ApJ, 859(1), 20
- Tomassetti, M., Porciani, C., Romano-Díaz, E., and Ludlow, A. D. (2015), *Simulating the H<sub>2</sub> content of high-redshift galaxies*, MNRAS, 446(4), 3330
- Tomita, K. (1967), *Non-Linear Theory of Gravitational Instability in the Expanding Universe*, Progress of Theoretical Physics, 37(5), 831
- Torrey, P., Vogelsberger, M., Marinacci, F., et al. (2019), *The evolution of the mass-metallicity relation and its scatter in IllustrisTNG*, MNRAS, 484(4), 5587
- Trainor, R. F., Steidel, C. C., Strom, A. L., and Rudie, G. C. (2015), *The Spectroscopic Properties of Ly $\alpha$ -Emitters at  $z \sim 2.7$ : Escaping Gas and Photons from Faint Galaxies*, ApJ, 809(1), 89
- Travascio, A., Zappacosta, L., Cantalupo, S., et al. (2020), *The WISSH quasars project. VIII. Outflows and metals in the circum-galactic medium around the hyper-luminous  $z \sim 3.6$  quasar J1538+08*, A&A, 635, A157
- Tumlinson, J., Peeples, M. S., and Werk, J. K. (2017), *The Circumgalactic Medium*, ARA&A, 55(1), 389
- Tumlinson, J., Thom, C., Werk, J. K., et al. (2011), *The Large, Oxygen-Rich Halos of Star-Forming Galaxies Are a Major Reservoir of Galactic Metals*, Science, 334(6058), 948

- Turner, M. L., Schaye, J., Steidel, C. C., et al. (2014), *Metal-line absorption around  $z \approx 2.4$  star-forming galaxies in the Keck Baryonic Structure Survey*, MNRAS, 445(1), 794
- Urry, C. M. and Padovani, P. (1995), *Unified Schemes for Radio-Loud Active Galactic Nuclei*, PASP, 107, 803
- Uzgil, B. D., Bradford, C. M., Hailey-Dunsheath, S., et al. (2016), *Constraining the ISM Properties of the Cloverleaf Quasar Host Galaxy with Herschel Spectroscopy*, ApJ, 832(2), 209
- Valentino, F., Magdis, G. E., Daddi, E., et al. (2018), *A Survey of Atomic Carbon [C I] in High-redshift Main-sequence Galaxies*, ApJ, 869(1), 27
- Vallini, L., Gallerani, S., Ferrara, A., et al. (2015), *On the [CII]-SFR Relation in High Redshift Galaxies*, ApJ, 813(1), 36
- van de Voort, F. and Schaye, J. (2013), *Soft X-ray and ultraviolet metal-line emission from the gas around galaxies*, MNRAS, 430(4), 2688
- van de Voort, F., Schaye, J., Altay, G., and Theuns, T. (2012), *Cold accretion flows and the nature of high column density H I absorption at redshift 3*, MNRAS, 421(4), 2809
- van de Voort, F., Schaye, J., Booth, C. M., et al. (2011), *The rates and modes of gas accretion on to galaxies and their gaseous haloes*, MNRAS, 414(3), 2458
- van de Voort, F., Springel, V., Mandelker, N., et al. (2019), *Cosmological simulations of the circumgalactic medium with 1 kpc resolution: enhanced H I column densities*, MNRAS, 482(1), L85
- van der Tak, F. F. S., Black, J. H., Schöier, F. L., et al. (2007), *A computer program for fast non-LTE analysis of interstellar line spectra. With diagnostic plots to interpret observed line intensity ratios*, A&A, 468(2), 627
- Venemans, B. P., Walter, F., Decarli, R., et al. (2017), *Molecular Gas in Three  $z \sim 7$  Quasar Host Galaxies*, ApJ, 845(2), 154
- Venemans, B. P., Walter, F., Neeleman, M., et al. (2020), *Kiloparsec-scale ALMA Imaging of [C II] and Dust Continuum Emission of 27 Quasar Host Galaxies at  $z \sim 6$* , ApJ, 904(2), 130
- Venemans, B. P., Walter, F., Zschaechner, L., et al. (2016), *Bright [C II] and Dust Emission in Three  $z \lesssim 6.6$  Quasar Host Galaxies Observed by ALMA*, ApJ, 816(1), 37
- Verhamme, A., Garel, T., Ventou, E., et al. (2018), *Recovering the systemic redshift of galaxies from their Lyman alpha line profile*, MNRAS, 478(1), L60
- Verhamme, A., Schaerer, D., and Maselli, A. (2006), *3D Ly $\alpha$  radiation transfer. I. Understanding Ly $\alpha$  line profile morphologies*, A&A, 460(2), 397

- Vidal-García, A., Falgarone, E., Arrigoni Battaia, F., et al. (2021), *Where infall meets outflows: turbulent dissipation probed by  $CH^+$  and  $Ly\alpha$  in the starburst/AGN galaxy group SMM J02399-0136 at  $z$  2.8*, MNRAS, 506(2), 2551
- Vizgan, D., Greve, T. R., Olsen, K. P., et al. (2022), *Tracing Molecular Gas Mass in  $z \sim 6$  Galaxies with  $[C II]$* , ApJ, 929(1), 92
- Vogelsberger, M., Genel, S., Sijacki, D., et al. (2013), *A model for cosmological simulations of galaxy formation physics*, MNRAS, 436(4), 3031
- Walch, S., Girichidis, P., Naab, T., et al. (2015), *The SILCC (Simulating the LifeCycle of molecular Clouds) project - I. Chemical evolution of the supernova-driven ISM*, MNRAS, 454(1), 238
- Walter, F., Bertoldi, F., Carilli, C., et al. (2003), *Molecular gas in the host galaxy of a quasar at redshift  $z = 6.42$* , Nature, 424(6947), 406
- Walter, F., Carilli, C., Bertoldi, F., et al. (2004), *Resolved Molecular Gas in a Quasar Host Galaxy at Redshift  $z=6.42$* , ApJ, 615(1), L17
- Walter, F., Weiß, A., Downes, D., et al. (2011), *A Survey of Atomic Carbon at High Redshift*, ApJ, 730(1), 18
- Wang, F., Wang, R., Fan, X., et al. (2019), *Spatially Resolved Interstellar Medium and Highly Excited Dense Molecular Gas in the Most Luminous Quasar at  $z = 6.327$* , ApJ, 880(1), 2
- Wang, R., Carilli, C. L., Neri, R., et al. (2010), *Molecular Gas in  $z \sim 6$  Quasar Host Galaxies*, ApJ, 714(1), 699
- Wang, R., Wagg, J., Carilli, C. L., et al. (2013), *Star Formation and Gas Kinematics of Quasar Host Galaxies at  $z \sim 6$ : New Insights from ALMA*, ApJ, 773(1), 44
- Wang, R., Wu, X.-B., Neri, R., et al. (2016), *Probing the Interstellar Medium and Star Formation of the Most Luminous Quasar at  $z = 6.3$* , ApJ, 830(1), 53
- Weinberger, R., Springel, V., Hernquist, L., et al. (2017), *Simulating galaxy formation with black hole driven thermal and kinetic feedback*, MNRAS, 465(3), 3291
- Weiß, A., Downes, D., Henkel, C., and Walter, F. (2005), *Atomic carbon at redshift  $\sim 2.5$* , A&A, 429, L25
- Weiß, A., Downes, D., Neri, R., et al. (2007), *Highly-excited CO emission in APM 08279+5255 at  $z = 3.9$* , A&A, 467(3), 955

- Weiss, A., Downes, D., Walter, F., and Henkel, C. (2007), *CO Line SEDs of High-Redshift QSOs and Submm Galaxies*, in *From Z-Machines to ALMA: (Sub)Millimeter Spectroscopy of Galaxies*, edited by A. J. Baker, J. Glenn, A. I. Harris, J. G. Mangum, M. S. Yun, volume 375 of *Astronomical Society of the Pacific Conference Series*, 25
- Wei, A., Henkel, C., Downes, D., and Walter, F. (2003), *Gas and dust in the Cloverleaf quasar at redshift 2.5*, *A&A*, 409, L41
- Whitaker, K. E., Franx, M., Leja, J., et al. (2014), *Constraining the Low-mass Slope of the Star Formation Sequence at 0.5  $z$  2.5*, *ApJ*, 795(2), 104
- Whitaker, K. E., van Dokkum, P. G., Brammer, G., and Franx, M. (2012), *The Star Formation Mass Sequence Out to  $z = 2.5$* , *ApJ*, 754(2), L29
- White, G. J., Ellison, B., Claude, S., et al. (1994), *CO and CI maps of the Starburst Galaxy M82.*, *A&A*, 284, L23
- White, M., Myers, A. D., Ross, N. P., et al. (2012), *The clustering of intermediate-redshift quasars as measured by the Baryon Oscillation Spectroscopic Survey*, *MNRAS*, 424(2), 933
- White, S. D. M. and Rees, M. J. (1978), *Core condensation in heavy halos: a two-stage theory for galaxy formation and clustering.*, *MNRAS*, 183, 341
- Wiersma, R. P. C., Schaye, J., and Smith, B. D. (2009), *The effect of photoionization on the cooling rates of enriched, astrophysical plasmas*, *MNRAS*, 393(1), 99
- Wilman, R. J., Geressen, J., Bower, R. G., et al. (2005), *The discovery of a galaxy-wide superwind from a young massive galaxy at redshift  $z \sim 3$* , *Nature*, 436(7048), 227
- Wisotzki, L., Bacon, R., Blaizot, J., et al. (2016), *Extended Lyman  $\alpha$  haloes around individual high-redshift galaxies revealed by MUSE*, *A&A*, 587, A98
- Wootten, A. and Thompson, A. R. (2009), *The Atacama Large Millimeter/Submillimeter Array*, *IEEE Proceedings*, 97(8), 1463
- Xu, H., Wise, J. H., Norman, M. L., et al. (2016), *Galaxy Properties and UV Escape Fractions during the Epoch of Reionization: Results from the Renaissance Simulations*, *ApJ*, 833(1), 84
- Xu, J. and Stone, J. M. (1995), *The Hydrodynamics of Shock-Cloud Interactions in Three Dimensions*, *ApJ*, 454, 172
- Yang, B., Stancil, P. C., Balakrishnan, N., and Forrey, R. C. (2010), *Rotational Quenching of CO due to H<sub>2</sub> Collisions*, *ApJ*, 718(2), 1062
- Yang, J., Venemans, B., Wang, F., et al. (2019), *Far-infrared Properties of the Bright, Gravitationally Lensed Quasar J0439+1634 at  $z = 6.5$* , *ApJ*, 880(2), 153

- 
- Yang, Y., Zabludoff, A., Jahnke, K., et al. (2011), *Gas Kinematics in Ly $\alpha$  Nebulae*, ApJ, 735(2), 87
- Zel'dovich, Y. B. (1970), *Gravitational instability: An approximate theory for large density perturbations.*, A&A, 5, 84
- Zhang, S., Cai, Z., Xu, D., et al. (2023), *Inspiraling streams of enriched gas observed around a massive galaxy 11 billion years ago*, Science, 380(6644), 494
- Zhang, Z., Shi, Y., Rieke, G. H., et al. (2016), *Distributions of Quasar Hosts on the Galaxy Main Sequence Plane*, ApJ, 819(2), L27



# Acknowledgements

First, I would like to thank my supervisor Guinevere for giving me the opportunity of working in her group, for her guidance and invaluable scientific advice throughout my research.

I am deeply thankful to my co-mentor, Fabrizio, who always found the time to guide and help me in this journey. His encouragement and support have been crucial in my growth as a researcher. Without him, all this work would not have been possible.

I also want to express my gratitude to the MPA secretaries, especially Maria and Gabi. Maria always offered her assistance with paperwork and administrative matters with a smile, and Gabi took care of all my visa issues. Without them, everything would have been much more stressful.

Now, I want to thank to all the friends I made during my PhD. To Oliver, for being my first friend in Germany, thanks for all his support, the hikes, drinking evenings, and loyalty. Oli, I am looking forward to visit you in your new life in the US. To Miha, for being always willing to help me in literally anything. Miha, I will never forget our Slovenia trip, you opened us the doors of your home, so remember you will always have a place in Chile. To Maria, Simon, Eva and Eli, for all the nice moments.

To my dear latin friends Jay, Monica, Joaquin, Simon and Yru, for making me feel at home even though I'm so far away. To Alejandra for her friendship, long talks, beer evenings and company. Ale, you have become an important pillar for me in this process.

To my family in Chile, I am immensely thankful for their unwavering support and love. A special thanks goes to my mom, Sandra, without whom none of this would have been possible. I love you mom, and I am looking forward to our next trip together. To my best friend Pallero for all the emotional support, patience and scientific discussions. Pallero, you know you are like my brother, thank you for always being there. I would also like to thank to the family I found in Munich: my cousin Camila. Camila, your presence brought many joyful moments, and I am grateful for everything.

Lastly, I want to express my deepest gratitude to my partner, Daniel, who has been by my side throughout this journey. Daniel, I don't have enough words to thank you for all the patience and unconditional support that you always provide me. You have helped me in so many ways. I could not have asked for a better partner, I love you, let's keep moving forward together.