Distribution and Evolution of Molecular Gas in Galaxies

Victoria Bollo Doizi

München 2025

Distribution and Evolution of Molecular Gas in Galaxies

Victoria Bollo Doizi

Dissertation der Fakultät für Physik
Dissertation of the Faculty of Physics

der Ludwig–Maximilians–Universität München
at the Ludwig Maximilian University of Munich

für den Grad des
for the degree of

Doctor rerum naturalium
Doctor of Natural Sciences

vorgelegt von
presented by

Victoria Bollo Doizi

aus
from

Santiago, Chile

München, den 01.09.2025

Erstgutachter (First Reviewer): Prof. Dr. Klaus Dolag

Zweitgutachter (Second Reviewer): Prof. Dr. Volker Springel

Tag der mündlichen Prüfung (Date of oral examination): 14.10.2025

This thesis has been carried out at the European Southern Observatory (ESO) under the supervision of Dr Céline Péroux and Dr Martin Zwaan.

A todas las mujeres que,
con coraje y valentía,
abrieron sendas que antes
ni siquiera existían.
Para que el cielo infinito
a todas nos convoque,
y el legado de sus luchas
en nuestra vida se invoque.
En especial a mi madre,
mi fuerza, mi gratitud,
mi eterna estrella guía,
mi raíz, mi gran Luz.

Contents

List of Figures

List of Tables

List of Abbreviations

ACA	ALMA Compact Array
ALMA	Atacama Large Millimeter/Submillimeter Array
ALMACAL	ALMA Calibrator Survey
AGN	Active Galactic Nuclei
ASPECS	ALMA Spectroscopic Survey in the Hubble Ultra Deep Field
ALPINE	ALMA Large Program to Investigate C^+ at Early Times
ASKAP	Australian Square Kilometer Array Pathfinder
ASPECS	ALMA Spectroscopic Survey in the Hubble Ultra Deep Field
CASA	Common Astronomy Software Application
CDM	Cold Dark Matter
CGM	Circumgalactic Medium
CO	Carbon Monoxide
CO SLED	Carbon Monoxide Spectral Line Energy Distribution
COLDz	CO Line Deep The CO Luminosity Density at High$-z$ Survey
CRISTAL	[C II] Resolved ISM in STar-forming galaxies with ALMA
CMB	Cosmic Microwave Background
Dec	Declination
DLA	Damped Lyman$-\alpha$ System
ELT	Extremely Large Telescope
EHT	Event Horizon Telescope
eROSITA	Extended ROentgen Survey with an Imaging Telescope Array
FAST	Five Hundred Meter Aperture Spherical Radio Telescope
FIR	Far-infrared
FWHM	Full Width Half Maximum
HIPASS	H I Parkes All Sky Survey

HUDF	Hubble Ultra Deep Field
HST	Hubble Space Telescope
IFU	Integral Field Unit
IGM	Intergalactic Medium
IR	Infrared
ISM	Interstellar Medium
JWST	James Webb Space Telescope
Lyα	Lyman α
MUSE/VLT	Multi-Unit Spectroscopic Explorer on the Very Large Telescope
MOST	Molonglo Observatory Synthesis Telescope
NIR	Near-Infrared
NOEMA	Northern Extended Millimeter Array
PdBI	Plateau de Bure interferometer
PHANGS	Physics at High Angular resolution in Nearby Galaxies Survey
PAH	Polycyclic Aromatic Hydrocarbon
PHIBSS	IRAM Plateau de Bure High$-z$ Blue Sequence Survey
QSOs	Quasar host galaxies
RA	Right Ascension
REBELS	Reionization Lensing Cluster Survey
RMS	Root mean square
SED	Spectral Energy Distribution
SFR	Star Formation Rate
SFRD	Star Formation Rate Density
SKA	Square Kilometer Array
SMGs	Submillimetre selected galaxies
SNR	Signal-to-noise ratio
SoFiA	Source Finding Application
ULIRGs	Ultra-luminous infrared galaxies
UV	Ultraviolet
VLA	Karl Jansky Very Large Array
VLBA	Very Long Baseline Array
WSRT	Westerbork Synthesis Radio Telescope
XMM-Newton	X-ray Multi-Mirror Mission Newton

Zusammenfassung

Diese Dissertation untersucht die Verteilung und Entwicklung von kaltem molekularen Gas – dem primären Treibstoff für Sternentstehung – sowie dessen Rolle bei der Gestaltung der Galaxienentwicklung. Die Arbeit basiert auf zwei großen Beobachtungskampagnen, die mit dem Atacama Large Millimetre/submillimetre Array Observatorium (ALMA) durchgeführt wurden: ALMACAL und MUSE–ALMA Haloes, die zusammen einen umfassenden und detaillierten Blick auf den Gasfluss in und aus Galaxien über die letzten 12 Milliarden Jahre ermöglichen.

Der erste Teil dieser Dissertation befasst sich mit ALMACAL, einer weiträumigen, tiefen Millimeter/Submillimeter durchmusterung, die aus neu aufbereiteten ALMA Kalibrationsscans erstellt wurde. Unter Nutzung von mehr als 2000 Stunden Integrationszeit über mehr als 1000 Kalibratorfelder zeigt die ALMACAL–22 Veröffentlichung, wie Kalibrationsdaten für die Erstellung hochwertiger Datenkuben wiederverwendet werden können. ALMACAL–22, das aus zufällig verteilten Pointings aufgebaut ist, minimiert die kosmische Varianz und stellt einen tiefen, wachsenden Datensatz zur Untersuchung von staubreichen Sternentstehungsgalaxien, extragalaktischen Absorptionslinien, der Physik aktiver galaktischer Kerne und der Entwicklung von molekularem Gas bereit.

Der zweite Teil nutzt die statistische Stärke von ALMACAL, um eine CO-selektierte Stichprobe von Galaxien aufzubauen, frei von optischen Vorselektionseffekten. Durch Anwendung einer Emissionslinien-Klassifikationsmethode werden neue Einschränkungen für die CO-Leuchtkraftfunktion und die kosmische molekulare Gasmassendichte ($\rho_{\mathrm{H_2}}$) abgeleitet. Die Ergebnisse zeigen einen Anstieg von $\rho_{\mathrm{H_2}}$ von $z \sim 0$ bis $z \sim 1.5$, gefolgt von einem Rückgang bei höheren Rotverschiebungen. Diese Entwicklung stimmt mit der kosmischen Sternentstehungsratendichteentwicklung überein, die in der Literatur durch eine Vielzahl von Messungen gut belegt ist. Diese Trends unterstützen ein Baryonenzyklus-Szenario, in dem kaltes Gas kontinuierlich nachgeliefert wird, während die Gasverarmungszeiten über

die Zeit relativ konstant bleiben, was auf stabile Sternentstehungseffizienzen auf kosmologischen Skalen hindeutet.

Der dritte Teil untersucht molekulares Gas in Galaxien, die durch Lyα-Absorption von neutralem Wasserstoffgas (H I) entlang der Sichtlinie ausgewählt wurden. Mithilfe von Beobachtungen aus einem ALMA Large Program (Cycle 10, PI: C. Péroux), das diese Systeme bei $z \sim 0.5$ ins Visier nahm, konnte molekulares Gas in Emission in etwa einem Viertel der Stichprobe nachgewiesen werden. Diese Galaxien weisen tendenziell niedrigere Sternentstehungsraten und längere Gasverarmungszeiten auf als typische Sternentstehungsgalaxien. Die Ergebnisse deuten darauf hin, dass sich diese Systeme in einer regulierten oder Übergangsevolutionsphase befinden, die durch Prozesse wie Gasakkretion, Feedback und niedrige Metallizitätsbedingungen im interstellaren Medium geprägt sind. Diese Bedingungen führen dazu, dass ihr Querschnitt für H I–Absorption größer ist, weshalb sie bevorzugt ausgewählt wurden.

Diese Dissertation liefert neue Erkenntnisse darüber, wie Galaxien molekulares Gas über die kosmische Zeit und in unterschiedlichen Umgebungen erwerben, behalten und verbrauchen. Sie setzt einen Maßstab für die Untersuchung von kaltem molekularem Gas in Galaxien und hebt wichtige zukünftige Projekte hervor. Gegen Ende gibt diese Arbeit einen Überblick über das derzeitige Verständnis von Gas in verschiedenen Phasen, einschließlich atomarer, ionisierter und molekularer Komponenten, unter Verwendung sowohl von Beobachtungen als auch von Simulationen. Die Schlussfolgerungen betonen die Bedeutung von Multiwellenlängendaten und kommender Einrichtungen sowie Upgrades und skizzieren einen klaren Weg, um die Rolle und Entwicklung von molekularem Gas in Galaxien über Milliarden von Jahren hinweg zu erforschen.

Abstract

This thesis investigates the distribution and evolution of cold molecular gas, the primary fuel for star formation, and its role in shaping galaxy evolution. This work is based on two major surveys conducted with the Atacama Large Millimetre/submillimetre Array observatory (ALMA): ALMACAL and MUSE–ALMA Haloes, which together provide a broad and detailed view of the gas flow in and out of galaxies over the last 12 billion years.

The first part of this thesis focuses on ALMACAL, a wide-area, deep millimetre/ submillimetre survey built from reprocessed ALMA calibration scans. Using over 2000 hours of integration time across more than 1000 calibrator fields, the ALMACAL−22 release demonstrates how calibration data can be repurposed to produce high-quality data cubes. ALMACAL−22, built from randomly distributed pointings, minimises cosmic variance and provides a deep, expanding dataset for investigating dusty star-forming galaxies, extragalactic absorption lines, active galactic nucleus physics, and the evolution of molecular gas.

The second part exploits ALMACAL's statistical power to build a CO-selected sample of galaxies, free from optical pre-selection biases. By applying an emission-line classification method, this work derives new constraints on the CO luminosity function and the cosmic molecular gas mass density ($\rho_{\mathrm{H_2}}$). The results reveal a rise in $\rho_{\mathrm{H_2}}$ from $z \sim 0$ toward $z \sim 1.5$, followed by a decline at higher redshifts. This evolution is consistent with the cosmic star formation rate density evolution, well established in the literature through a variety of measurements. These trends support a baryon cycling scenario in which cold gas is continuously replenished, with gas depletion times remaining relatively constant over time, implying stable star formation efficiencies on cosmological scales.

The third part investigates molecular gas in galaxies selected through Lyα absorption by neutral hydrogen gas (H I) seen along the line of sight. Using observations from an ALMA Large Program (Cycle 10, PI: C. Péroux) targeting these systems at $z \sim 0.5$, molecular gas

is detected in emission in about a quarter of the sample. These galaxies tend to have lower star formation rates and longer gas depletion times than typical star-forming galaxies. The findings suggest these systems are in a regulated or transitional evolutionary phase, shaped by processes such as gas accretion, feedback, and low metallicity conditions in the interstellar medium. These conditions cause their cross-section for H I absorption to be larger, which is why they were preferentially selected.

This thesis provides new insights into how galaxies acquire, retain, and deplete molecular gas across cosmic time and diverse environments. It sets a benchmark for the study of cold molecular gas in galaxies and highlights key future projects. Towards the end, this dissertation reviews the current understanding of multiphase gas, including atomic, ionised, and molecular components, using both observations and simulations. The conclusions look ahead to emphasise the importance of multi-wavelength data and upcoming facilities and upgrades, outlining a clear path for probing the role and evolution of molecular gas in galaxies over billions of years.

Chapter 1

Introduction

Galaxies are the fundamental building blocks of the observable Universe, composed of stars, gas, dust, and dark matter. They come in many shapes and sizes, from small, faint dwarfs to huge spirals like our Milky Way. Far from being still, galaxies grow, change and interact over time. They grow by pulling in gas, forming new stars, merging with other galaxies, and sometimes blowing gas back out into space. These ongoing processes shape how galaxies look and behave over billions of years. The following sections introduce key aspects of galaxy evolution: the mechanisms that govern their formation and growth (§1.1), the role of baryons in regulating this evolution (§1.2), the relevance of the circumgalactic medium (§1.3), the molecular gas as the primary fuel of star formation (§1.4), and finally, a summary of the specific surveys and science goals presented in this thesis (§1.5).

1.1 The formation and evolution of galaxies

1.1.1 The Cosmological Framework of Galaxy Formation

The prevailing framework for describing the cosmic structure formation and the evolution of galaxies is the Lambda Cold Dark Matter (ΛCDM) cosmological model (Blumenthal et al. 1984; Spergel et al. 2003; Planck Collaboration et al. 2020). While it has achieved remarkable success in explaining a broad range of observations, it remains incomplete. According to this framework, only about 5% of the Universe's matter-energy content is made of ordinary baryonic matter. As shown in Figure 1.1, the remaining 95% is attributed to two elusive components: the cold dark matter (CDM), which drives the gravitational

collapse of structures; and dark energy (Λ), responsible for the accelerated expansion of the Universe. The imbalance between the visible and invisible content of the Universe makes the study of galaxies particularly challenging.

After the Big Bang, the Universe was an extremely hot and dense plasma, a mixture of free electrons, protons, and photons. In this state, photons were constantly scattered by free electrons, keeping matter and radiation tightly coupled, so the Universe was opaque to light. As the Universe expanded and cooled, electrons and protons combined into neutral hydrogen atoms, a process known as recombination. This transition marked the decoupling of matter and radiation, allowing photons to travel freely through space for the first time. The relic radiation from this epoch is observed today as the Cosmic Microwave Background (CMB, Penzias & Wilson 1965). The CMB exhibits slight anisotropies (tiny fluctuations in temperature) which reflect density perturbations in the early Universe (Smoot et al. 1992). These perturbations served as the seeds for the formation of large-scale structures through gravitational instability (Peebles 1982).

During the so-called "dark ages" the Universe lacked sources of light, and gravity pulled together small irregularities in matter, leading to the hierarchical growth of structures (Peebles 1982; Springel et al. 2005). Dark matter began to collapse under its own gravity, forming the first non-linear structures. Baryonic matter, initially in the form of a nearly uniform, ionised gas, fell into the gravitational potential wells of dark matter. This led to the formation of the cosmic web, a vast network of filaments and voids that defines the large-scale structure of the Universe (Bond et al. 1996).

The first stars, known as Population III stars, formed from the pristine gas composed primarily of hydrogen and helium (Bromm & Larson 2004). Although still observationally elusive, they likely played a crucial role in the reionisation of the intergalactic medium (IGM) and enriched it with the first heavy elements through supernova explosions (Bromm et al. 2009; Salvadori & Ferrara 2012; Klessen & Glover 2023). These processes laid the groundwork for the formation of later generations of stars and galaxies, providing the chemical elements and radiation backgrounds that shaped the evolution of cosmic structures over time (Sibony et al. 2022; Riaz et al. 2022; Chowdhury et al. 2024).

The foundational framework for galaxy formation was established in the late 1970s (Rees & Ostriker 1977; White & Rees 1978). A two-stage model was proposed wherein dark matter collapses to form haloes, and baryonic gas subsequently falls into these potential wells. As the gas accretes, it heats up due to shocks and compression, forming a hot, diffuse

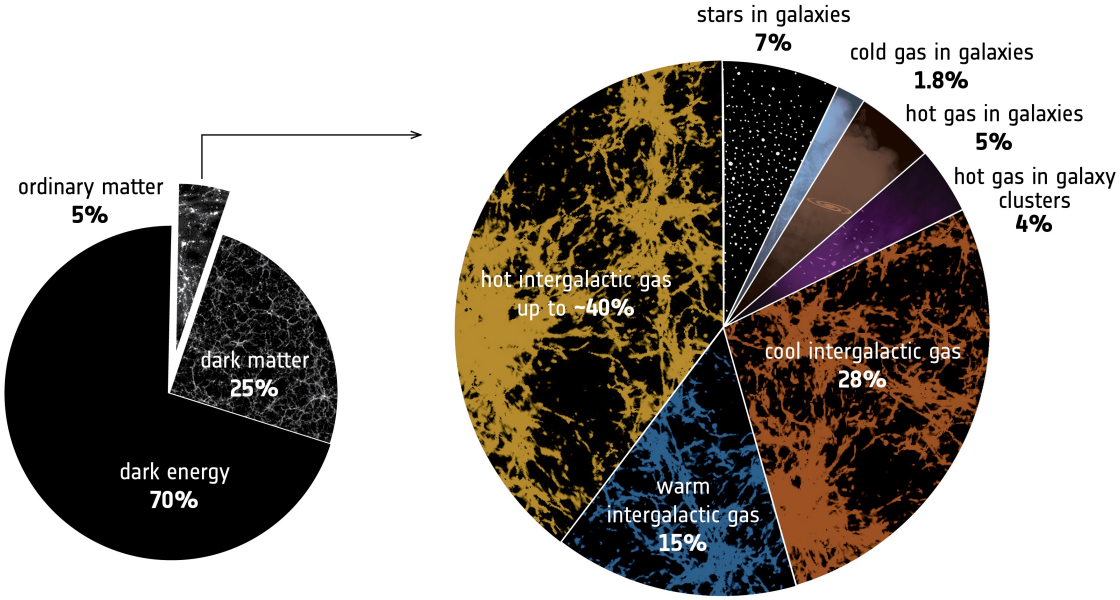

Figure 1.1: *Left:* It illustrates the overall composition of the Universe, dominated by dark energy (70%) and dark matter (25%), with only ∼5% consisting of ordinary (baryonic) matter. *Right:* It zooms into the baryonic matter, showing its breakdown across different baryonic phases. A large fraction of ordinary matter resides outside galaxies: cool (28%), warm (15%), and hot (up to 40%) intergalactic gas. Only a small portion is found in galaxies as stars (7%), cold gas (1.8%), and hot gas (5%), with additional contributions from hot gas in galaxy clusters (4%). This Figure, extracted from ESA/XMM-Newton findings, highlights the significance of diffuse intergalactic gas in the cosmic baryon budget, confirmed observationally by Nicastro et al. (2018).

atmosphere. Over time, the gas cools radiatively, condensing into the central regions of the halo where it can fragment and settle into a rotating disk or dense region, where stars and galaxies begin to form.

Numerical simulations over the last two decades have challenged this simplified view (e.g., Perivolaropoulos & Skara 2022; Efstathiou 2025). They show that gas accretion can occur in at least two different modes, depending on the mass of the haloes, redshift, and environment. First, the cold mode accretion suggests that filamentary streams from the cosmic web feed gas into galaxies without being shock-heated, allowing rapid and direct fuelling of star formation. This mode is particularly dominant in low-mass haloes. Conversely, the hot mode accretion occurs when gas is shocked-heated to a temperature exceeding $\sim 10^6$ K when entering massive haloes. This process forms a hot atmosphere that cools more slowly and provides a reservoir for later star formation (Kereš et al. 2005; Dekel et al. 2009; van de Voort et al. 2011; Tollet et al. 2022). These accretion modes have been confirmed and refined in cosmological simulations. The Millennium Simulation (Springel et al. 2005) played a foundational role in establishing the statistical framework of galaxy formation and large-scale structure evolution within the ΛCDM paradigm. Using high-resolution dark matter simulations combined with semi-analytic models, it set the standard for modern cosmological simulations (Kitzbichler & White 2007; Boylan-Kolchin et al. 2009). The simulation was run with GADGET-2, an N-body code for cosmological structure formation, whose successors GADGET-3 are used in simulations like Illustris and Magneticum (Springel et al. 2018; Dolag et al. 2016). The latest public release, GADGET-5, features improved hydrodynamics and scalability, enabling more accurate modelling of galaxy formation processes (Weinberger et al. 2020).

Building on these advances, a new generation of cosmological simulations now couples dark matter, gas dynamics, and baryonic physics with unprecedented detail. The IllustrisTNG project (Nelson et al. 2018; Springel et al. 2018) tracks dark matter, gas, stars, and black holes, including detailed prescriptions for stellar winds, feedback, and chemical enrichment, across volumes up to 300 Mpc per side. EAGLE (Hopkins et al. 2014; Crain et al. 2015) focuses on reproducing realistic galaxy stellar masses, sizes, and colours by carefully calibrating feedback processes to match low-redshift observables. FIRE (Schaye et al. 2015) takes a high-resolution "zoom-in" approach, directly resolving the multi-phase interstellar medium and stellar feedback without relying heavily on subgrid tuning. SIMBA (Davé et al. 2019) builds on the MUFASA code (Davé et al. 2016), with a strong empha-

sis on feedback modes and dust production. The Magneticum Pathfinder (Dolag et al. 2016) includes detailed modelling of hot gas in massive haloes and and the intracluster medium. More recently, the MillenniumTNG simulation (Hernández-Aguayo et al. 2023; Pakmor et al. 2023) combines the vast cosmological volume of Millennium with the sophisticated galaxy formation and feedback model from IllustrisTNG. These simulations have progressed from pure dark matter frameworks to full-physics galaxy formation models, capable of tracking dark matter, gas dynamics, star formation, black hole growth, chemical enrichment, and multi-phase feedback processes; providing a more complete picture of galaxy evolution across cosmic time.

The gas falling into galaxies is affected by different heating and cooling processes. As filaments fall into dense regions, they are heated by gravitational compression and shocks. At the same time, ultraviolet light from early stars and quasars heats and ionises the gas between galaxies (Haardt & Madau 1996; Schroetter et al. 2015; Meyer et al. 2015; Sharma et al. 2025). In denser areas, the gas can cool down by radiating energy, allowing it to clump together and form stars (Shangguan et al. 2020; Schinnerer & Leroy 2024). These interactions determine the efficiency of star formation and the morphological outcome of the galaxies that eventually form (Voit 2005; Pelupessy & Papadopoulos 2009; Gozaliasl et al. 2025; McClymont et al. 2025).

However, gas accretion alone does not explain the observed properties of galaxies (Silk & Mamon 2012). Feedback mechanisms, particularly from supernovae and active galactic nuclei (AGN), regulate the star formation process by heating or expelling gas from galaxies. Feedback processes prevent the overcooling of gas and are required to reproduce the observed stellar mass function, galaxy colours, and morphologies. Without feedback, simulations show galaxies would be much bigger and form stars too quickly compared to real observations (Nelson et al. 2019). Since many details happen on very small scales, cosmological simulations cannot directly model the physics of the interstellar medium. Instead, they use simplified "subgrid" models to represent the effects of feedback, such as radiative, kinetic, or thermal processes, based on local conditions (Springel & Hernquist 2003; Crain et al. 2015; Hopkins et al. 2018). The exact nature of feedback is still uncertain, which introduces some degree of uncertainty into the simulation results. Ongoing efforts to improve these models include an advanced treatment of stellar feedback, magnetic fields, and cosmic rays (Kim et al. 2018; Buck et al. 2020; Ji et al. 2020; Butsky et al. 2020, 2022; Dashyan & Dubois 2020; Farcy et al. 2022; Hopkins et al. 2023). These improvements aim

to address the theoretical challenges in current galaxy formation models (Naab & Ostriker 2017).

In summary, the ΛCDM cosmology presents a robust framework that predicts the existence and structure of the CMB, the expansion and the large-scale structure of the Universe, and the observed abundance of light elements, such as hydrogen and helium. Observational progress across multiple wavelengths, including X-ray, optical, and sub-millimetre, continues to refine and test theoretical models (e.g., Gruen et al. 2015, 2019; Bonavera et al. 2021; Gruen et al. 2023; Boylan-Kolchin 2023; Cook 2024; Cerardi et al. 2024; Peebles 2025; HSC Collaboration et al. 2025). For the purposes of this thesis, the focus is placed on optical observations and sub-millimetre interferometry, which together enable the study of galaxy formation and evolution.

1.1.2 Interstellar medium

The interstellar medium (ISM) is a complex mixture of gas and dust, affected by cosmic rays, magnetic fields and radiation (Saintonge & Catinella 2022). Despite representing a small fraction of the mass of the galaxies ($\sim 10 - 20\%$), the ISM plays a crucial role in regulating star formation. Stars are born from the ISM, and then they return energy and enriched material through stellar winds, supernovae, and other feedback mechanisms (Shapiro & Field 1976; Chevance et al. 2021; Ruffa et al. 2022; Michałowski et al. 2024). Understanding the ISM is therefore fundamental to studying the evolution of galaxies.

Hydrogen dominates the ISM, constituting about 70% of its mass, followed by helium at around 28%, with heavier elements, also known as "metals", making up the remaining 2% (Spitzer 1978; Klessen & Glover 2016). About 1% of the ISM mass is in solid form as dust grains composed primarily of silicates and carbonaceous compounds. These dust particles are well mixed with the gas and serve as key catalysts in the formation of molecules like H_2 (Draine 2011). They also shield molecular regions from dissociating ultraviolet (UV) radiation, thereby allowing molecular hydrogen to form. Dust grains play a crucial role not only in the chemistry of the ISM but also in its thermal balance and the appearance of galaxies in observations. Dust absorbs UV light and re-emits it in the far-infrared (FIR) range, which serves as a useful proxy for tracing star formation. This re-radiated FIR emission provides a powerful tool for tracing star formation activity, especially in dusty or heavily obscured regions where direct UV light is blocked (Calzetti et al. 2000).

The ISM is structured into several phases, each with different temperatures and chemical properties, but generally maintained in roughly pressure equilibrium with one another (Field et al. 1969; McKee & Ostriker 1977). These can be classified based on temperature and ionisation state (Ferriere 2001; Tielens 2005; Draine 2011), as follows:

- **Molecular gas clouds** are the coldest and densest component of the ISM, with typical densities $10^2 - 10^6\,\mathrm{cm}^{-3}$ and temperatures around $10 - 20\,\mathrm{K}$. They represent a significant fraction the total mass of the ISM, and are the primary site of star formation, despite using only a 1% of the galaxy's volume.

- **Cold Neutral Medium** (CNM) is characterised by atomic hydrogen at temperatures ranging from 50 to 100 K and densities of $20 - 50\,\mathrm{cm}^{-3}$. This phase is found in filaments aligned with the galactic disk and forms a transition between the molecular regions and more diffuse media.

- **Warm Neutral Medium** (WNM) consists of less dense atomic hydrogen with temperatures around $10^3 - 10^4\,\mathrm{K}$ and densities of $\sim 0.5 - 1\,\mathrm{cm}^{-3}$. It fills up to $\sim 30\%$ of a galaxy's volume and often lies at the interface between molecular clouds and ionised regions. This component is often traced through the emission of atomic hydrogen (H I).

- **Warm Ionised Medium** (WIM) corresponds to diffuse ionised hydrogen at similar temperatures of $\sim 10^4\,\mathrm{K}$ but lower densities ranging $\sim 0.1 - 1\,\mathrm{cm}^{-3}$. This gas is primarily photoionised by UV radiation from young stars, but shocks and collisional processes also contribute. The WIM can fill $\sim 10 - 30\%$ of a galaxy's volume and plays a key role in mediating energy and momentum transfer from stellar feedback.

- **Hot Ionised Medium** (HIM) is created by supernova explosions and stellar winds, being the most energetic phase. It contains gas at a temperature of $10^5 - 10^6\,\mathrm{K}$ and extremely low densities of $< 0.01\,\mathrm{cm}^{-3}$. This hot gas can fill up to 50% of a galaxy's volume.

These phases are not isolated, and there is a continuous cycle of matter among them (McKee 1995). Gas accretion from the intergalactic medium typically arrives in a warm state and cools to become neutral. This neutral gas may then condense into molecular clouds under favourable conditions. Conversely, energetic processes such as supernovae or

Table 1.1: Summary of Interstellar Medium (ISM) phases

Phase	T (K)	$n_{\rm H}$ (cm^{-3})	f_V
Molecular Gas (H_2)	10–20	10^2–10^6	$\sim 1\%$
Cold Neutral Medium (CNM)	50–100	20–50	1–5%
Warm Neutral Medium (WNM)	10^3–10^4	0.5–1	$\sim 30\%$
Warm Ionised Medium (WIM)	$\sim 10^4$	0.1–1	$\sim 20\%$
Hot Ionised Medium (HIM)	10^5–10^6	$\sim 10^{-4}$	$\sim 50\%$

AGN activity can heat or ionise gas, driving it into hotter, more diffuse states or even expelling it from the galaxy. Table 1.1 summarises the ISM phases, with the typical temperatures (T), hydrogen number density ($n_{\rm H}$), and approximate volume filling factor (f_V) within galaxies.

In the context of this thesis, particular emphasis is placed on the molecular gas and H I seen in absorption. These phases are not only fundamental to star formation but also critical for interpreting the gas conditions in galaxies across different cosmic epochs.

1.1.3 Galaxy growth and star formation

The growth of galaxies over cosmic time is primarily driven by their ability to convert gas into stars. Star formation is tightly correlated with the presence of cold molecular gas, predominantly molecular hydrogen (H_2), which forms in dense regions of the ISM (Kennicutt & Evans 2012). Observational studies have shown a power-law relationship between the surface density of star formation ($\Sigma_{\rm SFR}$) and the surface density of gas ($\Sigma_{\rm gas}$), commonly referred to as the Kennicutt-Schmidt relation (Schmidt 1959; Kennicutt 1989, 1998):

$$\Sigma_{\rm SFR} = A\Sigma_{\rm gas}^{N} \tag{1.1}$$

where A is the normalization and N the power-law slope. This relation, with an index of approximately 1.4, has been revisited multiple times and holds across a wide range of spatial scales and galaxy types at low-redshift, as shown in the left panel of Figure 1.2 (Kennicutt & Evans 2012; Heiderman et al. 2010; de los Reyes & Kennicutt 2019;

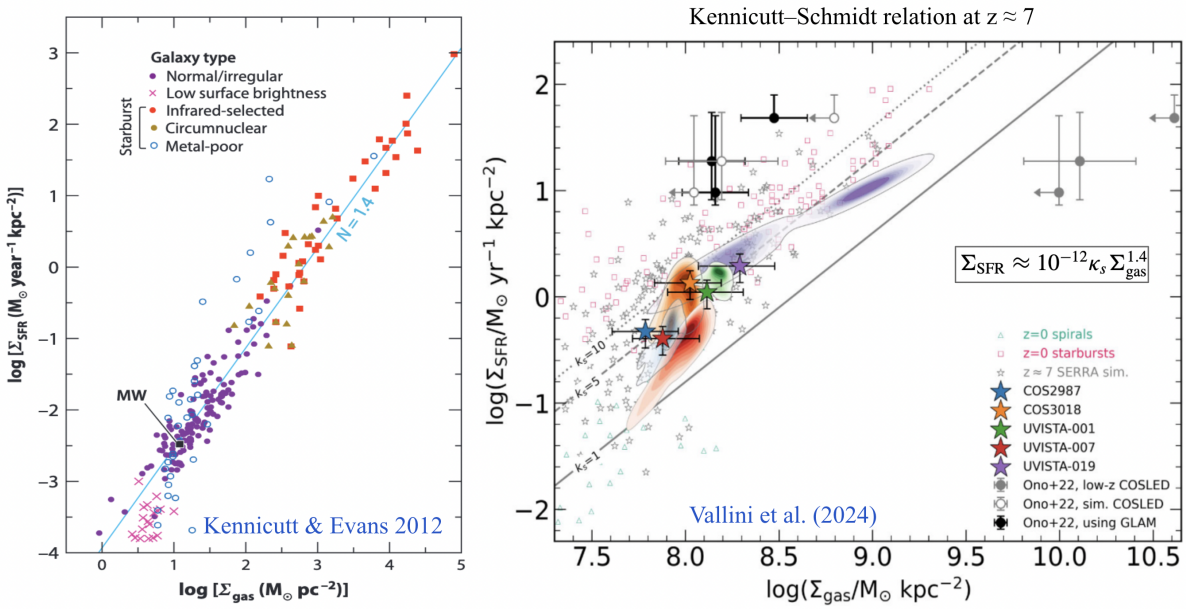

Figure 1.2: *Left:* It shows the relationship between star formation rate and gas surface density across different types of star-forming galaxies (normal spirals, starbursts including LIRGs and ULIRGs, circumnuclear starbursts, and low-metallicity galaxies). Most galaxies follow a common trend, including the Milky Way (MW), while low-metallicity systems systematically deviate (magenta crosses). The light blue line shows a slope $N = 1.4$ (not a fit). Extracted from (Kennicutt & Evans 2012). *Right:* It shows the relation between SFR and total gas surface densities, following the Kennicutt–Schmidt law scaled by the burstiness parameter $\kappa_s = 1, 5, 10$ (solid, dashed and dotted lines). Local spiral and starburst galaxies are plotted with green triangles and magenta squares. High-redshift ($z \sim 7$) galaxies from the SERRA simulation (Pallottini et al. 2017) are shown as grey stars, with pixel-by-pixel SFR–gas distributions represented by shaded regions and global values by colored stars. Three galaxies at $z \sim 6$ are shown by the black circles with different CO excitation and conversion factors. Figure extracted from Vallini et al. (2024).

Kennicutt & De Los Reyes 2021). At high redshift, many galaxies deviate from the local relation, often showing elevated SFRs relative to their gas surface densities, as shown in the right panel of Figure 1.2 for galaxies at $z \sim 7$. This deviation is quantified by the burstiness parameter κ_s, introduced by Ferrara et al. (2019), where $\kappa_s > 1$ indicates a starburst regime. Observations of massive, rare submillimetre galaxies at $z > 1$ have found

extreme values of $\sim 10-100$ (e.g., Hodge et al. 2014; Chen et al. 2017), suggesting intense, concentrated star formation. However, direct measurements of gas surface density in more typical galaxies remain difficult due to the faintness of emission lines in low-metallicity environments and the rising CMB temperature (Bolatto et al. 2013; Vallini et al. 2015). To overcome this, recent studies have proposed an indirect method to estimate the gas surface density, by linking it to the relative surface brightness of neutral (e.g. [C II] 158 μm) and ionised (e.g. [O III] 88 μm) line emissions (Ferrara et al. 2019; Vallini et al. 2021, 2024). These tracers can be spatially resolved with ALMA and the James Webb Space Telescope (JWST), enabling new constraints on star formation efficiency in the early universe (e.g., Akins et al. 2022; Herrera-Camus et al. 2022; Molyneux et al. 2022; Hsiao et al. 2024; Posses et al. 2023, 2025).

Galaxies grow not only through star formation but also by accreting cold gas from the intergalactic and circumgalactic media and merging with other galaxies, consistent with the hierarchical model of structure formation (e.g., White & Frenk 1991; Conselice 2014). Mergers and gas accretion can enhance star formation, drive morphological transformations, and fuel supermassive black hole growth (e.g., Springel et al. 2005; De Lucia et al. 2006; Avila-Reese et al. 2014). However, observations of galaxies from surveys like CANDELS and JADES show that major mergers are not the dominant mechanism for spheroid formation or star formation at $z > 1$ (Lofthouse et al. 2017; Puskás et al. 2025). Similarly, large cosmological simulations (e.g., Horizon-AGN, HELLO, NIHAO, Kaviraj et al. 2015; Wang et al. 2015; Waterval et al. 2024) confirm that while major mergers occur with moderate frequency, their contribution to stellar mass growth is limited, often accounting for only a few percent directly, with merger-driven starbursts representing a minor fraction ($\sim 2\%$) of star-forming galaxies and about 10% of the cosmic SFR density at $z \sim 2$. Instead, smooth cold gas accretion via filaments plays a critical role, fuelling high star formation rates ($> 100\,M_\odot\,\mathrm{yr}^{-1}$) in massive galaxies at $z \sim 2-4$, with simulations predicting that over 75% of the star-forming gas is supplied by this mode (Dekel et al. 2009; Kaviraj et al. 2015; Tollet et al. 2019, 2022; Aung et al. 2024; Waterval et al. 2025).

A central observational tool in studying galaxy evolution is the *galaxy main sequence* (MS), shown in the left panel of Figure 1.3. The main sequence represents a tight empirical correlation between stellar mass ($M_\star$) and star formation rate (SFR) observed in star-forming galaxies, suggesting a steady, self-regulated star formation (Brinchmann et al. 2004; Noeske et al. 2007; Daddi et al. 2007; Speagle et al. 2014; Whitaker et al. 2014; Leslie

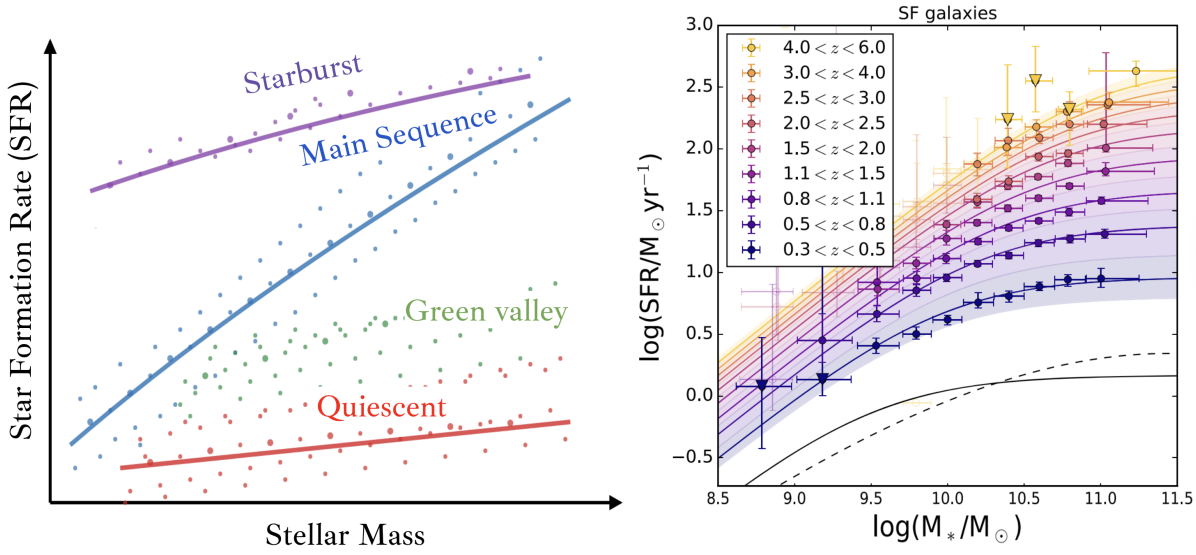

Figure 1.3: *Left*: A schematic diagram illustrating where galaxies tend to lie in the star formation rate versus stellar mass plane. Most galaxies lie along the star-forming main sequence (blue), while a significant population resides in the quiescent region (red), characterised by little to no ongoing star formation. A smaller number of galaxies occupy the starburst region (purple), forming stars at elevated rates, or lie in the green valley, representing a transitional phase between active and passive evolution. *Right:* The evolution of the main-sequence from low redshift ($0.3 < z < 0.5$) to high redshift ($4 < z < 6$), extracted from Leslie et al. (2020). The normalisation of the main sequence increases toward high redshift, implying that galaxies form more stars at a given mass. The slope of the main sequence flattens at high masses, especially at low redshift, possibly due to quenching by AGN feedback or other mechanisms.

et al. 2020; Thorne et al. 2021). Galaxies lying significantly above the MS are classified as *starbursts*, undergoing intense, short-lived star formation episodes, often triggered by mergers or disk instabilities (e.g., Sanders & Mirabel 1996; Hopkins & Beacom 2006; Kennicutt 1998). Below the MS lies the *green valley*, populated by galaxies transitioning from active to passive phases, likely quenching due to AGN feedback, morphological transformation, or environmental effects (Martin et al. 2007; Smethurst et al. 2015). At the low-SFR end are *quiescent galaxies*, often massive ellipticals with old stellar populations, having formed most of their stars at high redshifts ($z > 3$) and quenched through feedback and merging

(De Lucia et al. 2006; Croton et al. 2006; Whitaker et al. 2021; Gozaliasl et al. 2025). The scatter in the MS provides insights into the star formation histories: a small scatter (e.g., ∼0.1 dex) suggests smooth, continuous star formation, while a larger scatter (e.g., ∼0.4 dex) reflects more bursty or stochastic activity (e.g., Guo et al. 2013; Schreiber et al. 2015; Tacchella et al. 2016; Caputi et al. 2024).

The right panel of Figure 1.3 shows the evolution of the main sequence with redshift, extracted from Leslie et al. (2020). Toward low redshift ($z \lesssim 1$), the normalisation of the main sequence declines steadily, meaning that galaxies form fewer stars at a given mass. At higher redshifts ($z \gtrsim 2-4$), galaxies at a given stellar mass have much higher SFRs, indicating they form stars more actively. The slope of the main sequence is typically close to linear, but may flatten at high masses, especially at low redshift, possibly due to quenching by AGN feedback or other mechanisms (Schreiber et al. 2015; Whitaker et al. 2014; Thorne et al. 2021; Whitaker et al. 2021).

The integrated history of star formation is captured in the cosmic star formation rate density (SFRD), which traces the average SFR per unit volume across cosmic time. The SFRD peaked around redshift $z \sim 2$, a period also known as the "cosmic noon", and has since declined by over an order of magnitude. The SFRD was first presented by Madau et al. (1996) and Lilly et al. (1996), complemented further by UV dust-corrected and infrared (IR) observations (e.g., Steidel et al. 1999; Flores et al. 1999; Rowan-Robinson 2001). A widely used analytical fit was presented by Madau & Dickinson (2014), has since been updated with more recent measurements using UV, optical, and IR measurements (e.g., Bouwens et al. 2015; Smit et al. 2016; Katsianis et al. 2017; Bouwens et al. 2021; Katsianis et al. 2021; Zavala et al. 2021; Bollo et al. 2023; Covelo-Paz et al. 2024; Sun et al. 2025; Fu et al. 2025b). As shown in the left panel of Figure 1.4, the SFRD is made up of two components: the unobscured star formation (traced by UV/optical light) and the obscured star formation (traced by IR emission). At $z \lesssim 2.5$, the obscured star formation dominates, contributing around 70–80% of the total star formation. Beyond this point, dust-obscured star formation drops quickly, reflecting the strong decline in the number of infrared-luminous galaxies. By $z \sim 4$, the dust-obscured and unobscured (UV/optical) star formation rates become similar. At higher redshift ($z \gtrsim 4$), dust-obscured star formation becomes less dominant than unobscured star formation. This trend results from the relatively flat faint-end slope and rapid redshift evolution of the IR luminosity function compared to the UV one. By $z \sim 5$, the obscured star formation accounts for roughly $\sim 35\%$ of the total SFRD, dropping

to about $\sim 25\%$ at $z \sim 6$. Since the obscured contribution is mainly driven by massive, IR-bright galaxies, this decline likely reflects the scarcity of massive galaxies at these early epochs. (Zavala et al. 2021). Ultimately, the evolution of the SFRD reflects the declining availability of cold gas reservoirs, changing star formation efficiencies, and the evolving role of feedback processes.

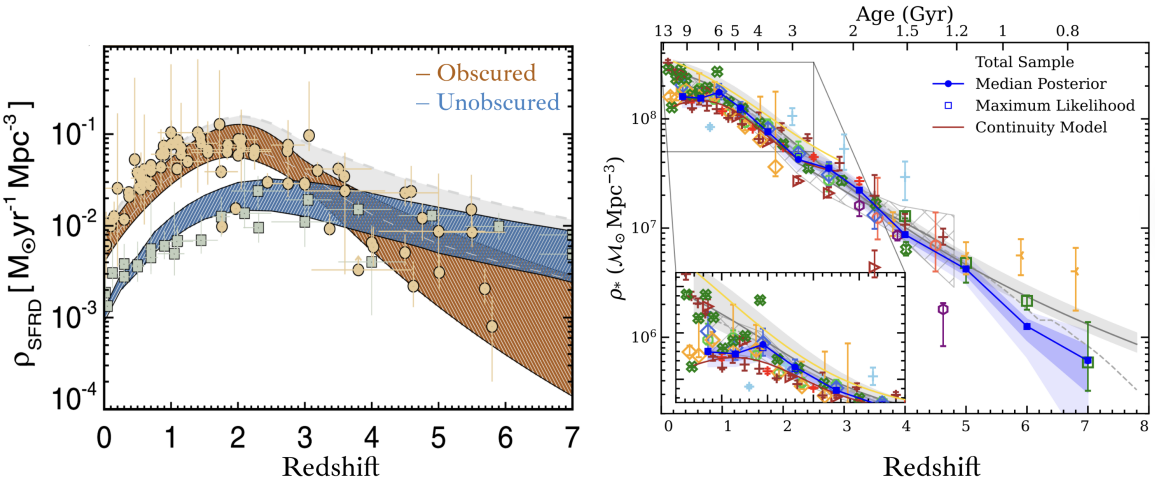

Figure 1.4: Cosmic star formation rate density (SFRD, left) and the galaxy stellar mass function (GSMF, right), extracted from Zavala et al. (2021) and Weaver et al. (2023). *Left:* The SFRD is separated into unobscured (UV-based, blue), obscured (IR-based, red) components. The UV observations trace star formation directly visible in the ultraviolet, while the IR curve accounts for star formation hidden by dust, which re-radiates absorbed UV light in the infrared. At $z \sim 0$–3, the dust-obscured star formation dominates the total SFRD. At higher redshifts ($z > 4$), the unobscured component becomes more significant. *Right:* Evolution of the cosmic stellar mass density from $0.2 < z \leq 7.5$, based on best-fit likelihood models (blue curve) integrated above $10^8\,M_\odot$ (Weaver et al. 2023). The GSMF represent the integral of the SFRD assuming a Chabrier initial mass function (IMF, Chabrier 2003) and a return fraction of $R = 0.41$, which is the fraction of the stellar mass that is immediately returned to the gas when massive stars explode.

A complementary statistical measure is the galaxy stellar mass function (GSMF), which quantifies the number density of galaxies as a function of stellar mass and redshift. The right panel of Figure 1.4 shows the derived best-fit galaxy stellar mass functions, integrated

above $10^8\,M_\odot$ from Weaver et al. (2023), and compares them to a range of observational results from mass-selected galaxy samples in the literature (Caputi et al. 2011; Santini et al. 2012; Ilbert et al. 2013; Muzzin et al. 2013; Tomczak et al. 2014; Grazian et al. 2015; Wright et al. 2018; Adams et al. 2021; McLeod et al. 2021b; Thorne et al. 2021). The abundance of massive galaxies ($M_\star > 10^{11} M_\odot$) remains nearly constant since $z \sim 1$ (Cedrés et al. 2025), while low-mass galaxies have increased in number toward lower redshifts, consistent with hierarchical mass assembly (Pozzetti et al. 2010; McLeod et al. 2021a). At higher redshifts ($z > 4$), the low-mass slope of the GSMF becomes steeper, indicating rapid growth of low-mass galaxies (González et al. 2011; Weaver et al. 2023).

An important trend in galaxy evolution is the "downsizing" phenomenon, wherein more massive galaxies form their stars earlier and faster than less massive ones (Cowie et al. 1996; Thomas et al. 2005; Mouri & Taniguchi 2006; Spitoni et al. 2020). This trend seems inconsistent with the bottom-up nature of ΛCDM structure formation, but may be explained by more efficient early cooling and higher gas accretion rates in massive haloes (Dekel et al. 2009; Dekel & Birnboim 2006).

Galaxies eventually transition from the star-forming sequence to a quiescent phase, where star formation ceases through various mechanisms. In massive haloes ($M_{\rm halo} \gtrsim 10^{12} M_\odot$), gas is shock-heated and cannot cool efficiently, this process is known as halo quenching (Birnboim & Dekel 2003; Zu & Mandelbaum 2016; Fu et al. 2025a). Additionally, AGN can suppress star formation by heating or expelling cold gas (Croton et al. 2006; Fabian 2012; Harrison et al. 2019; Lammers et al. 2023; Scharré et al. 2024). On smaller scales, morphological quenching may occur when the central bulge stabilises the disk against fragmentation (Martig et al. 2009; Lu et al. 2021; Oxland et al. 2024; Su et al. 2025; Gozaliasl et al. 2025). Environmental quenching, such as ram-pressure stripping in dense clusters, also contributes to the cessation of star formation (Peng et al. 2010; Hatfield & Jarvis 2017; Fillingham et al. 2018; Hamadouche et al. 2024).

Taken together, these observational and theoretical insights provide a detailed picture of how galaxies grow: through the regulated conversion of cold gas into stars, regulated by internal processes (e.g., stellar feedback, bulge growth) and external environmental factors (e.g., gas accretion, gas stripping). Galaxy growth is thus not uniform but shaped by a complex interplay between gas availability, cosmic time, and the galaxy's own mass and structure.

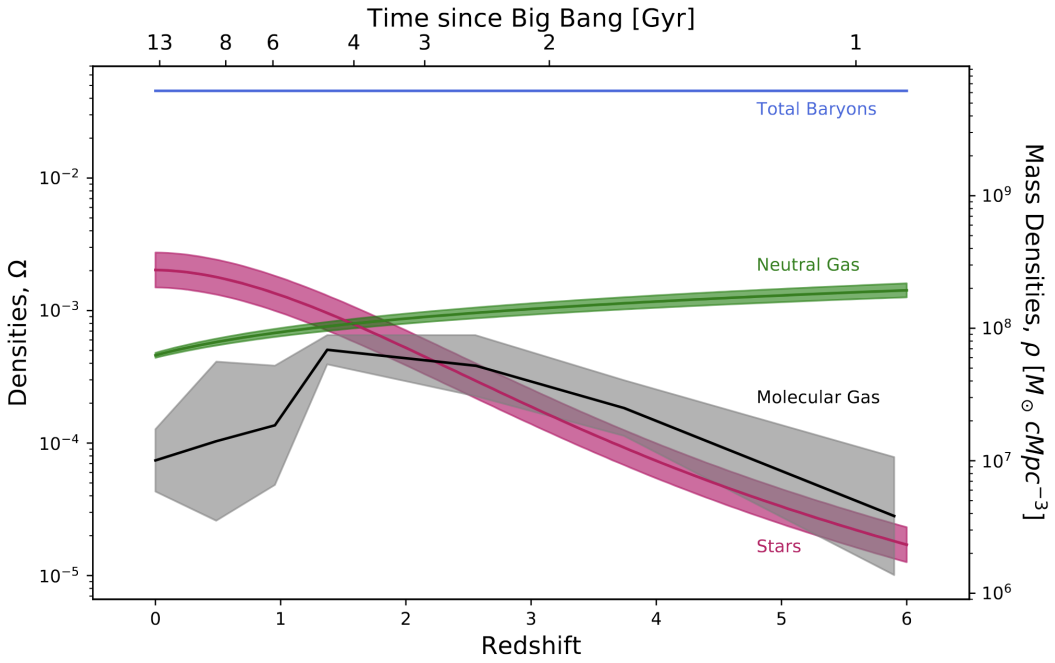

Figure 1.5: The cosmic evolution of neutral gas (green), molecular gas (grey) and stellar mass (pink) densities, extracted from Péroux & Howk (2020). While the neutral gas density declines steadily from high redshift towards the present day, the molecular gas increases from the early Universe, reaching a peak around $z \sim 2$ and declines towards $z \sim 0$. The evolution of the molecular gas resembles the shape of the cosmic star formation rate density, shown in the left panel of Fig. 1.4.

1.2 The baryon cycle

The baryon cycle refers to the ongoing exchange of baryonic matter, including gas and metals, between galaxies and their surroundings: the circumgalactic medium (CGM) and the intergalactic medium (IGM). This cycling is fundamental for understanding galaxy mass growth, star formation histories, metallicity evolution, the dark matter distribution on small scales, and the distribution of baryons in the Universe (Oppenheimer & Davé 2006; Somerville & Davé 2015; Tumlinson et al. 2017; Péroux & Howk 2020). Unlike early closed-box models of galaxy evolution, the baryon cycle framework emphasises that galaxies are open systems, shaped by continuous inflows and outflows of mass, energy, and metals (Davé et al. 2012; Lilly et al. 2013). Understanding this cycle is crucial to resolving the "missing baryons" and "missing metals" problems (Pettini et al. 1999; Pagel 1999), where

the observed baryonic and metal content in stars and cold gas is significantly lower than the cosmological predictions. Much of the missing material is thought to reside in diffuse, multiphase gas that is difficult to detect directly (Ferrara et al. 2005; Bouché et al. 2005, 2006, 2007; Michael Shull et al. 2014)

This section reviews the key components of the baryon cycle: the multiphase structure of the gas (§1.2.1), mechanisms of gas accretion and inflow, the role of feedback-driven outflows, and the processes that govern the recycling of gas back into galaxies (§1.2.2).

1.2.1 Multi-phase gas

The gas in galaxies exists in three main phases: atomic (H I), molecular (H_2), and ionised (H II or hotter). Each phase has distinct physical properties, observational techniques, and a different role in the overall context of cosmic evolution.

Atomic gas (H I)

Atomic hydrogen (H I) resides primarily in the outer galactic disk and inner circumgalactic medium ($\lesssim$ 50 kpc), and dominates the mass budget of the neutral ISM. It is typically found in two thermal regimes: the cold neutral medium (CNM), with temperatures ranging from 50 to 200 K and densities of $10 - 100$ cm^{-3}, and the warm neutral medium (WNM), with temperatures of $6\,000 - 10\,000$ K and densities of $\sim 0.1 - 1$ cm^{-3}. These regimes result from the balance between photoelectric heating and radiative cooling due to line emission and dust interactions (Wolfire et al. 2003).

H I is primarily observed through the hyperfine 21 cm spin-flip transition. At $z \sim 0$, the WALLABY survey on ASKAP will be a major leap forward, mapping H I over $\sim$ 14,000 deg^2 out to $z \lesssim 0.1$, and detecting hundreds of thousands of galaxies (Koribalski et al. 2020). Its high resolution and sensitivity allow for resolved imaging of gas disks and tidal features, providing a definitive local H I mass function. This effort builds on the legacy of surveys like ALFALFA on Arecibo (Haynes et al. 2011; Jones et al. 2018), which catalogued over 30,000 H I sources, and HIPASS on Parkes (Zwaan et al. 2005), which provided the first blind H I census in the southern hemisphere. Additional surveys such as APERTIF on WSRT and deep fields with FAST further enrich the H I landscape at $z \sim 0$ (Adams et al. 2018; van Cappellen et al. 2022).

Going beyond the local universe, direct detection of H I becomes more challenging due

to cosmological dimming. Nevertheless, several ongoing projects are now pushing into the $z \sim 0.5 - 1$ regime. For example, CHILES (VLA) is a deep pointing in COSMOS integrating 178 hours to detect H I out to $z \sim 0.45$ (Fernández et al. 2016). It has revealed galaxies with disturbed H I morphologies at $z \sim 0.12 - 0.17$, linking gas dynamics to group-scale interactions. In parallel, the MIGHTEE-H I survey on MeerKAT is a deep ($\sim$20 deg^2) map that reaches $z \sim 0.6$. Simulations predict MIGHTEE-H I will directly detect $\sim$ 3,000 galaxies at $0 < z < 0.4$ and statistically measure H I out to $z \sim 0.6$, tracing five billion years of H I evolution (Maddox et al. 2021). Similarly, the LADUMA project (also on MeerKAT) will integrate over 3,000 hours on a single deep field to probe H I emission all the way to $z \sim 1.45$ via direct detection and stacking (Baker et al. 2018). At $z \lesssim 1$, stacking techniques and 21-cm intensity mapping are also used (e.g., CHIME, MeerKAT, Masui et al. 2013; Switzer et al. 2013; Santos et al. 2015; Bandura et al. 2014).

At $z \geq 2$, H I is observed through absorption in the Lyα transition against background quasars, especially in damped Lyα systems (DLAs), which dominate the high-redshift H I budget (out to $z \sim 5$) (Péroux et al. 2003; Prochaska & Wolfe 2009; Crighton et al. 2015; Oyarzún et al. 2025). The SKA and pathfinders (e.g. MeerKAT, ASKAP) will significantly expand our ability to detect H I in absorption via 21-cm line against radio-bright background sources. These surveys will measure spin temperatures, covering factors, and probe the cold neutral medium at high redshift (Morganti et al. 2015b). SKA-Low aims to detect 21-cm fluctuations from the Epoch of Reionisation ($z > 6$), revealing the distribution of diffuse H I on cosmological scales (Koopmans et al. 2015).

As shown in the green curve of Figure 1.5, the comoving H I density evolves slowly from $z \sim 4$ to $z \sim 0$, decreasing by a factor of $\sim$ 2 dex, and remaining roughly constant from $z \sim 1.5$ onward (Péroux & Howk 2020; Walter et al. 2020). This supports the idea that H I acts as a long-lived intermediary reservoir in the galaxy fuelling process.

Molecular gas (H_2)

Molecular hydrogen exists in the coldest, densest environment of galaxies, predominantly within giant molecular clouds (GMCs). Temperatures in these regions are typically between 10 and 30 K, and volume densities span from 100 to 10 cm^{-3} in the cloud interiors. These conditions promote the formation of molecules on dust grains, enabling gravitational collapse.

Direct observations of H_2 are difficult in cold gas due to its lack of a permanent dipole

moment, and it only emits weak quadrupole rotational transitions, which are generally unobservable in typical cold ISM conditions. Consequently, molecular gas in galaxies is primarily traced through indirect methods, including carbon monoxide (CO) rotational transitions, dust continuum emission, and fine-structure lines (especially [C II] 158 μm) observed with facilities such as ALMA, NOEMA, and the VLA. These observatories have been crucial for establishing key scaling relations, such as the Kennicutt–Schmidt law, and for mapping molecular gas in galaxies, revealing structures like spiral arms, dense clouds, and bar-driven inflows that influence star formation (Leroy et al. 2009). The most widely used tracer is the CO(1−0) line at 115 GHz, which provides a basis for estimating molecular gas masses. However, converting CO luminosity to H_2 mass requires a CO-to-H_2 conversion factor (α_{CO}), which varies with metallicity, density, and excitation conditions (Bolatto et al. 2013). Further details will be introduced in Section 1.4. Beyond CO emission, H_2 can also be probed in other ways. In some cases, it is detected in absorption in DLAs along quasar sightlines, offering unique constraints on the molecular content and radiation fields of high-redshift galaxies (Noterdaeme et al. 2008; Balashev et al. 2019; Klitsch et al. 2021). Warm H_2 is directly observable via mid-infrared rotational lines, particularly in shocked or strongly irradiated regions. JWST/MIRI now enables observations of these transitions (e.g. S(0)–S(5)) out to $z \lesssim 3$, opening new windows on warm molecular gas in star-forming regions, AGNs, and galactic outflows (Appleton et al. 2023; Armus et al. 2023).

At low redshift ($z \sim 0$), the PHANGS–ALMA survey has mapped CO(2−1) at $\sim$ 100 pc resolution across 90 nearby star-forming galaxies, revealing how molecular cloud properties, star formation efficiency, and feedback vary with galactic environment (Leroy et al. 2021). The xCOLD GASS survey combined IRAM 30m and NOEMA data to build a representative sample of CO(1−0) measurements in $\sim$ 500 galaxies, establishing key scaling relations between molecular gas, stellar mass, and star formation (Saintonge et al. 2017). ALMaQUEST integrated ALMA CO imaging with MaNGA IFU spectroscopy to study the spatially resolved interplay between gas, star formation, and quenching across the galaxy population (Lin et al. 2020). The EDGE–CALIFA survey used NOEMA to link molecular gas to stellar populations and metallicity gradients in 125 galaxies (Bolatto et al. 2017). Together, these surveys provide a comprehensive view of the molecular ISM across diverse galaxy types and environments in the local Universe.

At intermediate redshift ($z \sim 1-3$), the ASPECS survey is a blind CO and continuum survey in the Hubble Ultra Deep Field (HUDF) at $z \sim 1-4$, offering the first census

of molecular gas mass density across cosmic time (e.g., Walter et al. 2016; Decarli et al. 2020; Aravena et al. 2020; Boogaard et al. 2020). The COLDz survey, a NOEMA and VLA effort, mapped CO(1−0) blindly at $z \sim 2-3$, constraining the CO luminosity function and the evolution of gas mass density (Pavesi et al. 2018; Riechers et al. 2019, e.g.,). GOODS-ALMA used deep 1.1 mm continuum imaging to detect dust-obscured galaxies at $z > 2$, many of which are missed in optical surveys (e.g., Franco et al. 2018). PHIBSS and its extension PHIBSS2 target tens to over one hundred main-sequence galaxies at $0.5 < z < 3$, using NOEMA to derive molecular gas fractions, depletion times, and excitation states (e.g., Tacconi et al. 2013; Freundlich et al. 2019; Barfety et al. 2025). The ALMA Calibrator (ALMACAL) survey used calibrator fields to search for molecular absorbers from $z \sim 0-2$, reducing biases from cosmic variance (Hamanowicz et al. 2022). More recent efforts include the ALMA Large program Cosmic Noon Disks with Observed Rotation (CONDOR, PI: Rizzo, Cycle 11), which will map CO gas in massive disk galaxies at $z \sim 1-3$ to study the roles of feedback, disk instabilities, and dark matter in shaping galaxies. In parallel, another ALMA Large Program, the ALMA Chemical Evolution (ACE, PI: Shivae, Cycle 11) survey will focus on low-metallicity galaxies at $z \sim 2$, observing CO and dust in 25 galaxies with metallicities as low as 35% of solar to trace how metals transition from gas to dust and to improve our understanding of the mass-metallicity relation at hiagh redshift.

At high redshift ($z > 4$), ALMA targets [C II] 158 μm emission as a tracer of the cold ISM. The ALPINE survey observed 118 UV-selected galaxies at $z = 4.4-5.9$, showing widespread [C II] emission often offset from UV light and revealing high gas fractions and rapid star formation cycles (e.g., Le Fèvre et al. 2020; Faisst et al. 2020; Dessauges-Zavadsky et al. 2020; Schaerer et al. 2020; Pozzi et al. 2024). The ALMA Reionisation Era Bright Emission Line Survey (REBELS) extended this to $z = 6.5-9$, detecting large amounts of dust and gas, which suggests early metal enrichment and rapid galaxy growth (e.g., Bouwens et al. 2022; Fudamoto et al. 2022; Bowler et al. 2024). More recently, the CRISTAL survey mapped [C II] in 19 galaxies at $z \sim 4.5$ with ~ 1 kpc resolution, revealing clumpy gas, rotation, signs of mergers, and evidence of early disk formation (e.g., Solimano et al. 2024; Villanueva et al. 2024; Herrera-Camus et al. 2025; Posses et al. 2025).

Overall, local and high-redshift surveys have mapped the evolution of molecular gas over cosmic time using different tracers. A compilation of the molecular gas mass density is shown in the pink curve of Figure 1.5. It peaks at $z \sim 1.5-2$, declining by a factor of ~ 6 towards the present (Péroux & Howk 2020). This decline resembles the drop in the

cosmic star formation rate density (Fig. 1.4, left panel), reinforcing the role of H_2 as the immediate fuel for star formation.

Ionised gas

Ionised gas spans a wide temperature range depending on its ionisation mechanism and environment. The warm ionised medium (WIM) is often photoionised by stellar UV radiation ($T \sim 10^4$ K, $n \sim 0.01 - 0.1$ cm^{-3}), while the hot ionised medium (HIM) is typically generated by supernovae and virial shocks ($T \sim 10^6 - 10^7$ K, $n \lesssim 10^{-3}$ cm^{-3}).

In the optical and near-infrared, ionised gas is observed via recombination lines (e.g., Hα, Hβ) and forbidden transitions (e.g., [O III] λ5007, [N II] λ6584, [S II] λ6717, 6731), which are used to estimate star formation rates, metallicities, electron densities, and excitation mechanisms. These lines also enable powerful diagnostic diagrams, such as Baldwin–Phillips–Terlevich (BPT diagram, Baldwin et al. 1981), which plots the ratios of specific emission lines, typically [O III] λ5007/Hβ versus [N II] λ6584/Hα. The diagram is widely used to distinguish between star formation and AGN-driven ionisation.

Among the most transformative instruments for studying ionised gas is the Multi Unit Spectroscopic Explorer (MUSE) on the Very Large Telescope (VLT). MUSE combines a wide field of view with integral field spectroscopy at high spatial resolution, enabling spatially resolved maps of ionised gas kinematics and line ratios across thousands of galaxies, from local systems out to $z \gtrsim 6$ (e.g., Bacon et al. 2015; Wisotzki et al. 2016; Inami et al. 2017). MUSE has been particularly effective at detecting faint Lyα emission from the CGM, mapping extended ionised haloes, and revealing galactic outflows and inflows with unprecedented detail. Further details on this technique will be addressed in §1.3.

Other key facilities include the Sloan Digital Sky Survey (SDSS; York et al. 2000) for large statistical samples, the Hubble Space Telescope (HST) for high-resolution line imaging, and ground-based IFU surveys such as CALIFA (Sánchez et al. 2012), SAMI (Bryant et al. 2015), MaNGA (Bundy et al. 2014), GAMA (Hopkins et al. 2013), all of which have mapped ionised gas properties in thousands of nearby galaxies. These surveys have uncovered trends in ionisation structure, metallicity gradients, and feedback signatures as a function of galaxy mass, morphology, and environment. In parallel, the Dark Energy Spectroscopic Instrument (DESI), operational since May 2021, is targeting nearly 30 million galaxies and 2.4 million quasars over one third of the sky, spanning redshifts from $z \sim 0.4$ (bright galaxies) to $z \sim 1.6$ (emission-line galaxies) and beyond (quasars out to $z \sim 3.5$)

(Napolitano et al. 2023). The Dark Energy Spectroscopic Instrument (DESI) (Collaboration et al. 2016, 2024) have collected high-resolution optical spectra of quasars, revealing the Lyα forest and metal absorption lines, tracing neutral and ionised intergalactic gas. Recent analyses using Year-1 data have begun stacking Mg II absorption in thousands of galaxy–quasar pairs at $0.75 < z < 1.65$, mapping the cool CGM in absorption with signal-to-noise sufficient to derive equivalent widths, column densities, and trends with galaxy mass and star-formation rate (Wu et al. 2024). Additionally, Euclid is executing a wide ($\sim 15{,}000\,\mathrm{deg}^2$) and deep ($\sim 50\,\mathrm{deg}^2$) survey in visible and near-infrared wavelengths with resolution $\sim 0.1''$ in VIS and $\sim 0.3''$ in NISP (Laureijs et al. 2011). Euclid will obtain slitless spectra and imaging for billions of galaxies out to $z \sim 2$, enabling large-scale studies of galaxy clustering, environment, and weak lensing maps (Collaboration et al. 2025).

At high redshift ($z > 4$), JWST's infrared spectra are opening a new window onto early galaxy formation, primarily through rest-frame optical emission lines like Hα, [O III], [O II], and Hβ. These lines, now redshifted into the near- and mid-infrared, are observable with JWST/NIRCam imaging, NIRSpec spectroscopy, and NIRCam grism data, offering key insights, such as ionisation state, metallicity, and star formation activity (Curti et al. 2023; Cameron et al. 2023). The COSMOS-Web survey, JWST's largest Cycle 1 program, uses NIRCam imaging and NIRISS grism spectroscopy to study ionised gas in galaxies from $z \sim 1$ to 7 (Casey et al. 2023). The JWST Advanced Deep Extragalactic Survey (JADES) offers the deepest spectroscopic view, detecting strong emission lines in galaxies up to $z \sim 10$, revealing highly ionised, metal-poor, and intensely star-forming systems near the epoch of reionisation (Eisenstein et al. 2023). The Cosmic Evolution Early Release Science (CEERS) survey combines imaging and spectroscopy at $z \sim 2 - 9$, uncovering compact, extreme [O III] emitters and complex nebular structures that demonstrate vigorous star formation and dynamic ionised gas (Finkelstein et al. 2023). These JWST surveys are transforming our understanding of ionised gas and galaxy evolution during and after the epoch of reionisation.

The hot ionised medium has been revealed by eROSITA through an all-sky X-ray survey, which has identified around 12,000 galaxy groups and clusters across a $z \sim 0.003 - 1.3$ (Predehl et al. 2021; Brunner et al. 2022). This dataset provides new insights into the distribution and properties of warm-hot gas in the outskirts of galaxy clusters and along the cosmic web's filaments. For the first time, eROSITA is directly imaging this diffuse intergalactic medium (IGM), which has long been hypothesised to harbour the "missing

baryons" eluded in cooler gas phases or stars (Nicastro et al. 2018; Tanimura et al. 2022). Surveys such as the eROSITA Final Equatorial-Depth Survey (eFEDS) and the ongoing all-sky survey continue to refine measurements of cluster masses, gas temperatures, and large-scale gas distributions, enabling a more complete baryon census and a deeper understanding of cosmic structure growth (Liu et al. 2022; Merloni et al. 2024).

In the far-infrared and sub-millimetre regime, ionised gas can be traced by fine-structure lines such as [N II] 122 μm, [N II] 205 μm, and [O III] 88 μm. These lines are critical at high redshift, where optical diagnostics are redshifted out of reach or obscured by dust. Instruments like ALMA and Herschel have been key in detecting and spatially resolving these lines in high-redshift galaxies and dusty star-forming systems (e.g., Ferkinhoff et al. 2010; Looze et al. 2014).

Ionised gas also plays a central role in galaxy feedback. Broad or blueshifted emission-line profiles reveal winds driven by massive stars and AGN, which inject energy and metals into the ISM and CGM, potentially regulating star formation and driving galaxy quenching (Veilleux et al. 2005; Schreiber et al. 2019; Veilleux et al. 2020). On cosmological scales, the presence of ionised metals in absorption (e.g., C IV, O VI) in the CGM and IGM reveals the large-scale impact of galaxies on their surrounding environments (Tumlinson et al. 2017).

1.2.2 Inflows, outflows and recycling of gas

Moving beyond closed-box models, galaxies are now understood within a framework of mass, energy and metal exchange with their environments (Bouché et al. 2010; Davé et al. 2012; Lilly et al. 2013; Peng & Maiolino 2014). In this modern view, galaxies grow at the centres of dark matter haloes through the condensation of baryons from the IGM and the CGM, either smoothly or through mergers, replenishing the fuel for star formation (Erb 2008; Tacconi et al. 2010; Putman et al. 2012; Sánchez Almeida et al. 2014).

However, a persistent issue in galaxy formation theory is the apparent inefficiency of star formation across most halo masses. As illustrated in the left panel of Fig. 1.6 (adapted from Silk & Mamon 2012), there is a systematic discrepancy between the predicted halo mass function from cold dark matter theory (red curve) and the observed galaxy luminosity function (blue curve). The number of galaxies with low and high luminosity is significantly lower than expected. This mismatch is widely interpreted as the consequence of feedback mechanisms that suppress star formation efficiency. In low-mass haloes, supernova-driven

winds expel gas before it can cool and form stars, while in high-mass haloes, AGN feedback heats or ejects gas, quenching star formation. The peak of this efficiency curve occurs at $L_* \sim 3 \times 10^{10} M_\odot$, where galaxies are most effective at turning baryons into stars. This suggests that galaxy growth is self-regulated by feedback from supernovae and a metallicity limit, which restricts growth in low-mass haloes, while hot accretion and black hole feedback do so in high-mass haloes (Behroozi et al. 2012).

This feedback-regulated framework forms the keystone of the gas regulator model, also known as the bathtub model of galaxy evolution (Bouché et al. 2010; Davé et al. 2012; Lilly et al. 2013; Peng & Maiolino 2014). This model, shown in the right panel of Fig. 1.6, illustrates how the gas reservoir of a galaxy (the bathtub) is governed by inflows of gas (the tap), outflows (the drain), and the conversion of gas into stars. It can be formalised by a mass continuity equation to describe the evolution of the gas reservoir in a galaxy, $M_{\rm gas}$, as a result of three primary processes: inflow, outflow, and star formation. The basic form of the differential equation is:

$$\dot{M}_{\rm gas} = \dot{M}_{\rm in} - (1 - R)\dot{M}_\star - \dot{M}_{\rm out} \tag{1.2}$$

where $\dot{M}_{\rm gas}$ is the time derivative of the galaxy's gas mass, $\dot{M}_{\rm in}$ is the inflow rate of cold gas (typically accreted or recycled from the IGM or CGM), $\dot{M}_\star$ is the star formation rate, and $\dot{M}_{\rm out}$ is the outflow rate of gas ejected via feedback processes. The term $(1 - R)$ represents the net gas consumption into long-lived stars, corrected by the mass return fraction R, which accounts for the recycled gas from short-lived stellar populations. The simplicity of this framework masks the complexity of its components: inflows are influenced by large-scale structure, cooling efficiency, and feedback from other galaxies; outflows are driven by stellar winds and AGN activity; and recycling of ejected material through galactic fountains closes the loop of the cycle.

Sustaining star formation relies on the accretion of pristine IGM gas (Erb 2008; Tacconi et al. 2010; Putman et al. 2012; Sánchez Almeida et al. 2014). While direct detection of inflows remains challenging, evidence of this phenomenon can be found in the Milky Way, for example, in the Magellanic Stream (Fox et al. 2014) and in high-velocity clouds (Zheng et al. 2017). In external galaxies, inflows are observed through redshifted absorption lines aligned with disk rotation, suggesting co-rotating accretion (Martin et al. 2012; Rubin et al. 2012; Bouché et al. 2016; Bowen et al. 2016; Ho et al. 2017; Rahmani et al. 2018a). A

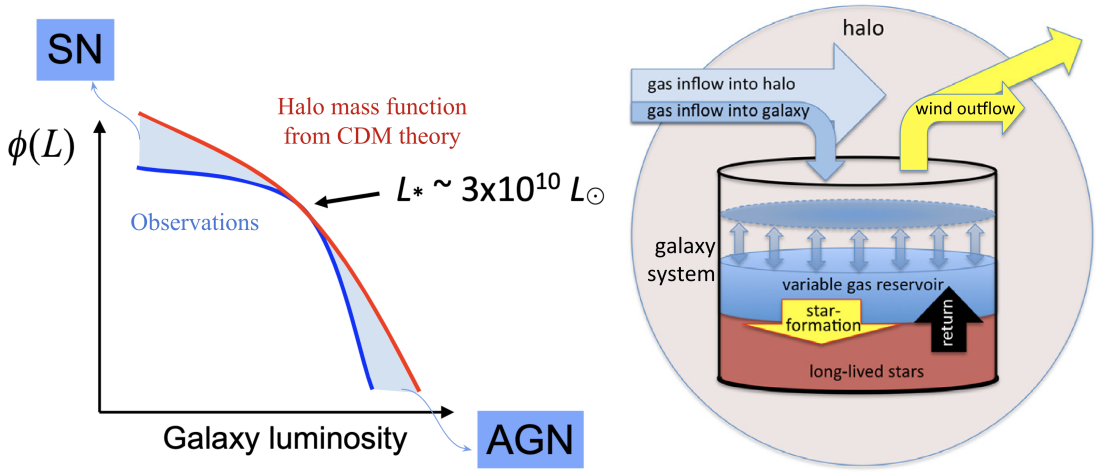

Figure 1.6: *Left:* The galaxy luminosity function, $\phi(L)$, versus galaxy luminosity, adapted from Silk & Mamon (2012). The red curve represents the expected number of galaxies from cold dark matter (CDM) theory halo mass functions, while the blue curve indicates actual observational data. The discrepancy at both low and high luminosities (blue shaded area) illustrates the "missing satellite" and "over-quenching" problems, respectively. Feedback from supernovae (SN) suppresses star formation in low-mass haloes, while active galactic nuclei (AGN) feedback regulates massive galaxy formation. The characteristic luminosity $L_* \sim 3 \times 10^{10} L_\odot$ marks the transitional regime where feedback processes balance baryonic cooling and accretion. *Right:* Bathtub model of regulation of galaxy evolution, extracted from Lilly et al. (2013). In this model, galaxies are systems in a state of evolving equilibrium between inflow, outflow, and star formation. While at early times, gas accumulates and star formation depends on the gas reservoir, at later times, star formation is regulated by the net accretion rate as galaxies have reached a steady state.

comprehensive book by Fox & Davé (2017) further details the multi-wavelength strategies used to trace accretion, including spectroscopy, integral field unit (IFU) and background quasars (Chapters 5 and 15). Yet, simulations indicate that inflow rates must scale with halo mass and redshift, reaching $10-50\,M_\odot\,\mathrm{yr}^{-1}$ at $z \sim 2$ and declining to $\sim 1-5\,M_\odot\,\mathrm{yr}^{-1}$ at $z = 0$ for Milky Way-type haloes (van de Voort et al. 2011; Nelson et al. 2013; Anglés-Alcázar et al. 2017; Mitchell & Schaye 2022; Barbani et al. 2023; Callingham & Helmi 2025). These inflows supply the gas required to maintain the observed star formation rates, especially in light of short depletion timescales in molecular gas reservoirs (Tacconi

et al. 2020).

Stellar feedback drives ubiquitous outflows across all redshifts (Martin 2005; Rupke et al. 2005; Veilleux et al. 2005; Weiner et al. 2009; Steidel et al. 2010; Martin et al. 2012; Newman et al. 2012; Chisholm et al. 2015, 2016; Heckman & Borthakur 2016; Pereira-Santaella et al. 2018; Fluetsch et al. 2019). These winds are efficient in removing metal-enriched gas from galaxies. Observations of UV and X-ray absorption lines in the CGM reveal that a substantial fraction of metals (up to $75-80\%$) is expelled from galaxies, with only $\sim 20-25\%$ remaining in the stellar and interstellar components (Tremonti et al. 2004; Peeples et al. 2014). Whether these metals leave the halo entirely or are retained in the CGM remains an open question with direct implications for chemical enrichment, cooling, and recycling (Oppenheimer & Davé 2008; Marasco et al. 2022; Barbani et al. 2023).

Many outflows have velocities below the halo escape speed, so a significant portion remains gravitationally bound and re-accretes via galactic fountains (Tumlinson et al. 2011; Rubin et al. 2012; Bouché et al. 2012; Stocke et al. 2013; Mathes et al. 2014; Bordoloi et al. 2014; Emonts et al. 2015; Pereira-Santaella et al. 2016; Zheng et al. 2017; Tremblay et al. 2018). This recycled accretion serves as a metal-enriched secondary inflow, prolonging star formation activity (Oppenheimer et al. 2010; Ford et al. 2014; van de Voort et al. 2017). Recycled gas tends to dominate over pristine IGM accretion at $z \lesssim 1$ in massive galaxies, and has been found to return on timescales of $\sim 0.1-1$ Gyr (Anglés-Alcázar et al. 2017; Hafen et al. 2019). This loop closes the baryon cycle, tightly coupling metal enrichment, star formation, and feedback on Gyr timescales.

These inflows, star formation, outflows, and recycling collectively define the full baryon cycle, a process now implemented in both cosmological and high-resolution zoom-in simulations (Brook et al. 2014; Shen et al. 2014; Mitra et al. 2015; Anglés-Alcázar et al. 2017; Tollet et al. 2019). The canonical picture thus sees galaxy evolution as the result of a self-regulating exchange of baryons, energy, and metals between galaxies and their gaseous environments. The challenge lies in capturing this complex, multiphase, multi-scale interaction: from cosmic web inflows to star-forming cores, from hot virialised gas to cold molecular clouds. A full understanding of galaxy formation and evolution ultimately depends on mapping this dynamic equilibrium in detail, both observationally and theoretically, throughout cosmic history.

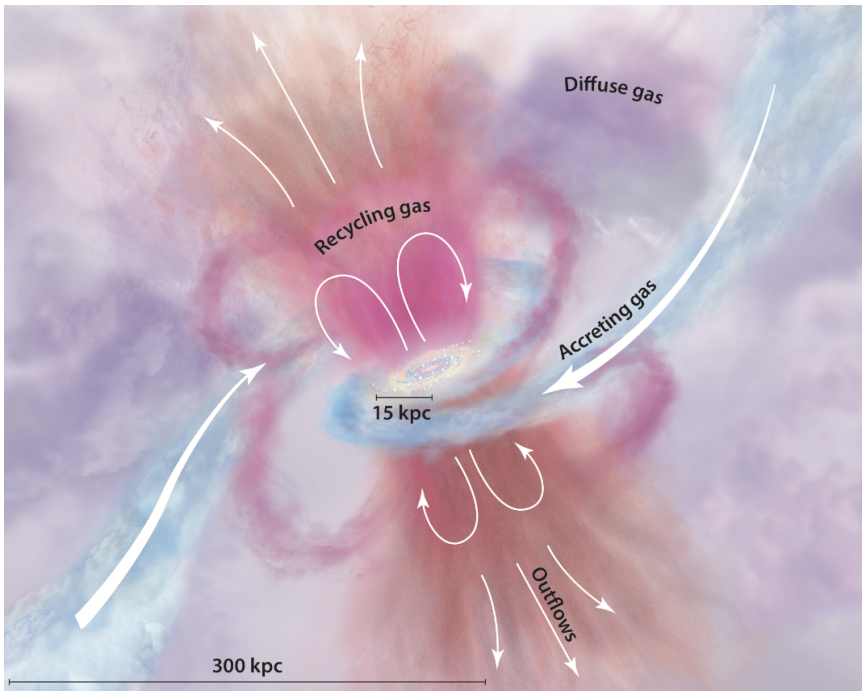

Figure 1.7: Schematic illustration of the circumgalactic medium (CGM) from Tumlinson et al. (2017). It shows a star-forming galaxy with a central bulge (red) and a gaseous disk (blue), embedded within an extended, multiphase halo. Cold filamentary inflow from the intergalactic medium feeds the disk (blue streams), while galactic outflows powered by stellar or AGN feedback emerge perpendicular to the disk plane (pink and orange). Some outflowing material is shown returning as recycled gas. The diffuse CGM halo (purple shades) represents a complex mixture of inflowing, outflowing, and recycled gas across a range of temperatures, densities, and ionisation states.

1.3 The circumgalactic medium

The circumgalactic medium (CGM) corresponds to the extended, diffuse, and multiphase gaseous halo that surrounds galaxies, typically out to their virial radii. Although not precisely bounded, the CGM generally spans from ten to several hundred kiloparsecs, serving as the interface between galaxies and the intergalactic medium (IGM). Nearly all the inflows and outflows of baryons pass through the CGM, making it a key regulator of galactic evolution (Tumlinson et al. 2017; Péroux & Howk 2020; Peroux & Nelson 2024). A schematic illustration from Tumlinson et al. (2017) is shown in Fig. 1.7, which provides

a conceptual overview: filamentary IGM accretion flows (in blue) feed a galaxy's disk and bulge, while feedback-driven outflows (in orange and pink) escape perpendicular to the disk. Previously ejected gas is shown recycling back, mixing with a diffuse halo of multiphase gas (in purple) that includes material from all these processes.

The CGM is now understood to be multiphase, containing cold (T $< 10^4$ K), warm (T $\sim 10^4 - 10^5$ K), warm-hot (T $\sim 10^5 - 10^6$ K), and hot (T$> 10^6$ K) gas, often co-existing at the same radii but with vastly different densities and ionisation states. Cold clouds are often embedded within a hotter diffuse medium (e.g., McCourt et al. 2018; Fielding et al. 2020). This complexity poses a major challenge for both simulations and observations.

1.3.1 Quasar absorption studies

Due to its low density and emissivity, the CGM is notoriously difficult to observe directly in emission, although progress has been made with deep integrations and stacking techniques (e.g., Cantalupo et al. 2014; Wisotzki et al. 2018). The most powerful tool for studying the CGM remains absorption-line spectroscopy against bright background sources, such as quasars and galaxies.

The technique exploits the absorption imprinted on the spectrum of a bright background object, typically a quasar, by intervening gas associated with foreground galaxies. As shown in Figure 1.8, light from the quasar passes through the cosmic web and multiple galaxy haloes before reaching the observer. Along this line of sight, neutral hydrogen (e.g. Lyα) and various metal ions (e.g. Si IV, C IV, N V) absorb specific wavelengths of light, creating a forest of absorption lines in the observed spectrum. These features trace the ionisation state, kinematics, and metallicity of the CGM and IGM over a broad redshift path.

Hydrogen absorption systems are typically classified by their neutral hydrogen column density:

- **Lyman Limit Systems (LLSs):** $10^{17.2} < N(\mathrm{H\,I}) < 10^{19}$ cm^{-2}; mostly ionised.
- **Sub-Damped Lyman-α Systems (sub-DLAs):** $10^{19} < N(\mathrm{H\,I}) < 2 \times 10^{20}$ cm^{-2}; partially ionised.
- **Damped Lyman-α Systems (DLAs):** $N(\mathrm{H\,I}) > 2 \times 10^{20}$ cm^{-2}; predominantly neutral.

DLAs are particularly valuable as they trace predominantly neutral gas and contain most of the Universe's cold H I reservoir (Wolfe et al. 2005), while sub-DLAs and LLSs, although more ionised, are sensitive to metal content of the CGM (Péroux et al. 2003, 2016; Fumagalli et al. 2011).

Two geometric configurations are commonly considered. First, the intrinsic or "down-the-barrel" absorption, where the background light source is the galaxy itself, and absorption occurs within its own halo. The velocity shift between absorption lines and the systemic redshift can then be interpreted directly as inflow or outflow, although the physical distance of the gas is unconstrained. Second, the intervening or transverse absorption, where the quasar is positioned behind another galaxy, and the gas is probed at a known projected distance (the impact parameter) from that galaxy. In this case, kinematic signatures are degenerate since redshifted or blueshifted absorption could result from inflows, outflows, or gas at various positions along the line of sight.

To break these degeneracies, additional diagnostics are used. For example, metallicity measurements help differentiate gas origins: high metallicity suggests enriched outflows, while low metallicity implies pristine infalling gas (Péroux et al. 2016). Another method is measuring the azimuthal angle, which is the orientation of the quasar sight line relative to the galaxy's major axis (Péroux et al. 2020). Simulations and some observations indicate that inflows preferentially align with the major axis (i.e., co-planar accretion), while outflows occur along the minor axis (Shen et al. 2012). However, recent studies find no strong correlation between azimuthal angle and metallicity, likely due to sample size limitations or projection effects (Weng et al. 2023b; Guo et al. 2023b).

A central challenge in interpreting CGM absorption systems is associating them with their host galaxies, particularly for intervening absorbers. While early work relied on indirect associations, more recent campaigns using instruments like SINFONI, MUSE, and MOSFIRE have led to the detection of several DLA host galaxies (Fynbo et al. 2010; Péroux et al. 2011; Rahmani et al. 2018b). These observations revealed that many absorbers trace gas in intra-group environments or are connected to multiple galaxies (Bielby et al. 2017; Péroux et al. 2019), challenging the single-host paradigm. Furthermore, stellar mass estimates have now been possible using spectral energy distribution (SED) modelling with HST photometry (Augustin et al. 2024).

High-metallicity DLAs often trace massive galaxies, but simulations suggest that DLAs arise preferentially in lower-mass galaxies and satellites, highlighting a selection bias in

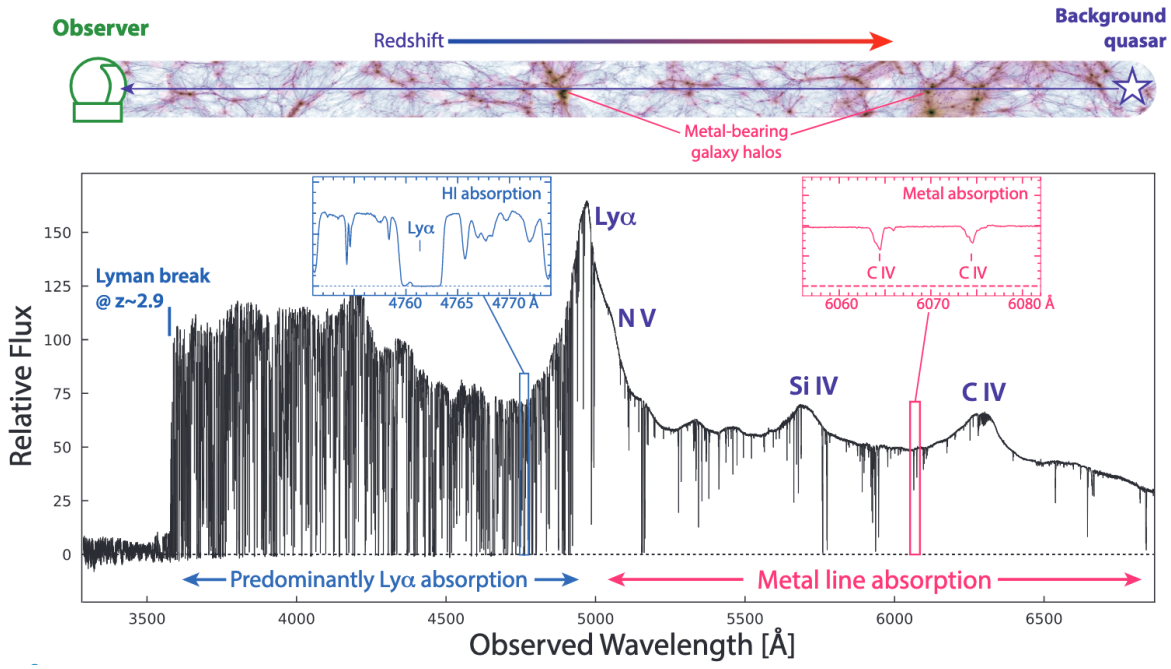

Figure 1.8: Quasar absorption-line spectroscopy of the CGM and IGM, extracted from Péroux & Howk (2020). Light from a background quasar passes through intervening cosmic structures, including galaxy haloes, imprinting absorption features from H I (e.g. Lyα) and metals (e.g. C IV, Si IV). The resulting spectrum reveals information about the ionisation, kinematics, and composition of the CGM along the line of sight. Insets show examples of Lyα forest and metal-line absorption.

current samples (Haehnelt et al. 1998; Pontzen et al. 2008; Fumagalli et al. 2014). Overcoming this bias will require future surveys targeting gas-rich, rather than metal-rich, environments.

It is important to note that each quasar spectrum probes only a single pencil-beam sightline through the CGM. Therefore, large statistical samples with secure galaxy–absorber associations are required to reconstruct the spatial and dynamical structure of the CGM and to draw robust conclusions about baryon flows around galaxies.

Cosmological hydrodynamical simulations, such as IllustrisTNG, Millennium TNG, EAGLE, FIRE, SIMBA, and Magneticum now complement and interpret CGM observations by modelling the interplay between gas dynamics and feedback (e.g., Nelson et al. 2019; van de Voort et al. 2019; Bogdan et al. 2023). These simulations consistently show that the CGM hosts a significant fraction of a galaxy's baryons and is shaped by complex processes

including gas accretion, feedback-driven outflows, and recycling flows, where material expelled by supernovae or AGN can re-accrete over Gyr timescales (e.g., Oppenheimer et al. 2010; Anglés-Alcázar et al. 2017). Recent simulation efforts incorporate additional physics, such as cosmic rays, magnetic fields, and radiation transport, resolving cold clumps, turbulence, and thermal instabilities (e.g. FIRE-2, FABLE, Bennett & Sijacki 2020; Butsky et al. 2020; Ponnada et al. 2022). However, full convergence on small-scale CGM structures remains limited by missing physics and resolution constraints, often necessitating the use of subgrid models and adaptive mesh refinement techniques (e.g., Hummels et al. 2019; van de Voort et al. 2021). Simulations also reveal that the CGM's structure depends on halo mass, redshift, and the strength of feedback. For example, in low-mass galaxies, stellar feedback more easily displaces gas from the disk into the halo without completely ejecting it. This leads to a CGM that is more massive and extended relative to the galaxy's baryonic content (e.g., Christensen et al. 2016).

Ultimately, multi-wavelength observations, particularly those targeting quasar absorbers and their associated galaxies, are key to studying the baryon cycle. Such observations will reveal the inflow and outflow processes that fuel star formation, regulate metal enrichment, and potentially quench galaxies across cosmic time. To fully interpret these processes, they must be combined with high-resolution cosmological simulations to directly compare the observed gas properties with theoretical predictions.

1.4 Molecular gas

Molecular hydrogen (H_2) is the dominant component of the interstellar medium in the dense, star-forming regions of galaxies. However, as mentioned above in §1.2.1, cold H_2 is very difficult to observe directly because it lacks a permanent dipole moment. Moreover, due to the low mass of hydrogen atoms, the lowest rotational transitions of H_2 occur at relatively high excitation energies: the first excited state ($J = 1$) lies 175 K above the ground state and is forbidden for reasons of symmetry, while the first allowed transition ($J = 2$) is excited at 510 K. These energy levels are rarely excited in typical molecular clouds with temperatures around 10−30 K (Dabrowski 1984).

As a result, astronomers often rely on other tracers of molecular gas, such as carbon monoxide (CO), atomic carbon ([C I]) and ionised carbon ([C II]), which can provide complementary insights into molecular cloud properties (Walter et al. 2011; Carilli & Walter

2013; Maiolino et al. 2015). CO, in particular, remains the most widely used and well-characterised tracer of molecular gas, especially at redshift $z \lesssim 6$ (Bolatto et al. 2013). This thesis primarily focuses on CO observations.

Table 1.2: CO rotational transitions

CO line	Rest frequency (GHz)
CO(1−0)	115.271
CO(2−1)	230.538
CO(3−2)	345.796
CO(4−3)	461.041
CO(5−4)	576.268
CO(6−5)	691.473

1.4.1 The CO line

CO is the second most abundant molecule in the ISM and forms easily in molecular clouds alongside H_2. CO is a diatomic molecule made up of one carbon atom and one oxygen atom, and has a permanent dipole moment. This allows it to emit efficiently via rotational transitions given by changes in the rotational quantum number J ($J = 0, 1, 2, ...$). The energy spacing of these transitions is much smaller than for H_2. The lowest rotational transition of ^{12}CO, $J = 1 \rightarrow 0$, occurs at 115 GHz ($\lambda \approx 2.6$ mm) and requires only 5.5 K to be excited. This transition is easily observable through Earth's atmosphere and is generally optically thick, making it a robust indicator of the presence of molecular gas even when not directly proportional to the H_2 column density. Table 1.2 lists all the emission frequencies of the lowest six rotational transitions of CO, spaced by ~ 115 GHz.

The CO molecule is used as a proxy for H_2 mass under the assumption that its integrated intensity, $W_{\rm CO}(J = 1 - 0)$, correlates with total molecular gas column density, $N(H_2)$, through $X_{\rm CO}$, defined as:

$$X_{\rm CO} = \frac{N(H_2)}{W_{\rm CO}(J = 1 - 0)} \, [{\rm cm}^{-2}/({\rm K\,km\,s}^{-1})] \tag{1.3}$$

where $N(H_2)$ is the H_2 column density in cm^{-2}, $W_{\rm CO}(J = 1 - 0)$ is the velocity-integrated CO intensity in units of K km s^{-1}, and $X_{\rm CO}$ is the conversion factor typically

taken as 2×10^{10} cm^{-2} (K km s^{-1})$^{-1}$ in the Milky Way disk (e.g., Bolatto et al. 2013).

The molecular gas mass can be written in terms of the CO luminosity as:

$$M_{\rm mol} = \alpha_{\rm CO} L'_{\rm CO} \tag{1.4}$$

where $M_{\rm mol}$ is the molecular gas mass in solar masses ($M_\odot$) and $L'_{\rm CO}$ is the CO luminosity defined as (Solomon et al. 1992):

$$L'_{\rm CO} = 3.25 \times 10^7 \times S_{\rm CO}\,\Delta v \frac{D_{\rm L}^2}{(1+z)^3 \nu_{\rm obs}}\,[{\rm K\,km\,s^{-1}\,pc^2}] \tag{1.5}$$

where $S_{\rm CO}\,\Delta v$ is the integrated line flux density in Jy km s^{-1}, $D_{\rm L}^2$ is the luminosity distance to the source in Mpc, and $\nu_{\rm obs}$ is the observed frequency in GHz. The factor $\alpha_{\rm CO}$ can be seen as a mass-to-light ratio.

The theoretical derivation of $X_{\rm CO}$ begins by considering a self-gravitating molecular cloud radiating an optically thick CO line. Under such conditions, the brightness of the line depends primarily on the excitation temperature ($T_{\rm ex}$) and the velocity dispersion, which in turn reflects the gravitational potential of the cloud.

The ground transition CO(1−0) is widely used in both local and high-redshift environments, since it is easily excited even in cold molecular clouds due to its low critical density ($n_{\rm crit} \sim 2 \times 10^3$ cm^{-3}). Observations of CO(1−0) have been calibrated against a variety of techniques to yield estimates of the total molecular gas mass. These calibrations include virial mass measurements of resolved Galactic giant molecular clouds (GMCs, Solomon et al. 1987; Scoville et al. 1987), column densities based on local thermodynamic equilibrium from optically thin isotopologues such as ^{13}CO (Goldsmith et al. 2008), dust extinction and emission modelling (e.g. Dame et al. 2001; Collaboration et al. 2011), and measurements of gamma-ray emission from cosmic-ray interactions with ISM nucleons (Strong & Mattox 1996; Grenier et al. 2005). All these methods converge toward a consistent value in the Milky Way disk: $X_{\rm CO} \approx 2 \times 10^{20}$ cm^{-2} (K km s^{-1})$^{-1}$, equivalent to $\alpha_{\rm CO} \approx 4.3 M_\odot$ (K km s^{-1} pc^2)$^{-1}$ when including helium.

However, the CO-to-H_2 conversion factor ($\alpha_{\rm CO}$) depends on the properties of the cloud. In low metallicity environments, there is less dust to shield CO from ultraviolet radiation, making CO more easily destroyed. As a result, a significant portion of H_2 remains undetectable in CO emission, which is known as CO-dark gas. The envelope of CO-dark molecular gas can account for up to 30% of the total molecular mass in a Galactic cloud

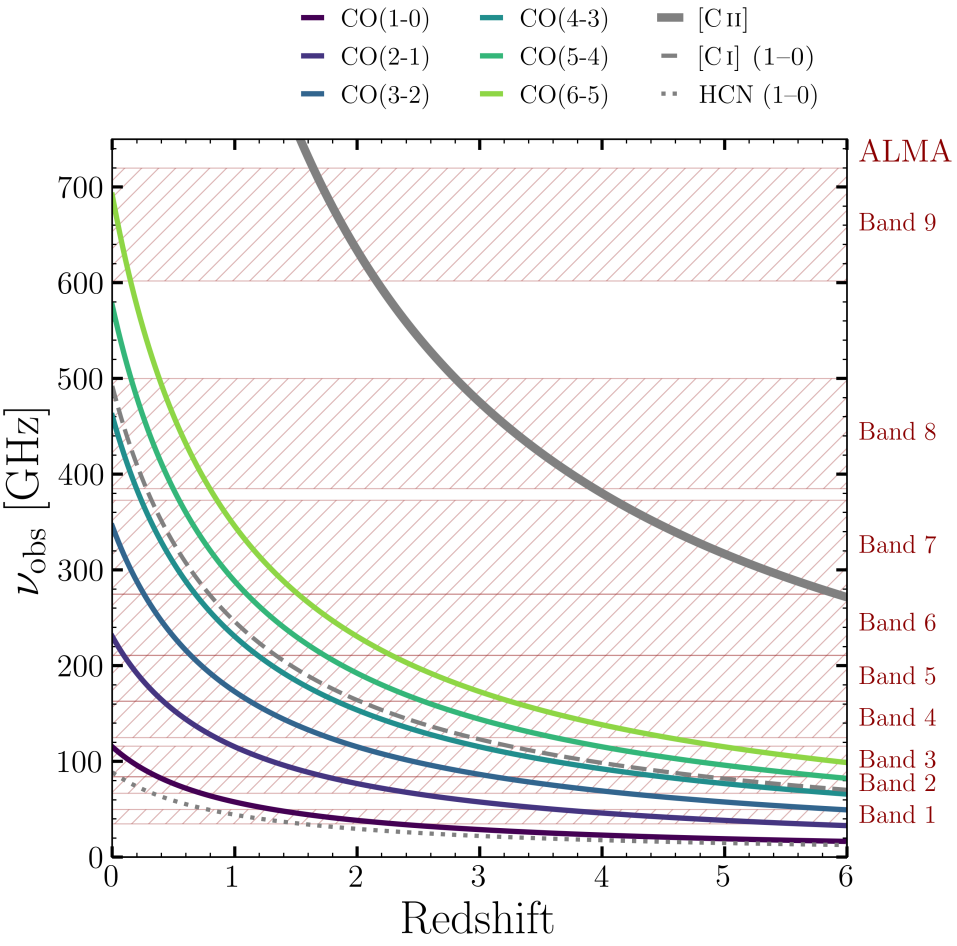

Figure 1.9: Observed frequencies of CO emission lines as a function of redshift. The grey shaded regions represent the frequency ranges covered by ALMA Bands. At low redshifts, low$-J$ CO lines are accessible, enabling studies of nearby galaxies. At higher redshifts, ALMA can observe higher$-J$ transitions, tracing the molecular gas content in earlier cosmic epochs. For redshifts $z \gtrsim 3$, the bright [C II] emission line enters ALMA's observable range, serving as an additional tracer of star-forming galaxies in the early Universe.

(Madden et al. 2006). Consequently, $\alpha_{\rm CO}$ increases with decreasing metallicity, as evidenced by measurements in Local Group galaxies (e.g., Blitz et al. 2007) and supported by dust-based mass estimates (e.g., Draine et al. 2007). In high-pressure environments, such as ultra luminous infrared galaxies (ULIRGs) and galaxy centres, molecular clouds may not be in virial equilibrium. Elevated temperatures, velocity dispersions, and turbulence

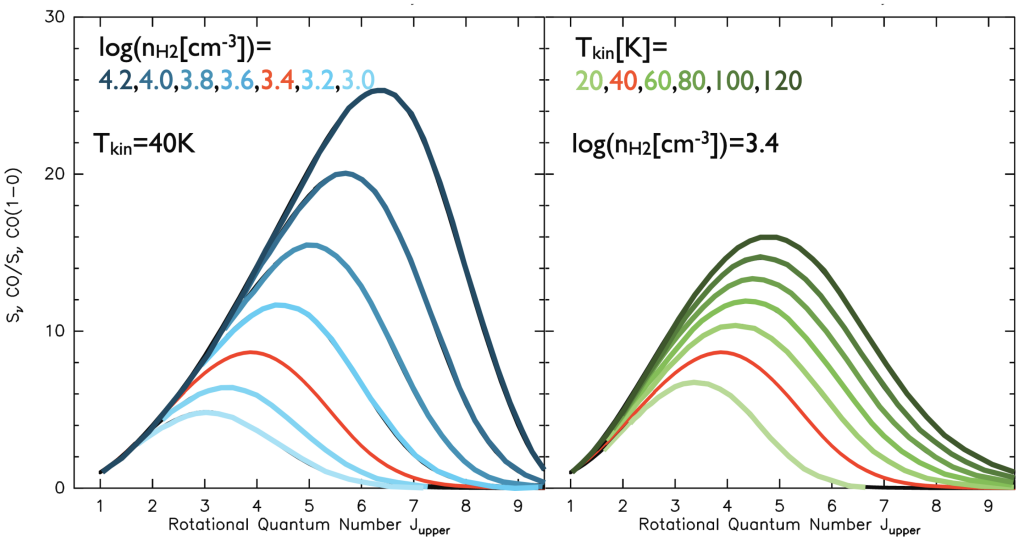

Figure 1.10: Flux ratio of CO rotational transitions relative to the CO(1–0) line, illustrating how the CO Spectral Line Energy Distribution (SLED) changes as a function of gas temperature and density, extracted from Carilli & Walter (2013). *Left*: Effect of varying molecular gas density at a fixed temperature of 40 K. *Right*: Effect of varying gas temperature at a fixed density of $\log n_{\mathrm{H_2}}[\mathrm{cm}^{-3}] = 3.4$. These model predictions demonstrate the degeneracy-breaking power of multi$-J$ observations for the determination of both the kinetic temperature and density of the emitting molecular gas.

can cause CO emission to originate from diffuse or non-self-gravitating gas. Under these conditions, α_{CO} drops significantly, reaching values as low as 0.8 $M_{\odot}$ (K km s^{-1} pc^2)$^{-1}$ (Downes & Solomon 1998).

At high redshift, the CO(1$-$0) line is shifted to longer wavelengths, so many studies instead observe higher$-J$ transitions (see Figure 1.9). Facilities such as the Karl G. Jansky Very Large Array (VLA, e.g., Riechers et al. 2020b), the Northern Extended Millimeter Array (NOEMA, e.g., Decarli et al. 2019), and most notably the Atacama Large Millimeter/submillimeter Array (ALMA, e.g., Hodge & da Cunha 2020) provide the sensitivity, spectral coverage, and angular resolution required to observe high$-J$ CO lines.

To convert high$-J$ lines into a total molecular gas mass, excitation corrections must be applied, typically by adopting average line ratios. Moreover, the line ratios offer a way to probe the excitation conditions of molecular gas by plotting the line luminosities of several CO transitions against J, these plots are known as the CO spectral energy distributions

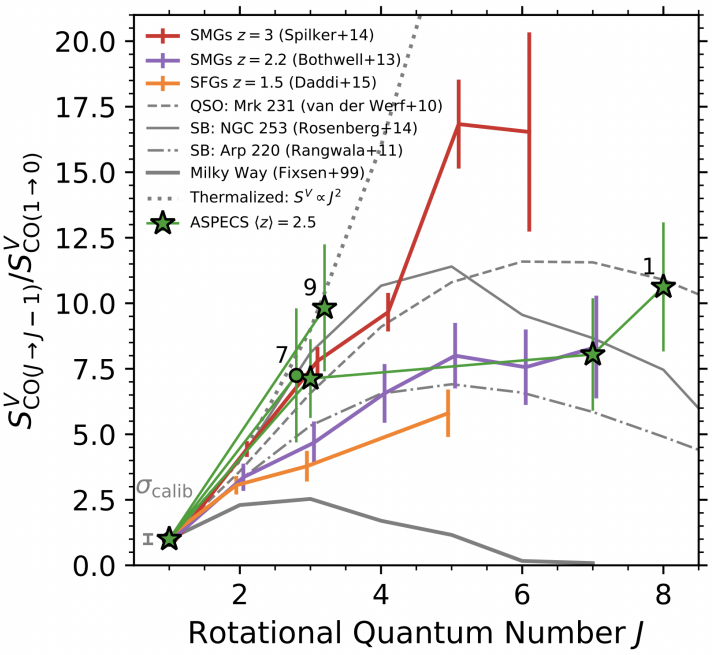

Figure 1.11: CO spectral line energy distributions (SLEDs) for ASPECS galaxies detected in CO, normalised to the CO(1−0) flux. Green markers represent detections, with stars marking AGN identified via X-ray. Stars denote X-ray sources classified as AGN (Luo et al. 2017). A grey error bar represents the combined calibration uncertainty for ALMA and VLA data. BzK-selected star-forming galaxies (Daddi et al. 2015), local starbursts (Rangwala et al. 2011; Rosenberg et al. 2015), and sub-mm galaxy (Bothwell et al. 2013). Figure extracted from Boogaard et al. (2020).

(CO SLEDs). The exact shape of the CO SLED depends on the temperature and density of the emitting gas and can, therefore, be used to study the ISM conditions in galaxies (e.g., Klitsch et al. 2022).

Figure 1.10, extracted from Carilli & Walter (2013), illustrates the change of the CO SLED for a fixed temperature of 40 K and changing densities (left panel), and for a fixed density and varying temperature (right panel). It is important to highlight that multi-line CO observations can constrain physical conditions on gas temperature, density, and thermal pressure.

Figure 1.11, from Boogaard et al. (2020), shows the CO SLEDs of the Milky Way, a starburst galaxy (M82), and high-redshift submillimeter galaxies (SMGs), demonstrating clear differences in excitation. Starbursts and quasar hosts tend to show highly excited

SLEDs, peaking at $J = 6$ or higher, while normal star-forming galaxies have SLEDs peaking at $J = 2$ or 3. The steep rise in the SLEDs for starbursts and SMGs indicates a warm, dense, and highly pressurised ISM, compared to the relatively flat SLED of a Milky Way cloud.

Overall, CO emission lines provide valuable insights into the molecular gas content of galaxies. Observing different CO transitions allows us to investigate the physical conditions of this gas. While both low− and high-J lines offer complementary information, even a limited set of transitions can reveal key details about different environments. High sensitivity and resolution are crucial to make this possible. In this context, ALMA has demonstrated unprecedented sensitivity and spatial resolution, enabling detailed studies of molecular gas throughout the Universe (Hodge & da Cunha 2020).

1.4.2 ALMA and Radio Interferometry

Table 1.3: Details of the ALMA Bands.

ALMA Band	Frequency [GHz]	Wavelength [mm]	Field of view [arcsec]
Band 1	35–50	6.0–8.6	147
Band 2	67–116	2.6–4.5	80
Band 3	84–116	2.6–3.6	58
Band 4	125–163	1.8–2.4	48
Band 5	163–211	1.4–1.8	35
Band 6	211–275	1.0–1.4	25
Band 7	275–373	0.8–1.0	17
Band 8	385–500	0.6–0.8	12
Band 9	602–720	0.4–0.5	9
Band 10	787–950	0.3–0.4	7

The Atacama Large Millimeter/submillimeter Array (ALMA) operates in the millimeter (mm) and submillimeter (sub-mm) wavelength regime, providing a unique window into the cool ($T < 100$ K), dusty, and often hidden components of galaxies. Unlike optical or infrared lights, which can be heavily obscured by dust, sub-mm/mm emission penetrates

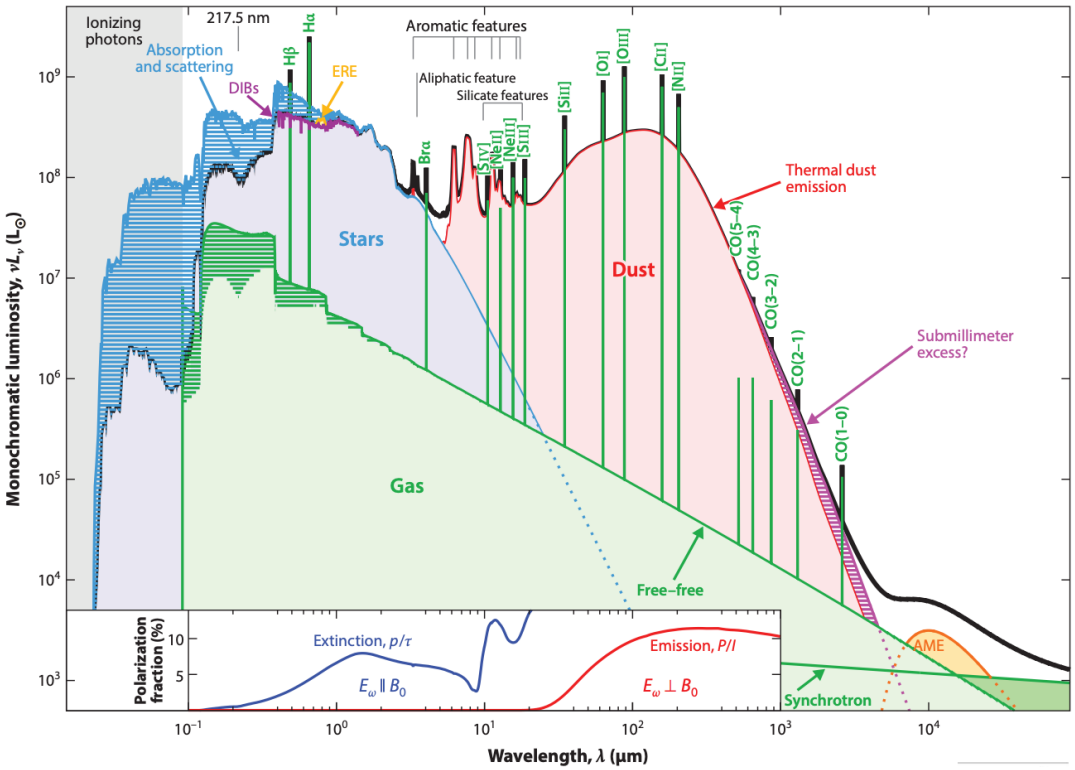

Figure 1.12: Schematic spectral energy distribution (SED) of a typical late-type galaxy, dominant continuum components (thermal dust, synchrotron, and free–free emission) and overlaid spectral lines, including multiple CO transitions and [C II]. The spectral locations of these features demonstrate the importance of sub-mm/mm observations for probing the cold ISM. ALMA's frequency coverage spans most of this window, enabling the detection of both line and continuum emission critical to studies of galaxy evolution. Further details on the model can be found in Galliano et al. (2018).

dense interstellar material, allowing us to trace cold molecular gas and dust, the raw ingredients for star formation. Sub-mm/mm observations capture both the thermal continuum emission from cold dust, heated by ongoing star formation or AGN activity, and key spectral lines such as CO and the [C II] 158 μm line. Figure 1.12 illustrates the long-wavelength spectrum of a typical late-type galaxy, highlighting the positions of CO transitions and [C II] atop the continuum shaped by dust, free-free emission, and synchrotron processes (extracted from Galliano et al. 2018).

Since the start of early science operations in 2011, ALMA has provided an exceptional

angular resolution through its long baselines, matching or exceeding that of the best optical telescopes. Its wide frequency and bandwidth coverage offer access to a wealth of spectral lines, along with its unmatched sensitivity, which has enabled the detection of fainter galaxies further down the luminosity function. These estimates have been crucial to build statistical studies of mass- and colour-selected samples as well as deep blind field surveys (e.g., ASPECS, COLDz, PHIBSS, Decarli et al. 2016; Riechers et al. 2019; Tacconi et al. 2013). This section provides a brief but detailed overview of the technical aspects that allow ALMA's operation.

Sensitivity in Radio Astronomy

Electromagnetic radiation in the radio regime behaves as a wave, but when the physical dimension of the observing system is much larger than the wavelength, radiation can be treated as rays travelling in straight lines (Wilson et al. 2013; Burke et al. 2019). A fundamental quantity is the specific intensity, I_ν, defined such that the infinitesimal power dP intercepted by an area element dA is:

$$dP = I_\nu(\theta, \phi) \cos\theta \, dA \, d\Omega \, d\nu \tag{1.6}$$

where I_ν is the specific intensity (or brightness) in units of W m^{-2} Hz^{-1} sr^{-1}, θ is the angle between the ray and the normal to the area dA, $d\Omega$ is the solid angle subtended by the source, and $d\nu$ is the frequency bandwidth. The specific intensity, similar to surface brightness in optical astronomy, is an intrinsic property of a source. It describes how much energy is emitted per unit area, per unit frequency, and per unit solid angle.

The flux density S_ν received from a source is obtained by integrating over its solid angle:

$$S_\nu = \int_{\Omega_s} I_\nu(\theta, \phi) \cos\theta \, d\Omega \tag{1.7}$$

Flux density is measured in units of W m^{-2} Hz^{-1}, but due to the faintness of astronomical radio sources, a practical unit called the Jansky is used, defined as $1\,\mathrm{Jy} = 10^{-26}\,\mathrm{W\,m^{-2}\,Hz^{-1}}$. For an unresolved source or small angles $\cos\theta \approx 1$.

The total power (P_ν) collected by the effective area (A_e) of a radio telescope per unit frequency is:

$$P_\nu = \frac{1}{2} A_e \cdot S_\nu \tag{1.8}$$

The smallest detectable signal is limited by system noise, often quantified using the system temperature T_{sys}, which includes receiver, sky, and ground contributions. The radiometer equation defines the root-mean-square noise fluctuations:

$$\Delta S_\nu = \frac{2k_B T_{sys}}{A_e \sqrt{\Delta\nu \cdot t_{int}}} \tag{1.9}$$

This relation shows the importance of a large collecting area, a wide bandwidth, and long integration times for improving the sensitivity.

Apeture Synthesis

Understanding the nature and limits of sub-mm/mm observations requires a fundamental shift in how we think about telescopes. Unlike optical wavelengths, where a single large mirror can deliver arcsecond-scale resolution, the much longer wavelengths of sub-mm/mm radiation need a larger aperture for comparable resolving power. The angular resolution for a single-dish telescope is approximated as follows:

$$\theta = 1.22 \frac{\lambda}{D} \tag{1.10}$$

where θ is the resolution, and D is the size of the antenna in meters. However, at mm/sub-mm wavelengths, achieving arcsecond resolution would require impractically large dishes. Interferometry overcomes this by combining the signals from an array of smaller telescopes spread over large distances.

Interferometry exploits the wave nature of light to synthesise a virtual telescope much larger than any single dish (Thompson et al. 2017). The principle relies on measuring the interference pattern formed by electromagnetic waves collected by pairs of antennas separated by a vector baseline, $\vec{b}$, which connects the phase centres of the two antennas, as shown in Figure 1.13[1]. When radiation from a distant source (S_0) reaches two antennas at slightly different times due to differences in the geometric paths, a path length difference is introduced, given by $\vec{b}{\cdot}S_0$. Then, the time delay can be written as $\tau_g = b{\cdot}S_0/c$, and the phase

[1]This figure is inspired by the figures in Chapter 9 of Burke et al. (2019), under the assumption that $w = 0$ for a 2D imaging approximation, which is the case of an ALMA-like interferometer according to the ALMA technical handbook in Cycle 12 (Cortes et al. 2025).

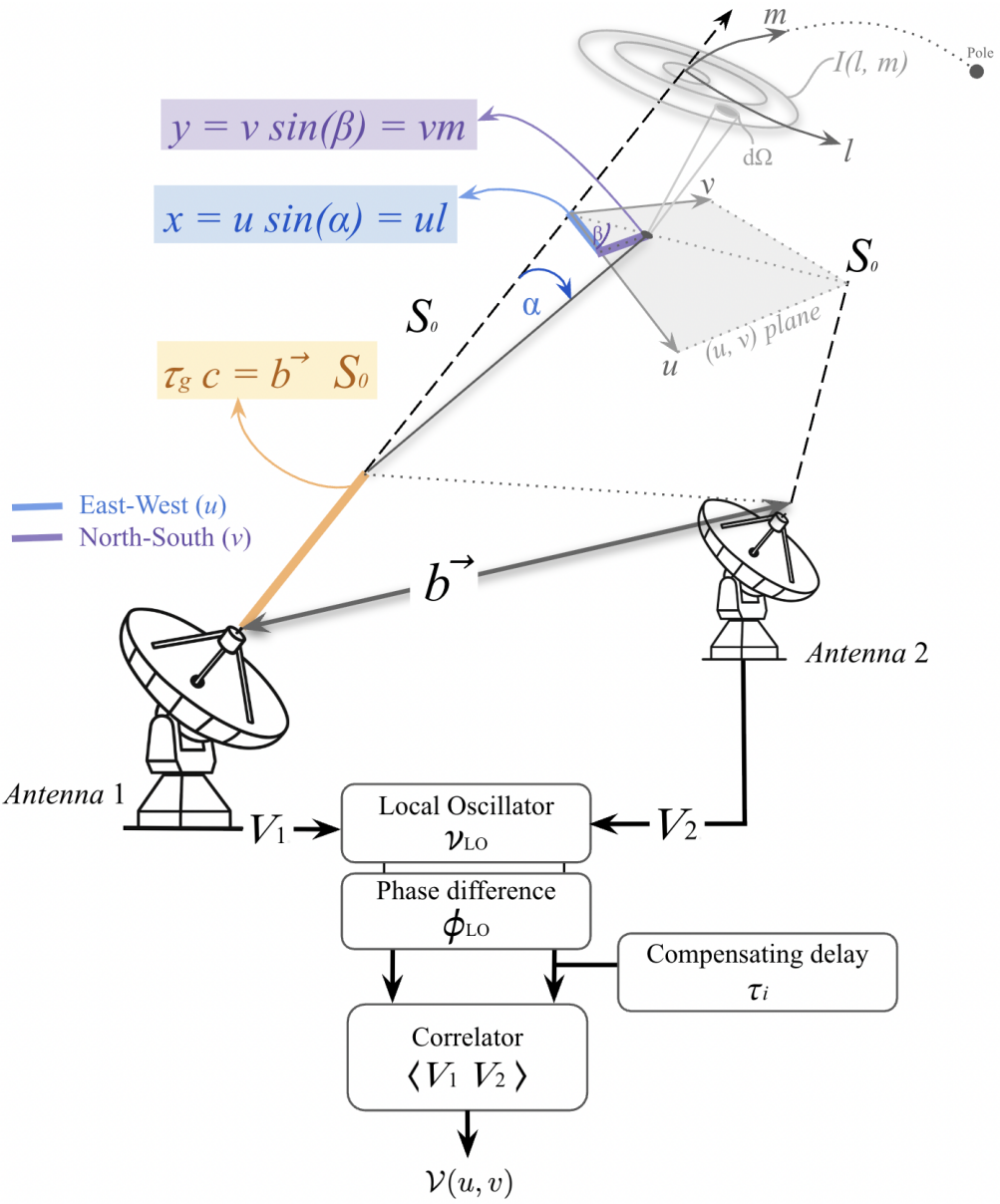

Figure 1.13: Schematic of a two-antenna radio interferometer. A radio source at direction vector S_0 is observed by two antennas separated by baseline vector $\vec{b}$. The source's signal arrives at slightly different times due to the geometric delay τ_g. This delay is compensated electronically, and the signals V_1 and V_2 are mixed with a local oscillator at frequency ν_{LO}, then cross-correlated. The resulting complex visibility $\langle V_1\, V_2 \rangle$, as a function of projected baseline components (u, v), samples the Fourier transform of the sky brightness distribution $I(l, m)$, building the visibility function $\mathcal{V}(u, v)$. The x and y coordinates in the (u, v) plane correspond to angular offsets in α and β relative to the baseline projection, and are scaled by spatial frequency components ul and vm. This framework forms the basis of aperture synthesis imaging used in radio interferometry.

difference is defined as $\phi = 2\pi\,\nu\,\tau_g$. This phase shift encodes angular information about the sky brightness distribution. To systematically analyse these data, the sky is projected onto a small patch using direction cosines l and m, defined relative to the pointing centre (typically $\hat{S}_0$). The baseline vector is projected onto units of wavelength (u, v, w), where (u, v) form the uv plane, the Fourier transform of the sky brightness distribution.

The observed quantity is the visibility function $\mathcal{V}(u, v)$, a complex number measured for each antenna pair, given by the van Citter-Zernike theorem (van Cittert 1934; Zernike 1938), which for a restricted field of view can be defined as:

$$\mathcal{V}(u,v) = \iint I(l,m)\; e^{2\pi i(ul+vm)}\; dl \cdot dm = A\, e^{i\varphi} \tag{1.11}$$

This is essentially the 2D Fourier transform of the sky brightness distribution $I(l, m)$, where (l, m) are direction cosines (see Fig. 1.13), and can be described simply by an amplitude A, and a phase φ. Each baseline at a given time samples one point in the uv plane. As the Earth rotates, the projected baselines change and sample many more points in the uv plane, densely paving the Fourier space (Rowson 1962).

Now, analogous to Eq. 1.9 the point-source sensitivity per visibility is given by:

$$\sigma = \frac{2k_B T_{\rm sys}}{A_{\rm e}\sqrt{t_{\rm int}\,\Delta\nu\,N_b}} \tag{1.12}$$

where N_b is the number of baselines, defined as $N_b = N \cdot (N-1)$, with N the number of antennas. Notably, the noise in the uv plane is additive and Gaussian.

To recover the sky brightness distribution, an inverse Fourier transform of the visibilities is required:

$$I(l,m) = \iint \mathcal{V}(u,v)\; e^{-2\pi i(ul+vm)}\; du \cdot dv \tag{1.13}$$

Then, the image can be reconstructed as the sum of the visibilities, where each visibility has an amplitude and phase representing the brightness and relative positions of the emission on a specific angular scale. The reconstructed image represents the convolution of the true sky brightness with the synthesised beam, which is the point spread function (PSF) of the array determined by the uv coverage.

The synthesised beam determines the image resolution, while the primary beam (anal-

ogous to the field of view in a single dish telescope) defines the area of the sky the interferometer is sensitive to, with sensitivity decreasing away from the pointing centre. Because the synthesised beam has a finite angular size, the flux density detected at any point in the image is typically quoted in Jy beam^{-1}, i.e., Janskys per synthesised beam area. This unit reflects the fact that the interferometer measures flux density distributed over the effective resolution element. For sources much smaller than the beam, the emission is beam-diluted, and the observed brightness underestimates the intrinsic surface brightness. If the source is comparable to or moderately larger than the beam, it is resolved, and the Jy beam^{-1} values approximate the true surface brightness. However, if the source is much more extended than the maximum angular scale recoverable by the array, large-scale emission is filtered out, and significant flux may be missed.

An additional effect to consider is beam smearing, a chromatic aberration caused by averaging visibility data over a finite bandwidth. It primarily affects sources far from the phase centre, causing a radial blurring that reduces peak brightness and artificially broadens the source in the image plane.

In cases where the source is optically thick or follows blackbody emission, it is common to express the intensity in terms of the brightness temperature, T_{B}, for the Rayleigh–Jeans limit ($h\,\nu \ll k\,T$):

$$I_\nu = \frac{2\,k\,\nu^2}{c^2}\,T_{\mathrm{B}} \tag{1.14}$$

This defines the temperature that a blackbody would need to have to produce the same specific intensity I_ν at frequency ν in GHz. To convert Jy beam^{-1} to brightness temperature (K), the synthesised beam must be known. If the beam is approximated as a 2D Gaussian with the major and minor axes (θ_{maj}, θ_{min}) of the resolution element in arcseconds, the conversion becomes:

$$\frac{T}{1\,\mathrm{K}} = \left(\frac{S_\nu}{1\,\mathrm{Jy}}\right)\left[13.6\left(\frac{300\,\mathrm{GHz}}{\nu}\right)^2\left(\frac{1''}{\theta_{maj}}\right)\left(\frac{1''}{\theta_{min}}\right)\right] \tag{1.15}$$

This relationship indicates that for a given flux density per beam, the brightness temperature increases if the beam size decreases, so the angular resolution is crucial when detecting compact or high-temperature features. Conversely, for extended sources that are not fully resolved, the brightness temperature can be underestimated because the emission

is spread over a larger beam area.

Figure 1.14: ALMA antenna array layout on the Chajnantor Plateau in Chile. The maximum baseline spans up to 16 km in the most extended configuration (C-10). The figure is obtained from the ALMA webpage.

ALMA Observatory and Technical Implementation

The Atacama Large Millimeter/submillimeter Array (ALMA) consists of 66 antennas located at 5,050 m elevation in the Atacama Desert in Chile (Fig. 1.14). Its location ensures minimal atmospheric water, critical for sub-mm/mm observations. The main 12-m array can use up to 50 antennas and be arranged in up to 10 different configurations (C-1 to C-10), providing baselines from 15 m to 16 km. The Atacama Compact Array (ACA) has 12 antennas of 7-m optimised for short baselines. The Total Power (TP) Array consists of 4 fixed 12-meter antennas used for single-dish measurements.

Each ALMA antenna is a high-precision parabolic reflector (Cassegrain design), designed to collect incoming radio waves and focus them onto a feed horn or receiver. The receiver, located in the antenna cabin, performs the first stage of signal processing. ALMA supports multiple receiver bands, from Bands 1 to 10, covering 35 GHz to 950 GHz (see

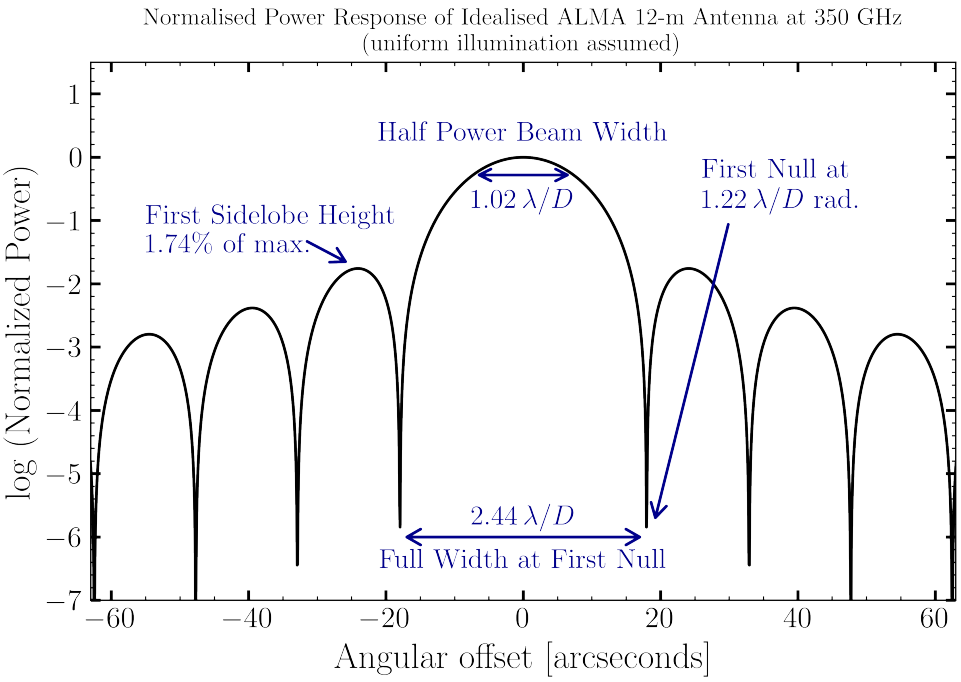

Figure 1.15: Normalised 1D antenna power pattern for a uniformly illuminated 12-meter dish at 350 GHz, shown on a logarithmic scale to highlight sidelobes. The half-power beam width (HPBW) of the main lobe is approximately 1.02 λ/D (i.e., the primary beam), and the full width between nulls (FWBN) is about 2.44 λ/D. The angular resolution, defined by the first null, is roughly 1.22 λ/D (i.e., 0.5× FWBN). In practice, ALMA's 12-m and 7-m antennas have a slightly broader HPBW of ∼1.13 λ/D due to non-uniform illumination.

table 1.3 for the detailed information). High-frequency signals (e.g., around 100 GHz) are challenging to amplify and process directly because electronics at these frequencies tend to be lossy and complex. Therefore, the receiver performs only minimal amplification and then uses the local oscillator (LO) to mix the incoming signal down to a much lower, more manageable intermediate frequency (typically a few GHz). This downconversion enables efficient, low-noise and cost-effective electronics to be used for further amplification, filtering and digital sampling. The LO creates a coherent, high-stability frequency tone that mixes the incoming signal in the receiver. The LO systems are synchronised across all antennas with a high-precision time reference, ensuring a common phase.

After digitisation, signals from all antennas are sent to the correlator, a specialised

supercomputer that performs cross-correlation of all antenna pairs (baselines) and auto-correlation. The correlator multiplies and cross-multiplies signals from antenna pairs and integrates over time and frequency, producing the complex visibilities. The output visibilities depend on the spectral resolution, which defines the velocity resolution of the observed lines; the bandwidth, which affects the sensitivity to the continuum; and the integration time, which impacts the noise and uv sampling. Figure 1.13 shows a schematic representation of this process for a two-antenna array, including the artificial delay inserted into the signal path of antenna 2 to compensate for the geometrical delay (τ_g), so that the signals from both antennas arrive at the correlator with the same phase. Further details can be found in the ALMA Technical Handbook (Cortes et al. 2025).

The antenna power response, or primary beam pattern, depends on the angular distance from the centre of the pointing. For electromagnetic radiation arriving exactly along the optical axis, wavefronts reflect off the dish with no path difference, resulting in a coherent sum at the focus and maximum power. However, for off-axis incident radiation, wavefronts arrive with increasing path differences across the dish surface. This self-interference reduces coherence and sensitivity as the angle increases.

To first order, the angular dependence of the antenna power response due to diffraction can be illustrated using a uniformly illuminated circular aperture. In this idealised case, the voltage response follows $V(\theta) \propto J_1(k\,D\,\sin\theta)/(k\,D\,\sin\theta)$, where J_1 is the Bessel function of the first kind, D is the antenna diamater, θ is the off-axis angle, and $k = 2\pi/\lambda$. The corresponding antenna power response, $P(\theta) \propto V^2(\theta)$, is also known as the Airy function, shown in Figure 1.15. It features a central peak (primary beam) and multiple sidelobes from alternating constructive and destructive interference. For a 12-meter antenna at $\lambda = 0.85$ mm ($\approx$ 350 GHz, ALMA Band 7), the Half Power Beam Width (HPBW) is given by $= 1.02\,\lambda/D$, and the first null in the power pattern occurs at $\theta = 1.22\,\lambda/D \approx 15''$, defined as half the full width between nulls (FWBN) of the primary beam, which is considered the Rayleigh resolution of the antenna.

In practice, real ALMA antennas are not uniformly illuminated. The feedhorns are designed to taper the illumination near the edges of the dish to reduce sidelobes. This results in a nearly Gaussian-shaped beam profile with a slightly broader HPBW of approximately HPBW $\approx 1.13\,\lambda/D$. The emission entering the sidelobes can contribute significantly to the total received signal if the off-axis source is bright. However, as shown in Table 1.3, at all standard ALMA bands, the sidelobes lie outside the typical field of view used in

imaging, making their contribution negligible. This justifies ignoring sidelobes for primary beam correction and supports using a Gaussian beam model that captures the main lobe behaviour only.

Imaging

Each baseline samples one visibility $\mathcal{V}_{ij}(u_k, v_k)$ per integration, so the challenge of image reconstruction arises because we sample discrete parts of the uv plane instead of the complete area. The inverse transform creates a dirty image, given by:

$$I_D(l, m) = I(l, m) * b(l, m) \ , \tag{1.16}$$

where I_D is the dirty image, $I(l, m)$ is the true sky distribution, and $b(l, m)$ is the dirty beam or the point spread function resulting from the inverse Fourier transform of the sampled (u, v) distribution.

Deconvolution is needed to recover the true sky image because the dirty image includes both the actual sky signal and artefacts caused by incomplete sampling. The most widely used algorithm is CLEAN (Högbom 1974), which iteratively: (1) identifies the peak in the dirty image, (2) subtracts a scaled version of the dirty beam, (3) records the subtraction as a CLEAN component, and (4) repeats until a threshold is reached.

CLEAN assumes a point-source model and works well for compact sources. However, it can also perform well for moderately extended emission under the right conditions. Extended emission can sometimes leave residuals or be underfit if the iteration parameters (`niter`, `threshold`) are not well chosen. To avoid this, it is important to set the `threshold` close to the image noise level to prevent noise being cleaned out, and choose `niter` high enough to deconvolve real emission but not so high that over-cleaning occurs. Over-cleaning happens when the algorithm starts removing noise fluctuations as a real signal, producing artefacts and biased fluxes. By carefully tuning these parameters, using prior knowledge of the noise and source structure, the algorithm can effectively handle both compact and extended emission. The final restored image is given by:

$$I_{\text{restored}}(l, m) = I_{\text{CLEAN}}(l, m) * b_{\text{restore}} + R(l, m) \tag{1.17}$$

where, I_{CLEAN} is the sum of CLEAN components, b_{restore} is the ideal restoring Gaussian beam, and $R(l, m)$ are the residuals.

To illustrate these concepts, Figure 1.16 presents simulated ALMA observations using a point source model in the top row and a point source plus an elongated Gaussian feature in the bottom row. Panel (a) shows the dirty beam, the synthesised PSF from the inverse Fourier transform of the (u, v) sampling. Panel (b) shows the (u, v) coverage, panel (c) is the true sky model (point source, and point source+Gaussian), (d) is the dirty image, which is the convolution of the true sky with the dirty beam. Panel (e) shows the restored image, obtained by convolving the image with the restoring beam using the CLEAN algorithm.

Interferometers filter out large-scale structure because they lack short baselines. The minimum baseline $B_{\rm min}$ sets the largest angular scale recoverable: $\theta_{\rm max} \approx \lambda/B_{\rm min}$. Similarly, the longest baseline determines the finest angular resolution $\theta_{\rm min} \approx \lambda/B_{\rm max}$. Thus, combining data from compact and extended configurations helps recover all spatial scales. To reproduce the imaging workflow for a radio interferometer (Fig. 1.16), both compact and extended configurations were included. In both examples (one with a pure point source and another with an additional elongated Gaussian feature), CLEAN successfully reconstructs the image. These figures were produced using CASA (Common Astronomy Software Applications, McMullin et al. 2007) package, developed to process VLA and ALMA data. The deconvolution method is implemented in CASA under the `tclean` task, an updated and flexible version of CLEAN. `tclean` incorporates advanced imaging capabilities such as multi-frequency synthesis, multi-scale deconvolution, wide-field imaging corrections, and automated masking, making it the standard tool for high-fidelity image reconstruction in ALMA data reduction.

Calibration

An accurate calibration process is essential for every interferometric observation, as it corrects for instrumental, atmospheric and antenna-based effects in time and frequency that could otherwise distort the incoming signal. In the context of ALMA, calibrators play a crucial role in correcting the phase, amplitude, flux density, and bandpass responses. These corrections are achieved through observations of dedicated calibrator sources, typically bright, compact quasars with known flux properties and stable spectral characteristics.

Observations are organised into execution blocks (EBs), each containing a single set of instructions for the telescope, including target position, frequencies, and calibration details, designed to be completed in one session. Multiple EBs are combined to complete a full observing project. Each EB includes several categories of calibrators, selected to correct

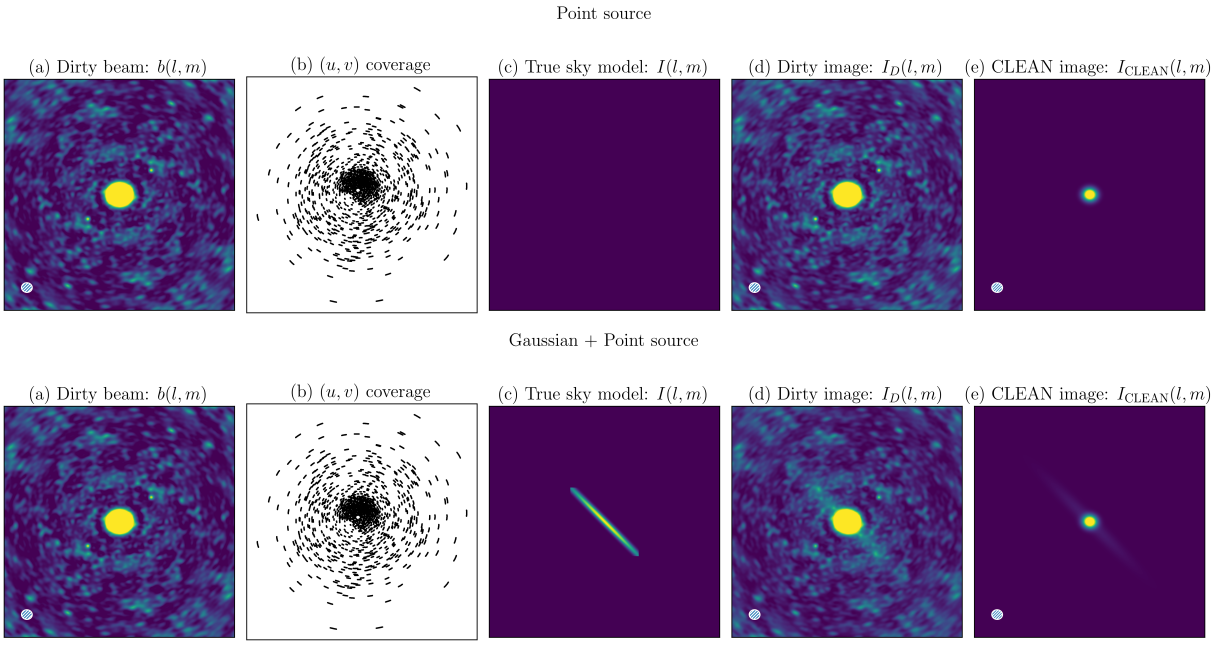

Figure 1.16: Simulated ALMA imaging pipeline for two synthetic models: a point source (top) and a point source plus a Gaussian elongated feature (bottom). (a) Synthesised dirty beam $b(l,m)$, showing the interferometer's PSF with sidelobes, derived from the uv sampling pattern. (b) The (u,v) coverage of the combined compact and extended array configurations, taking into account Earth rotation, populates the uv plane. (c) True sky brightness distribution $I(l,m)$ used in the simulation: a point source (top row), and a point source with an elongated Gaussian (bottom row). (d) Dirty image $I_D(l,m) = I * b$, showing the source convolved with the dirty beam. (e) CLEANed and restored image using a Gaussian restoring beam. These figures were produced using the CASA task `simobserve` to illustrate an example of the imaging workflow for a radio interferometer, including Fourier sampling, image formation, and deconvolution.

specific instrumental, atmospheric, or antenna-based effects (Cortes et al. 2025):

- **Flux calibrators** are used to establish the absolute flux scale of the observation. The correlator initially outputs amplitudes as a percentage of the system noise, which is converted to brightness temperature (K) using the system temperature measurements from ALMA's calibration devices. To convert from Kelvin to Jansky (Jy), an observation of a flux calibrator is needed. These models are accurate to ∼5% at

low frequencies, but the uncertainty increases to ∼20% in Bands 9–10. Most ALMA scheduling blocks include one scan, often serving both as a flux and bandpass calibrator. This scan is not shared between different observation runs, so it can accumulate a significant amount of telescope time.

- **Bandpass calibrators** correct the frequency-dependent response of the instrument across the observed bandwidth caused by the electronics and atmosphere. This calibration is essential for removing instrumental spectral features from the visibilities, particularly for accurate spectral line analysis and continuum subtraction. ALMA benefits from excellent temporal stability in its bandpass response (Kameno 2015)[2], so the bandpass calibrator is typically observed only once at the beginning of an execution block. These calibrators are usually bright, compact, and spectrally flat sources selected from a grid of well-characterised quasars located within about 30° of the science target. Depending on the desired accuracy and source strength, the scan duration ranges from 2 to 30 minutes. The resulting calibration table is applied to correct each frequency channel across the band.

- **Amplitude and phase calibrators** are used to correct amplitude fluctuations and atmospheric path length variations that introduce time-dependent variations in visibilities. Amplitude variations can occur due to changes in antenna efficiency with elevation, thermal effects on antenna structures, or pointing errors, especially at high frequencies and during the day. These can introduce flux variations of up to ∼10%. When the phase calibrator is bright enough, its known flux can be used to track and correct amplitude changes throughout the observation. Phase calibration is critical for image fidelity. It corrects rapid atmospheric and instrumental phase fluctuations by frequently switching between the science target and a nearby phase calibrator. The visits to the phase calibrator occur on timescales ranging from about 40 seconds for the most extended configurations at the highest frequencies, up to 600 seconds for the most compact configurations. The calibration is usually derived per spectral window and then interpolated to the science target. For reliable solutions, a stable, unflagged reference antenna, meaning an antenna that provides reliable data, located near the array centre is essential.

Most ALMA extragalactic calibrators are bright, compact, flat-spectrum radio sources,

[2]ALMA Cross Correlation Bandpass Stability (Kameno 2015)

such as quasars or blazars. These active galactic nuclei are ideal calibrators for several reasons. First, due to their compactness, most quasars used for calibrations are point-like at ALMA's resolution, minimising structural phase errors and simplifying the calibration model. Second, their intrinsic brightness ensures high signal-to-noise in short integrations. Third, their stability and monitoring, since ALMA keeps an updated ALMA calibrator source catalogue, regularly observed across multiple bands to track flux variability, spectral index, and structure.

Looking at astronomical objects in the same field as ALMA calibrators has a particular advantage, since they offer the unique possibility of self-calibration. Self-calibration is a post-processing technique that further improves the image quality by using the science target itself as a calibration source. This is possible when the target is sufficiently bright ($\gtrsim 50$ mJy) and compact (angular size smaller than the synthesised beam), typically with a peak signal-to-noise ratio exceeding $\sim 5 - 10$ in a given solution interval. In the initial model, the visibilities are first Fourier-inverted using standard calibration to produce a dirty image. The clean image generates a model of the target, which is then Fourier-transformed and compared to the observed visibilities to derive the gain solutions. The updated gains are applied to the visibilities, and the imaging process is repeated. Several iterations are performed, often starting with phase-only self-calibration and later incorporating amplitude self-calibration. The final image shows reduced phase errors, improved signal-to-noise ratio, and minimised artefacts such as sidelobes or spurious structures. Reviews on self-calibration methods are given by Pearson & Readhead (1984) and Cornwell & Fomalont (1989).

Self-calibration not only improves image fidelity, which results in sharper features and reduced sidelobes, but also enhances the dynamic range and signal-to-noise ratio of the final image. Nevertheless, it requires careful validation: over-calibration on a poor model can introduce artefacts like sidelobes or suppress real emission. Thus, each iteration must be guided by residual inspection, rms monitoring, and model stability.

Future Prospects

ALMA stands as a long-term facility, committed to continuous advancements. The observatory is currently undergoing transformative upgrades, including the ALMA Wideband Sensitivity Upgrade (WSU), scheduled for completion by 2030. This major enhancement will dramatically improve ALMA's capabilities. It will deliver a fourfold increase in spec-

tral line survey speed and boost continuum sensitivity by a factor of two, making ALMA more efficient. Looking ahead, the ALMA2040 initiative is currently evaluating potential upgrades, which will be defined in response to the evolving needs of the scientific community.

Beyond its standalone capabilities, ALMA is a cornerstone of the Event Horizon Telescope (EHT), a groundbreaking experiment that uses many of the world's most powerful radio observatories into a single, Earth-sized interferometer. This global VLBI network has, for the first time, imaged emission from material extremely close to black hole event horizons (Event Horizon Telescope Collaboration et al. 2019). The landmark images of M87 and Sagittarius A* stand as one of the most impressive achievements in interferometry, made possible by cutting-edge instrumentation, precise synchronization, and worldwide collaboration. Expanding the EHT network with next-generation facilities, such as the proposed Atacama Large Aperture Submillimeter Telescope (AtLAST), will open the door to imaging fainter targets and provide improved baseline coverage, enabling even sharper and more detailed studies of black hole environments.

In the future, planned upgrades and continued global collaboration will ensure that ALMA remains at the forefront of astronomical discoveries for decades to come.

1.5 This thesis

This thesis presents three related research projects that build upon two observational programs: ALMACAL and MUSE-ALMA Haloes surveys. These projects utilise archival and targeted data from ALMA to investigate the molecular gas and its connection to the circumgalactic medium of galaxies across cosmic time. Below is a summary of the key results already obtained from these surveys, providing context and motivation for the ongoing work.

1.5.1 The ALMACAL survey

The ALMACAL survey exploits calibration data from the Atacama Large Millimeter/submillimeter Array (ALMA) to conduct deep, blind observations across a wide range of scientific cases. Calibrator scans, originally used to correct for phase, amplitude, flux, and bandpass, are repurposed to produce deep, blind observations benefiting from large

telescope time. This strategy greatly increases the scientific value of ALMA's operations. The calibrators are chosen to be bright, compact extragalactic sources like flat-spectrum blazars, and near to the science targets. Although their observation schedule and frequency coverage are irregular, this randomness helps reduce environmental biases. Around 78% of these sources have known redshifts (mostly archival), with ongoing efforts to improve redshift accuracy and completeness through follow-up spectroscopy with X-Shooter on VLT (Weng et al. 2025).

Each ALMA execution block typically includes 2–3 calibrator scans, which are made immediately public and subsequently processed in ALMACAL through a systematic data reduction process: self-calibration, point-source subtraction, and imaging at uniform spectral resolution ($\sim$15.6 MHz). The dataset spans from ALMA Bands 3 to 10 ($84-950$ GHz) and covers a variety of spatial resolutions, enabling multi-scale analysis across diverse physical regimes. Since its inception in 2016, ALMACAL has demonstrated that calibrator fields can achieve sensitivities comparable to those of dedicated deep surveys, while covering much wider areas and reducing cosmic variance. This approach has opened new avenues to study dusty star-forming galaxies (DSFGs), active galactic nuclei (AGN), and the cold interstellar medium (ISM) across cosmic time (Oteo et al. 2016b, 2017; Klitsch et al. 2018; Bonato et al. 2018; Klitsch et al. 2019a, 2020; Hamanowicz et al. 2022; Chen et al. 2022; Klitsch et al. 2023; Chen et al. 2023; Bollo et al. 2024, 2025; Weng et al. 2025).

1.5.2 The MUSE-ALMA Haloes survey

The MUSE-ALMA Haloes survey was conceived to investigate the multiphase CGM through an absorption-selected approach, leveraging a powerful combination of instruments:

- HST: UV spectroscopy for H I selection using either FOS, COS, or STIS. High-resolution imaging is used for stellar morphology and spectral energy distribution fitting.
- VLT/MUSE: 3D optical Integral Field Unit (IFU) detects galaxies in emission around known quasar absorption systems.
- ALMA: Traces the molecular gas through the CO emission line in the galaxies associated with the absorbers.

- VLT/UVES/HIRES: High-resolution spectroscopy to characterise the neutral and metal phases of the absorber through quasar absorption lines.

Targeting strong H I absorbers ($\log N(\mathrm{H\,I})$ [cm^{-2}] $= 18.1 - 21.7$) at $0.2 < z < 1.4$, the survey has identified 32 absorption systems across 19 MUSE fields, with 79 associated galaxies within ± 500 km s^{-1}. The scientific development of the program proceeded through successive phases, each probing specific physical aspects of the galaxy-gas connection.

Early efforts of the survey used MUSE observations to demonstrate that a previously single-assigned absorber was, in fact, most often embedded within a galaxy group of four members (Péroux et al. 2017). The kinematics of the brightest galaxy's disk, traced via [O III] emission and the absorber's position and velocity, suggested that the absorption could arise from either a warped disk or intragroup gas, indicating a complex CGM environment beyond isolated host models. Follow-up high-resolution spectroscopy by Rahmani et al. (2018a) provides observational evidence supporting cold-mode accretion, identifying metal-poor absorption components co-rotating with the galaxy's disk but offset in velocity and angular momentum, consistent with warped inflowing gas. Conversely, other kinematic components in the same field aligned better with galactic outflows, illustrating that individual absorbers can encode multiphase flows with opposing dynamics.

The multiphase nature of the CGM was further studied by Klitsch et al. (2018), who combined ALMA and MUSE data to show that molecular gas, traced via CO(2−1) or CO(3−2), often extends beyond the optical ionised disk. Notably, in one system, CO emission was detected entirely off-nucleus, indicating that the molecular phase of the CGM is not necessarily confined to galactic interiors, and may participate in larger-scale exchange between galaxies and their surroundings.

Hamanowicz et al. (2020) provided the first pilot statistical insight into the relationship between galaxies and their surrounding absorbers, based on a subset of 14 out of 32 systems. Their study revealed a high galaxy detection rate (>85%), with most absorbers residing in group environments (3−11 members), and confirmed established anti-correlations between H I column density and impact parameter. They also found that absorber metallicities were systematically lower than those of the nearest galaxies, reinforcing the interpretation that CGM gas is either less enriched or originates from outside the immediate stellar disk.

The alignment of gas phases was explored in depth by Szakacs et al. (2021), who found compelling evidence of co-rotating molecular and ionised disks in a subset of absorbers. For example, they found a system where the molecular and [O III] kinematics were closely

aligned, and a low-metallicity absorption component in the same sightline suggested the presence of an infalling structure, likely cold gas from the CGM feeding the disk.

The survey's comprehensive phase was presented in Weng et al. (2022), which included results from the full MUSE-ALMA Haloes dataset, comprising 32 absorbers and 79 associated galaxies across 19 quasar fields. The study reinforced earlier trends: most absorbers have one or more galaxy associations within ±500 km/s, and their H I column densities decline with increasing normalised impact parameter. A consistent gradient between absorber and galaxy metallicities, especially at large distances, was found, providing new constraints on the metal transport and enrichment processes in galaxy haloes.

With morphological decompositions using HST and GALFIT, Karki et al. (2023) examined how galaxy structure modulates CGM properties. Their analysis of 66 galaxies revealed that stronger H I absorbers are statistically linked to smaller, more compact galaxies, and that both emission and absorption metallicities scale positively with stellar mass and specific star formation rate (sSFR). This suggests that metal-poor absorbers preferentially trace galaxies with modest past star formation and vigorous present-day gas consumption, indicative of evolving systems undergoing active baryon cycling.

A major advance in understanding CGM kinematics came from Weng et al. (2023a), who performed 3D forward modelling of absorber–galaxy configurations. Their analysis of 27 absorbers revealed a tentative bimodal distribution of azimuthal angles, with some absorbers aligned with disk planes (infall or co-rotation), and others along minor axes (outflows). Approximately 10% of absorbers showed kinematics consistent with infall, 30% with outflows, and the remainder with either extended disks or intragroup material. This provided the clearest evidence to date that H I absorbers encode both inflow and feedback signatures, crucial components of the baryon cycle.

Complementing the observational dataset, Weng et al. (2024) used the TNG50 cosmological simulation to trace the origins of H I absorbers at $z \sim 0.5$. They found that while some absorbers directly trace the CGM gas of central galaxies, a significant fraction, especially those at large galactocentric distances or with lower column densities, originate in satellite systems, nearby haloes, or the intergalactic medium (IGM). This theoretical insight stresses the importance of viewing the CGM as embedded within a dynamically connected cosmic web, rather than as an isolated halo phenomenon.

Finally, Augustin et al. (2024) derived stellar masses for the full galaxy sample and demonstrated a strong anti-correlation between stellar mass and H I column density. Galax-

ies with $\log M_\star > 11 M_\odot$ exhibited notably low surrounding H I content within ~ 120 kpc. This finding suggests that massive galaxies may efficiently deplete or expel neutral gas in their haloes, whether through star formation, feedback, or environmental processes.

Together, these foundational studies within the MUSE-ALMA Haloes project illustrate the power of combining absorption- and emission-based diagnostics to uncover the structure, kinematics, and evolution of the CGM across cosmic time.

1.5.3 Thesis Overview

The research presented in this thesis contributes to a deeper understanding of the physical processes that drive galaxy evolution. By combining wide-area, blind submillimetre observations from ALMACAL with the spatially resolved optical spectroscopy of MUSE-ALMA Haloes, this work addresses critical phases in the lifecycle of baryons: the accretion of gas onto galaxies, the conversion of this gas into stars, and the ejection of enriched material via feedback.

Built upon the ALMACAL legacy, Chapter 2 introduces the latest version of ALMACAL, dubbed as ALMACAL−22, and its calibration framework. Key questions addressed include:

- How can calibration observations be used to build a deep, wide-area sub-millimetre survey?
- How does ALMACAL compare with surveys in terms of depth, area, and cosmic variance?
- What can this dataset reveal about molecular gas, dusty galaxies, and AGN across cosmic time?

In Chapter 3, we move to one of ALMACAL's core cosmological applications: the blind search for CO emission lines from background galaxies. This work investigates the following questions:

- Can CO-selected galaxies from ALMACAL provide an unbiased view of cold gas across cosmic time?
- How much cold molecular gas existed at different times in the universe's history?

- What do these trends reveal about the baryon cycle and gas fuelling star formation in galaxies?

Chapter 4 presents a comprehensive statistical analysis of the molecular gas content in galaxies associated with H I absorbers. New, deep observations from an ALMA Large Program (Cycle 10, PI: Péroux) were specifically designed to detect CO emission lines, thereby enabling an unprecedented characterisation of the cold molecular phase within the CGM. The central scientific questions addressed are:

- What are the typical molecular gas properties of galaxies blindly selected through H I absorption, and what physical properties drive their gas content?
- How do the properties of gas-rich absorption-selected galaxies compare to the general population of star-forming galaxies at similar redshifts?
- What can we learn about star formation efficiency and cold gas cycling from the depletion timescales in these systems?

Together, Chapter 2 and Chapter 3 demonstrate how rethinking calibrator data as a scientific resource not only increases the scientific value of calibration strategies but also opens new windows into the cosmic estimates of the molecular gas mass evolution in the Universe. Simultaneously, Chapter 4 provides a window into the multiphase structure of the CGM, where galaxies interact with their environment through inflows, outflows, and recycling processes. Ultimately, these studies strengthen the empirical foundations of galaxy evolution models by linking the small-scale physics of cold gas with the large-scale dynamics of the baryon cycle. They also provide critical constraints for simulations aiming to capture the complexity of multiphase gas flows, the build-up of galaxy mass, and the interplay between galaxies and their cosmic environment. Chapter 5 concludes this thesis by summarising the key findings, placing them within the broader framework of galaxy evolution, and outlining future directions, including upcoming projects that will build upon and extend this work.

Chapter 2

ALMACAL–22. Data characterisation and products

The content of this chapter was presented in the published article Bollo et al. (2024), Astronomy & Astrophysics, 690, A258.

2.1 Introduction

Millimetre (mm) and sub-millimetre (sub-mm) observations offer a unique window into various astrophysical processes and phenomena in the Universe. At these wavelengths, key emission lines such as carbon monoxide (CO) and atomic and ionised carbon ([C I], [C II]) emit radiation that reveals essential information about molecular gas content, star-formation activity, and gas dynamics (Carilli & Walter 2013; Tacconi et al. 2020). Sub-mm observations help us explore the cosmic baryon cycle and the complex interplay between cold and dense molecular gas reservoirs, star formation activity, ionised gas, and dust emission (Péroux & Howk 2020).

Modern interferometers have produced large datasets consisting of hundreds to thousands of individual observations. In particular, the Atacama Large Millimeter/submillimeter Array (ALMA) is a pioneering observatory that provides high-resolution observations at mm and sub-mm wavelengths. Transforming the large volume of raw visibility (uv) data from ALMA into scientifically meaningful products is a significant task. This complex process includes several steps: calibration, imaging, and deconvolution.

ALMA has made remarkable progress in automating and ensuring high-quality cali-

bration of interferometric data. These advances have been driven by the diligent efforts of the observatory staff and the success of ALMA's sophisticated pipeline (Hunter et al. 2023), based on the Common Astronomy Software Applications (CASA) software (McMullin et al. 2007). As a result, ALMA provides its users with meticulously calibrated visibilities, laying the groundwork for further processing and scientific analysis. Despite the effective functionality of the pipeline, certain complications can arise during the calibration process, requiring manual intervention. These complications can include unexpected radio frequency interference, atmospheric anomalies, or misbehaved antennas, all of which affect data quality. In such cases, the expertise of astronomers becomes crucial to correct or flag these problems, ensuring the integrity and reliability of the final scientific products.

Over the years, several large programs (LPs) have exploited the capabilities of ALMA to study the cosmic evolution of gas and stars. In particular, the ASPECS survey in the Hubble Ultra Deep Field (HUDF; Walter et al. 2016; Decarli et al. 2016) aimed to detect CO and [C I] in galaxies without preselection at $z = 1 - 3$ using Bands 3 and 6. ASPECS covered an area of 4.6 arcmin^2 in Band 3 and 2.9 arcmin^2 in Band 6. PHANGS–ALMA (Leroy et al. 2021) is a survey of CO(2–1) emission from 90 nearby galaxies with a typical angular resolution of $\sim 1.5''$ and a total survey area of 1050 arcmin^2. At higher redshifts, the REBELS survey (Bouwens et al. 2022) targeted 40 UV-bright galaxies at $z > 6.5$, covering an area of 7 deg^2, and aimed to detect the [C II]-158μm and [O III]-88μm lines as well as dust-continuum emission. The ALPINE survey (Le Fèvre et al. 2020) was designed to study 118 star-forming galaxies (SFGs) at $4 < z < 6$. This survey targeted the [C II] line and continuum emission, covering an area of 25 arcmin^2. The CRISTAL survey (Solimano et al. 2024) selected 25 SFGs with available HST imaging and high stellar masses from the ALPINE sample ($\log M_{star}/M_\odot \geq 9.5$). CRISTAL targets [C II] with a resolution of $\sim 0.2''$, which is higher than its parent sample's resolution of $\sim 1''$. These surveys have played a key role in constraining the evolution of the molecular gas in galaxies across cosmic epochs. Using different molecular gas tracers shows that the evolution of the molecular gas mass density in the Universe aligns with the cosmic star formation history, providing insights into the process of gas accretion onto galaxies (Walter et al. 2020). In addition to the ALMA LPs, the ALMA Calibrator source catalogue has also been used by Audibert et al. (2022) on a representative sample of the NVSS (when flux limited to 0.4 mJy) to study the CO luminosity function up to redshift $z \sim 2.5$ and to assess the role of radioactivity in galaxy evolution. They found that most radio galaxies are more depleted and evolved

than the typical simulated halo galaxy.

Other facilities have also been used for surveys of cold gas tracers. The Plateau de Bure High$-z$ Blue Sequence Survey 2 (PHIBSS2) (Guilloteau et al. 1992), is a large observational campaign conducted with the Plateau de Bure Interferometer (PdBI), now part of the Northern Extended Millimeter Array (NOEMA) Observatory. With a frequency range spanning from 80 GHz to 350 GHz, Lenkić et al. (2020) searched for additional background emission lines in the fields of the original PHIBBSS survey. They explored the CO$(2-1)$, CO$(3-2)$, and CO$(6-5)$ emission lines in 110 main-sequence galaxies, covering a total area of ~ 130 arcmin^2. This survey has been used to derive the molecular gas mass density evolution by converting high$-J$ CO luminosity functions to CO$(1-0)$. Additionally, the COLDz survey looked directly for the CO$(1-0)$ emission at $z = 2-3$ and CO$(2-1)$ at $z = 5-7$ using more than 320 hrs of VLA time (Pavesi et al. 2018; Riechers et al. 2019, 2020b), covering a ~ 60 arcmin^2 area. However, the large uncertainties of these measurements reflect our limited understanding of the molecular gas content of galaxies across cosmic time.

Our current measurements of the molecular gas mass density ($\Omega_{\rm H_2}$) reach redshift $z \sim 7$ and show that the density increases from early times, peaks at $z \sim 1-3$, and decreases to the present day (e.g. Tacconi et al. 2010; Walter et al. 2020; Hamanowicz et al. 2022; Aravena et al. 2024). However, one of the main challenges for the molecular gas surveys is the effect of cosmic variance, which introduces uncertainty on how well the sampled volume represents the Universe (Decarli et al. 2020; Popping et al. 2020; Boogaard et al. 2023). Cosmic variance, caused by natural fluctuations in the Universe's large-scale structure, results in variations in the number density and distribution of cosmic objects across the sky (Driver & Robotham 2010). Consequently, surveys covering limited sky areas may inadvertently sample regions that are unusually rich or devoid of galaxies, leading to biased estimates of cosmic properties such as the number density of galaxies, clustering, and luminosity function. Small sample sizes exacerbate this issue by introducing statistical uncertainties, which increases the probability of sampling regions with atypical characteristics. Several studies have investigated the effects of field-to-field variance on observables such as the luminosity function and the number density of galaxies (Keenan et al. 2020; Lenkić et al. 2020; Gkogkou et al. 2022; Boogaard et al. 2023). Generally, there is a consensus on the significance of measuring the cosmic variance effect by comparing model predictions with observations and estimates from various sky regions (Popping et al. 2019b). This is

particularly crucial in deriving the molecular gas mass density of the Universe.

This chapter presents a new survey based on ALMA calibrator data accumulated up to May 2022, which we denominated as ALMACAL−22. The original ALMACAL project started in 2016 (Oteo et al. 2016a) and has produced several scientific outcomes dedicated to studying molecular gas, dusty star-forming galaxies (DSFGs), absorption lines along the line of sight of quasars, and active galactic nucleus (AGN) physics. ALMACAL−22 was built on the experience of previous pilot ALMACAL surveys, but it has now been expanded to include longer integration times. As part of the strategy to exploit this large dataset, we present the details of processing and imaging data, along with different tests, to ensure the best quality of the sample selection. Here, we provide the characteristics and properties of this new dataset. We compare the strengths of ALMACAL−22 to previous surveys. This new survey covers 1047 fields across the southern sky, intending to alleviate the limitations introduced by Poisson errors due to limited statistics and cosmic variance. We review various studies conducted since ALMA Cycle 1 that have used extragalactic calibration data to explore interesting scientific cases (e.g. Oteo et al. 2017; Klitsch et al. 2019b; Hamanowicz et al. 2022; Chen et al. 2022). We discuss how this new release, ALMACAL−22, can expand our understanding of molecular gas evolution, properties of DSFGs, extragalactic absorption lines, and AGN physics.

This chapter is organised as follows. In Sect. 3.2 we describe the ALMACAL−22 survey, the calibration process (Sect. 2.2.1), the selection of the pruned sample (Sect. 2.2.2), and the concatenation and imaging (Sect. 2.2.3). In Sect. 2.3 we explore the following properties: spatial distribution (Sect. 2.3.1), spatial resolution (Sect. 2.3.2), integration time (Sect. 2.3.3), and calibrator redshifts (Sect. 3.3.4). Section 4.5 compares the strengths of ALMACAL−22 with the previous surveys and discusses the scientific areas where ALMACAL−22 will significantly contribute. In Sect. 4.6 we summarise our key conclusions. Throughout this chapter, we use $H_0 = 70\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_\mathrm{M} = 0.3$, and $\Omega_\Lambda = 0.7$.

2.2 ALMA calibrator data

The ALMACAL−22 survey comprises archival data compiled from the calibration data used in every science project carried out by ALMA (Zwaan et al. 2022). Each PI-led scientific project involves several observations of a calibrator source that is close to the science

field. Most calibrators are bright sub-mm point sources classified as blazars (Bonato et al. 2018) — AGN galaxies with a jet pointing towards the line of sight (Urry & Padovani 1995). Blazars can be divided into two sub-classes: BL Lac objects, which are identified as radio galaxies, and flat-spectrum radio quasars (FSRQs), which are identified as quasars Padovani (2017). When targeting a source with an unknown structure and flux, calibration observations are crucial in interferometric astronomy. We used calibrator sources with well-known shapes and flux densities at (sub-)mm and radio wavelengths to adjust the bandpass response, set the flux density scale, and calibrate the amplitude and phase. This ensures we can correct for instrumental and atmospheric corruptions, providing accurate measurements of the target source. These calibration scans have exposure times and setups matching the project's PI requests. Repeated use of the calibrators for ALMA science observations effectively creates deep observations that often cover a significant fraction of the sub-mm spectrum. The most popular calibrators have data from multiple observations of different ALMA bands. Multiple observations of ALMA calibrators create a high-sensitivity dataset that covers a significant sky area (> 1000 arcmin^2).

This chapter presents 1047 calibration field data points accumulated from Cycle 1 (July 2012 – May 2022) from Band 3 through Band 10, resulting in a dataset of more than 30 Terabytes. These calibrator data are accessible to any user right after the main science dataset has passed quality assurance. In this section we explain the calibration process and subtraction of the calibrator at the centre of the pointing as well as the selection of the so-called pruned sample. Figure 2.1 shows the ALMACAL−22 field distribution on the sky for the full sample (top) and the pruned sample (bottom).

2.2.1 Calibration

A dedicated ALMACAL pipeline was created to automate the processing of all delivered datasets. The complete pipeline description is available in Oteo et al. (2016a), and the following is a brief overview of its essential components. This pipeline uses the `scriptForPI.py` script included in every delivered ALMA dataset, which generates fully calibrated data. While ALMA users use this script to create a final dataset, we used it to distinguish the calibrated calibrator data from the science observations. Next, we applied self-calibration to correct the short-time variability of the phase and amplitude during integration, improving the final image's dynamic range. This process enabled us to create a

model and remove the bright calibrator from the visibility data, resulting in the equivalent of blank sky or deep field observations.

During the execution of the ALMA calibration scripts, some bandpass calibrators needed to be consistently calibrated. Unlike with phase calibrators, calibration tables are not always applied to the bandpass calibrators. Most of the time, the bandpass calibrator is also the flux calibrator, so their solutions cannot be applied to themselves. To rectify this, the accurate flux density scale of the bandpass calibrators has to be recovered from the flux calibration table containing standardised values for each calibrator. Applying bandpass calibration tables to bandpass calibrators could, in principle, remove emission and absorption features from cubes that are constructed from these bandpass calibration observations. This could affect emission and absorption features at the phase centre. It should be noted, however, that the bandpass solution is spectrally averaged so that narrow spectral features are not calibrated out. Also, our ability to detect faint emission lines throughout the cubes is unaffected. Because emission lines on top of the quasar continuum are weak compared to the quasar continuum, they do not significantly affect the bandpass solution and do not affect the calibration quality of the data.

The calibration process extends to creating "pseudo-continuum" measurement sets, in which all channels within each spectral window are averaged. The averaging also boosts the signal-to-noise ratio (S/N) for the self-calibration. These files facilitate the implementation of two steps of self-calibration, initially focusing solely on phase and subsequently incorporating both amplitude and phase calibration. The solution interval for the calibration solution is set equal to the integration time. During the intervals of the self-calibration steps, a point source model is applied and fitted to the uv data. This approach offers the advantage of subtracting the point-source model independently for each observation, mitigating the impact of any flux variability in the calibration source. The final step involves deriving the calibrated visibilities and subtracting the continuum of the calibrator source to build line emission data cubes, thereby assembling the ALMACAL−22 dataset.

2.2.2 Pruned sample selection

While the central bright calibrator enables self-calibration, it introduces challenges to the scientific analysis. Most of the calibrators in our sample are blazars, but a few present extended structures that cannot be modelled as point sources. After subtracting the cal-

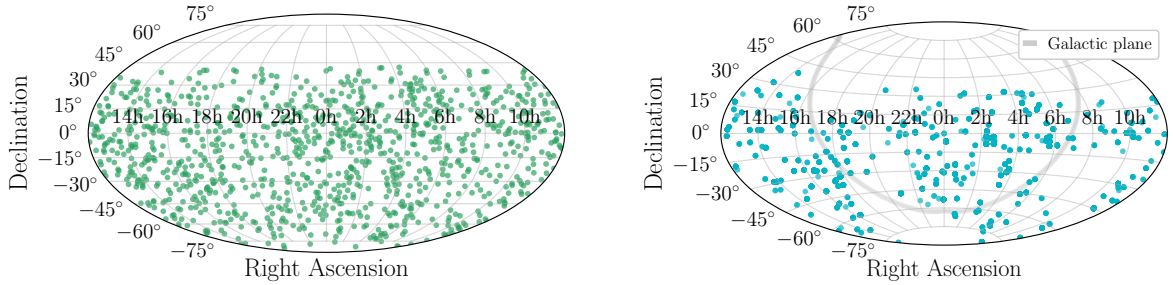

Figure 2.1: Spatial distribution of the calibrator fields for the full ALMACAL−22 (top) and pruned samples (bottom). The distribution of the full sample has no preferred direction, so the position of the calibrators in the sky can be considered homogeneous and only affected by the fact that ALMA observes the sky below $\delta < 40°$. In the pruned sample, two regions exhibit sparse data, possibly attributed to interference, the limited availability of deep fields, and observational difficulties stemming from their proximity to the galactic plane. Further details are provided in Sect. 2.3.1.

ibrator's continuum at the centre of the pointing as a point source, these structures may remain as residuals since they are not taken into account during the cleaning process. These residuals affect the quality of the calibration process, resulting in strong continuum artefacts, interferometric patterns, or extra features that produce high noise values. To overcome this problem, we implemented a sequence of pruning steps to prevent any negative impact on the final combined images.

We analysed key properties, such as the integration time, frequency coverage, bandwidth of each spectral window, spatial resolution, and baseline distribution. We aim to obtain a coherently combined homogeneous sample that will provide broad statistics of the Universe, minimise cosmic variance, and cover the maximum possible volume.

The pruned sample includes only observations made with the 12-metre array. To create the deep cubes, we summed the on-source integration time of observations that covered the same frequency range. We chose bins of 1 GHz to select all the files spanning the same frequency across the extent of the ALMA band's frequency coverage. To minimise contamination from artefacts, we used a multi-step approach. First, we created the continuum image for each measurement set using `CASA` and estimated the root-mean-square (RMS) noise. We chose to use the continuum image instead of the data cube because

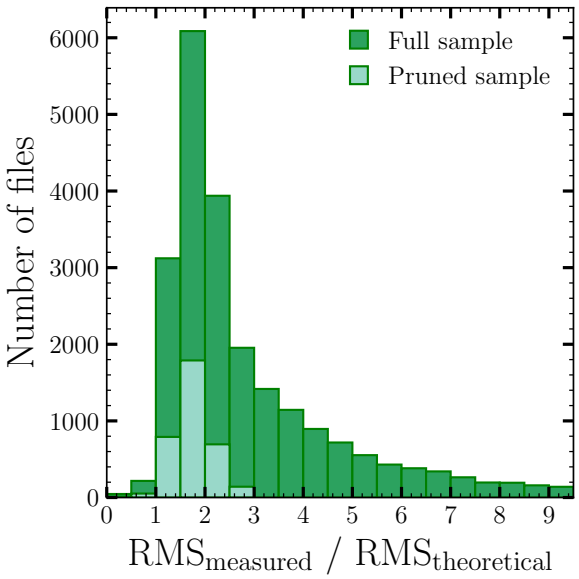

Figure 2.2: Distribution of the RMS of measured flux values in maps over the theoretical ones for each file in the full ALMACAL−22 and pruned samples. The pruned sample has a median of 1.8, reaching the imposed limit of 3, whereas the full sample's median is 2.4. The full sample also includes large values in the hundreds, although we truncate the figure at a ratio of 10. It is necessary and important to remove problematic files that may compromise the final product once these files are combined.

imaging the continuum is much faster and provides a good approximation of the quality of the files. We excluded files with an RMS value higher than 10^{-5} Jy, which introduced inconsistent noise patterns throughout the cube. Secondly, we estimated the theoretical sensitivity ($\sigma_{\rm theoretical}$) associated with each observation by using the `apparentsens` function of the `CASA` image `toolkit`. This function calculates the point source sensitivity in imaged cubes, accounting for image weights and visibility weights, also used in the ALMA Interferometric Pipeline (see Hunter et al. 2023). We insist that the measured RMS of each continuum imaged file compared with the theoretical sensitivity (ALMA pipeline) should not be greater than a factor of 3. We inspected a few files that were removed by the RMS criteria and found that most of them had strong interferometric patterns or errors in calibration. Figure 2.2 shows the distribution ratio of the measured RMS and the theoretical RMS. The pruned sample extends by up to a factor of 3 with a median value of 1.8, while the full sample has a median value of 2.4. For comparison, the full sample reaches RMS ratios of hundreds. These high-RMS observations could introduce significant noise artefacts in the final data product.

Despite our efforts with the sample selection, we still found files with strong interferometric patterns, such as image artefacts, baseline stripes, blurring, or distortions in the image. In some cases, these patterns cannot be easily identified from the RMS of the continuum. Several errors can arise when one performs self-calibration using the pipeline, such as poor phase and amplitude calibration. These artefacts are usually caused by baseline errors due to uncertainties in the antenna position when measuring the differences in signal phase and amplitude between pairs of antennas. Artefacts can also appear during imaging, resulting in sidelobes, noise bias, and spatial distortion. The errors in interferometric data are primarily identified in the uv plane during calibration and imaging processes. Considering the large amount of data that the ALMACAL−22 survey contains, it is not possible to visualise all the observations in the uv plane to flag the possible misbehaved antennas. Thanks to our large amounts of data, we can afford to discard corrupted data by checking the continuum map of each observation to determine if it presents strong patterns. We visually identified the presence of stripes, blurring, artificial elongation or shifting of sources, and symmetric or asymmetric patterns, then discarded observations exhibiting these issues to ensure data quality. After removing $\sim 15\%$ of the files, we combined the data with the same frequency coverage, selecting observations that contribute to a total integration time of at least 10 minutes.

We selected high-quality data to construct the survey. We started from a total of 34909 measurement sets (ms), that is, raw visibility files, of which 25594 add up to at least 10 minutes of integration time. Among these, 15270 files have acceptable RMS levels. After removing visual artefacts and reapplying the integration time criterion, we used 6494 ms files to build the pruned sample. Figure 2.3 displays the distribution of the number of files in each band for both the full ALMACAL−22 and pruned samples in the left panel. Bands 3, 6, and 7 dominate the full sample, while Bands 3, 4, and 6 dominate the pruned sample. There are no Band 10 data available in the pruned sample. The middle panel shows the distribution of data cubes in both samples, where Bands 3, 4, 6, and 7 dominate (details on how cubes are constructed are discussed in Sect. 2.2.3). The right panel shows the distribution of calibrators, with 1047 in the full sample and 401 in the pruned sample. For both samples, Band 3 has the most calibration fields, followed by 6 and 4.

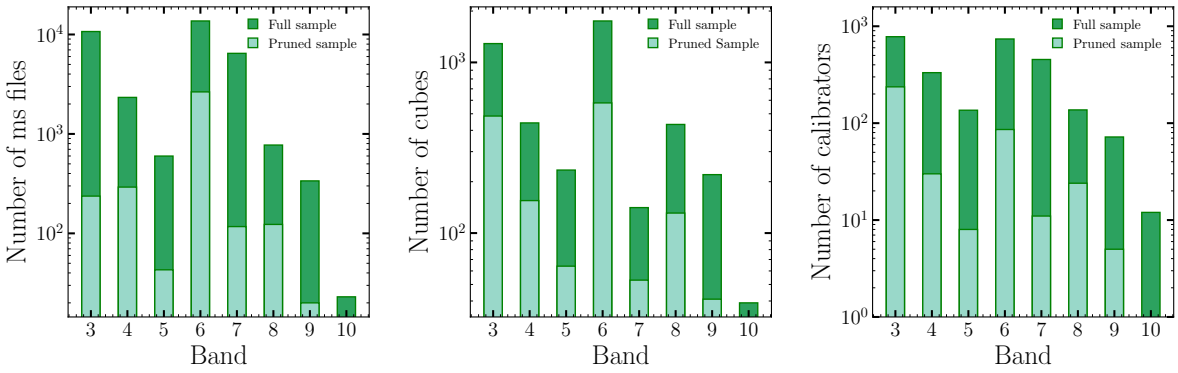

Figure 2.3: Distribution of the number of ms files (left), cubes (middle), and calibrators (right) for the full ALMACAL−22 and pruned samples across ALMA bands. The full sample contains 34909 ms files, 4547 cubes, and 1047 calibrators, while the pruned sample consists of 6494 ms files, 1508 cubes, and 401 calibrators. Band 6 comprises most files and cubes for the full ALMACAL−22 and pruned samples. Band 3 has the largest variety of calibration fields for both samples, followed by Bands 6 and 4. It should be noted that no data are available in Band 10 for the pruned sample, as the quality assurance discarded all the data that contributed sufficient integration time (10 minutes).

Table 2.1: Number of data cubes and calibrators in the pruned sample

Band	Frequency [GHz]	Nº data cubes	Nº calibrators
3	$84 - 116$	485	237
4	$125 - 163$	155	75
5	$163 - 211$	64	31
6	$211 - 275$	579	201
7	$275 - 373$	53	19
8	$385 - 500$	131	57
9	$602 - 720$	41	15

2.2.3 Building cubes

Data cubes have three dimensions, two of which contain spatial information and one spectral information given by a frequency range. To produce each data cube, we followed a sequence of steps to combine multiple observations. First, we concatenated all the uv observations into a single file to obtain a data cube. We selected the ms files with sequential frequencies to be included in the cube. Then we applied the `CASA` task `statwt` to the concatenated file to recalibrate the uv weighting of different observations based on the variance of data. We estimated the beam size using the `CASA` task `getsynthesizedbeam` from the `analysisUtils` package. We sampled each beam with 3 pixels (using a pixel scale = synthesised beam / 3) and picked the image size to be ~ 1.8 times that of the primary beam, selecting the number of pixels recommended by `CASA` to maximise the imaging efficiency. We defined the channel width for all cubes to be 31.2 MHz, a value selected to strike a balance between spectral resolution, S/N, and manageable data volume. Finally, we ran `tclean` in `CASA` for imaging using linear interpolation, natural weighting to optimise sensitivity, and 0.5 arcsec tapering. We parallelised `CASA` using eight cores and eight threads, which reduces by a factor of 4 the execution time of `tclean` for some cubes. The duration of the imaging process depends on how many observations we combine and the cubes' frequency coverage, ranging from a few minutes to several hours per cube.

The pruned sample is composed of 1508 cubes from Band 3 to Band 9. Table 2.1 details the number of cubes constructed in each band for the pruned sample and the number of

calibrators covered. The distribution of the number of cubes is shown in the middle panel of Fig. 2.3. The full sample exhibits three times more cubes than the pruned sample. The number of cubes in the full sample refers to the number of data cubes that can be built applying the integration time criterion only, that is, without taking into account the RMS selection and visual inspection that was used to build the pruned sample (see Sect. 2.2.2).

2.3 ALMACAL−22 properties

This section explores the inherent properties of the latest data from the ALMACAL−22 survey in both the full and pruned samples. We characterise the fundamental properties of this dataset to provide a clear understanding of the survey's scope and capabilities. These include the spatial distribution (Sect. 2.3.1), spatial resolution (Sect. 2.3.2), integration time (Sect. 2.3.3), and redshift of the calibrator sources (Sect. 3.3.4).

2.3.1 Spatial distribution

The ALMACAL−22 survey comprises calibrators randomly distributed across the southern sky, resulting in a widely dispersed area covering more than 1100 arcmin^2. The data collection strategy effectively captures diverse regions with a random distribution. This distribution is advantageous for serendipitous line and continuum detections, providing robustness against the effects of cosmic variance, a challenge often encountered in deep field surveys. The top panel of Fig. 2.1 shows the spatial distribution for the full ALMACAL−22 sample, and the bottom panel shows the pruned sample. In the pruned sample, two areas have relatively sparse data. These regions are near the projection of the galactic plane, likely resulting in fewer observations due to potential interference and the relative paucity of cosmological deep fields in those regions, or challenges associated with observing in those directions. The primary factor influencing the selection of the pruned sample is the minimum integration time, which determines how long observations need to be made to qualify for inclusion in the dataset.

2.3.2 Spatial resolution

The spatial resolution of the ALMACAL−22 dataset varies depending on the observing frequency and array configuration chosen by PIs based on their scientific goals. Higher-

frequency bands generally offer a finer spatial resolution, enabling the detection of more intricate structures and providing valuable insights into the morphology of the observed sources. Conversely, lower-frequency bands offer advantages regarding sensitivity and wider coverage but with a potential sacrifice of some spatial resolution.

Figure 2.4 shows how each band's spatial resolution is distributed. In the ALMACAL−22 full sample, the median value in Band 3 is 1.14 arcsec. This consistently decreases for higher-frequency bands, reaching a median spatial resolution of 0.13 arcsec in Band 10. In the case of the pruned sample, Band 4 shows the maximum value in the median spatial resolution of 0.87 arcsec. This value is lower in Band 3, then decreases to a mean value of 0.23 arcsec in Band 9.

2.3.3 Integration time

The integration time reflects the total on-source observation time obtained for the ALMA calibration data. The dataset's construction requires combining two or more observations if their frequencies overlap by at least 1 GHz. We added up the integration time in a 1 GHz moving window, selecting all the files spanning the same frequency. We imposed a lower limit of 10 minutes for the pruned sample across the whole frequency coverage. Variations in integration time in the cubes arise from the concatenation of short calibration pointings, each lasting several minutes with overlapping spectral coverage. From the varying depths reached, different sensitivity values can be achieved within a cube, but every frequency meets the minimum time criterion. The cube's frequency range within a cube is chosen to contain sequential frequencies. A calibration field observed several times in the same frequency band may have different frequency ranges observed, creating more than one cube for the same field and band, covering a distinctive frequency extent.

Figure 2.5 shows the distribution of total integration times for the full ALMACAL−22 dataset and the pruned sample. The median and maximum values of the total integration times for the full ALMACAL−22 dataset and the pruned sample are shown in Tables 2.2 and 2.3. The full sample comprises cubes that achieve more than 10 hours of integration time, while the pruned sample reaches an integration time of up to 7 hours. The median value in the full sample is 0.78 hours; in the pruned sample, it is 0.48 hours. The pruning process has only a slight impact on the median value, suggesting a small effect on the sensitivity. The mean sensitivity value reached in the pruned sample is ~ 0.78 mJy/beam

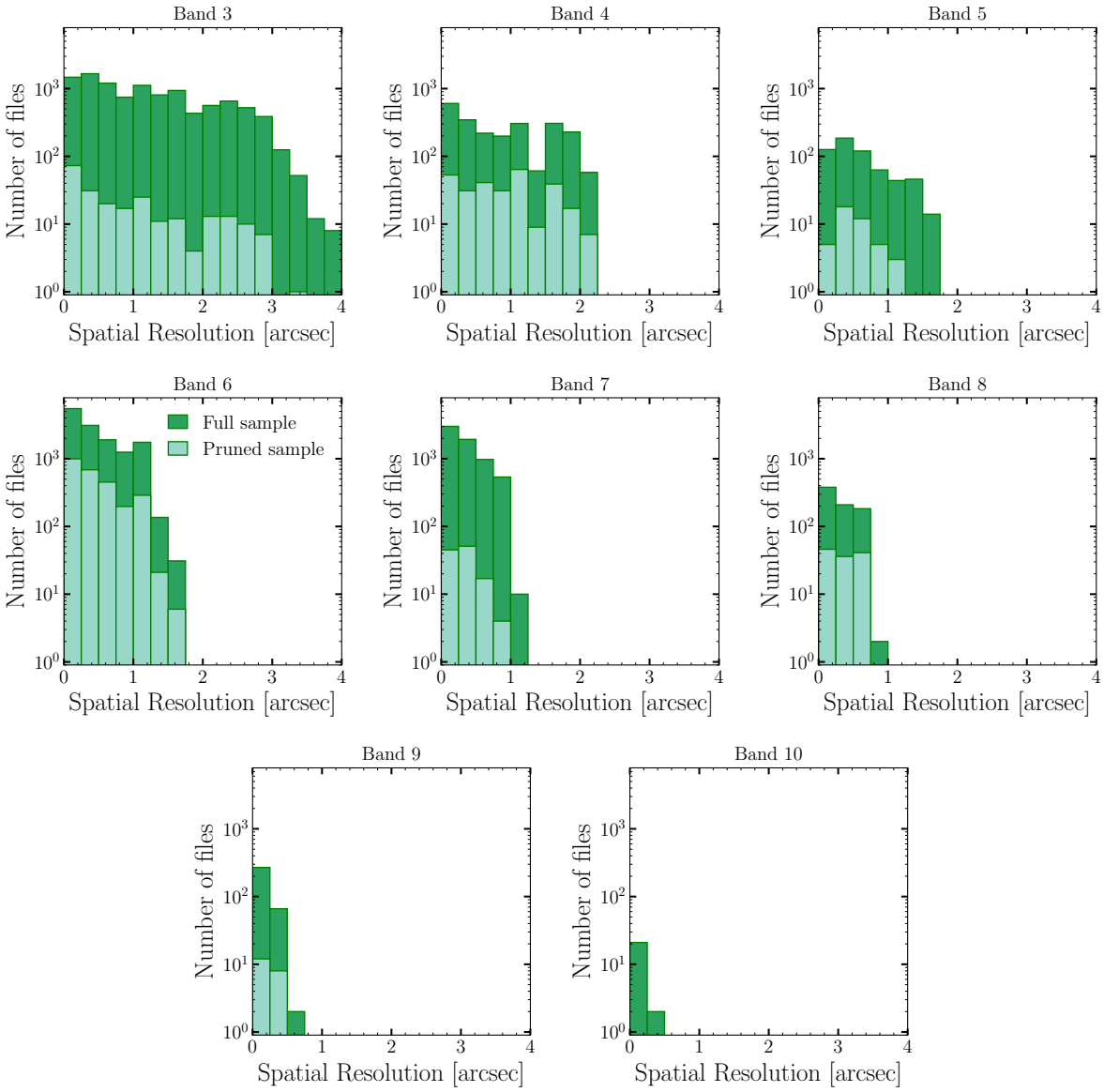

Figure 2.4: Distribution of the spatial resolution for the full ALMACAL−22 and pruned samples across ALMA bands. The mean values of the full ALMACAL−22 sample decrease consistently towards higher bands, from a median value of 1.15 arcsec in Band 3 to 0.10 arcsec in Band 10. In the pruned sample, Band 4 presents a median value of 0.87 arcsec, followed by Band 3 with 0.59 arcsec, and decreases in higher-frequency bands, reaching 0.23 arcsec in Band 9.

(for further discussion see Sect. 2.4.1).

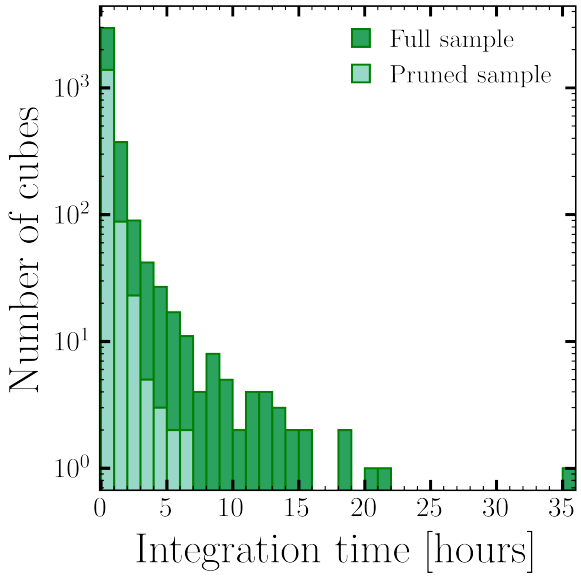

Figure 2.5: Distribution of the maximum integration time reached in each data cube for the full ALMACAL−22 sample and the pruned sample. The maximum and median values reached in each cube for each band are listed in Tables 2.2 and 2.3. The full sample has a mean value of 0.78 hours, while that of the pruned sample is 0.48 hours. The pruned sample achieves up to 7 hours of integration time, maintaining high sensitivity levels.

2.3.4 Redshift

The ALMACAL Redshift Catalogue relies primarily on a redshift database introduced by Bonato et al. (2018) for a substantial portion of the calibrator sample. This compilation was further enriched after cross-matching radio sources with optical sources from NED, SIMBAD (Wenger et al. 2000), and an optical catalogue of bright sources at gigahertz frequencies (Mahony et al. 2011). Team members individually checked each source, noting the redshift and provenance. This catalogue is also supplemented by VLT/X-Shooter observations of calibrator sources (ID 111.253L.001, PI: S. Weng and ID 0101.A-0528, PI: E. Mahony). Out of the initial 1047 sources, 675 have robust redshifts and spectra. For the pruned sample of this work, 390/635 (61%) calibrators have confirmed spectroscopic redshifts. Further details will be presented in a forthcoming paper (Weng et al., in prep.).

Table 2.2: Main properties of the full ALMACAL−22 sample.

Band	Median $t_{\rm int}$ [min]	Max $t_{\rm int}$ [h]	Median Res. [arcsec]	Area [arcmin^2]
3	19.7	21.2	1.15	1301.7
4	19.4	5.9	0.75	209.2
5	28.1	2.3	0.49	54.7
6	23.0	35.3	0.38	231.2
7	22.2	18.3	0.28	71.2
8	21.8	5.8	0.25	15.0
9	23.4	6.4	0.15	3.9
10	23.8	0.5	0.10	0.34

Notes. Columns: (1) ALMA band, (2) median integration time in the cubes of the full sample, (3) maximum integration time reached in the cubes of the full sample, (4) median spatial resolution of the full sample, (5) total covered area by all the cubes in the full sample.

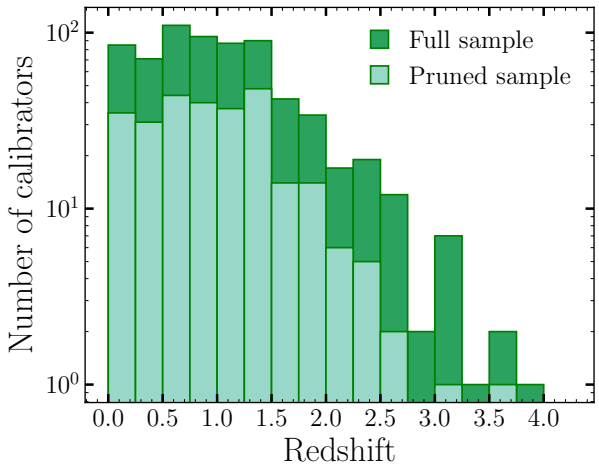

Figure 2.6: Redshift distribution of calibrators in the full ALMACAL−22 sample and the pruned sample (Weng et al., in prep.). The median redshift is $\langle z \rangle = 0.749$ in the full sample, and $\langle z \rangle = 0.797$ in the pruned sample. The highest redshift calibrator is at $z \sim 3.8$.

Table 2.3: Main properties of the pruned ALMACAL−22 sample.

Band	Median $t_{\rm int}$ [min]	Max $t_{\rm int}$ [h]	Median Res. [arcsec]	Area [arcmin^2]
3	19.0	6.1	0.59	693.1
4	14.8	2.7	0.87	250.8
5	13.5	1.0	0.48	63.1
6	19.4	6.9	0.39	103.6
7	14.6	2.0	0.29	16.2
8	19.2	5.2	0.34	26.1
9	20.2	0.78	0.23	3.2

Notes. Columns: (1) ALMA band, (2) median integration time in the cubes of the pruned sample, (3) maximum integration time reached in the cubes of the pruned sample, (4) median spatial resolution of the pruned sample, (5) total covered area by all the cubes in the pruned sample.

The redshift distribution for calibrator sources in the full ALMACAL−22 dataset and the pruned sample is shown in Fig. 2.6. The full sample of calibrators has a median redshift of $z = 0.749$, with a maximum of $z = 3.788$. In the pruned sample, the median redshift value is $z = 0.797$, and the maximum is $z = 3.591$. The sources excluded during pruning were likely more concentrated at lower redshifts. However, the redshift distribution in Fig. 2.6 shows that the higher redshift end is also less populated after pruning. Therefore, we do not expect any significant bias related to redshift in the pruned sample.

2.4 Discussion

This section delves into a comparative analysis of ALMACAL−22 and other surveys, emphasising key properties such as sensitivity and survey area. We provide an overview of the scientific goals the ALMACAL survey has achieved so far. We evaluate the unique characteristics of this extensive dataset compared to existing surveys, shedding light on its strengths and potential scientific contributions.

2.4.1 Survey area and sensitivity

We compared the sensitivity and total area covered by the ALMACAL−22 survey to those of previous large programs. Due to the primary beam response, the sensitivity in each cube decreases as we move away from the phase centre. We considered an area 1.8 times the primary beam size for each field, which we calculated using the `primaryBeamArcsec` function from the `analysisUtils` package in `CASA`. To estimate the sensitivity reached in each data cube, we used the `sensitivity` task from `analysisUtils`, which takes the integration time and central frequency as inputs.

ALMACAL−22's coverage is distinguished by its larger footprint than any previous survey, covering over one thousand square arcminutes. For comparison, the ALMA Spectroscopic Survey in the Hubble Ultra Deep Field (HUDF), ASPECS, (Decarli et al. 2016; Walter et al. 2016; Aravena et al. 2019; Boogaard et al. 2020), covers a 4.6 arcmin^2 area. The COLDz survey (Pavesi et al. 2018; Riechers et al. 2019, 2020b) covers a $\sim$ 60 arcmin^2 area. The 6 $\times$ 15m IRAM Plateau de Bure High−z Blue-Sequence Survey 2 (PHIBSS2) covers a total area of $\sim$ 130 arcmin^2 (Tacconi et al. 2018; Lenkić et al. 2020).

Regarding sensitivity, ALMACAL−22 data cubes achieve remarkable sensitivity levels in ALMA bands 3, 4, 5, 8, and 9, comparable to ASPECS and COLDz. Bands 6 and 9 offer a wider range of sensitivities, mainly due to the variety of total integration times obtained in each cube after combining data with overlapping frequencies. Additionally, Audibert et al. (2022) demonstrated that ALMA calibrator data can be used to estimate the molecular gas content of galaxies, reaching sensitivities around 0.482 mJy, comparable to surveys such as COLDz and GOODS-ALMA.

Figure 2.7 compares the total survey area in the pruned ALMACAL−22 sample reached by each band as a function of the sensitivity. For comparison, we also plotted the sky coverage and sensitivity values reached by other surveys (Lenkić et al. 2023; Riechers et al. 2019; Decarli et al. 2019; Gómez-Guijarro et al. 2022; Stach et al. 2019). Our pruned sample adds up to a total area of 1154 arcmin^2, shown as the light green dashed horizontal line, surpassing the total area covered by previous surveys. As the dark dashed horizontal line shows, the full sample area over all bands reaches 1887 arcmin^2. The total survey area for both the full and the pruned sample accounts for each calibration field once, prioritising the largest area. Overall, ALMACAL−22's combination of extensive area coverage and average sensitivity will complement findings from other large surveys, addressing the effect

of cosmic variance.

Despite these strengths, there are challenges to scientific interpretation. For example, since ALMACAL consists of pointed observations of specific sources, it cannot be considered a truly blind survey, raising concerns about potential biases in cosmic overdensities. However, Bonato et al. (2018) classified most calibrators as blazars ($\sim 90\%$). This property mitigates clustering significance since the jets' brightness is due to their orientation effects toward the observer rather than their mass. For instance, Sushch & Böttcher (2015) found that clustering effects of a significant number of galaxies (more than five) situated near the line of sight of the blazar beam are absent in the local universe, but they may be possible at higher redshifts ($z > 2$). Moreover, the lack of correlation in redshifts between continuum- or line-detected galaxies and the calibrators further undermines the impact of having the calibrator source at the centre of each pointing (see Hamanowicz et al. 2022). Additionally, Chen et al. (2023) explored an overdense region in ALMACAL similar to extreme proto-cluster cores and found the most likely explanation to be alignment effects. Furthermore, the smaller primary beams in higher ALMA bands probe smaller volumes at lower redshifts.

2.4.2 Science projects

The ALMACAL survey covers many scientific studies, focusing on four main areas: (1) molecular gas evolution, (2) properties of DSFGs, (3) extragalactic absorption lines, and (4) AGN physics. We briefly summarise what has been done in these areas and the impact that ALMACAL will have through further analysis of this dataset.

(1) The evolution of molecular gas has been investigated with ALMACAL through the redshifted CO emission line. ALMA observations cover the frequency range of CO lines at different redshifts. Klitsch et al. (2019b) searched for the CO line in absorption in gas-rich galaxies selected via quasar absorption lines. They found multiple CO transitions, revealing that galaxies were associated with optically identified AGN activity. They reported different factors when using the CO spectral energy distributions (CO SLEDs) as a proxy to estimate the amount of molecular gas compared to the widely used galactic values. These findings indicate the galactic values might overestimate the molecular gas masses for some absorption-selected galaxies. This difference highlights the need to construct CO SLEDs in different systems rather than assuming the values measured for typical SFGs. More

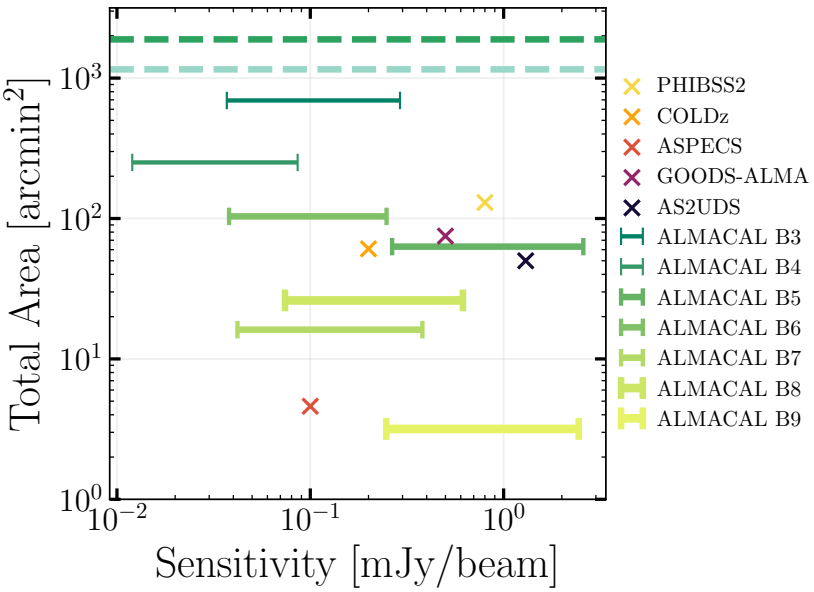

Figure 2.7: Total ALMACAL survey area vs the median sensitivity range reached in each band for the pruned sample. For comparison, we plot the area and sensitivity values from the previous surveys, PHIBSS2 (Lenkić et al. 2020), COLDz (Riechers et al. 2019), ASPECS (Decarli et al. 2016) in ALMA Bands 3 and 6, GOODS-ALMA (Gómez-Guijarro et al. 2022) in ALMA Band 6, and AS2UDS (Stach et al. 2019) in ALMA Band 7. The dashed light and dark lines show the total area accumulated by summing up all the bands for the pruned sample and the full sample. The total survey area counts every calibration field once, considering the largest area.

recently, Hamanowicz et al. (2022) developed an ALMACAL-CO pilot program to detect CO emission lines blindly over 38 calibrator fields, selected to have the longest integration times ($>$ 40 minutes). This pilot program aimed to probe the feasibility of using ALMA calibrator fields to look blindly for CO emitters. Eleven emission lines were detected, providing a consistent estimation of the evolution of molecular gas compared with previous surveys. ALMACAL's untargeted approach offers the advantage of being less sensitive to cosmic variance than previous deep surveys.

(2) Dusty star-forming galaxies represent more normal SFGs than conventional sub-mm galaxies; such faint systems are usually buried in the confusing noise of sub-mm wavelengths. Early on, Oteo et al. (2016a, 2017) exploited the high sensitivity levels achieved by ALMACAL by combining data from multiple visits to 69 ALMA calibrator fields in Bands 6 and 7. They found eight DSFGs and derived the number counts. They discovered systems so faint that even the deepest Herschel surveys would not have detected them. Klitsch et al. (2020) reported the first number counts in ALMA Band 8 over 81 calibrator fields, finding 21 DSFGs. Recently, Chen et al. (2022) extended the number counts estimation using 1001 calibration fields in the ALMA Bands 3, 4, 5, 6, and 7, covering a wavelength range from 3mm to 870 μm. They report the detection of 186 DSFGs with flux densities comparable to existing large ALMA surveys, but less prone to cosmic variance. Establishing the space density and contribution of DSFGs to the cosmic far-infrared can be a powerful way to validate galaxy formation and evolution models.

(3) The ALMACAL data provide the opportunity to study absorption lines due to galaxies along the line of sight of the calibrators. Klitsch et al. (2019a) reported several galactic absorption lines using 749 calibrators, but no intervening extragalactic molecular absorber was detected. They also used the cosmological hydrodynamical simulation IllustrisTNG (Naiman et al. 2018; Pillepich et al. 2018; Springel et al. 2018; Marinacci et al. 2018) to obtain new upper bounds on the molecular gas mass density. Their results are consistent with an increasing depletion of molecular gas in the present Universe compared to redshift $z \sim 2$. In a subsequent study, Klitsch et al. (2023) presented the first constraints on the molecular gas coverage fraction in the circumgalactic medium of low-redshift galaxies using estimates of CO column densities along the line of sight of quasars with intervening galaxies.

(4) As ALMACAL observations are repeated over the years, they allow for multi-year follow-up of the AGN variability. Bonato et al. (2018) examined 754 calibrator data using

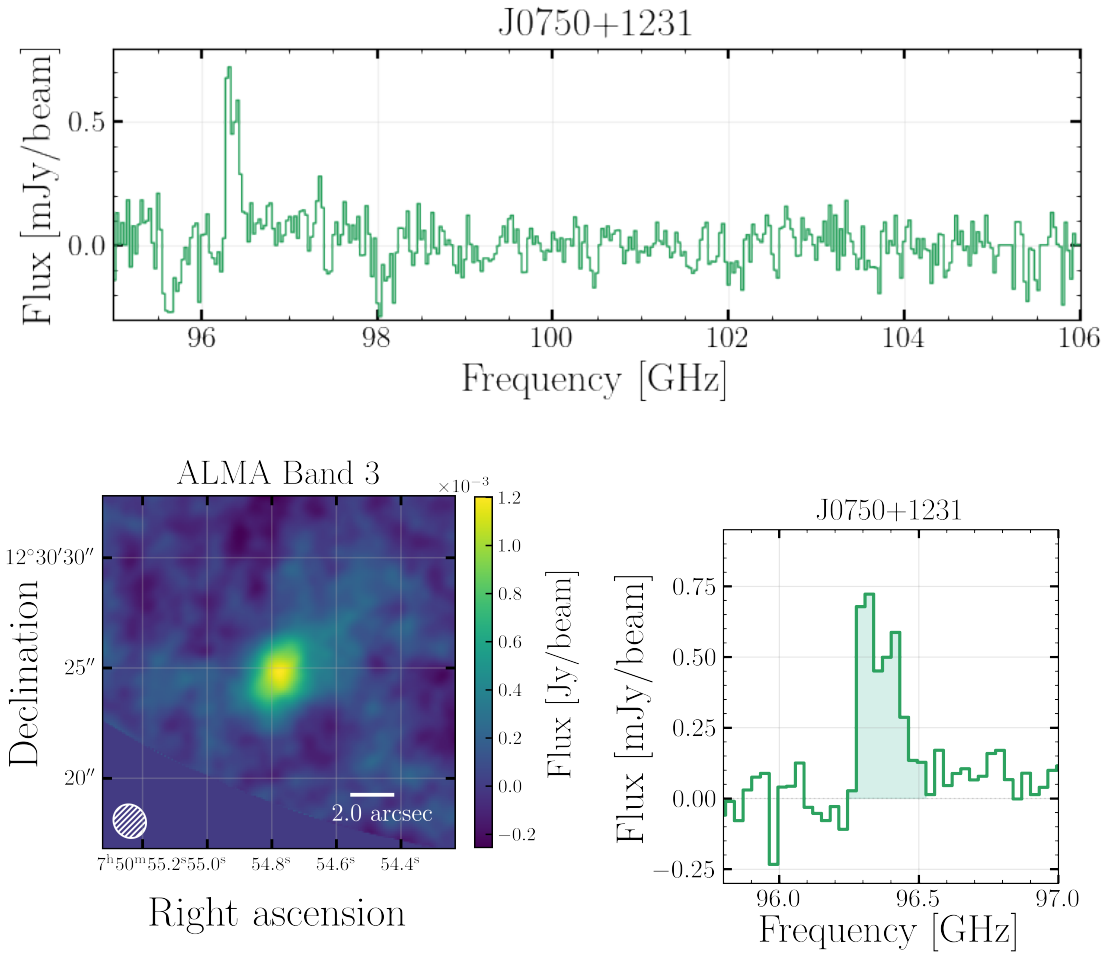

Figure 2.8: Example of an ALMACAL−22 data cube from the updated sample with a prominent emission line found in the calibrator field. The top panel shows the full spectral coverage of the region of the data cube where the emission line was found. The lower left panel displays a region of the continuum map of the calibrator field in Band 3, centred on the position of the emission line. The lower right panel shows the emission line, where the shaded region represents the line width reaching a S/N ~ 9.

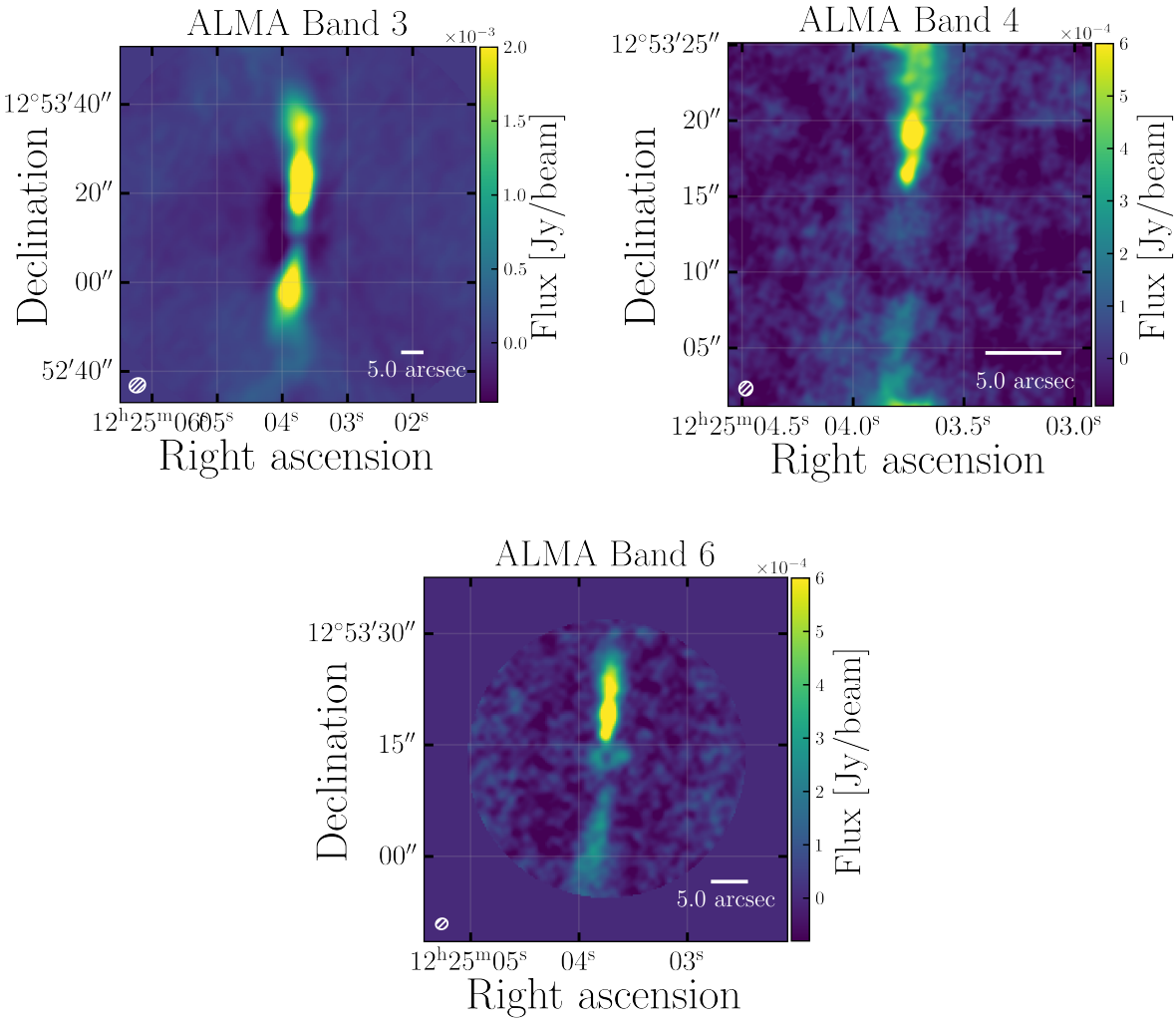

Figure 2.9: Example of three collapsed data cubes for the calibrator field J1225+1253, where jet emission emanates from the central quasar. ALMA Bands 3, 4, and 6 are shown in the top left, top right, and bottom panels. The central quasar in the centre has previously been subtracted in the uv plane.

Bands 3 and 6 in the time domain space, identifying most of them as blazars based on their flat spectrum and low-frequency spectral index. They constructed the light curve of the blazars and found that the median variability index increases steadily with increasing source-to-time lag from 100 to 800 d. Husemann et al. (2019) then studied the morphology and kinematics of the gas surrounding the calibrators. They detected a CO$(1-0)$ emission arc structure around the AGN of the quasar 3C 273. This arc morphology of the molecular gas is completely different from that of the ionised gas. This raises the question of whether the molecular gas is bound to a stellar overdensity formed from a recent galaxy interaction or is currently forming in situ due to a density wave and increased ambient pressure caused by the expanding outflow, as predicted in simulations for luminous AGNs (Mukherjee et al. 2018). Komugi et al. (2022) detected extended mm emission associated with the host galaxy of a prototypical radio-loud quasar to investigate the QSO-host interstellar medium (ISM) interaction.

The new ALMACAL−22 dataset, comprising 1047 calibrator fields, promises significant advancements in our current understanding of these science cases. First, the increased calibrator fields will alleviate the field-to-field variance when determining molecular gas evolution. The volume probed by each CO transition will increase by a factor of ~ 10 compared to previous surveys, reducing the uncertainties and expanding at different redshift bins. Figure 2.8 shows an example of a blind search across the calibrator fields, where a prominent emission line was detected. Second, by adding 46 calibrator fields to the last search for DSFGs (Chen et al. 2022) and new observations of the same fields, it will be possible to achieve deeper sensitivity levels, which could improve the estimates of the number counts by detecting the faintest systems in the Universe. Thirdly, there are over 300 calibrator fields where the CO absorption lines have yet to be searched for. This updated version of the ALMACAL survey provides a significantly larger dataset than the data analysed in Klitsch et al. (2019b). Finally, ALMACAL−22 will allow us to study the variability of specific AGN types such as blazars, including both BL Lacs in the nearby Universe and FSRQs at larger distances. Calibrators are usually chosen to be point-like, but some have shown extended large-scale jet structures. Figure 2.9 illustrates the same calibration field showing extended structures in ALMA Bands 3, 4, and 6. Studying the effect of these powerful jets on star formation could help us to understand the high-energy processes better.

2.5 Conclusion

ALMACAL−22 is a large survey of ALMA calibrator observations, collecting over 30 TB of data and covering 1047 calibrator fields across the southern sky. We have presented all the calibrator data from science projects taken until May 2022, accumulating over 1000 square arcmin and more than 2000 hours of observing time. Here, we provide the characteristics of the survey.

We have presented the selection of a pruned sample, a subset of the highest-quality data. We outline the data processing details from Band 3 to Band 9 to obtain data cubes as final products. The pruned sample contains 401 calibrator fields and 1508 data cubes. We provide an overall review of the main properties of both the full and pruned samples, including the spatial distribution, spatial resolution, integration time, and redshift of the calibrator sources.

In a forthcoming paper, we will revisit the pruned sample to investigate serendipitous detections of the CO emission lines in the calibrator fields. In this way, this survey will provide clues to the evolution of molecular gas in the Universe through an untargeted approach. As ALMACAL−22 is one of the largest surveys to date, it will allow us to make statistical estimates that are less sensitive to potential cosmic variance effects.

Overall, ALMACAL is an ever-growing project, as every scientific project requires calibrator data, and the size of the dataset will continue to grow over the years. A diverse range of catalogues will be established, including the redshift catalogue, extended jets catalogue, and confirmed molecular line catalogue, showcasing the valuable insights ALMACAL−22 will contribute to the scientific community.

Chapter 3

ALMACAL. Evolution of the CO luminosity function and the molecular gas mass density out to $z \sim 6$

The content of this chapter was presented in the published article Bollo et al. (2025), Astronomy & Astrophysics, 695, A163.

3.1 Introduction

The baryon cycle is a key driver of galaxy evolution, involving the inflow and outflow of the gas that regulates star formation over cosmic time (Péroux & Howk 2020; Walter et al. 2020). Galaxies form and grow within the framework of the baryon cycle, whereby gas is accreted from the inter- and circum-galactic media, primarily in the form of neutral atomic hydrogen (H I), which eventually cools and condenses into molecular hydrogen (H_2), the raw material for star formation (Tumlinson et al. 2017; Tacconi et al. 2020).

Understanding the baryon cycle is crucial for interpreting two fundamental cosmic quantities: the cosmic star formation rate density (SFRD) and the stellar mass function. The cosmic SFRD, $\psi_\star(z)$, tracks the rate at which stars are formed across the Universe as a function of redshift. It peaks around two billion years after the Big Bang ($z \sim 2$), then declines towards the present day (Madau & Dickinson 2014). The cosmic stellar

mass function, $\rho_\star(z)$, represents the cumulative mass of stars within galaxies, providing insights into the efficiency of star formation and how galaxies build up their stellar content over time. Together, these two observables are key to understanding the history of star formation in the Universe.

Cold gas – atomic and molecular – plays a critical role in the evolution of these cosmic quantities (Carilli & Walter 2013; Saintonge & Catinella 2022). Atomic hydrogen (H I) is the largest, more diffuse gas reservoir in galaxies, which must cool and condense to form molecular clouds, from which stars ultimately form. At low redshift, H I is typically observed through its 21−cm emission (e.g. Zwaan et al. 2005; Jones et al. 2018), while at high redshift it is detected through absorption lines (e.g. Péroux et al. 2003). These methods provide a direct way to map the distribution of atomic gas and its role in galaxy evolution. However, molecular gas is a more direct precursor to star formation, which takes place in the dense core of cold gas clouds (Kennicutt & Evans 2012; Tacconi et al. 2020). Because H_2 lacks a permanent dipole moment, we generally rely on carbon monoxide (CO) as its tracer (Carilli & Walter 2013; Bolatto et al. 2013). Observatories such as the Atacama Large Millimeter Array (ALMA), the Northern Extended Millimeter Array (NOEMA) operated by the Institute for Radio Astronomy in the Millimetre Range (IRAM), the Very Large Array, the Atacama Pathfinder EXperiment (APEX), and the Owens Valley Radio Observatory (OVRO), among others, have enabled the detection of CO across a wide range of redshifts, allowing us to estimate the molecular gas content in galaxies (Hodge & da Cunha 2020).

Another key observable is the CO luminosity function (CO LF), which quantifies the distribution of CO luminosities in galaxies. It provides a valuable diagnostic for studying the molecular gas properties and their role in galaxy evolution. By constructing the CO LF across different cosmic epochs, previous works have traced the evolution of the molecular gas mass density across different redshifts, $\rho_{H_2}(z)$. For instance, the ALMA Spectroscopic Survey in the Hubble Ultra Deep Field (ASPECS) survey (Decarli et al. 2016) detected CO in galaxies across $z = 1$–3 using ALMA bands 3 and 6, covering areas of 4.6 and 2.9 arcmin^2, respectively (Decarli et al. 2019, 2020). Boogaard et al. (2023) combined the results of ASPECS with NOEMA observations to constrain the bright end of the CO LF up to $z \sim 6$, based on observations of the Hubble Deep Field North (HDFN), spanning 8.5 arcmin^2. Lenkić et al. (2020) explored CO emission lines in 110 main-sequence galaxies using fields from the original PHIBBSS survey (Guilloteau et al. 1992), covering a total area

of ~ 130 arcmin^2. The COLDz survey detected CO in galaxies at $z = 2$–3 and at $z = 5$–7 using more than 320 hr of time with the Very Large Array (Pavesi et al. 2018; Riechers et al. 2019, 2020b), spanning ~ 60 arcmin^2. Audibert et al. (2022) exploited the ALMA calibrator catalogue to study the CO LF up to redshift $z \sim 2.5$ and to examine the role of radio emission in galaxy evolution. In addition to CO-based measurements, molecular gas masses have also been inferred from dust emission using scaling relations based on stellar mass or star formation rates (Berta et al. 2013; Scoville et al. 2017; Magnelli et al. 2020) and from the [C II] emission line at higher redshifts (e.g. Aravena et al. 2024; Dessauges-Zavadsky et al. 2020; Zanella et al. 2018). Despite these advancements, the molecular gas mass density at several redshift bins in these surveys exhibit significant measurement uncertainties, sometimes exceeding 1 dex, highlighting ongoing challenges in accurately measuring molecular gas content.

Accurate measurements of the cosmic molecular gas mass density face two primary obstacles: cosmic variance and Poisson uncertainties. Cosmic variance reflects variations in galaxy properties across different regions of the Universe, while Poisson errors arise from the low number of detected galaxies in each luminosity and redshift bin. For instance, bright sources – which are sparsely distributed – can significantly increase shot noise. On the other hand, faint sources contribute to clustering noise, where their uneven distribution over large cosmic structures impacts measurements of the matter distribution in the Universe.

The choice of the observed region can introduce significant uncertainties when measuring cosmic properties, such as the number density of galaxies and LF. For example, Driver & Robotham (2010) used the Sloan Digital Sky Survey (SDSS) with galaxies at redshift $z = 0.03$–0.1 to study how the cosmic variance changes with survey volume and field shape, providing a formula to quantify these variations. Similarly, Moster et al. (2011) used N-body simulations to calculate the excess variance caused by cosmic effects, showing that it significantly exceeds Poisson variance, which only accounts for random sampling. Both studies concluded that surveys that only probe limited volumes inevitably provide biased estimates of cosmic statistics. This makes it challenging to derive accurate and representative measurements of molecular gas over cosmic time.

Simulations play a crucial role in overcoming observational limitations by providing theoretical predictions of CO emission and molecular gas distribution in galaxies. Hydrodynamical simulations and semi-analytical models (SAMs) have successfully reproduced a wide range of galaxy properties such as UV-to-FIR LFs of galaxies, the number counts,

and the redshift distribution of sub-millimetre galaxies (SMGs; Narayanan et al. 2012; Somerville & Davé 2015).

Current cosmological simulations do not directly include a cold gas component due to the complexity of the underlying physics. We therefore rely on post-processing to derive molecular gas properties. There has been a growing interest in using simulations to model the CO LF and the molecular gas density at several redshifts (Lagos et al. 2012, 2015; Popping et al. 2016; Vallini et al. 2018; Lagos et al. 2018; Popping et al. 2019a,b; Maio et al. 2022; Béthermin et al. 2022; Bisigello et al. 2022; Garcia et al. 2023; Guo et al. 2023a; Lagos et al. 2024; Ragone-Figueroa et al. 2024). However, reproducing the bright end of the CO LF remains difficult, and molecular gas density predictions have not yet converged.

Recent studies have explored the impact of cosmic variance on observables like the CO LF using simulated data. For instance, Popping et al. (2019b) compared observational data from the ASPECS survey with predictions from cosmological simulations, including post-processed IllustrisTNG (Pillepich et al. 2018) and the Santa Cruz semi-analytical model (Somerville et al. 2001). They found that cosmic variance has a stronger impact at lower redshifts ($z < 1$) since the same area of sky at higher redshifts has a larger volume. Similarly, Keenan et al. (2020) used simulations to find that for a survey of ~ 50 arcmin^2 with CO luminosities of around 10^{10} K km s^{-1} pc^2, the uncertainties from Poisson noise and cosmic variance become comparable. Also, Gkogkou et al. (2022) found that using only Poisson variance underestimates the total uncertainty by up to 80%, particularly for lower luminosity sources and larger survey areas. They reported that at redshifts $z < 3$, cosmic variance can introduce up to 40% uncertainty in molecular gas density estimates in small surveys (e.g. 4.6 arcmin^2, similar to ASPECS), but this drops to below 10% for surveys covering more than 43.2 arcmin^2.

Observations have begun to acknowledge cosmic variance as a key factor in explaining discrepancies between molecular gas density estimates across different surveys (Lenkić et al. 2020; Boogaard et al. 2023). This marks a shift from previous studies, which often neglected to include cosmic variance in uncertainty estimates, resulting in optimistic error estimates.

This paper aims to alleviate the limitation introduced by cosmic variance when estimating the molecular gas mass density by exploiting ALMA calibrator data, as part of the ALMACAL−22 survey (Zwaan et al. 2022). We make use of the sample selected and described in a previous paper of this series (Bollo et al. 2024), consisting of the highest-quality and deepest data compiled from 2016 to 2022. This study builds upon the methodology

and results of the previous ALMACAL-CO survey (Hamanowicz et al. 2022). The proof-of-concept pilot of this earlier work successfully demonstrated the feasibility of using ALMA calibration data to characterise CO-selected galaxies and estimate their molecular gas content. We use the statistical approaches and classification techniques introduced in that study in this extended analysis. With the larger dataset now available, we can refine our results and place stronger constraints on the CO LF and the evolution of the molecular gas mass density across cosmic time, fully aligning with the expectations set by the previous work.

This chapter is organised as follows. In Section 3.2, we describe the ALMACAL−22 survey and summarise the calibration process, the selection of the sample, and the imaging. Section 3.3 details the steps taken to search for serendipitous detections of CO emission lines, including the source finding algorithm used (§ 3.3.1), the completeness correction (§ 3.3.2), the estimation of the fidelity (§ 3.3.3) and the redshift (§ 3.3.4), and the determination of the final catalogue (§ 3.3.5). Section 3.4 presents the volume estimates (§ 3.4.1), the CO LF (§ 3.4.2), the Schechter fits (§ 3.4.3), the molecular gas mass density (§ 3.4.4), the results of adopting the lowest CO transition possible (§ 3.4.5), and the uncertainties of our results (§ 3.4.6). Section 4.5 compares our findings for the CO LF (§ 3.5.1) and the molecular gas mass density (§ 3.5.2) with observations and simulations, discusses the effect of cosmic variance (§ 3.5.3), and provides a complete census of the baryon cycle (§ 3.5.4). In Section 4.6, we summarise our key conclusions. Throughout this chapter, we use $H_0 = 70$ $\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_\mathrm{M} = 0.3$, and $\Omega_\Lambda = 0.7$. The present-day cosmological critical density is $\rho_{0,\mathrm{crit}} \simeq 277.4\ h^2\ \mathrm{M_\odot\ kpc^{-3}}$.

3.2 ALMACAL−22 and sample selection

ALMACAL-22 exploits the archival ALMA calibrator sources and their fields observed since 2016 to produce scientific outcomes (Zwaan et al. 2022). ALMA calibrators are typically bright quasars, observed for a few minutes in every ALMA PI-led observing project. Each calibration scan has a setup that matches the PI requirements, so the full ALMACAL survey has a diverse range of configurations in terms of spatial and spectral resolution, sensitivity and integration time. Most of the calibrators in ALMACAL are bright sub-millimetre point sources classified as blazars (Bonato et al. 2018; Weng et al. 2025).

This work uses the latest data release, which comprises the ALMA calibrator data taken up until 2022 May, so-called ALMACAL−22 (Bollo et al. 2024). The full ALMACAL survey comprises more than 30 TB of calibrator data from ALMA Cycle 1 to Cycle 10, over more than 1000 fields. Bollo et al. (2024) defined a pruned sample to represent the highest quality data, where each entry exceeds ten minutes of accumulated integration time. We focus the analysis on this pruned sample, which includes only observations made with the 12-metre array and files with low root-mean-square noise values (RMS $<$ 0.01 mJy per channel). We summarise the main calibration steps, sample selection and data processing below.

A dedicated ALMACAL pipeline (Oteo et al. 2016a) automates the processing of calibrator data and produces images of delivered ALMA datasets. The pipeline calibrates using `scriptForPI.py`, extracts calibrator data, applies self-calibration to correct phase and amplitude variability, and removes the bright calibrator. A point source model is applied during self-calibration, resulting in calibrated visibilities and the creation of data cubes for the ALMACAL−22 dataset.

Due to the uneven nature of the ALMACAL data, a pruning procedure was applied to select the highest-quality data. The pruning steps are based on key properties like integration time, frequency coverage, spatial resolution and RMS noise. The on-source integration of observations covering the same frequencies was set to accumulate more than 10 minutes to reach a meaningful sensitivity level. Only observations with noise levels consistent with theoretical sensitivity were selected. From an initial 34909 measurement sets (ms), only the $\sim$ 20% with the highest quality remain in the pruned sample. The total number of data cubes and calibrators in the pruned sample is listed in Bollo et al. 2024, Table 1. In addition to the pruning, there is potentially a bias related to clustering effects around the calibrators, as these objects are often located in massive haloes and associated with large overdensities. The pruning fields and data in this work explicitly address this concern by removing CO lines coinciding with the redshifts of the calibrators. A detailed discussion is provided in Section §3.3.4.

Data cubes – consisting of two spatial dimensions and one spectral dimension – were created by combining multiple observations. The uv observations were concatenated into a single file to re-calibrate the uv weights using CASA's `statwt` (CASA Team et al. 2022). The choices made to optimise the imaging process can be summarised as follows. The beam size samples 3 pixels, and the image size is circular, with a ratio 1.8× that of the primary

beam. The channel width is 31.2 MHz, a value that balances spectral resolution, S/N and data volume. Imaging was performed with CASA's `tclean`, using linear interpolation, natural weighting, and 0.5″ uv tapering. We took special care in examining the shape of the synthesised beam when combining data from different arrays. By comparing the PSF shape to a 2D Gaussian fit, we found minimal deviations. As a result, the contributions to the residual map are negligible, ensuring that the flux measurements remain accurate and free from significant beam distortions. The pruned sample contains 1508 data cubes covering 401 different calibrator fields. Most of the cubes are in ALMA bands 3 and 6, with a maximum integration time of ~ 7 hours and a mean sensitivity of ~ 0.78 mJy per channel (see Bollo et al. 2024 for further details).

Table 3.1: Properties of the CO transitions in the ALMACAL−22 survey

Transition	$\nu_{\rm rest}$ [GHz]	Redshift	Volume log[cMpc]^3	$N_{\rm ind}$ [1]	CV[2] %
$CO(1-0)$	115.27	$0-0.37$	4.85	173	4.85
$CO(2-1)$	230.54	$0-1.70$	6.07	240	1.75
$CO(3-2)$	345.54	$0-3.15$	6.30	259	1.51
$CO(4-3)$	461.04	$0-4.40$	6.39	274	1.20
$CO(5-4)$	576.27	$0-5.80$	6.44	287	1.07
$CO(6-5)$	661.47	$0-7.10$	6.47	284	1.40
Total			7.06	299	

Notes. $N_{\rm ind}$ is the number of independent sightlines in which a given transition can be detected. CV is the percentage of cosmic variance estimated using Eq. 3.7 (see § 3.5.3 for more details).

3.3 Line search

Detecting emission lines in more than 10^3 cubes with significant noise variations is challenging. To address this, we used the Source Finding Algorithm SOFIA-2 (Westmeier et al. 2021), to identify potential candidates. We assessed these candidates' reliability based on

[1] $N_{\rm ind}$ is the number of independent sightlines in which a given transition can be detected.
[2] CV is the percentage of cosmic variance estimated using Eq. 3.7 (see § 3.5.3 for more details).

the S/N and line width detected, and we estimated the completeness factor of the entire line search process. In this section, we describe the identification of CO candidates (§ 3.3.1), the estimation of the completeness (§ 3.3.2), reliability (§ 3.3.3), the determination of redshifts (§ 3.3.4), and the final compilation of a catalogue of candidates (§ 3.3.5).

3.3.1 Source searching algorithm

SOFIA-2 is a fully automated pipeline designed to find 3D sources in the WALLABY survey conducted on the Australian SKA Pathfinder (ASKAP). Based on the open access source finder SOFIA 1 (Serra et al. 2015), SOFIA-2 has been re-written in C with multi-threading for speed, being considerably faster and more efficient than the previous version. It allows for searching on the spectral and spatial axes, offering two different algorithms: a simple threshold finder and the smooth and clip (S+C) finder. The threshold finder applies a flux threshold to the data, which is useful mainly for data adjusted for noise variations. The default S+C finder, described in Serra et al. 2012, iteratively smooths the data cube across various spatial and spectral scales to detect significant emission signals above a set threshold.

We input the processed ALMA data cubes after cleaning, calibrating, and based on multiple observations combined in uv space. We performed the search on data cubes that remained uncorrected for the shape of the primary beam response function. Before the primary beam correction, the data preserves its original sensitivity pattern to apply a uniform detection threshold to the emission line search. We flagged spectral channels for which the noise deviated from the median by more than 5σ. In SOFIA-2, we applied spatial and spectral smoothing based on the median absolute deviation (MAD) with a 3σ threshold. This helps to automatically flag corrupted data, like channels with radio-frequency interference or pixels with residual continuum emission. SOFIA-2 measures and adjusts for the local noise level using a running window of 25 pixels and 15 channels in the spatial and spectral domains.

We chose the smooth and clip (S+C) algorithm to identify sources, while smoothing to suppress noise and outliers. The algorithm iteratively rejects data points that deviate beyond the 3σ threshold, so extreme values do not affect the smoothing process. Then, the 'linker' combines detected pixels in the binary mask into coherent detections using a friends-of-friends algorithm. It links all pixels within a merging radius of three, treating

sources as three-dimensional collections of pixels. After the source identification, SOFIA-2 delivers the properties of the sources, such as their position, and size. Three co-authors (VB, CP, MZ) visually inspected the spectra and moment maps of the candidates to ensure that they were not confused with noise peaks. The final catalogue contains 87 emission lines (see § 3.3.5 for further details about the selection).

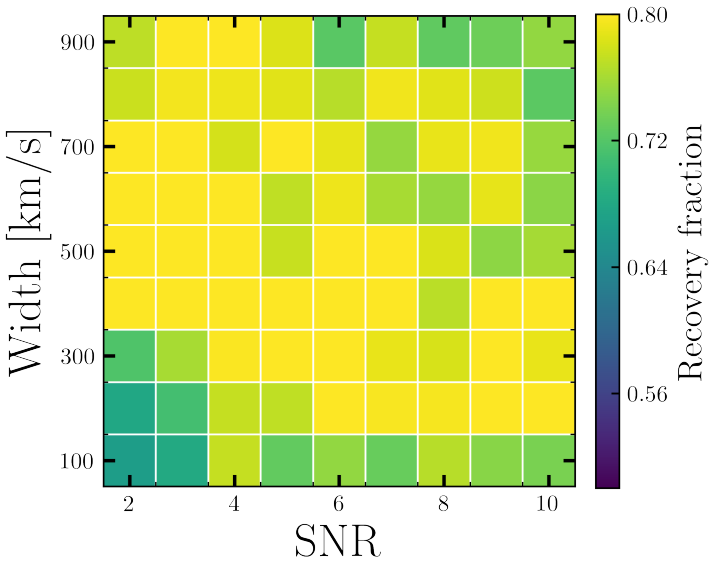

Figure 3.1: Completeness fraction of mock sources. The grid shows several combinations of the injected peak signal-to-noise ratio (S/N) and line width. The heat map represents how successfully the algorithm, SOFIA-2 (Westmeier et al. 2021), detects mock sources injected into data cubes, with the recovery fraction indicated by the colour. For sources with S/N 4 and line width 300 $\mathrm{km\,s^{-1}}$ we reach a completeness above $\sim 75\%$.

3.3.2 Completeness

We performed a completeness test to assess the quality of the cubes and the efficiency of the source-finding algorithm in retrieving artificial sources. To estimate the completeness factor, we injected mock sources into each data cube. We added artificial sources to each data cube and ran the source finder with the same parameters used to look for real emissions to quantify the recovery fraction. Iterating this process helps us to determine the quality of the cubes and algorithm's robustness under different conditions.

A set of mock sources was generated with known flux densities, positions, and sizes. The S/N values of the synthetic sources range from 2 to 18, and the line widths vary between 100 and 1600 $\mathrm{km\,s^{-1}}$. These ranges were chosen to explore the parameter space broadly. In each cube, 20 mock sources were randomly injected with varying properties. We ran SOFIA-2 on the data cube containing both the real and injected sources. We estimated the percentage of recovered sources by comparing the position of detected sources with the injected mock sources.

We determined whether SOFIA successfully recovers the injected sources and records both true positives (correctly identified sources) and false negatives (missing sources). The process was repeated ten times per cube to build up statistically significant samples, with each iteration using different mock sources to simulate varying conditions.

We analysed the detection rate, or completeness factor (c_i), based on the S/N and line width of the mock sources across two categories: high-completeness and low-completeness cubes. Some cubes consistently showed a recovery fraction of zero across all iterations. Further investigation revealed that these cubes were either extremely narrow (significantly less than 4 GHz) or exhibited structured noise. Since real detections in these cubes would likely be missed and their missing data cannot be corrected, we excluded them from the sample. However, cubes with low but non-zero completeness are kept, as their incomplete information can be accounted for using the completeness factor, and they still contribute to the statistical inference in the analysis. Ultimately, we kept 1107 cubes, removing 401 cubes with zero completeness.

Fig. 3.1 displays the completeness grid for our sample, indicating the fraction of successfully recovered injected sources as a function of the signal-to-noise ratio (S/N) and line width within feasible detection limits. We find that synthetic sources with S/N $\gtrsim 4$ and line widths of $\gtrsim 300$ $\mathrm{km\,s^{-1}}$ are successfully recovered in over 70% of cases. Although the recovery fraction becomes roughly homogeneous after these values, there is a lack of monotonic improvement towards the top right corner. To investigate why the highest bin in the grid does not reach full completeness, we examined several mock sources with high S/Ns and broad velocity components. In most cases, these sources fell into narrow spectral cubes, where the continuum level is insufficient for comparison, leading the algorithm to miss them.

3.3.3 Fidelity

We estimated the reliability of emission lines based on the significance of the detections compared to the noise distribution in the data cubes. We inverted the cubes and ran the source finder (SOFIA-2) with the same parameters, again searching for emission lines (which are, of course, absorption features in the physical cube). As we do not expect absorption features, any emission lines in the inverted cubes must represent random noise peaks.

We examined the spatial and spectral extent of the identified sources, estimating their S/N and line width. This distribution of noise peaks can be defined through the fidelity coefficient, given by:

$$F(\mathrm{S/N}, \sigma) = 1 - \frac{N_{\mathrm{negative}}(\mathrm{S/N}, \sigma)}{N_{\mathrm{positive}}(\mathrm{S/N}, \sigma)}, \tag{3.1}$$

where N_{negative} and N_{positive} are the number of negative and positive detections for a certain combination of S/N and line width.

Fig. 3.2, top panel, shows the 2D histogram of the positive candidates. The x axis represents S/N, while the $y$axis is the number of detection channels. We note that over 400 positive detections are initially found, but not all are real (see § 3.3.5). The bottom panel of Fig. 3.2 shows the fidelity grid, built as is indicated by Eq. 3.1. There are cases where a high fidelity value is found for low S/N and small kernel widths (bottom left corner). This may result from a limitation of the source-finding algorithm, which can favour narrow lines and lead to poorly constrained statistics in that range. However, these values do not affect our estimates, since the majority of our detections lie well above this region. Emission line candidates spanning at least seven channels with S/N > 7 reach the highest fidelity values, as is indicated beyond the grey lines in the top and bottom panels of Fig. 3.2. Most of the high-fidelity sources are outside the grey region.

3.3.4 Redshift estimation

In ALMACAL−22, identifying emission line candidates to determine redshifts poses significant challenges. Some calibrator fields are observed multiple times using different ALMA bands, but the archival calibrator data provides uneven spectral coverage across different fields. When an emission line was detected in fields with coverage in multiple ALMA bands, the likelihood of covering the spectral range for a second CO emission line increases. By evaluating the probability of each detected line being a specific CO transition, we checked

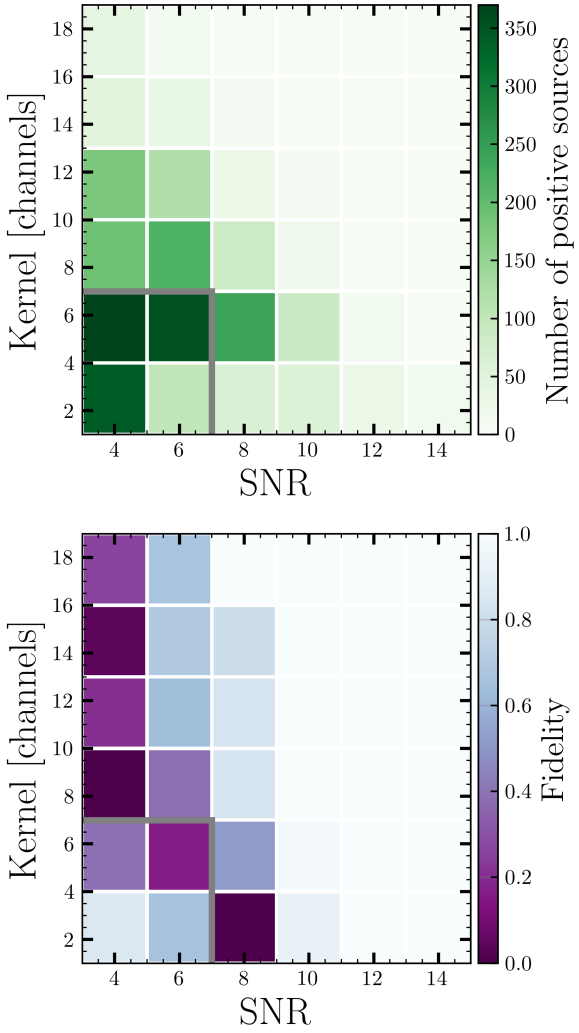

Figure 3.2: *Top*: 2D histogram of the number of positive sources detected in all the data cubes when searching for emission lines using SOFIA-2. The x axis represents the S/N of the line and the y axis shows the number of channels spanned by each detection. *Bottom*: Fidelity grid derived from Eq. 3.1 considering the number of positive and negative sources found for a given combination of S/N and number of channels. Values close to unity represent the highest fidelity, while zero represents the least reliable parameters. Lines with S/N above seven and spanning about seven channels usually have high fidelity values, i.e. beyond the grey lines of both panels.

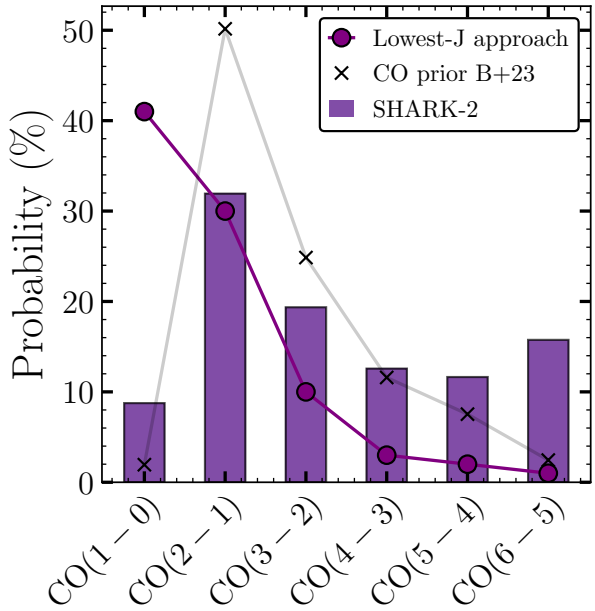

Figure 3.3: Probability distribution of all the CO detections in our sample. Given the measured properties of each emission line, we estimated redshifts using a probabilistic approach, based on the semi-analytical model, `SHARK-2` (Lagos et al. 2024) (see § 3.3.4). Overall, the CO(2–1) transition is the most likely to be detected in our sample, while the CO(1–0) transition is the least likely. This probability is carried into the further analysis. For comparison, we also include the CO prior redshift distribution association of sources with *HST* counterparts reported in Boogaard et al. (2023), written as B+23, and the probability distribution obtained adopting the lowest-J approach explained in Section §3.4.5.

if the spectral coverage includes another transition in any of our data cubes. This analysis revealed a total of 37 fields where a second transition could have been potentially covered. However, in most cases (29 of those 37 fields), potential detections fell outside the field of view due to the shrinking primary beam in higher ALMA bands. For the other eight sources, the sensitivity is too low to detect fainter transitions. Ultimately, we did not identify any potential transitions in these fields, so we are reliant on detections of one emission line.

We took a two-step approach to estimate the redshift. First, we employed a probabilistic method to estimate the redshift (§ 3.3.4). Second, we utilised known spectroscopic or photometric information (§ 3.3.4).

Probabilistic approach

We created a redshift probability distribution using the semi-analytical model, `SHARK-2` (Lagos et al. 2024), in the same way as is presented in Hamanowicz et al. 2022 with a previous version, `SHARK-1` (Lagos et al. 2018). `SHARK-2` includes improvements in physical models tracking the properties of supermassive black holes and active galactic nuclei feedback, and the environmental effects of satellite galaxies.

Based on the galaxy population of the SAM, we created a 2D histogram that maps CO transition flux against redshift. We then calculated a probability coefficient for each combination of flux and redshift by dividing the number of objects in each bin by the total number of galaxies in the simulated sample. This approach allowed us to determine the likelihood of a given CO flux originating at a specific redshift. High-level transitions ($J_{\rm up} = 7, 8, 9, 10$) are rare since very high kinetic temperatures and densities are needed to excite the CO molecule sufficiently to populate these higher rotational levels. The most likely detections are transitions $J_{\rm up} = 1, 2, 3, 4, 5$, and 6, placing our candidates between redshifts 0 and 6. Instead of simply selecting the $J-$transition with the highest probability for each detection, we adopted a more comprehensive approach. We included all plausible J transitions, each weighted according to the probabilities evaluated by `SHARK-2`. This method allows us to account for the uncertainty in identifying the exact J transition.

Fig. 3.3 shows the probability distribution of all the CO candidates in our sample. `SHARK-2` assigns the CO(2–1) transition most frequently in our catalogue, and CO(1–0) as the least frequent. This distribution closely matches the shape of the CO line prior used for redshift associations in sources without counterparts seen in Boogaard et al. (2023).

Two main factors influence the shape of this distribution. First, the CO molecule emits radiation at different frequencies depending on its rotational energy level (J level). Some transitions, like CO(2–1), are naturally more common, or stronger, while others are rarer or weaker. Second, the volume probed by our observations for each transition, especially CO(1–0), also plays a role, as lower-frequency transitions probe smaller volumes.

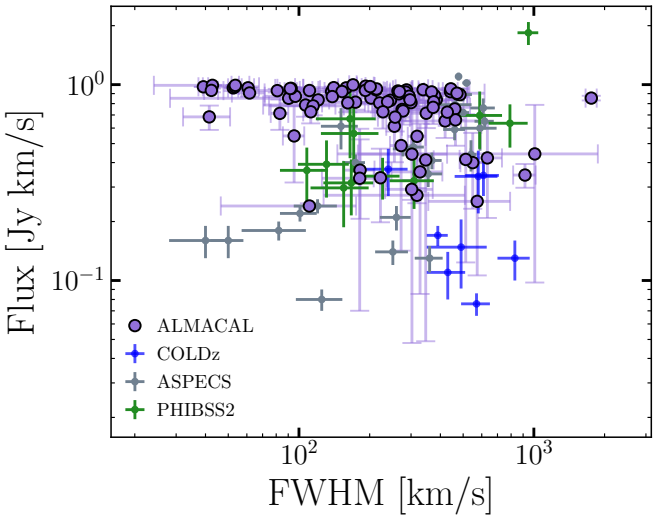

Figure 3.4: Integrated flux and width of the CO emitters detected in ALMACAL−22 (purple). For comparison, we include detections from other surveys: ASPECS (Decarli et al. 2020), COLDz (Riechers et al. 2020b) and PHIBSS (Lenkić et al. 2020). Our candidates span a similar range of line widths as previous studies, and are generally brighter, as is expected from the difference in depth and volume.

Existing spectroscopic and photometric data

We inspected NED, as well as MUSE and *HST* archives, to find the candidates' optical/NIR counterparts. This search yielded seven MUSE cubes. From the MUSE cubes, we confirmed the spectroscopic redshift of three objects, all of them at $z < 1$. We fixed the known spectroscopic redshift for three sources and set the fidelity parameter to one. The other objects are not detected in any of the possible emission lines that MUSE could have covered. However, MUSE spectra do not cover any emission lines of sources at $1.4 \lesssim z \lesssim 2.8$. Another reason for not detecting these sources could be heavy dust obscuration, making a galaxy faint in the rest-frame UV range.

Photometric data are available for seven sources. We performed SED fitting with `EAZY` (Brammer et al. 2008) to estimate their photometric redshifts (photo$-z$). In three cases, the photo$-z$ values match one of the highest probability redshifts given by `SHARK-2`; we adopted the photometric redshift for these sources. For the remaining objects, there are fewer than five photometric bands available, resulting in broad photo$-z$ probability distributions that cannot reliably estimate the redshift or assign a redshift probability greater than 10%. For these cases, we kept the redshift probability described earlier (§ 3.3.4).

3.3.5 Final catalogue

The final catalogue was created after a thorough visual inspection of all detections, along with an analysis of the spectra and the integrated emission-line map across the spectral range of the detected lines. We selected candidates with S/N > 4 and a clear peak in the moment map compared to the background noise. Duplicate detections – for example, those offset by only a few pixels but originating from the same region – were excluded. The final catalogue of CO emitters includes 87 sources.

We extracted the spectra using the aperture from SOFIA's output catalogue. To ensure the accuracy of the number of channels spanned by the emission line, we re-calculated the S/N estimate by dividing the peak flux by the RMS of the moment map. The moment map was created by integrating across the frequency range of the emission and collapsing it into a single image. To measure the emission line flux, we used the function `scipy.optimize.curve_fit` to fit a Gaussian profile to the emission lines. We repeated this procedure 100 times for each candidate, perturbing the flux using the RMS of the cube. From the best fit, we use the mean values of the standard deviation and amplitude to estimate the flux and the full width at half maximum (FWHM). Our final catalogue of CO sources contains the flux and FWHM with their uncertainties, as well as the completeness and fidelity factors estimated from the S/N and line width of each object, reported in Table A.1 of the Appendix. Generally, we see that in fields where a lower rms noise level is reached, more detections are found. However, in some cases, fields with lower sensitivity still exhibit multiple detections, which can be attributed to the stochastic nature of source distribution within the field of view. Also, there are a few sources with a FWHM $\gtrsim$ 1000 km/s, such as J0334-4008.4, most likely explained by blended sources for which the achieved resolution is insufficient to resolve them.

Fig. 3.4 shows the measured CO fluxes against the FWHM of each emission-line candidate. The FWHM values measured in ALMACAL−22 are comparable with previous surveys, while the fluxes are generally higher. This trend is expected due to the differences in the sensitivity and volume probed by different surveys. The ALMACAL−22 survey benefits from a larger observational volume, but it has an uneven sensitivity distribution among data cubes. As a result, we can identify and measure the flux of more luminous sources, while detecting fainter systems is more challenging.

As an additional verification of the final catalogue, we estimate the probability of detecting high-redshift interlopers with bright [C II] emission, at 1900.53 GHz, instead of a CO transition. The [C II] emission line can be detected in one of the ALMA bands for redshifts $z > 4$, where it enters ALMA band 7. We estimate the volume covered by the [C II] line in band 7 between redshifts $z = 4$ and $z = 6$, to be 5858.16 cMpc^3, using the same methodology employed to estimate the probed volume for CO transitions (see § 3.4.1). We calculated the expected number of sources within this volume based on the [C II] LF in this redshift range (Yan et al. 2020; Casavecchia et al. 2024). We found that a maximum of two sources can be expected in a survey covering the volume probed by the [C II] line. We further discuss the effect of removing two band 7 detections in § 3.4.6.

To confirm that our detections are independent of the calibrator, we verify that the redshifts of our sources differ from those of the calibrator. The calibrator redshift catalogue ($z_{\rm cal}$) in ALMACAL−22 is based on a database from Bonato et al. (2018), which is extended through cross-matching with optical catalogues (NED, SIMBAD and Mahony et al. 2011) and supplemented by 70 VLT/X-Shooter spectra (ID 111.253L.001, PI: S. Weng and ID 0101.A-0528, PI: E. Mahony). Our sample includes 87 CO emitters spread across 46 quasar fields, with redshift information available for 37 fields (Weng et al. 2025). For each emission line, we compare the calibrator's redshift against all probable redshifts. Additionally, we verify that the estimated photometric redshift (§ 3.3.4) is distinct from that of the calibrator. Four candidates were excluded from our sample due to probable redshifts within 2000 km s^{-1} of the quasar redshift. Thus, we assert that our candidates are independent of the calibrator. Despite using data centred on calibrators, this approach does not introduce bias into our survey.

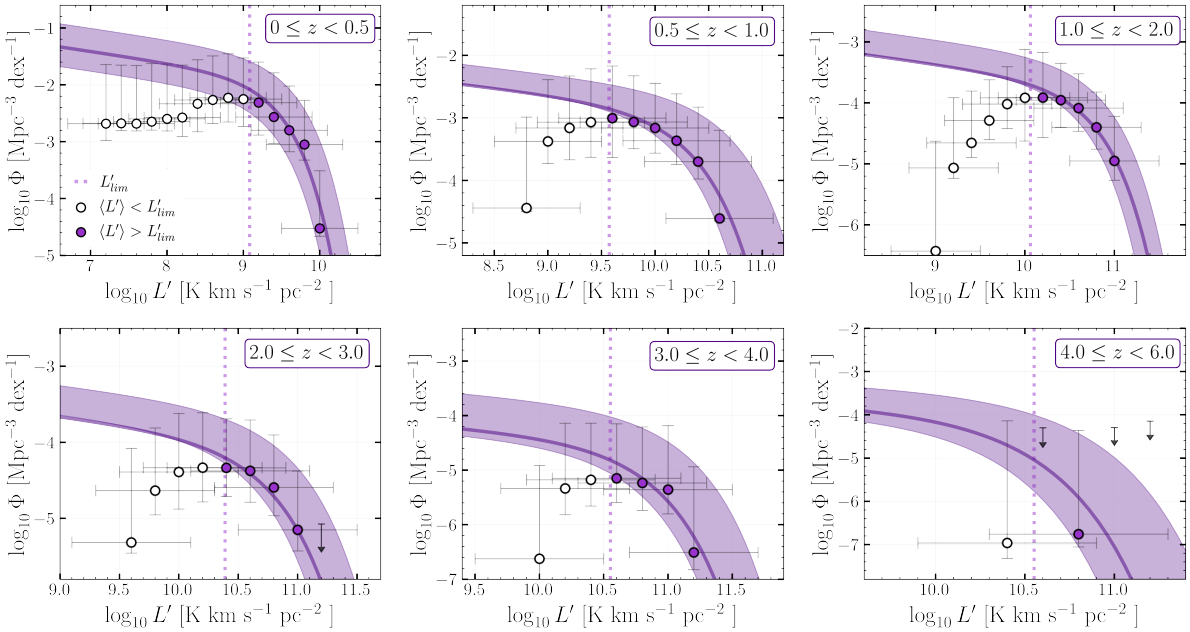

Figure 3.5: CO LF across redshift bins: $z = 0$–0.5, $z = 0.5$–1, $z = 1$–2, $z = 2$–3, $z = 3$–4 and $z = 4$–6. Detections are shown in circles and non-detection as arrows. The vertical dotted line in each panel indicates the detection limit, representing the faintest luminosity detectable at a 50% reduction in primary beam sensitivity. The CO LFs are fitted using only the filled points that lie above this luminosity threshold. The best-fit Schechter function (Schechter 1976) is plotted, with Φ^* and L'^* allowed to vary freely, while the faint-end slope is fixed at $\alpha = -0.2$. The shaded region represents 1σ confidence intervals. Overall, the CO LF reveals a decreasing normalisation factor, while the 'knee' of the function shifts to higher luminosities as redshift increases.

Table 3.2: Schechter best-fit parameters of the CO LF

Redshift	$\log_{10}\Phi^*$ [Mpc^{-3} dex^{-1}]	$\log_{10}L'^*$ [K km s$-$1 pc^{-2}]
$0.0-0.5$	$-1.51^{+0.32}_{-0.37}$	$9.31^{+0.18}_{-0.09}$
$0.5-1.0$	$-2.48^{+0.51}_{-0.31}$	$10.12^{+0.17}_{-0.51}$
$1.0-2.0$	$-3.32^{+0.50}_{-0.17}$	$10.58^{+0.26}_{-0.18}$
$2.0-3.0$	$-3.61^{+0.59}_{-0.08}$	$10.59^{+0.30}_{-0.17}$
$3.0-4.0$	$-4.10^{+0.91}_{-0.29}$	$10.62^{+0.50}_{-0.18}$
$4.0-6.0$	$-3.54^{+0.40}_{-0.16}$	$10.16^{+0.33}_{-0.17}$

Notes. Fixed $\alpha = -0.2$

3.4 CO luminosity function and molecular gas mass density

A LF describes the statistical distribution of the number density of sources over a range of luminosities. In particular, the CO LF provides insight into the abundance of molecular gas in galaxies by revealing the variation in the number density of CO emitters. This section describes the methodology used to derive the CO LF and the results. First, we estimate the volume covered by the ALMACAL$-$22 survey in § 3.4.1. We build the CO LF in § 3.4.2 and fit Schechter functions in § 3.4.3. We present the molecular gas mass density estimates at different redshift bins in § 3.4.4. Finally, we explore the lowest$-J$ approach in § 3.4.5 and the uncertainties associated with our measurements in § 3.4.6.

3.4.1 Volume estimation

We estimated the co-moving volume covered in our sample using the same methodology described in Hamanowicz et al. (2022). Here, we briefly summarise the steps for this calculation.

First, we defined the sky area covered by each ALMA cube, which depends on the frequency range and the field of view of the observations. For each CO transition ladder

(from $J = 6$ to $J = 0$), we calculated the co-moving volume by integrating over both the redshift element and the solid angle subtended by the detectable emission lines.

Second, we integrated the sensitivity variation across ALMA data cubes. The sensitivity decreases from the centre of the primary beam, following a pattern that resembles a Gaussian function. We estimated the lower limits of detectable volume, modelling an emission line with CO luminosities ranging from 10^5 to 10^{13} K km s^{-1} pc^2. When the frequency coverage of the cube constrained the volume, we integrated the volume elements through concentric rings, each defined by the observed frequency. The resulting co-moving volumes for each CO transition are detailed in Table 3.1.

3.4.2 CO LF

We calculated the CO luminosity from CO fluxes using the following equation, from Solomon et al. (1997),

$$L'_{\rm CO} = 3.25 \times 10^7 \frac{S_{\rm CO}\Delta V}{(1+z)^3} \left(\frac{D_{\rm L}}{\nu_{\rm obs}}\right)^2 \ [\text{K km s}^{-1}\text{pc}^2] \,, \tag{3.2}$$

where $L'_{\rm CO}$ is in units of K km s^{-1} pc^2, $\nu_{\rm obs}$ is the observed frequency of the CO line in GHz, $D_{\rm L}$ is the luminosity distance of the galaxy in Mpc, z is the redshift, and $\rm S_{CO}\Delta V$ is the integrated flux in Jy km s^{-1}. We converted the luminosity measured for mid- and high-J CO into CO(1–0), scaling by the empirical conversion factors from Boogaard et al. (2020): $r_{j\to1}$ ={3.33, 5.20, 4.76, 2.70, 0.53} at $z < 2$ and $r_{j\to1} = \{4.09, 8.24, 12.21, 14.68, 13.86\}$ at $z > 2$, for $J = 2, 3, 4, 5, 6$.

To construct the CO LF, we sampled over the uncertainties of the integrated flux and the conversion factor used to derive the CO(1–0) flux, the completeness and fidelity factor, and the redshift probability. The completeness factor depends on whether the emission lines were found in a cube with a high or low completeness factor (§ 3.3.2). The fidelity depends on the S/N and the line width of the emission lines. As was noted previously, if the line has been confirmed by a counterpart with a matching spectroscopic or photometric redshift, we assume the fidelity is equal to one.

We performed 1000 realisations of the CO LF over five independent bins, shifted by 0.2 dex, where each bin spans one dex in luminosity. We adopted the same methodology presented by Decarli et al. (2016), using independent shifted bins. This approach allows us to evaluate the dependence of the re-constructed CO LF on the bin definition. In each

iteration, the redshift of each source will depend on the redshift probability calculator described earlier (§ 3.3.4). Hence, the same source will have different redshift values according to the weight of their probability in each realisation.

We binned galaxies by their CO luminosities, counted the number in each bin, and then normalised these counts by the co-moving volume. This gave us the number density of galaxies at different CO luminosities. We built the normalised LF, $\Phi(L)$, also known as the number density of objects per unit luminosity, per unit co-moving volume, using the following expression:

$$\Phi(\log L') \cdot d(\log L') = \sum_i \frac{F_i}{V_i \cdot c_i} \ [\mathrm{Mpc}^{-1}\mathrm{dex}^{-1}] \,, \tag{3.3}$$

where $\Phi(L')$ is the luminosity density per cubic megaparsec per dex, $d(\log L')$ is the luminosity bin, V_i is the comoving volume accessible for each transition, F_i is the fidelity and c_i is the completeness factor estimated in § 3.3.2 and Eq. 3.1.

The final CO LF was divided into six redshift bins: $0 \leq z < 0.5$, $0.5 \leq z < 1$, $1 \leq z < 2$, $2 \leq z < 3$, $3 \leq z < 4$, and $4 \leq z < 6$. Fig. 3.5 shows the median values with their uncertainties. The error bars show the natural spread of data derived from 1000 realisations. We calculated the 16th and 84th percentiles and the 1σ Poisson confidence intervals for each bin, which accounts for the statistical fluctuations, following the method described in Gehrels (1986). When a particular luminosity bin contains fewer than one source on average, we provide a 1σ upper limit to reflect the lower reliability of the data in that bin.

3.4.3 Schechter fits

We fitted the CO LF using an analytical Schechter function (Schechter 1976) in logarithmic scale as follows:

$$\log \Phi(L') = \log \Phi^* + \alpha \log\left(\frac{L'}{L'^*}\right) - \left(\frac{L'}{\ln(10) L'^*}\right) + \log(\ln 10) \,, \tag{3.4}$$

where $\Phi(L')$ is the number of galaxies per co-moving volume with a CO luminosity between $\log L'$ and $\log L' + d(\log L')$. Φ^* is the normalisation factor, α is the faint-end slope, and L'^* indicates the luminosity at which the LF changes from a power law to an exponential function, also known as the 'knee' of the LF.

Fig. 3.5 shows the best-fit Schechter functions, with the shaded regions representing 1σ confidence intervals. Vertical lines indicate the luminosity limits, which were determined by considering the noise level of each ALMA cube and a 50% drop-off across the Gaussian-shaped primary beam. To ensure consistency, we adopted the limits from the shallowest cubes in the sample for each redshift bin and CO transition. The highest redshift panel contains only one data point above the luminosity limit with significant statistical weight. Nevertheless, we display the fit for completeness. To derive the Schechter function fits, we performed 500 realisations, sampling the CO LF within its uncertainties. We only included bins where the median luminosity exceeds the calculated luminosity limit (indicated by filled points in Fig. 3.5). We allowed the parameters Φ^* and L'^* to vary freely during the realisations. However, due to the limited number of data points at the faint end, we adopted a fixed slope at $\alpha = -0.2$, following the approach used by Boogaard et al. (2023). The median values and the uncertainties obtained from the best fits are reported in Table 3.2 for different redshift bins.

3.4.4 Molecular gas mass density

The molecular gas mass density refers to the total mass of molecular gas per unit volume in a given region of the Universe. It is often denoted as $\rho_{\rm H_2}$ or $\Omega_{\rm H_2}$ when normalised to the critical density of the Universe. We derive the molecular gas mass of the CO emitters using the CO(1–0) luminosity and a conversion factor, $\alpha_{\rm CO}$, as follows:

$$M_{\rm H_2} = \alpha_{\rm CO} \cdot L'_{\rm CO(1-0)} \ \ [M_\odot] \,, \tag{3.5}$$

where we used $\alpha_{\rm CO} = 3.6 M_\odot$ (K km s^{-1} pc^2) for all detections. Adopting this value makes the comparison with previous studies straightforward, since they use the same $\alpha_{\rm CO}$ (Decarli et al. 2019; Riechers et al. 2019; Lenkić et al. 2020; Riechers et al. 2020b). A different constant value affects our results linearly.

We derive the molecular gas mass density, $\rho(M_{\rm H_2})$, by integrating $L'_{\rm CO}\Phi(L'_{\rm CO})$ up to the luminosity limit and applying the conversion factor, $\alpha_{\rm CO}$. We exclude the same data points as before when fitting Schechter functions. Fig. 3.6 shows the estimates from ALMACAL−22 for each redshift bin, where the right y axis shows the unitless density parameter for molecular gas, $\Omega_{\rm H_2} = \rho_{\rm H_2}/\rho_{0,\rm crit}$. The associated uncertainties were estimated in the same way as for the CO LF, explained above.

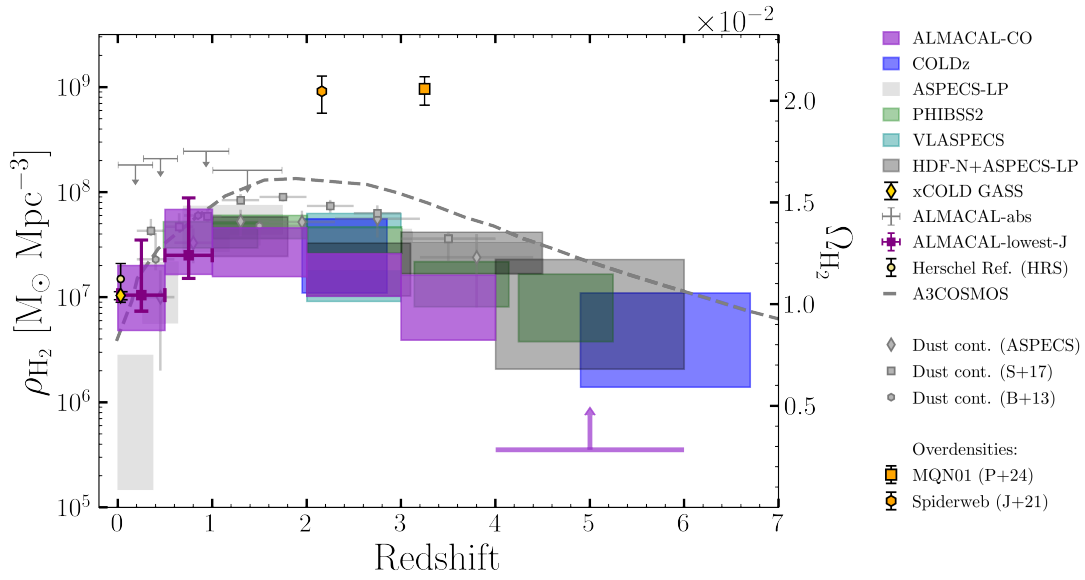

Figure 3.6: Cosmic molecular gas mass density evolution measured from ALMACAL−22 in purple. The right y axis represents the unitless density parameter for molecular gas, $\Omega_{H_2} = \rho_{H_2}/\rho_{crit,z=0}$. For comparison, we include estimates from other surveys: COLDz (Riechers et al. 2020b), ASPECS CO (Decarli et al. 2020) and dust measurements (Magnelli et al. 2020), PHIBSS (Lenkić et al. 2020), VLASPECS (Riechers et al. 2020a), HDF-N+ASPECS-LP (Boogaard et al. 2023), xCOLD GASS (Saintonge et al. 2017; Fletcher et al. 2021), ALMACAL in absorption (Klitsch et al. 2019b), ALMACAL−22 adopting the lowest−J approach (this work, see § 3.4.5), from the K−band of the Herschel Reference Survey (HRS) (Andreani et al. 2020), A^{3}COSMOS (Liu et al. 2019), other estimates from dust continuum (Scoville et al. 2017; Berta et al. 2013), and from recent works reporting over-densities at $z \sim 2$ in the Spiderweb galaxy protocluster (Jin et al. 2021) and at $z \sim 3$ in the MUSE Quasar Nebula 01 (MQN01) field (Pensabene et al. 2024). We plot consistent estimates scaled to the same cosmology (H_0) used in this work and without helium contribution. We find overall consistency with the trend reported by other blind surveys, although slightly lower values are found at high redshift ($z > 1$).

Table 3.3: Cosmic molecular gas density from ALMACAL

Redshift	$\log \rho_{H_2}$ $[M_\odot\ Mpc^{-3}]$
$0.0 - 0.5$	$6.92^{+0.35}_{-0.37}$
$0.5 - 1.0$	$7.44^{+0.35}_{-0.38}$
$1.0 - 2.0$	$7.35^{+0.21}_{-0.51}$
$2.0 - 3.0$	$7.18^{+0.13}_{-0.47}$
$3.0 - 4.0$	$6.84^{+0.34}_{-0.35}$

3.4.5 Lowest$-J$ possible

We now explore the conservative approach of assuming the lowest$-J$ CO transition detectable for all sources. Given the frequency of each line, we classify the CO emission line at the lowest possible redshift. This method focuses on transitions from $J = 1$ to $J = 4$, which are expected to be the brightest and the most likely to be observed. We performed 1000 realisations using the same luminosity and redshift bins. In each realisation, we perturbed the luminosity and the empirical scaling factor to convert high$-J$ into CO(1–0) within linear uncertainties. We applied the same estimates of the volume and the correction factors (completeness and fidelity) described in § 3.4.2.

We built the CO LF using Eq. 3.3 and derived the molecular gas mass density using Eq. 4.2. Our results are displayed in Fig. 3.6 as purple boxes. These estimates match the previous results using the redshift probability calculator built from the semi-analytical model, `SHARK-2`, described in § 3.3.4. The main difference is the uncertainty, which is larger for the estimates derived using a probabilistic approach. Using the lowest$-J$ approach, most of our detections would be classified as CO(1–0), considering that most were found in band 3. This likely skews the distribution of CO across cosmic epochs, since it does not consider the volume probed by each transition. As a result, we may be over-estimating the molecular gas content through this approach. Still, it is a conservative way to validate the use of models to estimate redshifts based on more complex CO line ratios, as was previously noted by Hamanowicz et al. (2022).

3.4.6 Possible biases and uncertainties

We identified the potential effects of the measured uncertainties on interpreting our results. One significant factor is the Poisson error, particularly in bins with very few objects ($\lesssim 5$), as we are probing nearly one dex in luminosity at the bright end of the LF. Additionally, the redshift uncertainties, particularly in converting high$-J$ CO transitions into CO(1–0), could lead to the misclassification of transitions. Misinterpreting a higher$-J$ transition as CO(1–0) would imply lower gas masses.

Another uncertainty stems from the uneven sensitivity of the data cubes. While we treat sensitivity on average, the data combined multiple observations with varying configurations, leading to changes in the noise across different spectral channels. Despite the efforts to correct the presence of false positives or missing detections through the completeness and reliability factors, we still allow for the possibility that some uncertainties are not fully understood. In some cases, we had to exclude data cubes where no mock sources were retrieved, resulting in a zero completeness factor. Low completeness and low reliability, especially for sources with low signal-to-noise ratio and narrow line width, add further uncertainty, particularly for the brightest sources.

We also explored the possibility that we are detecting other lines than CO, particularly the bright [C II] emission line, but the probabilities are quite low. This line is detectable in ALMA band 7 at high redshifts ($z > 4$), where the observed volume suggests a maximum of two detections. We observe five detections in band 7, and even if two of them in the high-redshift bin were removed, it would not affect the overall results, as the values in that bin are lower limits.

3.5 Discussion

This section presents our findings in the context of cosmic evolution. In § 3.5.1 and § 3.5.2, we compare our CO LF and the molecular gas mass density estimates with previous works using observations and simulations. In § 3.5.3, we delve into the effects of the field-to-field variance of ALMACAL−22 compared with previous surveys. In § 3.5.4, we investigate the evolution of baryons associated with galaxies averaged over cosmic time and space, including the new estimates from ALMACAL−22.

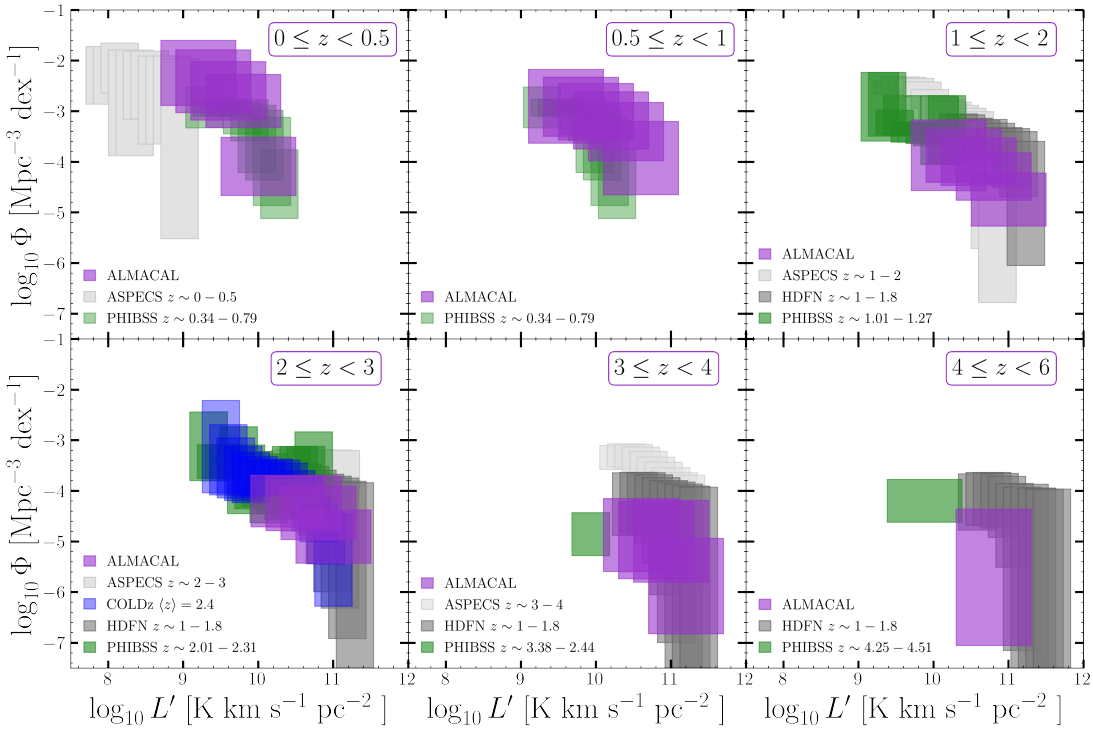

Figure 3.7: CO LF in different redshift bins, from $z = 0$ to $z = 6$, as is indicated at the top right of each panel. The measurements from ALMACAL−22 are shown in purple, including the uncertainty of each bin given by the extension of the boxes. We derived the CO LF based on the CO(1–0) luminosity, calculated as is explained in § 3.4.2. For comparison, we include observational constraints found by previous surveys at similar redshift ranges: ASPECS (Decarli et al. 2019, 2020), PHIBSS2 (Lenkić et al. 2023), COLDz (Riechers et al. 2019), HDFN (Boogaard et al. 2023). In some cases, the estimates from these surveys were derived from different CO transitions, here we converted the luminosities $L'_{\mathrm{CO(1–0)}}$ consistently to our methodology. The CO LF derived from the ALMACAL survey aligns well with findings from other surveys, while probing a wider redshift range.

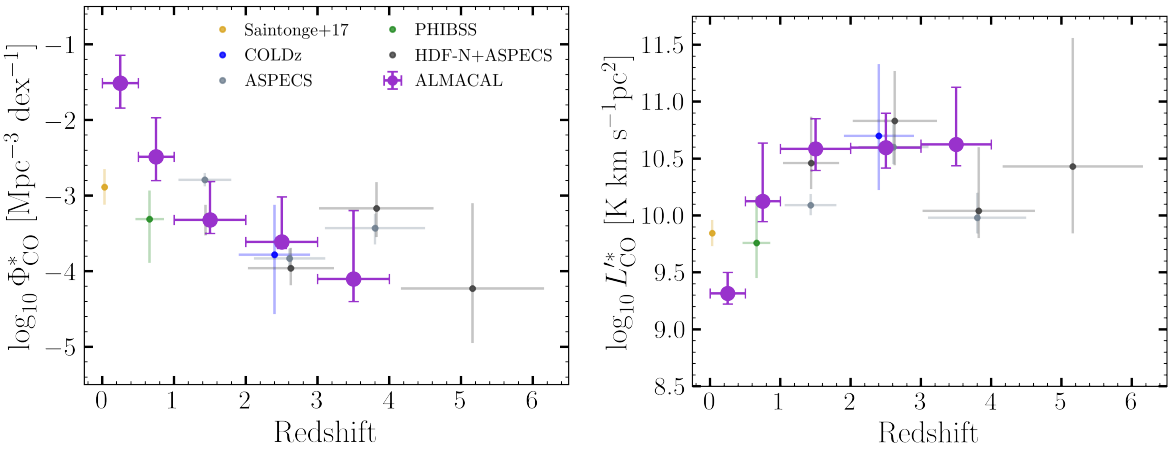

Figure 3.8: Evolution of the Schechter best-fit parameters for the CO LF across $z \sim 0$ to $z \sim 6$. *Top*: The evolution of Φ^* shows a decrease with redshift from $z = 0$. *Bottom*: The evolution of $L^*_{\rm CO}$ shows a consistent increase from $z \sim 0$ to $z \sim 2$, and remains roughly constant at higher redshifts. For the fits shown in Fig. 3.5, the faint-end slope was fixed at $\alpha = -0.2$.

3.5.1 CO LF

Comparison with other observations

Fig. 3.7 shows the redshift evolution of the CO LF derived from $z = 0$ to $z = 6$ from the ALMACAL−22 survey. Our estimates are based on the CO(1–0) luminosity. For comparison, we also include measurements from previous surveys at similar redshifts. Findings from the COLDz survey (Riechers et al. 2019) ASPECS survey are also based on the $L'_{\rm CO(1-0)}$. Estimates from PHIBSS2 (Lenkić et al. 2023), ASPECS (Decarli et al. 2019, 2020), and HDFN surveys (Boogaard et al. 2023) are derived from higher−J CO transitions, scaled to the CO(1–0) consistently with our measurements as is detailed above (§ 3.4.2).

At low redshift, $z \sim 0$–1.0, our estimates indicate higher values of the number density of sources at a given luminosity, compared to findings from ASPECS and PHIBSS. At intermediate redshift, $z \sim 1$–2, we find values slightly lower than the results from PHIBSS but consistent with ASPECS and HDFN. At higher redshift, $z \sim 2$–6, our results are in good agreement with COLDz and HDFN, but slightly lower than ASPECS. At these redshifts, the estimates from PHIBSS are derived from high−J CO transitions, and the uncertainties converting them CO(1–0) may be larger, so they were included just for reference.

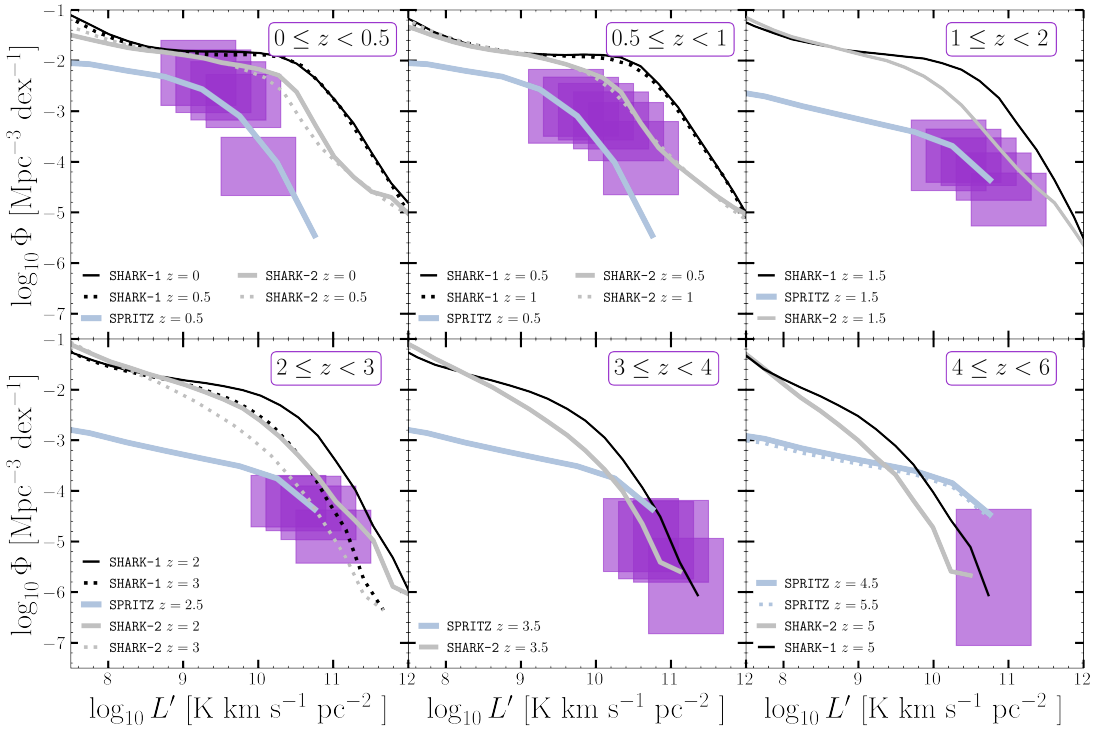

Figure 3.9: CO LF from redshift $z = 0$ to $z = 6$ derived from ALMACAL−22 in comparison with simulations. We include the predictions from `SHARK-1` (Lagos et al. 2018) at $z = 0$–3, from `SHARK-2` (Lagos et al. 2024) at $z = 0$–5.5, and from `SPRITZ` (Bisigello et al. 2022) at $z = 0.5$–5.5. We report consistency in most of the redshift ranges, and we recall the need for simulations to expand their range towards the bright end of the CO LF.

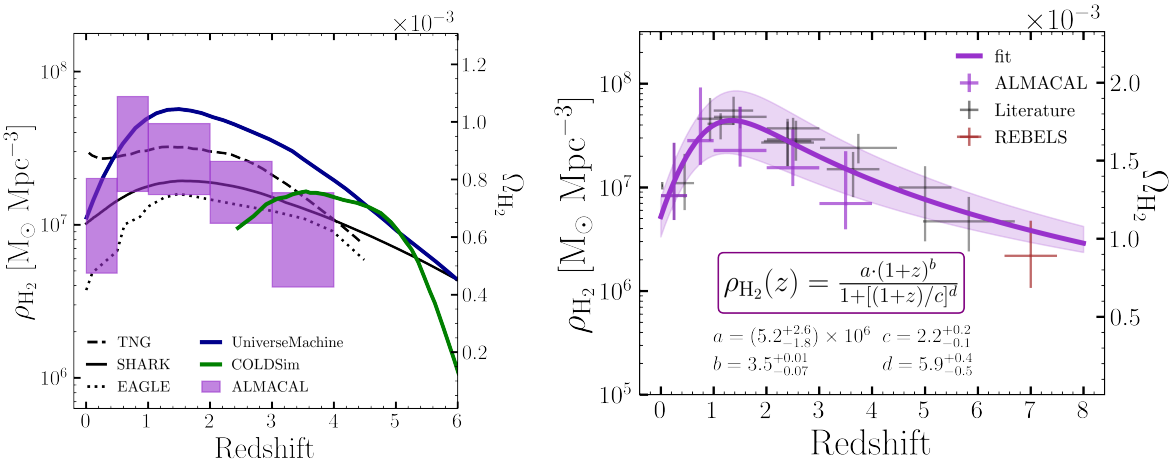

Figure 3.10: *Top*: Evolution of the molecular gas mass density obtained from ALMACAL−22 (purple) in comparison with predictions from simulations. We include the results from Illustris TNG (Popping et al. 2019b), EAGLE (Lagos et al. 2015), `SHARK-2` (Lagos et al. 2018), UniverseMachine (Guo et al. 2023a), and COLDSim (Maio et al. 2022). *Bottom*: We compiled several measurements from previous CO surveys, as is shown by the grey points, including ALMACAL−22 (see text). We present a new fit of the functional form used in Walter et al. 2020, including the ALMACAL−22 estimate, shown by the purple line, with the 1σ uncertainty shown by the shaded region. For reference, we add the estimates of ρ_{H_2} at the highest redshift measured up to date by the REBELS survey as the dark red point (Aravena et al. 2024), which is in excellent agreement with the extrapolation of our fit.

Another interesting way to explore the evolution of the CO LF is through their Schechter parameters. In Fig. 3.8, we show the evolution of the Schechter parameters derived in § 3.4.3. We compare our results with the surveys mentioned above, including the parameters of the well-known CO LF in the local Universe reported by xCOLD (Fletcher et al. 2021). The normalisation factor (top panel), Φ^*, shows a decreasing trend towards higher redshift, from $z \sim 0$ to $z \sim 6$ by a factor of $\sim 3\times$. The characteristic luminosity (bottom panel), L'^*, increases from redshift $z \sim 0$ to $z \sim 2$ by a factor of $\sim 1.15\times$ and has a constant value towards higher redshift. This evolution suggests the CO luminosity in high-redshift galaxies is dominated by galaxies of $\log L'_{\rm CO} \lesssim 10.5$ up to cosmic noon, after which brighter systems become more common. Overall, there is a good agreement between the parameters estimated from the ALMACAL−22 survey and previous studies. The largest difference lies in the low-redshift bins, for which the assumption of a fixed faint slope at $\alpha = -0.2$ may not be appropriate, as low-z studies have reported a different value ($\alpha \sim 1.2$, Fletcher et al. 2021). Further, examination of the evolution of the Schechter parameters of the CO LF is needed, particularly in the low-mass regime ($\log L'_{\rm CO} \lesssim 9$), where the number density values are more difficult to constrain due to observational limitations.

Comparison with simulations

Several approaches have been used to model the physics of the CO emission line through hydrodynamical simulations and semi-analytical models. Hydrodynamical simulations simultaneously model the physics of gas, dark matter and stars within galaxies, providing detailed predictions about the formation of molecular gas like CO and H_2 (e.g. Pelupessy & Papadopoulos 2009; Obreschkow et al. 2009; Narayanan et al. 2012; Katz et al. 2017; Vallini et al. 2018; Casavecchia et al. 2024). These simulations capture the small-scale physical processes, but their high computational cost limits the number of galaxies that can be simulated, especially over a wide range of properties. Also, their resolution is often insufficient to fully resolve small-scale processes like star formation (Lagos et al. 2012). On the other hand, semi-analytical models (SAMs) use simplified assumptions to describe the evolution of galaxies, such as symmetries in galaxies and haloes. These models are usually run on halo catalogues which are made beforehand from an available dark-matter-only simulation (e.g. `SURFS` Elahi et al. 2018), and then simplified models for gas physics are applied. The main advantage of SAMs is their computational efficiency, allowing them to explore a wide range of values in the parameter space. SAMs have been successful

in modelling the interstellar medium (ISM), star formation, and CO emission across different transitions, as well as considering factors like gas metallicity and radiation fields (e.g. UV and X-ray Lagos et al. 2012; Popping et al. 2014, 2016). While hydrodynamical simulations provide detailed insights into physical processes, SAMs offer broader statistical power, complementing both approaches in understanding the physics of CO emission and galaxy evolution.

We display the predictions of the CO LF from simulations in Fig. 3.9. We compare the results from ALMACAL−22 with the findings from the semi-analytical model `SHARK-1` (Lagos et al. 2018), `SHARK-2` (Lagos et al. 2023) and the phenomenological simulation `SPRITZ` (Bisigello et al. 2022), which is based on empirical and theoretical relations. At low redshift, $z \lesssim 1$, `SPRITZ` produces values that are very much consistent with ALMACAL, particularly at the bright end of the LF. In this range, both `SHARK` versions produce values slightly above our estimates. In the rest of the panels, simulations are in good agreement with our findings, particularly with `SHARK-2`, which effectively constrains the bright end of the CO LF. Although we use the redshift probability distribution derived from SHARK-2 as a prior in our analysis, the final results on the molecular gas mass density and luminosity function are decoupled from the simulation. This independence is demonstrated by the differences observed in the CO luminosity functions and the results from the lowest$-J$ approach, explained in Section §3.4.5 and further discussed in §3.5.2.

`SHARK-2` implements a new AGN feedback model, compared to `SHARK-1`, which reduces gas content and enhances quenching in massive galaxies, especially at lower redshifts, driving differences seen at higher luminosities. At lower masses, the updated model accounts for additional environmental processes like ram pressure and tidal stripping, affecting gas content and causing variations in the faint-end slope of the CO LF, particularly at lower redshifts where satellite galaxies are more common. We recall the need for hydrodynamical simulations to expand the simulated galaxy population towards the high-mass regime and probe fainter regimes of the LF. So far, the CO LF, particularly at high redshift, is still not available in a cosmological context; in other words, big boxes (Bisigello et al. 2022; Garcia et al. 2023).

3.5.2 Molecular gas mass density evolution

Comparison with other observations

Fig. 3.6 compares the molecular gas mass density estimates, $\rho_{\rm H_2}$, from ALMACAL−22 with previous works. We include the estimates from ASPECS (Decarli et al. 2019, 2020), COLDz (Riechers et al. 2020b), PHIBSS (Lenkić et al. 2020), VLASPECS (Riechers et al. 2020a), ALMACAL in absorption (Klitsch et al. 2019a), dust measurement from Berta et al. (2013); Scoville et al. (2017); Magnelli et al. (2020), estimates in the local Universe from xCOLD GASS (Saintonge et al. 2017; Fletcher et al. 2021) and Andreani et al. (2020), estimates at redshift $z \sim 2.2$ and $z \sim 3.2$ for over-dense systems from (Jin et al. 2021; Pensabene et al. 2024). We also included the estimates from the A^{3}COSMOS dataset (Liu et al. 2019), which uses different calibration methods to derive the molecular gas mass, including CO lines, SED-fitted dust mass and Rayleigh-Jeans(RJ)-tail dust continuum.

In the lower redshift bins ($z \sim 0$–1), our estimates are consistent with the findings from xCOLD GASS and Andreani et al. (2020), and slightly above the estimate by ASPECS. The lowest−J approach, explained in Section §3.4.5, shows excellent agreement with the main methodology, which relies on redshift estimated trough `SHARK-2`, demonstrating that our results remain independent. At $z \sim 1$–2, we agree with the results from PHIBSS, but are below HDFN and ASPECS. At $z \sim 2$ our estimates are consistent with VLASPECS, while PHIBSS, while HDFN and ASPECS report higher values. At $z \sim 3$–4 we find consistency with PHIBSS, but the other surveys show higher values. In the highest redshift panel, the integration of the CO LF does not provide fully consistent measurement, as none of the individual measurements is completely above the luminosity limit, so we provide a lower limit. Estimates of the molecular gas mass density from dust continuum Berta et al. (2013); Scoville et al. (2017); Magnelli et al. (2020) are usually slightly above our estimates but still follow the same evolutionary trend. High-redshift estimates of regions classified as over-densities are reported by Jin et al. (2021) and Pensabene et al. (2024). These estimates are ~ 1.5 dex above the trend followed by blind surveys, including ALMACAL, as one might expect. In particular, Pensabene et al. (2024) provides estimates, based on ALMA bands 3 and 6, of a sky region comparable in size to the ASPECS large program (~ 4 arcmin^2 vs. 4.2 arcmin^2). Their detection of over-densities at such high redshifts further highlights the significant impact of cosmic variance on cosmic measurements. This emphasises the critical need for surveys that cover multiple sky regions, rather than the

common practice of focusing on a single contiguous area.

Comparison with simulations

Cosmological hydrodynamical simulations have also predicted the evolution of the molecular gas mass density, $\rho_{\rm H_2}$. These simulations infer the molecular gas mass density by post-processing the outputs of cosmological boxes. First, they simulate the large-scale distribution of galaxies and gas in cosmological boxes. Then, molecular gas content is estimated a posteriori by models of density or temperature that link the simulated gas properties to the presence of molecular gas. We compare our estimates with predictions from simulations in the top panel of Fig. 3.10. We include the predictions from the semi-analytical model `SHARK-1` (Lagos et al. 2018), the cosmological simulation TNG (Popping et al. 2019a), EAGLE (Lagos et al. 2015), COLDSim (Maio et al. 2022; Casavecchia et al. 2024), and UniverseMachine Guo et al. (2023a). Among these, COLDSim stands out by uniquely integrating time-dependent non-equilibrium chemistry with cosmological hydrodynamics, enabling direct modelling of molecular gas. We find a relatively good agreement between our estimate and the predictions by `SHARK`, EAGLE and COLDSim, while TNG and UniverseMachine have higher values.

We followed the model presented in Walter et al. (2020) to fit the molecular gas mass density evolution with redshift, compiling the current estimates. We included all the CO observations compiled in that work together with the new estimates from ALMACAL−22. We performed 500 realisations of the following functional fit proposed by Madau & Dickinson (2014):

$$\rho_{\rm H_2} = \frac{a \cdot (1+z)^b}{1 + [(1+z)/c]^d} \, . \tag{3.6}$$

We find $a = (5.2^{+2.6}_{-1.8}) \times 10^6$, $b = 3.5^{+0.01}_{-0.07}$, $c = 2.2^{+0.2}_{-0.1}$, and $d = 5.9^{+0.4}_{-0.5}$. The bottom panel of Fig. 3.10 shows our results with a solid purple line, and the 1σ uncertainty is shown in the shaded region. The contribution of ALMACAL−22 to the molecular gas mass density estimation indicates lower values at high redshift. If we extrapolate our model to high redshift, we find an excellent agreement with the estimation of $\rho_{\rm H_2}$ at redshift $z \sim 7$ measured by the REBELS survey (Bouwens et al. 2022) from the [C II] emission line reported by Aravena et al. (2024), shown in the dark red point in the bottom panel of Fig. 3.10.

3.5.3 Cosmic variance

We explored the effect of the field-to-field variance in ALMACAL−22 in comparison with the previous surveys. The public simulation, SIDES (Béthermin et al. 2022), provides an estimate of the evolution of variance in the molecular gas mass density as a function of survey size, assuming $J = 2$ and a redshift range $z = 0.5$–6. The estimate of the cosmic variance is affected by the redshift slice, since high-redshift bins are less prone to cosmic variance effects. We compare the effect of cosmic variance only considering the area covered by ALMACAL−22 and previous surveys in the top panel of Fig. 3.11.

Aside from the total area covered, the number of independent sight lines is key to estimating cosmic variance. We adapted Equation 4 from Driver & Robotham (2010) for a conical survey, replacing the transverse lengths A and B by πR as follows:

$$\begin{aligned}\zeta_{\rm CV}(\%) =&[219.7 - 52.4\log_{10}(\pi R^2 \cdot 291.0) \\ &+ 3.21(\log_{10}(\pi R^2 \cdot 291.0))^2]/\sqrt{N \cdot C/291}\,,\end{aligned} \tag{3.7}$$

where R is the transverse length at the median redshift, C is the radial depth, and N is the number of independent sight lines, all in units of $h_{0.7}^{-1}$. We performed this calculation for each CO transition. We considered unique calibrator fields and data cubes where each line could be detected to account for the number of independent sight lines. Table 3.1 provides the estimates of the redshift range for each CO transition, the number of independent sight lines (N_{ind}), and the percentage of cosmic variance.

The bottom panel of Fig. 3.11 compares the cosmic variance from different surveys. We use the values of cosmic variance compiled in Lenkić et al. (2020) (Table 1 of their paper). We see that for all CO transitions, ALMACAL−22 has the lowest cosmic variance seen to date by at least a factor $\sim 5\times$, reaching values lower than 5%. The main reason that ALMACAL−22 has such low values of cosmic variance is that the survey comprises many independent calibrator fields instead of one contiguous area. Previous studies have acknowledged this issue, particularly for the ASPECS survey. Popping et al. (2019b) estimated the H_2 cosmic density in boxes covering the same volume covered in ASPECS and simulated the field-to-field variations. They found that cosmic variance can lead to variations up to a factor of $3\times$. This variance is particularly significant due to the small survey area of ASPECS, 4.6 arcmin^2.

The formula provided by Driver & Robotham (2010) for estimating cosmic variance was

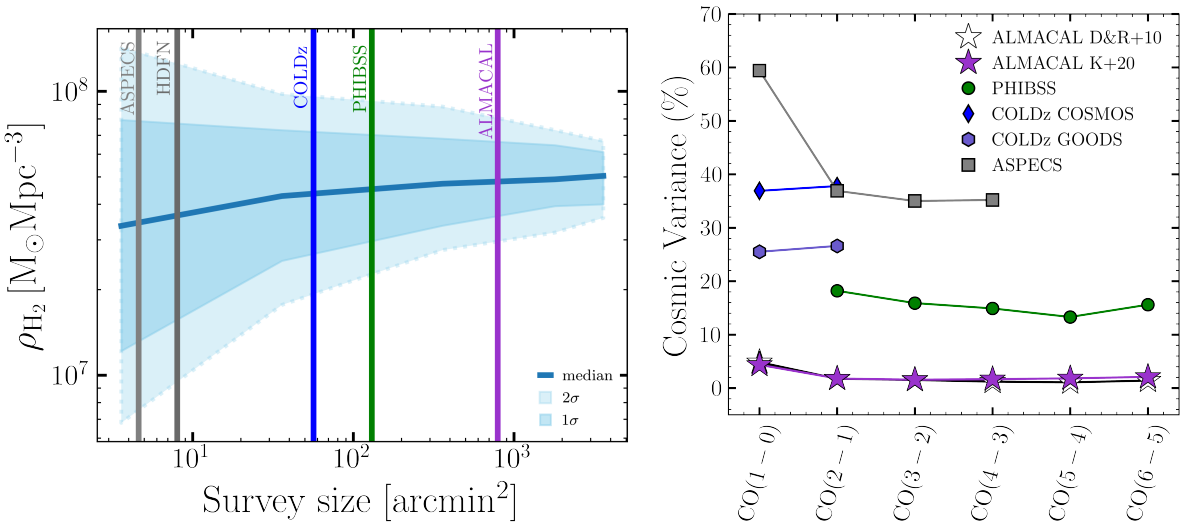

Figure 3.11: *Top*: Effect of cosmic variance based purely on the survey size. We show the results of the molecular gas mass density variation derived from the SIDES simulation (Béthermin et al. 2022; Gkogkou et al. 2022), publicly available for $J = 2$ at redshifts $z = 0.5$–6. The solid blue line indicates the median value of ρ_{H_2}, and the dark and light areas represent the 1σ and 2σ variation due to the cosmic variance. We plot the area covered by ALMACAL−22 along with other CO surveys that have measured the molecular gas mass density. Overall, ALMACAL−22 is the survey less affected by cosmic variance to date. *Bottom*: Cosmic variance estimation of ALMACAL−22 survey in purple, based on Driver & Robotham 2010 (D&R+10), for different CO transitions (Table 3.1). We also plot the variance derived from the formula presented by Keenan et al. (2020) (K+20), which agrees with the previous prescription. The figure shows estimates from ASPECS (Decarli et al. 2019, 2020), PHIBSS (Lenkić et al. 2020), COLDz GOODS and COSMOS (Pavesi et al. 2018; Riechers et al. 2019), previously reported in Lenkić et al. (2020). For every CO transition, ALMACAL−22 has the lowest effect of field-to-field variance ($\lesssim 5\%$).

originally intended for surveys measuring the variance in galaxy counts in fairly uniform galaxy populations. However, it has been noted by Keenan et al. (2020) that it does not take into account the variance in luminosity moments since they use galaxies in a restricted range of magnitude at $z \sim 0$. This issue becomes important for CO line surveys or molecular gas surveys, particularly for biased populations like bright CO emitters. A new method to estimate cosmic variance was proposed by Keenan et al. (2020) using simulated data. This new prescription includes the influence of galaxy number counts and how galaxies are distributed and clustered. We use the formula presented in Appendix A of Keenan et al. (2020) for mean brightness temperature to calculate the fractional uncertainty in the cosmic variance of different area and redshift intervals. We use the mean redshift, the redshift interval and the area covered by each field for a given CO transition to estimate the fractional uncertainty. We then combine the uncertainty of all the fields as independent samples dividing by the square root of the number of independent sight lines ($\sqrt{N_{\rm ind}}$). We obtain values consistent with the estimates derived from the formula proposed by Driver & Robotham (2010), as is shown in the bottom panel of Fig. 3.11.

Gkogkou et al. (2022) used the Uchuu cosmological simulation to analyse the impact of survey size on the variance of the luminosity function. They modelled the errors using both Poisson statistics and clustering contributions. The total variance is decomposed into Poisson and clustering components. They reported the field-to-field variance of simulations that replicate the redshift range and sizes of observational data from ASPECS and COLDz. The field-to-field variance introduced in the molecular gas density ranges from 77% to 81% in an ASPECS-like survey and from 57% to 73% for a COLDz-like survey. They also found a shift between the mean $\rho_{\rm H_2}$ of their model and the one from Keenan et al. (2020), which could be mainly explained by differences in the integration limits, the cosmological simulation used, the scaling relations, and the scatter to assign CO luminosities to galaxies. The predictions from Keenan et al. (2020) may be under-estimated by an order of magnitude according to the predictions from Gkogkou et al. (2022) at redshifts $z \sim 2$–4. However, both approaches agree we need survey sizes of at least ~ 70 arcmin^2 to prove the evolution at redshifts, $z \gtrsim 3$.

Overall, cosmic variance plays a significant role in the measurement of galaxy number density and LF, with uncertainties reaching as high as 70% for the smallest fields and decreasing to about 25% for the largest ones probed until now. Accounting for cosmic variance is essential to improve the accuracy of observational constraints on the cosmic

star formation rate density and to deepen our understanding of the large-scale structure of the Universe. Figure 3.6 shows that the ALMACAL results are consistent with previous surveys; however, the uncertainties due to cosmic variance are not displayed. In Figure 3.11, we compare the impact of cosmic variance and we observe that this issue dominates other surveys. In contrast, due to its survey strategy, ALMACAL has a minimum value. To optimise survey strategies, it is beneficial to utilise multiple small fields instead of a single contiguous area, as this approach effectively mitigates cosmic variance. Future surveys are encouraged to consider the impact of cosmic variance when reporting uncertainties, moving beyond the traditional focus on Poisson uncertainties. The literature has shown that relying solely on Poisson uncertainties fails to capture the true levels of uncertainty (Popping et al. 2019b; Keenan et al. 2020; Gkogkou et al. 2022).

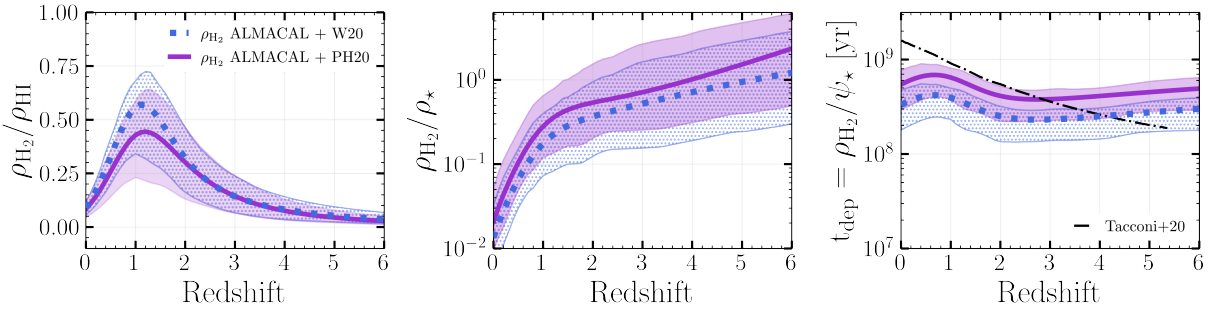

Figure 3.12: Redshift evolution of baryonic component in the Universe. *Top*: Ratio of cosmic molecular-to-atomic gas density as a function of redshift. *Middle*: Ratio of molecular gas-to-stellar mass density as a function of redshift. *Bottom:* Cosmic gas depletion timescale, is defined as the density of molecular gas over the cosmic star formation rate density. We include the molecular gas depletion times commonly derived for individual galaxies as the ratio M_{H_2}/SFR, as is reported in Tacconi et al. (2020). We use the fit of $\rho_\star$ presented by Madau & Dickinson (2014). We used the values of ρ_{HI} and $\psi_\star$ fitted in the compilation done by Péroux & Howk (2020) in dotted blue and Walter et al. (2020) in purple. For the star formation rate density and stellar mass function, the former assumes a Chabrier IMF (Chabrier 2003) and a return fraction $R = 0.41$, while the latter assumes a Salpeter IMF (Salpeter 1955) and $R = 0.27$. The ρ_{H_2} corresponds to the fit presented in § 3.5.2, which includes the results from previous surveys that have measured CO along with ALMACAL.

3.5.4 Census of the baryon cycle

In this section, we use the new $\rho_{\rm H_2}$ estimates provided by ALMACAL to revisit the cosmic baryon cycle. Specifically, we investigate how the cosmic baryon density has evolved over time, focusing on the contributions from stars, atomic hydrogen (H I), and molecular gas (H_2).

Madau & Dickinson (2014) compiled data of two key observables: the cosmic star formation rate density, $\psi_\star$, and the stellar mass density, $\rho_\star$. Péroux & Howk (2020) used a Chabrier IMF (Chabrier 2003) in their review, and a return fraction of $R = 0.41$, which is the fraction of the stellar mass that is immediately returned to the gas when massive stars explode, whereas Walter et al. (2020) assumed a Salpeter IMF (Salpeter 1955) and $R = 0.27$. For the atomic gas mass density, Péroux & Howk (2020) modelled the evolution of neutral gas with redshift using a two-parameter power law, which uses fewer free parameters. On the other hand, Walter et al. (2020) adopted a hyperbolic tangent function, originally suggested by Prochaska & Neeleman (2018), to fit the atomic gas. In summary, we adopt the best-fit models for $\psi_\star(z)$, $\rho_\star(z)$ and $\rho_{\rm HI}(z)$ from both Péroux & Howk (2020) and Walter et al. (2020). For $\rho_{\rm H_2}(z)$, we use the updated fit described in §3.5.2, which includes estimates from ALMACAL−22.

We analysed the gas density ratios of these three components (stars, atomic gas and molecular gas) in Fig. 3.12. In each panel, the solid purple lines indicate the ratios derived using $\rho_{\rm H_2}(z)$ from this work, and the fits for $\rho_{\rm HI}$, $\psi_\star$ and $\rho_\star$ provided in Walter et al. (2020). The dotted blue lines show the ratios using $\rho_{\rm H_2}(z)$ including ALMACAL−22 and the fits for $\psi_\star$, $\rho_\star$ and $\rho_{\rm HI}$ presented in Péroux & Howk (2020). For both models, the 1σ uncertainty is shown by the shaded regions.

The top panel shows the molecular-to-atomic gas density ratio as a function of redshift ($\rho_{\rm H_2}/\rho_{\rm HI}$). It peaks at around $z \sim 1.5$, indicating a period of high star formation activity, characterised by the conversion of H I to H_2 being particularly efficient. This efficiency is slightly lower than the one predicted by Walter et al. (2020) but still consistent with their predicted scenario of cosmic twilight, where the star formation activity in galaxies declines, likely due to the shutdown of the inflow and the accretion of gas.

The middle panel shows the ratio of molecular gas-to-stellar mass density ($\rho_{\rm H_2}/\rho_\star$) as a function of redshift. This quantity evolves significantly from $z \sim 1.5$ down to the local universe ($z = 0$), indicating that the molecular gas for star formation becomes more limited

relative to the existing stellar mass. At high redshift ($z > 1.5$), our fits indicate a lower molecular gas, which means a lower capacity for star formation compared to that observed by Walter et al. (2020). The Chabrier IMF predicts about 30% lower total stellar mass for a given population than the Salpeter IMF, since fewer low-mass stars are included. This would suggest that galaxies still have significant molecular gas reservoirs available for star formation, particularly at high redshifts, which may not be true at low redshifts.

The bottom panel shows the cosmic gas-depletion timescale, defined as the density of molecular gas over the cosmic star formation rate density ($\rho_{H_2}/\psi_\star$). The dot-dashed line shows the functional fit to the molecular depletion timescale based on M_{H_2}/SFR ratio for a sample of massive galaxies (Tacconi et al. 2020), which roughly agrees with the global estimates from Péroux & Howk (2020). At redshifts above $z \sim 2$, we find a relatively constant depletion timescale, which indicates that molecular gas is consumed by star formation steadily during this period. The depletion timescale increases from $z \sim 2$ to $z = 0$, so the consumption rate is slowing down relative to the available gas, leading to the quenching of star formation. A Salpeter IMF would imply that galaxies deplete their gas more quickly, while a higher ratio (Chabrier IMF) would suggest galaxies have longer depletion timescales, consistent with galaxies at high redshift (Scoville et al. 2017). This trend is consistent with the findings from Péroux & Howk (2020) and their claim regarding the need to continuously replenish molecular gas to sustain star formation.

These findings align with predictions from the gas regulator model, also known as the 'bathtub model', as is described in studies like Bouché et al. (2010); Davé et al. (2011); Lilly et al. (2013); Peng & Maiolino (2014). This model uses continuity equations to track the flow of baryons into and out of galaxies and picture them as systems in a dynamic equilibrium between gas inflow, outflows and star formation. As has previously been noted by Péroux & Howk (2020), this model provides an effective framework for understanding how the continuous cycling of baryons drives galaxy evolution. In the early stages, gas accumulates, and star formation is primarily constrained by the available cold gas reservoir. As galaxies evolve, they reach a steady state where the net gas accretion rate regulates the overall star formation rate.

3.6 Conclusions

In this paper, we have presented the ALMACAL-CO project, built upon the experience of the previous pilot survey presented by Hamanowicz et al. (2022). We searched for serendipitous detections of CO emission lines in the fields of ALMA calibrators. We used the highest available quality of data, selected as part of the ALMACAL−22 data release presented in Bollo et al. (2024). We performed a search for CO emitters using the source-finding algorithm, SOFIA-2 (Westmeier et al. 2021). We detected 87 CO emitters over 299 calibrator fields. We confirmed the redshift of three sources with optical spectroscopic counterparts and estimated the photometric redshift of three sources with optical and near-IR counterparts. We determined redshift probabilities for the remaining 81 candidates based on the semi-analytical model, `SHARK-2` (Lagos et al. 2023). We built the CO LF through realisations that sample the properties of the CO emitter, such as their completeness, fidelity and redshift probability. We constrained the CO LF and calculated the molecular gas mass density up to redshift, $z \sim 6$. We summarise our main findings as follows:

1. We have probed the bright end of the CO LF, from $z \sim 0$ to $z \sim 6$. We find good agreement with existing observations, but with slightly higher values of the number density of sources in the low redshift bins ($z \lesssim 1$) and slightly lower values in the higher redshift bins ($z \gtrsim 1$). Simulations of the CO LF (Lagos et al. 2018; Bisigello et al. 2022, `SHARK, SPRITZ`) agree with our findings, but we stress the need to expand their prediction to the bright end (Figs 3.7 and 3.9).

2. We found an evolution of the Schechter parameters of the CO LF with redshift. The normalisation factor, $\Phi^*_{\rm CO}$, decreases with redshift. The characteristic luminosity, defined as the knee of the luminosity function, shows an increasing trend from the local Universe to $z \sim 2$ (Fig. 3.8).

3. The molecular gas mass density reported by ALMACAL−22 shows higher values than previous observations in the lower redshift bins ($z \lesssim 1$). It presents a slight tension at higher redshift ($z \geq 1$), where ALMACAL−22 finds lower values (Fig. 3.6). We found a good agreement with the simulations predicting $\rho_{\rm H_2}$ from `SHARK`, EAGLE and COLDSim (Fig. 3.10, top panel).

4. We present a new analytical molecular gas mass density fit, compiling previous estimates based on CO surveys and the new ALMACAL−22 results. Irrespective of

what CO line ρ_{H_2} is based on, the effect of cosmic variance is less than 5% (Fig. 3.10, bottom panel).

5. ALMACAL−22 is the survey that best addresses the uncertainty introduced by cosmic variance at the brighter end (Fig. 3.11).

6. We re-visit the redshift evolution of the baryon components, updating the molecular gas mass density with the values provided by ALMACAL. We use previous fits of the atomic gas mass density, stellar mass function and the cosmic star formation rate density from Péroux & Howk (2020) and Walter et al. (2020). Overall, we report consistency with their claims regarding the evolution of the density ratios, though with smaller molecular gas content (Fig. 3.12).

Our findings serve as a valuable stepping stone for future work that conducts surveys less prone to cosmic variance. We emphasise the need for additional telescope time to confirm the redshift of our sources, particularly those that may be obscured at optical wavelengths, where sub-millimetre observations would be crucial.

Probing the molecular gas density across different environments is a collective effort by the scientific community. The more data we gather, the clearer the picture becomes. Each survey brings unique strengths and limitations, and to fully understand the evolution of molecular gas in the Universe, we also need simulations with diverse prescriptions. Together, these approaches will help unravel the complexities of molecular gas evolution over cosmic time.

Chapter 4

MUSE-ALMA Haloes XIII. Molecular gas in $z \sim 0.5$ H I–selected galaxies

The content of this chaper was submitted by Bollo et. al to Astronomy & Astrophysics on July 21st, 2025.

4.1 Introduction

The evolution of galaxies is fundamentally regulated by the availability and cycling of gas, which serves as raw material for star formation. Cold gas plays a crucial role in this process, as it forms stars that enrich the interstellar medium (ISM) with metals and dust over time (Kennicutt & Evans 2012; Carilli & Walter 2013; Tacconi et al. 2020). Feedback from supernovae (SNe) and active galactic nuclei (AGN) injects mechanical and radiative energy into the surrounding environment, influencing the efficiency of star formation and driving large–scale outflows (Oppenheimer & Davé 2008; Lilly et al. 2013; Faucher-Giguère & Oh 2023). These processes collectively shape the baryon cycle, which describes how gas is accreted, converted into stars, and ejected into the circumgalactic medium (CGM), with the potential for subsequent re-accretion (Tumlinson et al. 2017; Péroux & Howk 2020; Peroux & Nelson 2024).

A complete understanding of the baryon cycle requires studying the CGM, which acts as a reservoir of gas regulating galaxy growth (Werk et al. 2014; Nelson et al. 2020; Naab

& Ostriker 2017; Oppenheimer et al. 2021). The CGM has been extensively probed using a multi-wavelength approach. Observations have revealed that the CGM is multiphase, comprising gas at a range of temperatures, densities, and ionisation states. This includes cold neutral gas, warm ionised gas, and hot, X-ray emitting plasma. Optical and ultraviolet absorption-line spectroscopy, particularly with the Cosmic Origins Spectrograph on the Hubble Space Telescope (HST) (e.g., Muzahid et al. 2015; Richter et al. 2017) has been crucial in studying ionised gas phases. X-ray observations with Chandra, XMM-Newton, and eROSITA trace the hot component of the CGM (e.g., Strickland et al. 2004; Li 2013; Bogdán et al. 2013; Merloni et al. 2012; Zhang et al. 2025; Lin et al. 2025). Atomic hydrogen (H I) has been investigated both in emission and absorption across cosmic time. Emission observations using the 21 cm line have been key to studying the star formation efficiency, gas kinematics, and disk structure in nearby galaxies (e.g. Haynes & Giovanelli 1984; Zwaan et al. 2005; Bigiel et al. 2008; Leroy et al. 2008; Walter et al. 2008; Catinella et al. 2018; Yu et al. 2022; Sharma et al. 2023; Wang et al. 2023, 2025). However, direct detection of 21 cm emission becomes increasingly difficult beyond $z \sim 0.1$, and even harder for individual galaxies at $z > 0.4$ (Fernández et al. 2016; Xi et al. 2024). To probe higher redshifts, statistical techniques such as spectral stacking and intensity mapping have been adopted (Masui et al. 2013; Rhee et al. 2018; Bera et al. 2019; Chowdhury et al. 2020), and efforts are ongoing supported by SKA precursors and pathfinders (e.g., Blyth et al. 2016; Adams & van Leeuwen 2019; Koribalski et al. 2024; Maddox et al. 2021). Given the emission limitations, 21 cm absorption studies provide a powerful alternative, as sensitivity is independent of the galaxy's luminosity distance. FAST has recently enabled discoveries of new low-redshift absorbers (Su et al. 2023; Yu et al. 2024), while the Australian Square Kilometre Array Pathfinder (ASKAP) and MeerKAT have begun probing intermediate redshifts through surveys like The First Large Absorption Survey in H I (FLASH) and the MeerKAT Absorption Line Survey (MALS) (e.g., Allison et al. 2020; Hotan et al. 2021; Glowacki et al. 2022; Deka et al. 2023; Yoon et al. 2024).

The H I absorption selection method, in particular, has proved to be a powerful tool for detecting neutral gas in the CGM (Wolfire et al. 1995; Wolfe et al. 2005). By observing the Lyα absorption lines against bright background sources, this method enables the identification of diffuse, cold gas that may be difficult to detect in emission. It is highly sensitive to low column densities, independent of redshift, luminosity, or star formation, and provides detailed kinematic information about the gas distribution (e.g., Péroux et al.

2006; Zwaan et al. 2005; Bouché et al. 2010, 2016; Frank et al. 2012; Kacprzak et al. 2011; Stern et al. 2016).

The use of 3D integral field spectroscopy (IFS) has revolutionised our ability to study the connection between galaxies and the CGM, especially in systems selected via H I absorption, such as damped Lyman-α (DLA, $\log N(\mathrm{H\,I})/\mathrm{cm}^2 \gtrsim 20.3$) and sub-damped Lyman-$\alpha$ systems (sub-DLA; $19.0 \lesssim \log N(\mathrm{H\,I})/\mathrm{cm}^2 < 20.3$, Péroux et al. 2003). These absorption-selected systems probe some of the most gas-rich environments in the Universe and have long been used to study gas metallicities (e.g., Rafelski et al. 2012; Kulkarni et al. 2005, 2015), kinematics (e.g., Prochaska & Wolfe 1998), gas temperatures (e.g., Kanekar et al. 2014), and chemical enrichment histories (e.g., Dessauges-Zavadsky et al. 2007; Quiret et al. 2016).

Multiple optical surveys have used the IFS VLT/MUSE (Bacon et al. 2010), such as MUSE-QuBES (e.g., Muzahid et al. 2020; Dutta et al. 2024, 2025), and MEGAFLOW (e.g., Bouché et al. 2016; Schroetter et al. 2016; Langan et al. 2023) that established statistical links between CGM absorbers and their associated galaxies. Klimenko et al. (2023) report results from an HST-COS survey with the opposite approach of determining the CGM properties of galaxies observed with IFS at $z < 0.1$ in the MaNGA survey. Parallel programs including MAGG (e.g., Dutta et al. 2020; Lofthouse et al. 2020; Fossati et al. 2021), the Cosmic Ultraviolet Baryon Survey (CUBS) (e.g., Boettcher et al. 2021; Cooper et al. 2021; Zahedy et al. 2021), and studies with the Keck/KCWI optical IFS (e.g., Martin et al. 2019; Nielsen et al. 2020) have extended these insights to higher redshifts, broader dynamic ranges, and higher spectral resolution.

While significant progress has been made in studying the ionised and neutral atomic components of the CGM, the molecular gas content remains less explored. Molecular gas is a critical ingredient for star formation, and is closely connected with gas accretion and star formation (Tacconi et al. 2020; Walter et al. 2020; Saintonge & Catinella 2022). Observations with the Atacama Large Millimetre/submillimeter Array (ALMA), Northern Extended Millimeter Array (NOEMA), and Jansky Very Large Array (JVLA) have opened new avenues to detect molecular gas in and around galaxies. Previous works have hinted at long gas depletion timescales and suppressed star formation efficiencies in some H I–selected systems, suggesting a distinct phase of galaxy evolution not well represented in emission–selected surveys (e.g., Kanekar et al. 2018, 2020; Neeleman et al. 2017, 2018, 2019; Klitsch et al. 2019b; Péroux et al. 2019; Szakacs et al. 2021; Kaur et al. 2021, 2024).

However, a systematic study of the molecular gas content of H I–selected galaxies with multiwavelength data is still lacking.

The MUSE-ALMA Haloes survey was designed to probe the multiphase nature of the circumgalactic medium by targeting 32 H I Lyα absorbers with column densities in the range $\log N(\mathrm{H\,I})/\mathrm{cm}^2 = 18.1 - 21.7$ at redshifts $0.2 < z < 1.4$. It combines new observations from the Multi Unit Spectroscopic Explorer (MUSE) on the Very Large Telescope (VLT), ALMA, and HST. The scientific development of the survey builds on previous work and first results based on smaller subsamples of H I absorbers. These early studies, using precursor observations, investigated the environments of absorbers, revealing intragroup associations and multi-galaxy systems (Péroux et al. 2017, 2019), evidence for cold accretion (Rahmani et al. 2018a), outflows (Rahmani et al. 2018b), and extended molecular gas beyond the ionised disks (Klitsch et al. 2018). Kinematic alignment between molecular and ionised gas phases was explored by Szakacs et al. (2021), while Hamanowicz et al. (2020) reported absorber–galaxy correlations based on 14 systems. The full statistical analysis of the MUSE data was presented by Weng et al. (2022), who identified 79 galaxies within ±500 km/s of the absorbers in 19 MUSE fields, each covering a 1×1 arcmin^2 field of view, providing a first comprehensive view of CGM galaxy environments in this sample. These 79 galaxies form the basis of the MUSE-ALMA Haloes survey sample (Péroux et al. 2022), on which subsequent studies have built. Karki et al. (2023) found that these absorption-selected galaxies lie on the star-forming main sequence with a 2σ scatter, and that higher H I column densities tend to be associated with more compact galaxies. They also showed that both emission and absorption metallicities correlate with stellar mass and specific SFR, suggesting that metal-poor absorbers trace galaxies with lower past star formation and more rapid current gas consumption. Weng et al. (2023a) applied 3D forward modelling of the ionised gas kinematics to explore the physical origins of the absorbers. In parallel, Weng et al. (2024) used the TNG50 simulation to track H I absorbers around galaxies at $z = 0.5$, finding that many absorbers at large impact parameters or low column densities arise from satellites, neighbouring haloes, or the IGM. Augustin et al. (2024) examined the stellar mass distribution of the sample, reporting an anti-correlation between stellar mass and H I column density, suggesting that more massive galaxies are surrounded by less neutral hydrogen. Finally, Karki et al. (submitted) investigated 20 interacting galaxies associated with quasar absorbers, showing that tidal features trace galaxy-CGM interactions, revealing enhanced star formation, distinct kinematics, and extended CGM

properties compared to non-interacting systems.

This paper presents the first results of an ALMA Large Program for MUSE-ALMA Haloes survey (Cycle 10, PI: C. Péroux), targeting galaxies selected as H I absorbers. The science goals of this survey include identifying the role of the molecular gas in H I-rich galaxies, comparing with the canonical scaling relations, and exploring their connection to other galaxy properties. By comparing their depletion timescales, mass-metallicity relations, and physical diversity with other galaxy populations, we aim to better understand the physical conditions that govern the efficiency of gas conversion into stars.

This paper is organised as follows. In Section 4.2, we describe the MUSE-ALMA Haloes ALMA Large program and summarise the science goals achieved with this multiwavelength dataset. Section 4.3 details our analysis and results. Section 4.4 presents the properties of H I–selected systems and compares our findings with existing literature. Section 4.5 discusses the nature of H I–selected galaxies. In Section 4.6, we summarise our key conclusions. Throughout this paper, we use $H_0 = 70\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_\mathrm{M} = 0.3$ and $\Omega_\Lambda = 0.7$. The present-day cosmological critical density is $\rho_{0,\mathrm{crit}} \simeq 277.4\ h^2\ \mathrm{M_\odot\ kpc^{-3}}$. We adopt $\log Z = 12 + \log(\mathrm{O/H})_\mathrm{PP04}$ following Pettini & Pagel (2004) and refer to it simply as $12 + \log(\mathrm{O/H})$.

4.2 Multi-wavelength data

The analysis in this paper builds upon the extensive dataset assembled by the MUSE-ALMA Haloes survey (Péroux et al. 2022). In this section, we provide a brief overview of the ancillary datasets from VLT/UVES, VLT/MUSE, HST, and ALMA, and introduce the new ALMA observations obtained as part of a Cycle 10 Large Program (PI: C. Péroux, ID: 2023.1.00127.L). This rich dataset enables a comprehensive characterisation of the stellar, ionised, atomic, and molecular gas properties of absorption-selected galaxies, offering new insights into the baryon cycle and the physical processes governing galaxy evolution.

The MUSE-ALMA Haloes survey has measured absorber redshifts, H I column densities from Voigt profile fitting of high-resolution quasar spectra, and metal abundances of the absorbing gas. Within the MUSE field of view, a total of 3658 extra-galactic sources were found across all the fields, of which 79 galaxies were identified within ±500 km/s of the absorbers and > 100 kpc in projected separation Weng et al. (2023b). From galaxy morphologies, position angles and inclinations were used to calculate azimuthal angles

between galaxies and background quasars (Hamanowicz et al. 2020; Péroux et al. 2022; Weng et al. 2023a).

4.2.1 Ancillary optical observations

VLT observations

Observations from the Very Large Telescope (VLT) Multi-Unit Spectroscopic Explorer (MUSE) in 19 quasar fields allowed to identify 79 galaxies associated with strong H I absorbers at intermediate redshifts ($z \sim 0.5$) (Weng et al. 2022) (PIs: Péroux; ESO programs 96.A-0303, 100.A-0753, 101.A-0660, 102.A-0370, and Klitsch 298.A-0517).

Spatially resolved MUSE spectroscopy provided rest-frame optical emission lines ([O II], Hβ, [O III], Hα, [N II]), yielding spectroscopic redshifts, star formation rates, and gas-phase metallicities via strong-line diagnostics calibrated by Curti et al. (2017) and Maiolino et al. (2008).

Complementary new high-resolution quasar spectroscopy from VLT/UVES (PI: Péroux, ID: 113.A-0369), along with archival Keck/HIRES and VLT/X-Shooter data, provides detailed measurements of metal absorption lines (e.g., Fe II, Si II, Zn II, C IV). Column density estimates offer dust and ionisation corrections, and CGM metallicity measurements in neutral gas absorption from Voigt profile fitting (Weng et al. 2023a; Karki et al. 2023; Augustin et al. 2024; Halley et al., in prep).

HST observations

The HST ancillary data comes primarily from broad-band imaging observed with the Wide Field Camera 3 (WFC3) during Cycle 27 under GO Program ID 15939 (PI: Péroux) using both optical (UVIS) and infrared (IR) detectors, as well as archival data from the Wide Field and Planetary Camera-2 (WFPC2) and WFC3 from programs 5098, 5143, 5351, 6557, 7329, 7451, 9173, and 14 594 (PIs: Burbidge, Macchetto, Bergeron, Steidel, Malkan, Smette, Bechtold, Bielby). For a comprehensive description of the observational setup, data reduction, PSF subtraction, photometric measurements, and the master table of all targets, see Péroux et al. (2022).

The high spatial resolution of HST imaging (resolution ∼0.04”) provided a crucial complement to MUSE observations, enabling the identification of structural features such as tidal tails and galaxy morphologies (Karki et al. 2023).

Augustin et al. (2024) performed Spectral Energy Distribution (SED) fitting on the multi-band HST photometry using the Le Phare code (Arnouts et al. 1999; Ilbert et al. 2006). The fitting assumed Bruzual & Charlot (2003) SED templates, a Calzetti et al. (1994) dust extinction law, and a Chabrier (2003) initial mass function. The resulting stellar mass measurements span a wide range from $\log(M_\star/M_\odot) = 7.8 - 12.4$, highlighting the diversity of galaxies associated with H I absorbers. The most robust mass determinations (34 galaxies) had at least two detections in different HST filters, with higher-mass galaxies ($\log(M_\star/M_\odot) > 10$) typically having more reliable measurements due to their stronger detections across multiple bands.

4.2.2 ALMA observations

Previous ALMA data

In addition to the rich optical and UV dataset, we compile sub-millimetre observations from archival data. Kanekar et al. (2018) observed one absorber associated with one galaxy included in the MUSE-ALMA Haloes survey in the quasar field J0138-0005. The ALMA observations in Band 4 (program IDs: 2013.1.01178.S, 2015.1.00561.S, PIs: J. Prochaska, N. Kanekar) targeted the CO(2–1) line. They performed data reduction in the Common Astronomy Software Applications (CASA, McMullin et al. 2007), using `tclean` with a natural weighting, producing cubes with a spectral resolution of 100 km s^{-1} and a rms noise of $\sim$ 0.20 mJy beam^{-1} in the data cubes. They placed an upper limit on the CO luminosity for the galaxy associated with the absorber.

Klitsch et al. (2018) used ALMA calibrator data from the ALMACAL survey (Zwaan et al. 2022; Bollo et al. 2024) for quasar J0423-0130, linked to four galaxies from the MUSE-ALMA Haloes survey. They analysed CO(2-1) and CO(3-2) lines from Band 4 and Band 6 data (1333 s and 605 s total), reduced with CASA (Briggs weighting), achieving velocity resolutions of 33 and 22 km s^{-1}, and the rms noise levels in the data cubes are $\sim$0.28 and $\sim$0.37 mJy beam^{-1}, respectively. One detection was reported, with two upper limits and one source outside the ALMA field of view.

Four quasar fields (Q1232-0224, Q0152-2001, Q1211+1030 and Q1130-1449) were observed with ALMA to cover CO(2–1) or CO(3–2) under the programs 2016.1.01250.S and 2017.1.00571.S (PI: C. Péroux) and 2018.1.01575.S (PI: A. Klitsch), presented in Péroux et al. (2019) and Szakacs et al. (2021). The programs targeted seven galaxies included

in the MUSE-ALMA Haloes survey, which had been identified by MUSE and associated with absorbers at $z \sim 0.4$ with impact parameters ranging from 8 to 82 kpc. The primary spectral window was centred on the redshifted CO(3−2) frequency (345.796 GHz) in high spectral resolution mode, complemented by three additional spectral windows in low resolution mode (31.250 MHz). They performed data reduction in CASA, using `tclean` with a Briggs weighting scheme, reaching a spectral resolution of 50 km s^{-1} per channel and an RMS sensitivity of 0.28 mJy beam^{-1}. They reported two detections out of the seven systems that are part of the MUSE-ALMA Haloes survey.

In total, we compiled ALMA archival data for eleven galaxies from the MUSE–ALMA Haloes Survey, all of which were observed in the CO(2−1) or CO(3−2) transitions. By design of the MUSE-ALMA Haloes survey, they have MUSE and HST data, as mentioned earlier in the text.

MUSE-ALMA Haloes ALMA Large Program

The ALMA data presented in this paper were obtained from the ALMA Large Program (2023.1.00127.L, Cycle 10, PI: C. Péroux). The observations were conducted between December 2023 and March 2024 using ALMA configurations C43-2, C43-3, and C43-4, which provided angular resolutions ranging from 0.6″ to 1.1″ at the observed frequencies.

This program targeted CO emission lines (specifically CO(2−1), CO(3−2), CO(4−3) transitions) using Band 4 ($125 - 163$ GHz) and Band 6 ($211 - 275$ GHz), reaching a rms of 0.16 mJy beam^{-1} over 50 km s^{-1}. Our sample consists of 39 H I-rich galaxies associated with 17 quasar absorbers with known H I column densities ($\log N(\mathrm{H\,I}) > 10^{18}$ cm^{-2}) at redshifts $0.3 \lesssim z \lesssim 1.2$ (Péroux et al. 2022).

The data were calibrated and imaged using the standard ALMA pipeline in CASA (version 6.5.6), with phase and/or amplitude self-calibration applied to five fields where the background quasar was bright enough (> 5 mJy) to achieve a signal-to-noise ratio of five. The images were built using `tclean` with pixel sizes adjusted according to the synthesised beam and a natural weighting scheme. Following these calibrations, we also performed continuum subtraction using `uvcontsub` with a polynomial order of 2, and the final data cubes were produced and corrected for primary beam effects using `impbcor`. Final data cubes were created with ~ 60 km/s spectral binning for optimal line sensitivity. Full details of the observations, calibration procedures, and data quality assessments are presented in a separate paper (Péroux et al., in prep.). All subsequent analyses are based

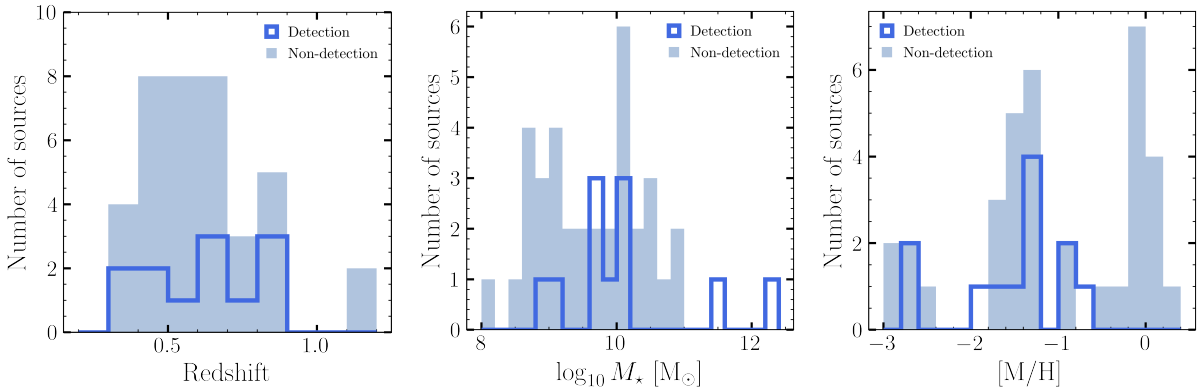

Figure 4.1: Redshift (left), stellar mass (middle) and absorption neutral gas metallicity (right) distribution of the CO-detections (blue) and non-CO-detections (light blue) sources from the MUSE-ALMA Haloes survey, including all the sources from the new ALMA Large Program and previous ALMA data (see Section 4.2.2). The middle and right panels have fewer numbers because only those with reliable $M_\star$ and [M/H] determinations are included. The median values for the detections are 0.6, 9.75 and -1.23 for the redshift, stellar mass and absorption metallicity, while the non-detections have median values of 0.5, 9.43, and -1.20. Overall, we report no clear correlation between the detection rate and any of these properties, suggesting that the detection of CO emission may be governed by a combination of global galaxy properties, rather than just one property alone.

on data from the complete MUSE-ALMA Haloes survey.

4.3 Analysis and Results

4.3.1 Emission line search identification

We systematically searched for CO emission lines within the calibrated and cleaned ALMA data cubes. The search used two complementary approaches to maximise completeness while maintaining reliability.

First, we conducted a targeted search at the expected positions of 39 H I–selected galaxies previously identified in Weng et al. (2022). For each galaxy, we applied a systematic aperture optimisation procedure to maximise the likelihood of detecting CO emission. We

used elliptical apertures with semi-major and semi-minor axes ranging from 1 to 4 times the synthesised beam size. For each aperture size, we rotated the ellipse from 0° to 90° in steps of 5°, systematically exploring different orientations. At each combination of size and angle, we extracted spectra and computed the signal-to-noise ratio (SNR), estimated from the integrated line flux over the beam size, to identify the configuration that maximised the signal significance. This method allowed us to capture the full extent of potentially resolved CO emissions. We defined emission line candidates where the signal exceeded 3σ spanning at least ~ 50 km/s, where σ represents the RMS noise level measured in line-free channels, which is the minimum expected line width based on detection for the previous ALMA data (§4.2.2). The aperture optimisation technique allowed us to successfully detect CO emission lines in nine sources from our targeted sample.

In parallel, we performed a blind search across all the fields using the Source Finding Application SOFIA-2 (Serra et al. 2015; Westmeier et al. 2021) that uses a number of different source detection algorithms to find emission lines in radio data cubes. In this process, we used the data cubes, not corrected from the primary beam, to model the noise level homogeneously across the field of view. We use the `Smooth + Clip Finder`, which smooths the data in both spatial and spectral directions using several 3D smoothing kernels. Our setup consisted of a detection threshold of 3σ, a minimum requirement of emission spanning one channel, and a reliability estimated by SOFIA-2 of 0.9, which compares the distribution of positive and negative sources in parameter space, defined by signal-to-noise ratio and line width. We found six sources that matched the positions of the primary targets, which were also detected using the methodology explained above. The remaining three detections, which were not identified by SOFIA-2 compared to the targeted search, showed a lower signal-to-noise ratio (SNR < 5) in regions of reduced sensitivity, such as the edges of the primary beam. Additionally, this approach allowed us to identify potential CO line emitters that may not have been previously detected in optical observations. These serendipitous detections will be discussed in subsequent papers.

In total, we detected CO emission from nine galaxies at a significance of at least 3σ out of our targeted sample of 39. For the remaining 30 sources, we calculated 3σ upper limits based on the sensitivity reached in the data cube and assuming a line width of 200 km/s, typical for galaxies in our sample. These are stringent limits due to the depth of our survey.

When combined with previous CO observations from the MUSE-ALMA Haloes survey

(Kanekar et al. 2018; Klitsch et al. 2018; Szakacs et al. 2021), which added three detections and eight upper limits, our comprehensive analysis provides a total of 12 CO detections out of 50 galaxies at $0.3 < z < 1.2$, as shown in the left panel of Figure 4.1. Before the MUSE-ALMA Haloes survey, only a handful of H I at $z \sim 0.5$ had been targeted in CO (e.g., Kanekar et al. 2018). Our observations increased the number of targeted galaxies by a factor of 5 and nearly doubled the number of CO detections ever previously known. Figure 4.2 shows the resulting moment map and extracted spectrum for each detected source.

4.3.2 Non-detections stacking

We performed spectral cube stacking for the 30 sources with no individual CO detection. For each galaxy, we extracted a subcube centred on the expected position of the galaxy, using a spatial size of 50 kpc at its redshift. The spectral axis spanned 1000 km s^{-1}, centred on the expected frequency of the CO line. All subcubes were then rebinned to a common spatial and spectral resolution, adopting the average beam size and spectral bin width across the sample.

The cubes were stacked, averaging each subcube weighted by the square of the luminosity distance to account for flux dimming with redshift. This weighting helps recover a quantity closer to the intrinsic luminosity by correcting for the cosmological decrease in observed flux with distance. We also applied a weight given by the sensitivity of each data cube. We examined the resulting stacked cube for potential CO emission at the central position.

To extract the spectrum, we used the same methodology presented in §4.3.1, optimising the aperture size to maximise the signal-to-noise ratio. The same aperture was used to estimate the significance of the emission line in the stacked cube. To quantify the noise, we extracted spectra from 100 random apertures placed across the stacked cube, excluding the central region. These apertures matched the size determined during optimisation. The per-channel error was then computed as the standard deviation across all 100 spectra. The resulting moment map and spectrum of the stacked cube are presented in the bottom right panel of Figure 4.2. The spectrum exceeds 3σ in two consecutive channels. We report a stacked CO emission line detected at a significance level of $\sim 4.2\,\sigma$.

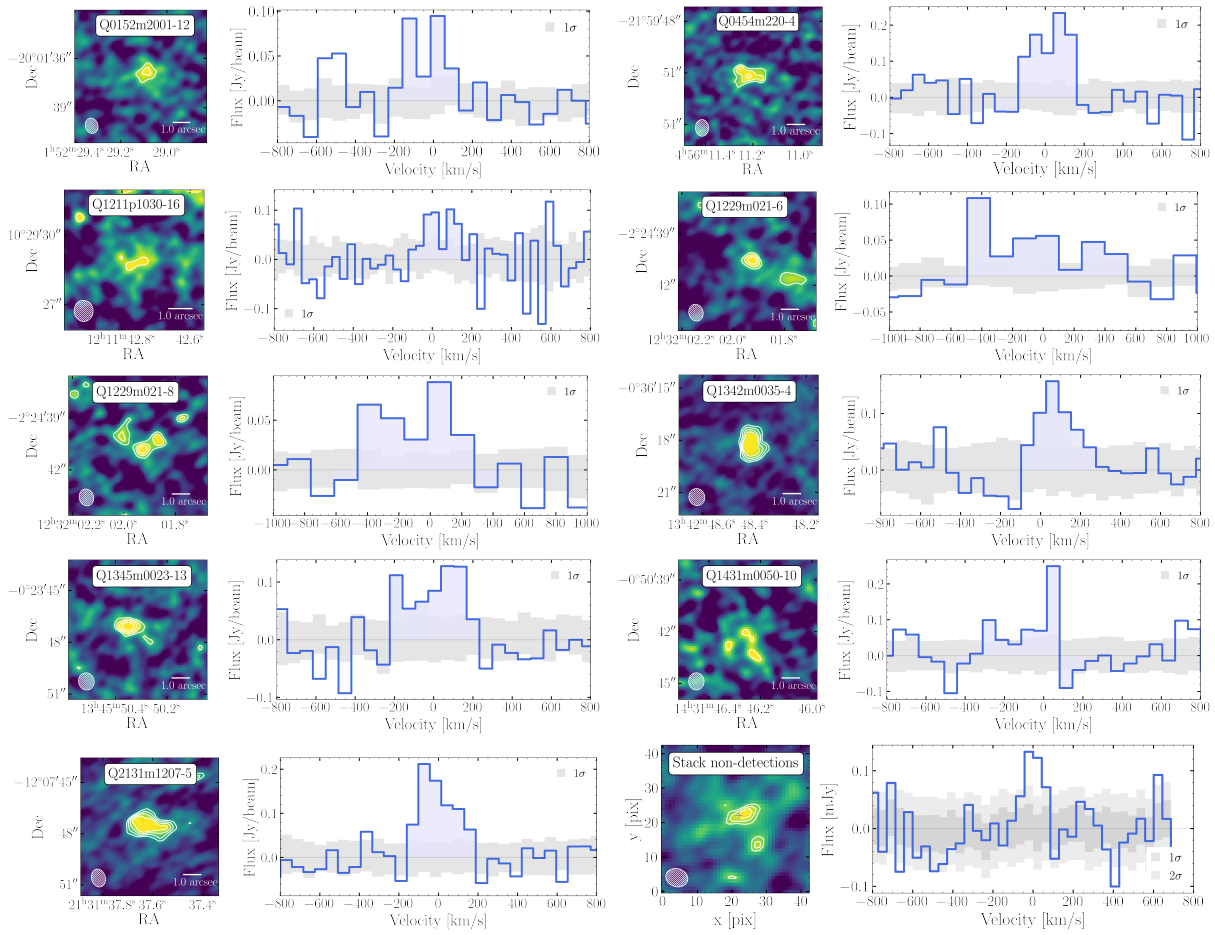

Figure 4.2: CO detections and spectral profiles of the new nine detected sources. Left panels show the moment zero maps with the 3, 4, and 5σ level contours, while the right panels display the extracted CO spectra. The grey shaded regions around each spectrum represent the $\pm 1\sigma$ RMS noise level estimate from the individual data cubes, and the blue shaded region indicates the velocity range used to create the moment maps. The final row, right figure, presents the rest-frame stacked spectrum of non-detections with $\pm 1\sigma$ and $\pm 2\sigma$ levels. A summary of the integrated CO fluxes and associated uncertainties for all detected sources is provided in Table 4.1.

Table 4.1: Measured properties of the new nine detected sources from the MUSE-ALMA Haloes ALMA Large Program.

Galaxy	$z_{\rm gal}$	$\log N({\rm H\,I})$	b	CO line	$F_{\rm peak}$	$F_{\rm CO}$	FWHM	$\log L'_{\rm CO}$	$\log M_{\rm H_2}$
		$\rm cm^{-2}$	kpc		mJy	mJy km $\rm s^{-1}$	km $\rm s^{-1}$	K km $\rm s^{-1}pc^2$	$M_\odot$
Q0152m2001_12	0.78	18.87	55.0	CO(2−1)	0.61	122 ± 40	241 ± 71	9.13 ± 0.15	9.71 ± 0.13
Q0454m220_4	0.48	18.65	107.6	CO(2−1)	1.10	256 ± 31	198 ± 37	9.09 ± 0.05	9.73 ± 0.05
Q1211p1030_16	0.89	18.50	77.4	CO(4−3)	0.90	90 ± 39	112 ± 46	9.11 ± 0.24	9.94 ± 0.24
Q1229m021_6	0.83	18.84	182.3	CO(2−1)	0.39	197 ± 43	419 ± 191	9.53 ± 0.08	10.33 ± 0.08
Q1229m021_8	0.83	18.84	125.5	CO(2−1)	0.36	172 ± 16	464 ± 204	9.36 ± 0.05	10.19 ± 0.05
Q1342m0035_4	0.53	19.78	43.6	CO(2−1)	1.44	273 ± 39	169 ± 33	9.23 ± 0.06	9.86 ± 0.06
Q1345m0023_13	0.61	18.85	55.6	CO(2−1)	0.93	201 ± 43	224 ± 75	9.09 ± 0.12	9.75 ± 0.12
Q1431m0050_10	0.61	19.18	45.4	CO(2−1)	3.13	116 ± 47	77 ± 13	9.07 ± 0.14	9.74 ± 0.15
Q2131m1207_5	0.43	19.50	48.5	CO(2−1)	0.85	210 ± 25	225 ± 40	8.85 ± 0.06	9.48 ± 0.06

Notes. b is the impact parameter, and $F_{\rm peak}$ is the measured peak flux

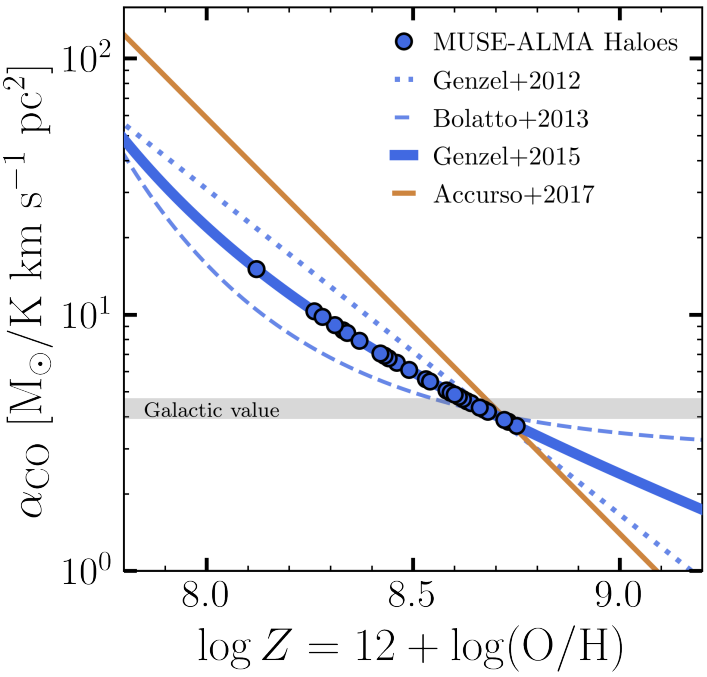

Figure 4.3: Different α_{CO} models across varying metallicites. In this work, we adopt the model proposed by Genzel et al. (2015), shown in the blue points, which corresponds to the geometric mean of the prescriptions of Bolatto et al. (2013) and Genzel et al. (2012), as indicated in Eq. 4.3. As a reference, we also include the model proposed by Accurso et al. (2017), which would lead to higher values for α_{CO} for the galaxies in our sample. To prevent extrapolation into poorly constrained regions of parameter space, we impose an upper limit on α_{CO}, set to the lowest emission metallicity measured in our sample, which leads to an $\alpha_{CO} \sim 15$.

4.3.3 CO Luminosity

We calculated the CO luminosity from CO fluxes using the following equation, from Solomon et al. (1997),

$$L'_{\rm CO} = 3.25 \times 10^7 \frac{S_{\rm CO}\Delta V}{(1+z)^3} \left(\frac{D_{\rm L}}{\nu_{\rm obs}}\right)^2 \ [\rm K\ km\ s^{-1} pc^2]\ , \tag{4.1}$$

where $L'_{\rm CO}$ is in units of K km s^{-1} pc^2, $\nu_{\rm obs}$ is the observed frequency of the CO line in GHz, $D_{\rm L}$ is the luminosity distance of the galaxy in Mpc, z is the redshift, and $\rm S_{CO}\Delta V$ is the integrated flux in Jy km s^{-1}. We converted the luminosity measured for mid- and high$-J$ CO into CO(1$-$0), scaling by the empirical conversion factors from Fixsen et al. (1999); Weiß et al. (2005): $r_{J\rightarrow1}$ ={2.5, 3.5, 2.5}, for $J = 2, 3, 4$. These values were also assumed in previous ALMA observations of the MUSE-ALMA Haloes survey published by Kanekar et al. (2018); Szakacs et al. (2021), and we scaled the values from Klitsch et al. (2018) accordingly. Table 4.1 shows the measured CO fluxes and full width at half maximum (FWHM) for all the detected sources. The FWHM was calculated from a Gaussian fit to the line profile, using the relation FWHM $= 2.355 \times \sigma$, where σ is the standard deviation obtained from the fit.

4.3.4 Molecular gas mass

We derive the molecular gas mass of our sources using the CO(1–0) luminosity and the empirical conversion factor, $\alpha_{\rm CO}$, following Tacconi et al. (2010):

$$M_{\rm mol\,gas}/M_\odot = \alpha_{\rm CO} \cdot L'_{\rm CO(1-0)} \quad , \tag{4.2}$$

The conversion factor $\alpha_{\rm CO}$ increases with decreasing metallicity Z, since lower metallicity environments allow ultraviolet radiation to penetrate deeper into molecular clouds, leading to more extensive photodissociation of CO (Wolfire et al. 2010; Bolatto et al. 2013). From the different metallicity corrections proposed in the literature, we adopt the geometric mean of the recipes by Bolatto et al. (2013) and Genzel et al. (2012) as adopted by Genzel et al. (2015) and Tacconi et al. (2018):

$$\begin{aligned}\alpha_{\rm CO} = 4.36 \times \Big[0.67 \times \exp\left(0.36 \times 10^{-(12+\log({\rm O/H})-8.67)}\right) \\ \times\ 10^{-1.27\times(12+\log({\rm O/H})-8.67)}\Big]^{1/2}\ [M_\odot/({\rm K\,km\,s^{-1}\,pc^2})].\end{aligned} \tag{4.3}$$

where 12+log(O/H) is the metallicity on the Pettini & Pagel (2004) scale. Figure 4.3 shows the comparison of the different prescriptions proposed by Genzel et al. (2012); Bolatto et al. (2013); Accurso et al. (2017) and the adopted model for our sources from Genzel et al. (2015).

The metallicities for our targets were derived by Weng et al. (2023b) using the R_3 strong-line calibration (Curti et al. 2017, 2020), defined as $R_3 = \log([\mathrm{O\,III}]\,\lambda5007/\mathrm{H}\beta)$. This method is applicable up to redshift $z \sim 0.85$ and is largely unaffected by dust obscuration. For galaxies at $z < 0.4$, the O3N2 diagnostic (Pettini & Pagel 2004), given by $\log([\mathrm{O\,III}]\,\lambda5007/\mathrm{H}\beta) - \log([\mathrm{N\,II}]\,\lambda6584/\mathrm{H}\alpha)$, was also measured to cross-check R_3-based metallicities and, in some cases, resolve the degeneracy in the R_3 indicator. Errors were estimated by propagating flux uncertainties and using Monte Carlo sampling to solve the metallicity calibration polynomial.

Out of the 39 galaxies included in the new ALMA observations, we have direct metallicity measurements for 24 from the MUSE data based on the optical emission lines to get the R_3 auroral strong line calibration. For the remaining 15, stellar mass estimates are available for 13 galaxies, allowing us to infer their metallicities using the mass-metallicity relation (Pettini & Pagel 2004). To maintain consistency with previous studies (e.g., Saintonge et al. 2017; Tacconi et al. 2018), we use the following mass–metallicity relation:

$$\log Z = 8.74 - 0.087 \times (\log(M_*) - b)^2, \tag{4.4}$$

with $b = 10.4 + 4.46 \times \log(1+z) - 1.78 \times (\log(1+z))^2$ (Genzel et al. 2015; Erb et al. 2006; Maiolino et al. 2008; Zahid et al. 2014; Wuyts et al. 2014). For the two galaxies without reliable stellar mass estimates (IDs: Q0454p039_15, Q0454p039_57), we investigated the metallicity estimated in absorption along the line of sight of a quasar background. The metal content of the absorbing gas is very poor ($[\mathrm{M/H}] \lesssim -1.0$), so we take a lower limit in the metallicity set to the lowest observed value in our sample. This approach is more conservative than using an extrapolation of the prescriptions mentioned before, as the α_{CO} quickly increases toward low metallicities. Using the lowest observed metallicity, we adopt an $\alpha_{\mathrm{CO}} \sim 15$ for the metal-poor regime. For sources from the literature, we reanalysed the $M_{\mathrm{H_2}}$ estimates using an α_{CO} value consistent with our methodology.

4.4 Properties of H I–selected galaxies

The selection of H I absorbers through absorption features in a background quasar spectrum allows the detection of diffuse, low-surface brightness neutral gas that often eludes direct emission observations due to the low density of gas (Wolfe et al. 2005; Prochaska et al. 2017). This method offers redshift-independent sensitivity and yields accurate measurements of metallicity that are largely insensitive to temperature or ionisation conditions, unlike emission diagnostics (Pettini et al. 1999; Kulkarni et al. 2005; Lehner et al. 2013). In this section, we revise the properties of absorption-selected galaxies and relate them to their cold gas content.

4.4.1 Molecular gas detection rate

Despite the high neutral hydrogen column densities of the targeted absorbers in the MUSE-ALMA Haloes survey ($\log\,[N(\mathrm{H\,I})/\mathrm{cm}^2] \gtrsim 18$), CO emission is detected in 12 out of 50 galaxies, corresponding to a detection rate of $\sim 24\%$. Although this detection rate is lower than those reported in some previous studies (e.g., Klitsch et al. 2018), it represents a substantial number of detections drawn from a broader and less biased sample. This represents a crucial step toward a more comprehensive understanding of molecular gas in absorption-selected galaxies.

Previous studies have reported high CO detection rates ($> 60\%$), particularly in systems pre-selected for high metallicity or H_2 absorption. For instance, Neeleman et al. (2016) reported CO(1−0) emission from a $z \sim 0.101$ absorber-associated galaxy with a molecular gas mass of $4.2 \times 10^9\,M_\odot$ and low star formation rate, indicating a long gas depletion timescale. Similarly, Møller et al. (2018) detected CO(2−1) in a highly metal-rich DLA galaxy at $z = 0.716$, finding a large molecular mass ($\sim 2.3 \times 10^{10}\,M_\odot$) but suppressed star formation, deviating from canonical SFR–$M_{\mathrm{H_2}}$ scaling relations. These examples suggest that molecular gas-rich absorption-selected galaxies may undergo inefficient star formation, possibly due to environmental quenching. Kanekar et al. (2018) extended this trend with a sample of high-metallicity H I absorbers at $z \sim 0.5$–0.8, detecting CO in five out of seven cases (detection rate $> 70\%$). These galaxies have molecular gas masses from 0.6 to $8.2 \times 10^{10}\,M_\odot$, again linked with modest star formation rates, reinforcing the idea of gas-rich but low-efficiency star-forming systems. Klitsch et al. (2021) found CO emission in five of six H_2-bearing absorbers (detection rate $> 80\%$), further supporting a strong link between H_2

absorption and CO-rich galaxies, particularly in group environments or overdensities. They found no clear correlation between CO detection and absorber metallicity and concluded that H_2 absorbers trace diffuse molecular gas in the CGM or intragroup medium, rather than the central disk. Further analysis by Klitsch et al. (2022) of the CO excitation in absorption-selected galaxies showed a broad range of ISM conditions, highlighting the heterogeneous nature of these systems and the complexity of relating CO luminosity to total H_2 content.

Collectively, these findings suggest that absorption selected galaxies at $z \lesssim 1$ showing CO emission, particularly at high metallicities, may represent a distinct population, characterised by substantial molecular gas reservoirs but often exhibiting suppressed or inefficient star formation activity relative to typical star-forming galaxies at similar epochs. However, these higher detection rates likely benefit from sample selections that favour CO-rich systems.

The MUSE-ALMA Haloes survey provides a complementary perspective. By selecting galaxies solely based on their H I absorption properties, without bias toward metallicity or known H_2 content, the survey probes a broader galaxy population. Our ∼24% detection rate thus provides a more representative estimate of molecular gas occurrence in H I-rich systems at $z \lesssim 1$, offering a statistically robust view of cold gas in the CGM and galaxy environments traced by quasar absorbers.

To further explore possible factors influencing CO detection, Figure 4.1 shows the distribution of CO detections and non-detections as a function of redshift, stellar mass, and absorption metallicity. We find no clear dependence of detection rate on any of these parameters, suggesting that the presence of CO-emitting molecular gas is not simply governed by global galaxy properties, like the metallicity alone. This is consistent with recent results indicating that local conditions and environmental factors, such as gas density, pressure, or group dynamics, may play a more significant role in regulating cold molecular gas content in H I–rich systems (Lee et al. 2022).

4.4.2 Absorber-galaxy connection

The absorbers in the MUSE-ALMA Haloes survey span a range in H I column density of $\log [N(\mathrm{H\,I})/\mathrm{cm}^2] \sim 18 - 22$, placing them in the sub-DLA to DLA regime. The detection rate of CO emission appears to be independent of the column density as illustrated in

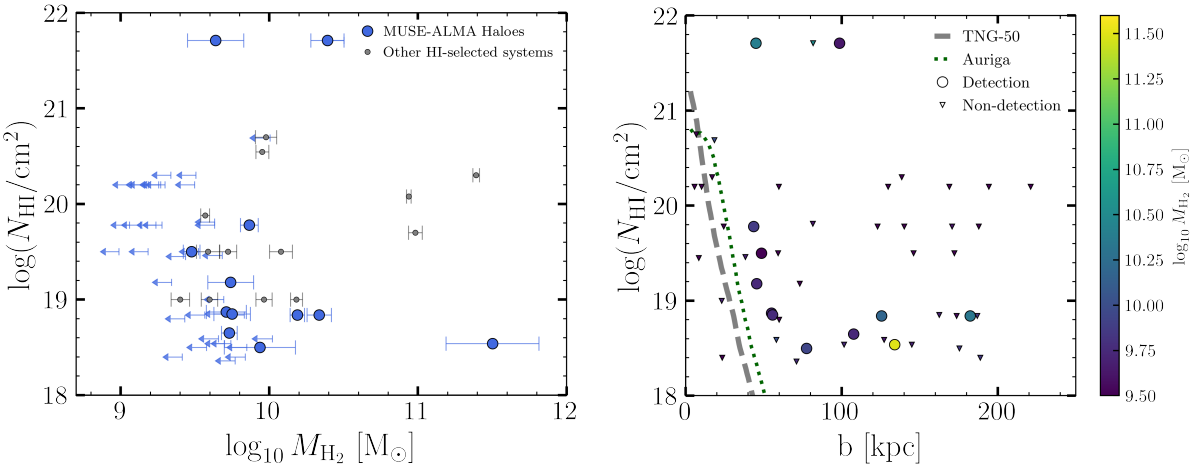

Figure 4.4: *Left:* Molecular gas mass (M_{H_2}) as a function of H I column density (N_{HI}). Blue circles represent our sample in comparison with previous studies on H I-absorption selected systems from Kanekar et al. (2018) and Klitsch et al. (2021), shown in grey. *Right:* N_{HI} and impact parameter, colour coded by the molecular gas masses. In this plot, we include all impact parameters for the sources in our sample, rather than limiting to the smallest b as done in previous studies (e.g. Péroux et al. 2016; Hamanowicz et al. 2020; Weng et al. 2022). The circles represent CO detections, and the triangle represents non-CO detections. We also include the median radial profiles for the neutral hydrogen column density from the simulations by van de Voort et al. (2019) and Nelson et al. (2020) in the dashed and dotted lines, respectively. We report no clear correlation between the column density and impact parameter for different molecular gas masses.

the left panel of Figure 4.4. Systems with high $N(\mathrm{H\,I})$ ($\geq 10^{21}$ cm^{-2}) are not necessarily associated with large molecular reservoirs, and conversely, galaxies with substantial $M_{\mathrm{H_2}}$ can show relatively modest $N(\mathrm{H\,I})$ values in absorption. This suggests that the absorbing H I gas, while indicative of the presence of neutral material, does not reliably trace the spatially integrated cold molecular gas content. This conclusion is consistent with the findings of Klitsch et al. (2021) in a sample of H_2-bearing absorbers, where despite the high CO detection rate, they also report no clear correlation between $N(\mathrm{H\,I})$, molecular gas fraction, or impact parameter and the host galaxy's total $M_{\mathrm{H_2}}$.

Our detections span a broad range of impact parameters (from ∼40 to 180 kpc), and again, we observe no significant trend between impact parameter and $M_{\mathrm{H_2}}$ or $N(\mathrm{H\,I})$. Some systems with large molecular reservoirs are found at relatively large projected separations, while others close to the quasar sightline show little molecular gas. Moreover, when examining $N(\mathrm{H\,I})$ versus impact parameter color-coded by $\mathrm{M_{H_2}}$, shown in the right panel of Figure 4.4, we find no coherent trend. As a reference, we also include the results from the Auriga zoom-in simulations of a Milky Way-mass galaxy (van de Voort et al. 2019), and from the post-processing of TNG50 around massive haloes ($\sim 10^{13.5} M_\odot$) at $z \sim 0.5$ (Nelson et al. 2020). Systems with similar impact parameters exhibit large scatter in both $N(\mathrm{H\,I})$ and molecular mass, supporting the notion that the location and kinematics of absorbing gas are not strongly tied to the total molecular reservoir of the galaxy. Note that, unlike typical comparisons of these two quantities (e.g. Péroux et al. 2016; Hamanowicz et al. 2020; Weng et al. 2022), we do not restrict the plot to only the smallest impact parameter. In contrast, Klimenko et al. (2023) identified a strong correlation between $N(\mathrm{H\,I})$ and impact parameter, found that absorber metallicities matched the emission metallicity gradients derived from integral field spectroscopy (IFS), and reported evidence of co-rotation between absorbing gas and galaxy disks extending out to ∼10 effective radii. These results suggest that, in some systems, absorbing gas can trace structured, rotating components of the CGM. However, such coherence is not observed across our full sample. Additionally, Augustin et al. (2024) found an anticorrelation between the stellar mass of host galaxies and the $N(\mathrm{H\,I})$ of associated absorbers. These findings indicate an evolutionary trend in the CGM composition with stellar mass, where lower-mass galaxies tend to host haloes rich in cool, dense H I gas, while higher-mass systems show a depletion of such gas.

Interestingly, out of the 12 galaxies with detected CO emission, all show consistent redshifts with the associated H I absorber ($\Delta v < 500$ km s^{-1}), and all but one (ID

Q0152m2001_12) are embedded in environments where other galaxies or absorption systems are present at similar redshifts. This highlights the idea that molecular gas is not only a feature of individual galaxies, but also a product of their interaction with the surrounding cosmic web and galaxy associations, consistent with earlier findings Klitsch et al. (2019b); Hamanowicz et al. (2020). This environmental connection may facilitate the accumulation or retention of molecular gas through processes such as galaxy-galaxy interactions or group-scale gas accretion. Gravitational interactions can funnel and compress gas, increasing the molecular gas reservoir, while group environments enable continuous gas inflow and help prevent gas loss due to their deeper potential wells. This can lead to galaxies that harbour large amounts of molecular gas but exhibit relatively low star formation rates, suggesting that environmental factors may promote gas buildup without immediately triggering efficient star formation. Moreover, gas flow geometry derived from MUSE data (Weng et al. 2022) reveals that some of these CO-rich systems are linked to inflows (e.g., Q0152m2001_12 and Q2131m1207_5 at ∼50 kpc), while others, like Q1229m021_8 at ∼120 kpc, show outflow-like kinematics, indicating that molecular gas can be present in galaxies that are both accreting and expelling material, depending on the galaxy's location and dynamics within its environment.

4.4.3 H I absorption selected-galaxies compared to star-forming populations

To contextualise our sample, we compare the properties of the host galaxies associated with the absorbers with the well-established galaxy scaling relations. We used a subsample of 89 galaxies at $0.3 < z < 1.2$ from the PHIBSS survey (Tacconi et al. 2018) ranging in stellar mass from $10^{9.8}$ to $10^{11.8}$ $M_\odot$ and the full sample from xCOLD GASS (Saintonge et al. 2017) at $z \sim 0$ ranging in stellar mass from 10^{9} to $10^{11.3}$ $M_\odot$, which is considered representative of the local galaxy population. In Figure 4.5, the left panel shows star formation rate (SFR) as a function of molecular gas mass (M_{H_2}), while the right panel shows the stellar mass against M_{H_2}. We distinguish the MUSE-ALMA Haloes (blue circles) from other literature H I–selected galaxies (grey points) from Kanekar et al. (2018) and Klitsch et al. (2019b), while the red and orange circles represent the PHIBSS and xCOLD GASS samples, respectively.

Our H I–selected galaxies lie in a different location of the parameter space from emission-

selected galaxies. At fixed M_{H_2}, they exhibit systematically lower SFR than PHIBSS and xCOLD GASS (Figure 4.5, left panel). While the normal star-forming galaxies at the same redshift range follow the expected scaling relation (Tacconi et al. 2018), shown by the solid red line, our MUSE-ALMA Haloes sample falls below this relation by ~ 0.6 dex on average. This offset implies that H I–selected galaxies are less efficient in converting their molecular gas into stars. Even accounting for dust corrections (shown by the blue downward arrow), which is ~ 0.24 dex on average, the offset remains.

We investigate the stellar mass–molecular gas mass ($M_\star - M_{H_2}$) relation in the context of the MUSE–ALMA Haloes survey. As shown in the right panel of Figure 4.5, there is a systematic offset in this scaling relation for galaxies selected via H I absorption, compared to the main sequence of star-forming galaxies at similar redshifts. H I–selected systems lie approximately 0.5 dex below the reference relation derived from emission-selected galaxy samples (Saintonge et al. 2017; Tacconi et al. 2018), indicating that they host disproportionately large molecular gas reservoirs relative to their stellar masses.

We examine the molecular gas mass-metallicity relation for our H I–selected sample, as shown in the left panel of Figure 4.6. We use the derived gas-phase metallicities from nebular lines presented in Weng et al. (2022), and compare these with molecular gas masses for the CO-detected systems. The CO-detected galaxies in our sample generally follow the same locus as normal star-forming galaxies in the molecular gas mass–metallicity plane. Most of the non-detections lie at lower metallicities, falling below the expected trend for their gas content. This suggests that emission metallicity may influence the detectability of molecular gas via CO lines, since CO is easily photo-dissociated in low metallicity environments, and there may be a lot of molecular gas that is not traced by CO (e.g. Madden et al. 2020). To quantify this, we construct an emission metallicity distribution for both CO detections and non-detections as shown in the right panel of Figure 4.6. A Kolmogorov–Smirnov (KS) test reveals a statistically significant difference between the two populations, with a p-value of 0.004. This indicates that the metallicity distributions of the CO-detected and CO non-detected galaxies are unlikely to be drawn from the same parent population.

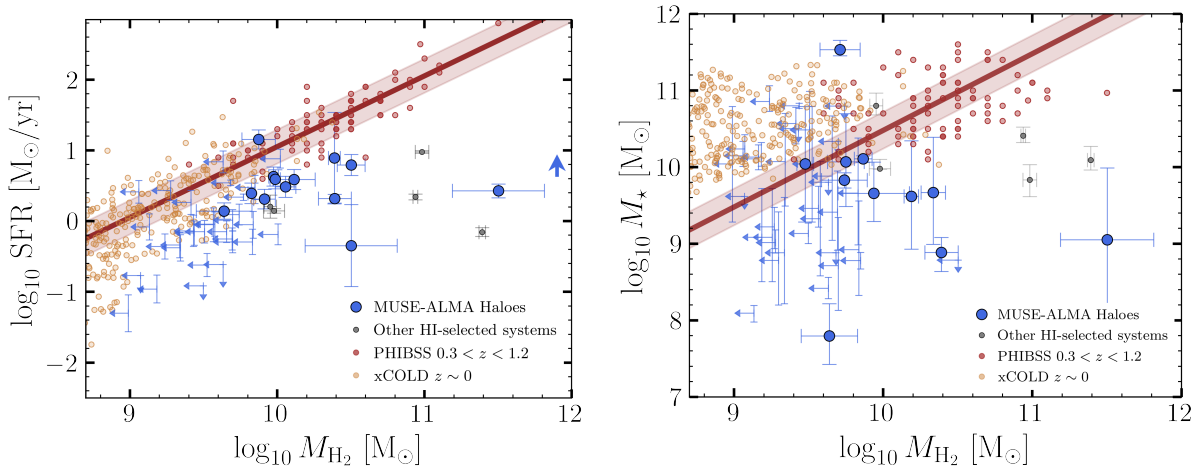

Figure 4.5: *Left:* Star formation rate (SFR) as a function of molecular gas mass (M_{H_2}). *Right:* Stellar mass ($M_\star$) as a function of molecular gas mass (M_{H_2}). Our sample is shown in blue circles, and the blue arrows represent a 3σ upper limit. We also include the mean dust correction for SFR depicted by a large blue arrow on the right side of the left panel. For comparison, we include data from the xCOLD GASS survey (Saintonge et al. 2017), which represents low-redshift galaxies, and the PHIBSS survey (Tacconi et al. 2018), in the redshift range $z \sim 0.3-1.2$. The solid line represents the molecular gas main sequence scaling relation from Tacconi et al. (2018). We also included estimates from Kanekar et al. (2018) and Klitsch et al. (2021) in both panels, labelled as other H I–selected systems. Our sources lie ~ 0.7 dex below the expected SFR$-M_{H_2}$ relations for 'normal star-forming galaxies' at the same redshift range. This trend suggests that H I–selected systems exhibit low star formation efficiency, likely because they are still actively accreting gas from the intergalactic medium or interacting within group environments, and have not yet reached the equilibrium conditions characteristic of main-sequence galaxies.

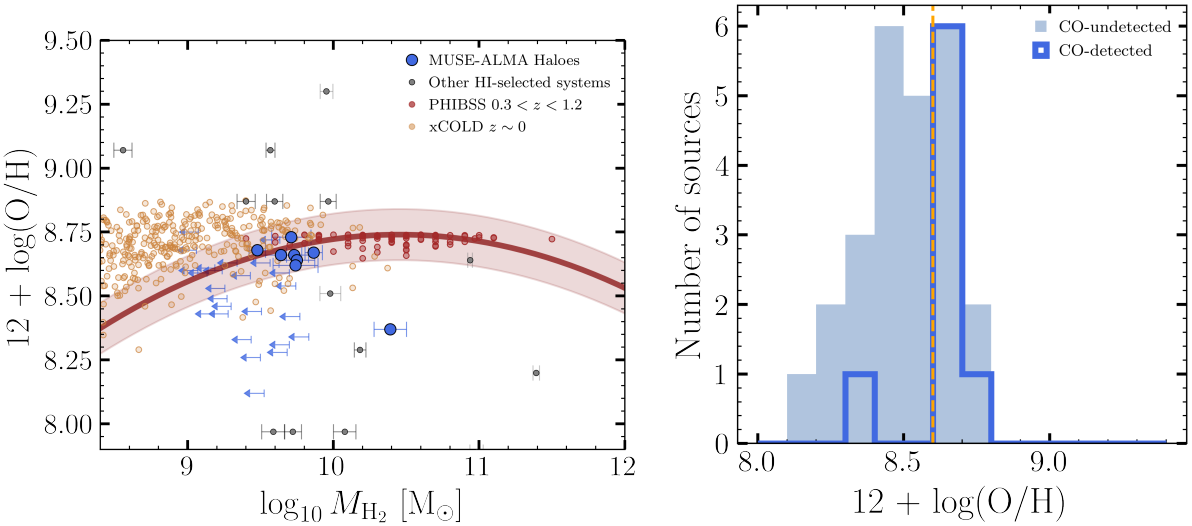

Figure 4.6: *Left:* Nebular emission metallicity, $12 + \log(\mathrm{O/H})$, as a function of molecular gas mass ($M_{\mathrm{H_2}}$) for our H I–selected galaxies. Blue circles indicate CO detections, and blue leftward arrows show 3σ upper limits for CO non-detections. The solid line indicates the mass-metallicity relation from Genzel et al. (2015) for the PHIBSS sample (Tacconi et al. 2018). A positive correlation emerges between metallicity and CO detectability, with CO-detected galaxies tending to occupy the higher-metallicity regime ($\log Z \gtrsim 8.5$), as found previously in many studies (e.g., Bolatto et al. 2013). *Right:* Histogram of emission metallicities for all galaxies, with CO detections in blue and non-detections in light blue filled histogram. Although the distributions overlap significantly, a two-sample Kolmogorov-Smirnov test yields a p-value of 0.004, indicating a statistically significant difference in the underlying metallicity distributions. This reinforces the interpretation that metallicity plays a key role in governing the observability of molecular gas in H I–selected systems.

4.5 Discussion

4.5.1 Implication of CO detections for H I–selected samples

Understanding the possible relation between H I-absorption selected galaxies and their CO emission is essential to interpret their physical properties in a wider context. Several studies have suggested a trend between high-metallicity absorption systems and CO emitters at $z < 1$ (e.g., Kanekar et al. 2018; Møller et al. 2018; Klitsch et al. 2021). Interestingly, this trend at high redshift ($z \geq 2$) has also been explored. CO emission has been detected in ~ 7 out of ~ 20 DLA fields surveyed to date, with detection predominantly arising in systems with relatively high metallicities ([M/H]≥ -0.72) and large molecular gas masses ($M_{\rm mol} \geq 5 \times 10^{10} M_\odot$). For instance, Neeleman et al. (2018) reported the detection of CO(3–2) and far-infrared continuum emission from a galaxy associated with a DLA at $z = 2.192$. Located at an impact parameter of 30 kpc from the quasar sightline, the galaxy exhibits a dust-corrected star formation rate (SFR) of $\sim 110\, M_\odot$ yr^{-1} and a molecular gas mass of $\sim 1.4 \times 10^{11} M_\odot$. Kanekar et al. (2020) conducted a systematic ALMA survey of 12 high-metallicity DLAs ($z \sim 1.7$–2.6), detecting CO emission in five fields. The galaxies have molecular gas masses ranging from 1.3×10^{10} to 2.1×10^{11} M$_\odot$ and impact parameters between 5.6 and 100 kpc. The study confirmed a strong correlation between DLA metallicity and the likelihood of CO detection, with high-metallicity systems more frequently associated with massive, gas-rich galaxies. In another example, Kaur et al. (2022) used NOEMA to detect CO(3–2) emission from the galaxy DLA0201+365g at $z = 2.4604$. The galaxy, located $\sim$7 kpc from the quasar sightline, was found to contain a molecular gas mass of $\sim 5 \times 10^{10}$ M$_\odot$. Despite its gas richness, the galaxy had an SFR upper limit of just 2.3 M$_\odot$ yr^{-1}, suggesting either significant dust obscuration or a long gas depletion timescale. Further support for large molecular gas reservoirs in H I–selected galaxies at high redshift comes from JVLA detections of CO(1–0) and CO(3–2) in two DLAs at $z \approx 2.193$ and $z \approx 2.356$ (Kaur et al. 2024, 2025). These galaxies exhibited CO velocity widths of $\sim$500–600 km s^{-1} and molecular gas masses approaching 10^{11} M$_\odot$. Interestingly, while DLA1228–113g showed near-thermal excitation of mid-J CO transitions, DLA1020+2733g showed sub-thermal excitation, indicating a relatively low star formation surface density and possible differences in gas excitation and star formation efficiency. More recently, Neeleman et al. (2025) reported a [C II] detection rate of $> 70\%$ for galaxies at $z \sim 4.1 - 4.5$ selected to have an absorption metallicity [M/H]> -1.5.

These studies reveal that H I–selected galaxies at low and high redshift, particularly those associated with metal-rich DLAs, often harbour massive molecular gas reservoirs and exhibit complex, multiphase interstellar and circumgalactic media. However, their star formation activity is not always balanced with their gas content, confirming that selection via H I absorption preferentially traces systems at particular evolutionary stages not typically represented in flux-limited, emission-selected samples (e.g., Kulkarni et al. 2005)

The stringent detection limits in our sample probe the intermediate molecular gas mass range ($M_{\rm mol} \lesssim 10^9 - 10^{10} M_\odot$), and notably, several non-detections occur in galaxies with high stellar masses ($> 10^{10} M_\odot$). The lack of CO detections in most absorber-selected galaxies does not necessarily imply the absence of molecular gas. Instead, it likely reflects the presence of "CO-dark" molecular gas, defined as regions where H_2 exists without detectable CO emission due to a variety of factors. This interpretation is supported by several factors. First, in low metallicity environments, reduced dust shielding allows ultraviolet radiation to penetrate more deeply into molecular clouds. As a result, substantial reservoirs of H_2 may exist without CO emission. As shown by Madden et al. (2020), this effect is particularly significant in dwarf and metal-poor galaxies, which are frequently associated with high H I column density absorbers (e.g., York et al. 1986; Kulkarni et al. 2000). Second, physical conditions such as sub-critical gas densities or beam dilution in extended, low-surface-brightness structures can further suppress observable CO. Simulations by Li et al. (2018) found that a large fraction of H_2 resides in diffuse regions where gas densities are below the critical threshold required to excite CO transitions, where CO emission may be weak or absent. Observations by Smit et al. (2014) also demonstrated that CO-dark gas frequently occurs in filamentary structures, which can be overlooked in high-resolution surveys. If molecular gas is spatially extended with low surface brightness, the CO emission may be spread over a large area and diluted below the detection limits of the observing beam. Given that absorber-selected galaxies are biased toward tracing diffuse, metal-poor, or extended gas phases, the absence of CO detections is best explained not by a true lack of molecular gas but by its predominantly CO-dark nature. Overall, the environmental conditions and observational limitations strongly support the scenario in which CO-dark molecular gas dominates the molecular content in these systems.

4.5.2 Physical diversity of H I absorbers

The wide range of galaxy–absorber configurations in our sample emphasises the physical diversity inherent to H I–selected systems. 11 of the 12 CO-detected galaxies are associated with additional galaxies or absorbers at the same redshift. This implies that absorption selection is particularly sensitive to complex environments, such as group-scale structures or filaments, rather than isolated field galaxies. Consistent with this, Klitsch et al. (2021) found that H_2-bearing absorbers often trace overdensities or galaxy groups. Kinematic classifications from Weng et al. (2022) further reveal that the absorbing gas in several of these systems likely originates from outflows or inflows, rather than rotationally supported disks. Thus, the cold gas reservoirs probed via H I absorption are embedded in a diverse range of physical contexts, and absorption-selected galaxies do not represent a uniform or simple population.

While some of these galaxies are actively star-forming and embedded within overdense regions, others appear more quiescent, yet still retain significant molecular gas masses. The H I–selected galaxies in our sample show a mismatch between SFR and M_{H_2}, implying that they lie well above the star formation "main sequence" in depletion time, as shown in Figure 4.7. Therefore, they form stars less efficiently than expected for their molecular gas mass by a factor of ~ 0.7 dex, placing them in a slower, more quiescent evolutionary phase. As noted by Kanekar et al. (2018) and Klitsch et al. (2022), this apparent disconnect between molecular gas content and star formation activity points to environmental or dynamical suppression of star formation. Several mechanisms may contribute to the observed inefficiency of star formation in these systems. One possibility is that they are in an early phase of gas accretion, where molecular gas has recently condensed, but star formation has not yet fully ignited. Alternatively, internal feedback or environmental influences, such as turbulence, ram pressure, or low enrichment, may act to suppress or delay the conversion of gas into stars. In low-metallicity environments, an overestimated CO-to-H_2 conversion factor (α_{CO}) could also contribute to the apparent excess in molecular gas, though this alone is unlikely to explain the consistent suppression across multiple systems.

At the same time, our H I–selected galaxies have stellar masses generally lower than their molecular gas mass, being more gas-rich than typical emission-selected field galaxies of the same stellar mass. This offset also supports the scenario of H I–selected systems being at an early stage of stellar mass buildup, possibly following recent gas accretion

from filamentary structures or group environments. A similar offset was also reported by Kulkarni et al. (2022) in the H I versus stellar mass relation, which they interpreted as evidence for recent accretion fuelling the gas reservoirs in absorption-selected galaxies. These trends align with findings from other absorption-selected samples (e.g., Neeleman et al. 2016, 2019; Kanekar et al. 2018), reinforcing the notion that H I absorbers trace a distinct population of gas-rich galaxies. Their properties may bridge the gap between typical star-forming galaxies and low-surface-brightness, gas-dominated systems that are often missed by emission surveys.

Finally, the strong segregation in emission metallicity between CO-detected and non-detected galaxies adds further nuance to this picture. The detectability of CO is clearly influenced by metallicity, which governs the abundance and survival of CO relative to H_2 in diffuse, unshielded regions (Bolatto et al. 2013; Schruba et al. 2012). Even when molecular gas is present, low-metallicity environments may have CO emission too faint to detect, particularly in the outskirts or CGM of galaxies. The fact that stellar mass and redshift distributions are comparable between CO detections and non-detections, while metallicity differs, suggests that local ISM conditions, such as enrichment, shielding, and density, are the primary regulators of molecular gas observability. This is consistent with both theoretical expectations and recent observational evidence emphasising the role of small-scale environmental factors in shaping cold gas content (e.g., Stern et al. 2016; Diemer et al. 2019).

4.5.3 Role in cosmic baryon cycle

Our analysis of H I–selected galaxies provides valuable insights into the cosmic baryon cycle, particularly regarding gas accretion, star formation efficiency, and the conversion of atomic to molecular gas. These systems, often overlooked in emission-selected surveys, provide a unique window into the processes governing galaxy evolution.

In the context of the equilibrium gas regulator model, which describes the balance between gas inflow, star formation, and outflows, our H I–selected galaxies appear to occupy a distinct regime. The model developed by Bouché et al. (2010), Forbes et al. (2014), and Lilly et al. (2013) suggests that, at redshifts $z < 2$, the gas consumption time is typically shorter than the stellar mass doubling time $t_\star$, meaning that star formation is not driven by cosmic accretion and gas availability. However, in our sample, the molecular gas depletion

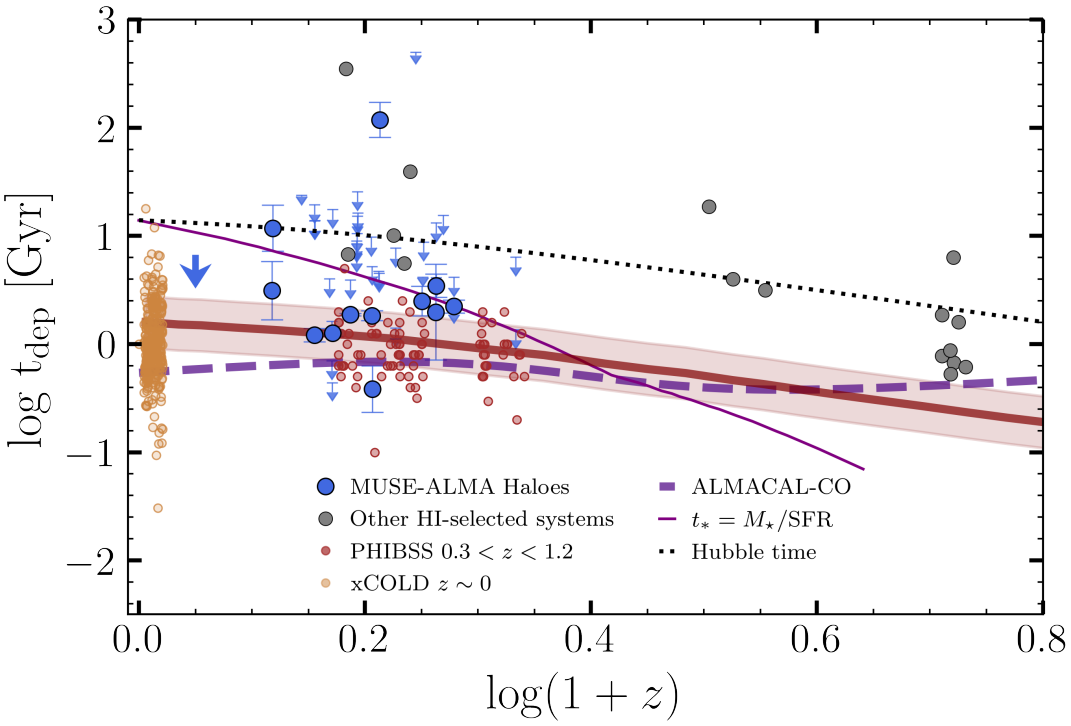

Figure 4.7: Molecular gas depletion time ($t_{dep} = M_{H_2}$/SFR) as a function of redshift for our H I–selected sample (blue circles), compares with literature scaling relations and emission-selected samples (Saintonge et al. 2017; Tacconi et al. 2018). The solid red curve shows the empirically derived t_{dep} evolution for main-sequence galaxies from Tacconi et al. (2018), the dotted line marks the Hubble time at each redshift, the thin solid line represents the stellar mass doubling time ($t_* = M_\star$/SFR) from Tacconi et al. (2020), and the purple dashed line is from the cosmic depletion timescales derived from the ALMACAL survey (Bollo et al. 2025). H I–selected galaxies, both from our sample and from high-redshift literature sources (gray dots; Klitsch et al. 2021; Neeleman et al. 2018, 2020, 2025; Fynbo et al. 2018; Kanekar et al. 2020, 2018; Kaur et al. 2022, 2025), consistently lie above the main-sequence relation, indicating lower star formation efficiencies. Several galaxies exhibit t_{dep} approaching or exceeding the Hubble time, suggesting inefficient molecular gas consumption and possible bottlenecks in the H I–H_2 transition or suppressed star formation.

times often exceed $t_\star$, shown in the thin purple line in Figure 4.7, placing these systems in the inverse regime described by Lilly et al. (2013) and Tacchella et al. (2016). In this regime, gas accumulates faster than it is consumed, and star formation becomes primarily governed by the depletion time. This supports the idea that H I–selected galaxies may be undergoing early stages of gas buildup or entering a temporarily inefficient star-forming phase.

The observed trend of depletion timescale with redshift in our sample appears to persist across cosmic time. Several H I–selected galaxies at high redshift also show unusually long depletion timescales, as shown in the grey points of Figure 4.7. We included estimates from Kaur et al. (2025), who report gas depletion timescales of $\gtrsim 0.6$ Gyr for a galaxy at $z = 2.357$ (DLA 1020+2733g), where the star formation rate (SFR) is only weakly constrained. Similarly, other $z > 2$ H I–selected galaxies, such as DLA1228-113g at $z \approx 2.193$ with $t_{\rm dep} \approx 1.2$ Gyr (Neeleman et al. 2018; Kaur et al. 2022), DLA 0918+1636g at $z \approx 2.5832$ with $t_{\rm dep} \approx 0.5$ Gyr (Fynbo et al. 2018; Kaur et al. 2022), and DLA 0817+1351g at $z \approx 4.258$ with $t_{\rm dep} \approx 0.8$ Gyr (Neeleman et al. 2020), also exhibit star formation efficiencies lower than those typical of main-sequence galaxies at similar redshifts, which often have depletion times as short as ∼0.2 Gyr. From the DLAs studies by Neeleman et al. (2025) at $z > 4$ and assuming that the [C II] luminosity is tracing the molecular gas reliably in normal star-forming galaxies (Zanella et al. 2018), we estimated their depletion timescales to range from $t_{\rm dep} \sim 0.53$ to 1.8.

The efficiency of converting atomic hydrogen to molecular hydrogen is a crucial bottleneck in the star formation process. Our observations might be suggesting that the H I-to-H_2 transition is delayed in these systems, potentially due to sub-solar metallicities that reduce dust abundance and limit the shielding required for H_2 formation. This is consistent with theoretical predictions by Krumholz et al. (2009) and Sternberg et al. (2014), which show that in low-metallicity environments, the H I-to-H_2 transition can be suppressed or delayed, with timescales extending to tens or hundreds of Myr. Additionally, turbulent gas conditions can affect molecule formation. Simulations by Glover & Mac Low (2007) demonstrate that while turbulence can enhance local densities and promote H_2 formation, it can also increase mixing and disrupt the formation process in lower-density regions, leading to longer conversion times. The presence of detectable CO emission in some of these systems suggests that at least partial molecular gas formation has occurred. However, the inefficiency in converting H I to H_2 is likely due to low metallicity, a low dust-to-gas ratio,

and disrupted interstellar medium conditions, leading to subdued star formation despite significant gas reservoirs.

Although our sample shows high molecular gas masses, this does not necessarily indicate unusually gas-rich galaxies. Since these systems are selected based on strong H I absorption, a significant atomic gas component is expected (O'Beirne et al., in prep.). Supporting this, Messias et al. (2024) reports that the H I-to-H_2 mass ratio remains $\sim 1-3$ over a wide redshift range, implying that the molecular gas we observe likely coexists with comparable H I reservoirs. This suggests that H I–selected galaxies occupy the upper range of total gas content, with large supplies of gas that are not yet fully ignited for star formation.

4.5.4 Implications for molecular gas simulations

The low star formation efficiencies observed in H I-selected galaxies deviate from expectations based on standard galaxy scaling relations, presenting new challenges and opportunities for theoretical modelling. These findings highlight persistent open questions that are difficult to address within current cosmological simulations, particularly in predicting the resolved CO emission in galaxies (Popping et al. 2019a). While large-volume simulations (e.g., Keating et al. 2020; Inoue et al. 2020; Olsen et al. 2021; Pallottini et al. 2022) can reproduce global CO trends using sub-resolutions and post-processed radiative transfer models, they often lack the resolution and detailed physics required to track the formation and distribution of molecular gas self-consistently. To overcome these limitations, several recent studies have begun incorporating non-equilibrium chemistry solvers into high-resolution simulations of isolated or zoom-in galaxies (e.g., Richings & Schaye 2016; Lupi et al. 2018, 2020; Richings et al. 2022), enabling more realistic predictions of the CO-to-H_2 abundances and star formation efficiencies across a range of galactic environments.

For example, Thompson et al. (2024) uses non-equilibrium chemistry with FIRE-2 simulations of isolated disk galaxies that can reproduce the observed CO-H_2 and X_{CO}-metallicity relations in the Milky Way and the xCOLD GASS sample. Similarly, Richings et al. (2022) shows that the H I-to-H_2 transition is sensitive to local stellar radiation and dust depletions. Lupi et al. (2018, 2020) provide further insights through GIZMO+KROME simulations of isolated galaxies and the post-processing of cosmological boxes with CLOUDY. They demonstrate how local radiation fields and shielding strongly regulate molecular gas con-

tent and its link to star formation. This supports the interpretation that the CGM–ISM interface in our absorption-selected galaxies may host molecular gas not yet forming stars efficiently, owing to unresolved microphysics or radiative feedback.

On cosmological scales, Maio et al. (2022) uses high-resolution simulations (ColdSim) with time-dependent non-equilibrium chemistry, capturing the evolution of H_2 and depletion times from $z \sim 7$ to 2. Their models reproduce the observed H I and H_2 densities and show that even low-metallicity gas can form significant molecular fractions when self-shielded. Meanwhile, Inoue et al. (2020) models molecular gas cloud populations within IllustrisTNG galaxies and computes CO line emission without relying on a fixed α_{CO}. Their approach reproduces the CO luminosity functions at $z \sim 0$ but finds that dwarf galaxies, often CO-dark, contribute non-negligibly to the molecular mass budget, a feature also observed in the metal-poor galaxies in our survey. Lastly, Olsen et al. (2021) presents SÍGAME3, which models FIR emission from cosmological simulations with more detailed radiative transfer. Their models generally agree with observations but still overpredict CO luminosities, highlighting how emission-based diagnostics may be biased at low SFRs or in CGM-dominated regimes like those in our sample.

In this context, the MUSE-ALMA Haloes survey's empirical constraints on molecular gas content, star formation efficiency, and CGM structure offer a valuable dataset for informing and calibrating future simulations. By revealing that absorption-selected galaxies are often inefficient at converting molecular gas into stars, despite having sizeable CO reservoirs, our results motivate new theoretical efforts to explore how feedback, gas accretion, and environmental effects shape the baryon cycle in and around galaxies.

4.6 Conclusions

Understanding the molecular gas content of galaxies selected via H I absorption offers a critical piece in the broader puzzle of galaxy evolution. In this paper, we present new ALMA observations of CO(2–1), CO(3–2), or CO(4–3) emission from galaxies at $0.3 < z < 1.2$ observed in the ALMA Large Program MUSE-ALMA Haloes survey (Cycle 10, PI: C.Péroux). These galaxies were initially selected through H I absorption along the line of sight of bright optical quasars (Weng et al. 2022). Notably, this study represents the first non-metallicity-biased survey of molecular gas in H I–selected galaxies, as the targets were selected based solely on neutral hydrogen absorption.

Our main findings are as follows:

1. We report a CO detection rate of 24% for our sample, highlighting the heterogeneous nature of H I–selected galaxies. While some absorbers are associated with metal-rich, actively star-forming galaxies with substantial molecular reservoirs, others correspond to gas-rich but molecule-poor systems potentially in early evolutionary stages. This diversity reflects the wide range of physical conditions accessible through absorption selection, from chemically enriched disks to diffuse CGM structures.

2. While the CO-detected galaxies exhibit gas-phase metallicities consistent with the mass–metallicity relation, their star formation rates and stellar masses are generally lower than expected from the scaling relations followed by emission-selected star-forming galaxies at similar redshifts. This suggests that absorption-selected systems do not fully align with the main sequence of galaxy growth and may be at different stages of evolution.

3. The inferred depletion timescales often exceed 0.5 Gyr, indicating inefficient star formation despite the presence of significant cold gas reservoirs. This is consistent with prior studies of absorption-selected galaxies (e.g., Neeleman et al. 2020; Kaur et al. 2022) and reinforces the idea that these systems occupy a phase where gas accretion, feedback, and ISM conditions hinder the efficient buildup of stellar mass.

4. The properties of H I–selected systems (elevated gas fractions, low star formation efficiency, and agreement with expected metallicities) stress their importance in the baryon cycle. H I–selected galaxies may be actively assembling gas from the intergalactic medium or interacting within group environments (e.g., Péroux et al. 2019; Hamanowicz et al. 2020), but have not yet reached the equilibrium conditions typical of main-sequence galaxies, or they are on their way to be quenched.

H I–selected systems, often missed in emission-selected surveys, serve as crucial laboratories for studying the interplay between cold gas accretion, feedback, and star formation regulation. By combining absorption-line selection with molecular gas observations, we can access lower-luminosity, lower-SFR systems and build a more complete picture of the processes shaping galaxy evolution.

Interpreting the stringent CO non-detections afforded by these deep data remains challenging: it is unclear whether they result from genuinely low molecular content, low excitation conditions, or suppressed CO emission due to low metallicity. To disentangle these scenarios, it is essential to explore alternative tracers such as neutral carbon ([C I]) and dust continuum, which are less sensitive to metallicity and excitation than CO. Such complementary observations will help refine our understanding of the molecular gas content in diffuse, metal-poor, or quiescent systems.

Chapter 5

Conclusions and Outlook

5.1 Summary

This thesis examines the distribution and evolution of cold molecular gas, the primary fuel for star formation, and its role in shaping the evolution of galaxies. The different phases of the baryon cycle at galaxies are explored by integrating wide-area, blind submillimetre observations from the ALMACAL survey with spatially resolved optical spectroscopy from the MUSE–ALMA Haloes survey. Ultimately, shedding light on the accretion of gas onto galaxies, the subsequent conversion of this gas into stars, and the ejection of enriched material through feedback processes.

The ALMACAL survey uses calibration data from the Atacama Large Millimeter/ submillimeter Array (ALMA) to perform deep, blind observations across a wide range of astrophysical topics, without requiring additional telescope time. Using data from 1047 calibrator fields spanning ALMA Bands 3 to 10 and covering more than 1000 square arcminutes, ALMACAL transforms routine calibration scans into a wide and deep sub-millimetre survey. These calibrators are typically bright, compact extragalactic sources, such as blazars, of which $\sim 78\%$ have known redshifts.

The MUSE–ALMA Haloes is a multiwavelength observational program designed to investigate the multiphase nature of the circumgalactic medium (CGM) by targeting 19 quasar sightlines containing 32 H I Lyα absorbers with column densities in the range $\log N(\mathrm{H\,I})/\mathrm{cm}^2 = 18.1 - 21.7$ at redshifts $0.2 < z < 1.4$. The survey combines CO emission line detections (ALMA) with ionised gas (VLT/MUSE), stellar populations (HST), and neutral gas seen in absorption (VLT/UVES). The survey builds on earlier work that

identified absorber environments featuring galaxy groups, cold accretion, outflows, and extended molecular gas. A statistical analysis of 19 MUSE fields revealed 79 galaxies near the absorber redshifts, forming the core sample for follow-up studies.

Chapter 2 presents the ALMACAL−22 data release, which compiles ALMA calibrator data until May 2022. It includes over 30 TB of data covering 1047 calibrator fields across the southern sky and spanning more than 2000 hours of observing time. A selection of the highest quality data is defined as the pruned sample, comprising 401 fields and 1508 data cubes. This work outlines the methodology used to combine different observations and quality assessments to create science-ready data cubes across ALMA Bands 3 to 9. In terms of depth and area, ALMACAL's values are comparable with legacy submillimetre surveys (e.g., ASPECS, PHIBSS, COLDz), and the cosmic variance effects are at least five times less, thanks to the sampling of thousands of independent lines of sight. The survey enables robust statistical studies of dusty star-forming galaxies, active galactic nuclei, absorption systems, and molecular gas across cosmic time. The chapter outlines the data processing methodology and previews future work, including serendipitous CO emission line detections. The survey also identifies an unbiased subset of calibrators with radio jets, offering a rich and largely untapped sample for future studies of AGN activity and jet evolution. While targeted surveys remain essential for addressing specific questions, ALMACAL demonstrates a powerful and efficient alternative: repurposing multi-band, wide-field calibration data to enable a broad range of extragalactic science without dedicated telescope time. Its legacy provides strong synergies with facilities like JWST and SKA, and serves as a model for leveraging calibration data at other observatories.

Chapter 3 focuses on exploiting the unique statistical advantages of ALMACAL to constrain the cosmic molecular gas mass density ($\rho_{\mathrm{H_2}}$), by relating CO line luminosity to H_2 content through a conversion factor (α_{CO}). Using a source finding algorithm to identify CO emission lines across the pruned ALMACAL sample, the final CO-selected sample avoids biases associated with traditional optical pre-selection or targeted deep surveys. The resulting $\rho_{\mathrm{H_2}}$ rises from the present day and peaks at $z \sim 1.5$, followed by a gradual decline of by $\sim 1\,\mathrm{dex}$ toward higher redshifts. This trend mirrors the evolution of the cosmic star formation rate and highlights the central role of cold molecular gas in driving galaxy growth. The ratio of molecular gas to stellar mass remains consistent with predictions from baryon cycling models, particularly the so-called "bathtub" framework, where continuous accretion and outflows regulate star formation. Moreover, the global gas depletion timescale appears

nearly constant across cosmic time, suggesting that the efficiency of star formation per unit gas mass is universal. These results position ALMACAL as a key dataset for quantifying the cold gas supply that fuels star formation in the Universe.

Chapter 4 turns to galaxies identified through H I absorption through the Lyα line, offering a complementary view of molecular gas content in galaxies not selected via their stellar light or line emission. In the MUSE–ALMA Haloes survey, new ALMA observations of such galaxies revealed 12 CO detections among 50 galaxies, effectively doubling the number of known CO-emitting H I-selected galaxies at $z \sim 0.5$. These galaxies span a wide range of stellar masses and metallicities, but they generally lie below the star-forming main sequence. Despite having substantial molecular gas reservoirs, these systems show relatively low star formation rates and long gas depletion timescales ($\gtrsim 0.5$ Gyr). This suggests that they may be in a regulated phase, where environmental conditions or feedback processes are suppressing star formation. By comparing these systems to emission-selected galaxies, we find that H I-selection reveals a population that may otherwise be underrepresented in galaxy surveys, i.e., galaxies with abundant gas but inefficient star formation. These results further our understanding of the baryon cycle by tracing the inflow of gas into the CGM, its storage there, and its gradual conversion into stars.

Together, these three studies provide a comprehensive overview of the molecular gas content in galaxies throughout cosmic history. They demonstrate the value of combining emission and absorption line techniques, as well as wide-field and multi-epoch surveys. This approach mitigates the limitations of small samples and cosmic variance. The findings support the idea that galaxy evolution is driven by the cycling of baryons through inflow, star formation and feedback across different gas phases. By tracing cold gas from the early universe to intermediate redshifts and across a range of star formation rates, this thesis sheds new light on how galaxies build and regulate their gas reservoirs over time.

5.2 Outlook

5.2.1 ALMACAL

The unique nature of ALMACAL, characterised by self-calibration, high dynamic range, and multi-band coverage, provides a novel opportunity to have both a robust statistical sample and high sensitivity. This section now looks toward how this rich resource can be

exploited further to investigate different scientific cases. The following projects are based on ideas that leverage available data and new observational proposals submitted by the author as Principal Investigator (PI).

ALMACAL–Jets

Many calibrator fields host bright, radio-loud quasars exhibiting clear jet-like structures. As shown in recent studies (e.g., Husemann et al. 2019; Komugi et al. 2022), these fields provide a unique opportunity to study the jet-ISM and jet-CGM interaction, thanks to their exceptional dynamic range and multi-frequency depth. The ALMACAL sample is therefore a valuable resource for investigating jet-driven feedback throughout cosmic history. One promising approach is to use these jet-hosting calibrators to investigate feedback processes using three key approaches: cross-matching with optical and radio jet catalogues, building broadband radio SEDs, and conducting targeted searches for molecular gas using the CO line.

The first step is to systematically cross-match ALMACAL calibrators with established optical and radio catalogues of quasars jets, such as those from the MOJAVE program in VLBA (Lister et al. 2019), VLASS in VLA (Lacy et al. 2020), or the RACS survey in ASKAP (Duchesne et al. 2024). By identifying sources where the morphology of the jet is resolved in multiple bands, it is possible to examine whether the orientation of the jet correlates with the distribution and properties of the continuum and cold gas in the host galaxy. Such comparisons are crucial to test how feedback operates in different geometries, as previously observed in systems like IC 5063 (Morganti et al. 2015a), 3C 293 (Oosterloo et al. 2019), and PSO J352–15 (Walter et al. 2025). Figure 5.1 shows that in some of these systems, the CO emission is closely aligned with the radio jet axis (panels A, B, D), while in others, misalignments or clumpy structures indicate more complex dynamical interactions (panel C). Cross-identification also enables accounting for orientation biases in the classification of blazars, and helps to filter out those that lack signatures of resolved jets or interaction with the ISM. The ultimate goal is to determine whether the presence, direction or size of jets is related to cold gas reservoirs or signs of disturbed kinematics.

A second major avenue is building broadband radio spectral energy distributions (SEDs) for these sources. The multi-band ALMA data in ALMACAL combined with lower-frequency radio data from surveys such as NVSS in VLA (Condon et al. 1998), SUMSS in MOST (Mauch et al. 2013), and RACS in ASKAP (McConnell et al. 2020), allow for well-

sampled radio SEDs across several orders of magnitude in frequency. From these SEDs, one can derive key physical parameters including the synchrotron spectral index, spectral breaks, and turnover frequencies. These metrics provide constraints on the age and energy content of the radio jets. Previous works have shown that sources with steep spectral indices and clear spectral ageing are likely to be remnant or fading AGNs (e.g. Harwood et al. 2013; Brienza et al. 2021; Qu et al. 2024; Jurlin et al. 2021), whereas flatter spectra indicate ongoing jet activity (Tadhunter 2016; Mahatma et al. 2023). By comparing radio SED shapes to the presence and distribution of molecular gas, one can investigate whether younger jets are more commonly associated with rich gas reservoirs and active feedback or whether molecular gas is depleted as the AGN ages. Mukherjee et al. (2021) performed 3D hydrodynamic simulations to study how relativistic jets interact with turbulent gas disks before breaking out into the galactic halo. They found that jets carve cavities and drive outflows in the inner kiloparsec, while compressing the outer disk and sometimes triggering inflows. Inclined jets have a stronger impact, launching wide-angle, sub-relativistic outflows and increasing turbulence throughout the disk. These results match observed features in sources like PKS 1549–79 (Ruffa et al. 2022), where the presence of disturbed CO is associated with a recently triggered radio outburst. However, a systematic search of the relation between jet properties and molecular gas is still lacking.

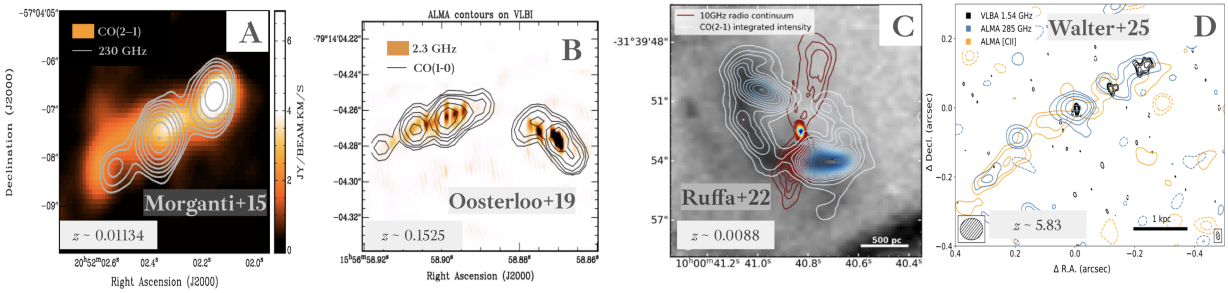

Figure 5.1: Examples of molecular gas and radio jet morphologies in AGN host galaxies. Panels show systems with varying degrees of alignment between CO emission and radio jet structures: IC 5063 (Morganti et al. 2015b), 3C 293 (Oosterloo et al. 2019), PKS 1549-79 (Ruffa et al. 2022), and PSO J352.4034-15.3373 (Walter et al. 2025). These cases illustrate the diversity of jet-ISM interactions, from closely aligned molecular gas and jet emission (A, B, D) to more complex or misaligned configurations (C) at different redshifts.

The third key component of this strategy involves a systematic search for CO emission lines in the ALMACAL survey. By targeting low$-J$ and mid$-J$ transitions such as CO(1$-$0), CO(2$-$1), and CO(3$-$2), one can efficiently trace the molecular gas content in these jet hosts. These lines are sensitive to the bulk of the cold molecular phase, which serves as the fuel for star formation and AGN activity. The detection of such lines allows us to measure the gas mass, map its kinematics, and determine whether it is inflowing, outflowing, or rotating. This information can be correlated with jet morphology and age, as inferred from the radio SEDs. Previous studies have shown that CO-emitting gas near radio jets often exhibits disturbed velocities and asymmetries, which suggests the presence of outflows or turbulence (Feruglio et al. 2015; Oosterloo et al. 2019; Husemann et al. 2019; Ruffa et al. 2022, 2025). In some cases, fast jet-driven molecular outflows extend on kiloparsec scales (Mahony et al. 2013), while in others, the cold gas appears to be compressed along the jet axis, potentially triggering star formation (Bicknell et al. 2000; Salomé et al. 2018). Recent works have proposed that molecular gas may form in situ via rapid cooling of shocked ionised gas (e.g., Richings & Faucher-Giguère 2018), while other scenarios suggest that jets can lift existing cold gas from the disk into large-scale outflows (e.g., de Arriba et al. 2023). The ALMACAL calibrators span a wide range in redshift and luminosity, making them ideal to statistically evaluate these competing models across different cosmic epochs and environments.

These three steps will establish a comprehensive strategy for exploiting the ALMACAL-Jets sample to its full scientific potential. The combination of high-dynamic-range continuum imaging, broad frequency coverage, and archival depth makes ALMACAL an exceptional laboratory for testing how radio jets influence the molecular gas reservoirs of galaxies and regulate the processes of star formation and black hole growth.

CO emitters in ALMACAL

A natural continuation of the work on ALMACAL molecular gas emitters presented in Chapter 3, is the spectroscopic follow-up, of the CO emitters identified in ALMACAL, in the optical and near-infrared. Many of these galaxies are likely to be massive, dusty, and star-forming systems, but their true nature remains uncertain without rest-frame optical diagnostics. To address this, the ALMACAL team was awarded observing time with the VLT/X-shooter (ID: 115.285Y.001, PI: V. Bollo) to obtain broad spectral coverage across UVB, VIS, and NIR ranges for a carefully selected subset of CO emitters. This

was optimised to enable the high signal-to-noise detection of nebular emission lines, such as Hα, Hβ, [O II], [O III], and [N II]. This program is strategically designed to confirm the redshifts of CO emitters, explore their gas-phase metallicity, dust attenuation, star formation activity, and ultimately tie together the molecular gas content from ALMA with the ionised gas properties observable in the optical regime.

A key goal of this follow-up is to accurately derive the molecular gas mass. These observations aim to mitigate uncertainties in converting high$-J$ CO transitions to CO(1$-$0), by quantifying the fluxes of different CO transitions and comparing them to theoretical models. Although initial redshifts were inferred by matching CO line frequencies with predictions from semi-analytic models (`SHARK-2`, Lagos et al. 2018), there is still a risk of misidentifying a CO transition. Redshift confirmation via Hα or [O III] ensures a secure line identification and enables a reliable measurement of galaxy properties. This step is crucial, as it provides the fundamental basis for calculating molecular gas mass from CO. Ultimately, these new observations have the potential to reduce the uncertainty in estimates of the molecular gas mass out to redshift $z \sim 2.5$ by a factor of 5.

Once the redshift has been confirmed, the SFR can be derived from the Hα luminosity (Kennicutt 1998; Kennicutt & Evans 2012), which directly traces ionising photons from young stellar populations. Combined with the molecular gas masses derived from ALMA CO fluxes (Daddi et al. 2010; Tacconi et al. 2020), one can calculate gas depletion timescales and star formation efficiencies. These quantities are key for understanding how efficiently galaxies convert gas into stars, how this efficiency evolves with redshift and stellar mass, and how these trends relate to the findings presented in Chapter 3. The link between SFR and H_2 content underpins models of galaxy growth and regulation (e.g. Péroux & Howk 2020; Peroux & Nelson 2024), and this sample offers the opportunity to investigate that relationship in galaxies that bridge the observed discrepancy between obscured and unobscured star formation at intermediate redshifts ($z \sim 0.5 - 1.5$, Bouwens et al. 2020).

Additionally, the gas-phase metallicities can be estimated using strong-line diagnostics such as O_3N_2, R_{23}, and the [O III]/[O II] ratio (Maiolino et al. 2008; Kewley & Ellison 2008; Curti et al. 2017). These metallicity indicators allow placing these galaxies on the mass–metallicity relation. The metallicity is a key evolutionary tracer, linked to past star formation, gas accretion, and outflows. In particular, understanding whether CO-rich galaxies are metal-poor or enriched provides insight into whether their gas is pristine (inflow-driven) or processed (recycled via stellar evolution). Furthermore, the observed

Hα/Hβ ratio (the Balmer decrement) will be used to estimate dust attenuation (Calzetti et al. 1994; Groves et al. 2012), an essential correction for accurate SFR and metallicity estimates.

Beyond gas properties, stellar masses can also be inferred through SED fitting, combining X-shooter optical/NIR data with ALMA observations. Using models such as BAGPIPES (Carnall et al. 2018), CIGALE (Boquien et al. 2019), or pPXF (Cappellari 2023), one can fit both continuum and absorption features to constrain stellar mass, population age, and star formation histories. These fits can be used to derive gas-to-stellar mass ratios and to explore whether CO-rich galaxies represent the early stages of assembly, or whether they are already massive systems that are gas-rich. The stellar mass is a key parameter in evaluating the role of these galaxies in building up the cosmic stellar mass density, whether they lie on the star-forming main sequence or diverge toward starburst or quiescent phases.

With all these measurements (redshift, SFR, metallicity, stellar mass, and dust content), the cosmic molecular gas mass density (ρ_{H_2}) can be reconstructed in the redshift range $z \sim 0.03 - 1.65$. By calculating the CO-to-H_2 conversion factor (α_{CO}) informed by metallicity (e.g., Bolatto et al. 2013; Accurso et al. 2017), one can improve the fidelity of gas mass estimates and refine the integrated cosmic H_2 mass density across time. This is a critical step toward explaining the decline of the cosmic star formation rate since $z \sim 2$ (Madau & Dickinson 2014), and assessing whether this decline is due to diminished gas supply or inefficiencies in star formation.

The combination of ALMACAL and the VLT/X-Shooter spectra enables the interplay between gas and stars to be explored in unprecedented detail. By focusing on a sample selected solely by molecular gas content, rather than optical brightness or infrared luminosity, a particular slice of the galaxy population can be probed These galaxies may hold the key to understanding the obscured, dusty star formation that dominates the Universe at intermediate redshift (Elbaz et al. 2011; Tacchella et al. 2021). This program will provide a benchmark for combining CO surveys in ALMA and future instruments like JWST and ELT, as it will unravel the molecular gas properties and their connection to star formation, metal enrichment, and dust production. Ultimately, these insights will refine the picture of how galaxies assemble their stars, metals, and dust across cosmic history.

5.2.2 Legacy value of the MUSE-ALMA Haloes Survey

The MUSE-ALMA Haloes survey provides a multi-wavelength dataset that combines deep VLT/MUSE integral-field spectroscopy, ALMA CO observations, HST imaging and high-resolution UV spectroscopy with VLT/UVES to probe the CGM of H I-selected systems. In this survey, probes of neutral atomic gas, ionised gas, molecular gas, and stellar components have been assembled for a sample of 79 galaxies at $z \sim 0.5$. This rich dataset allows a baryon census in and around galaxies, bridging emission-line diagnostics (star formation rates, metallicities) with absorption-based estimates (H I column densities) and molecular gas masses (CO emission). The results presented in Chapter 4 already show that H I–selected galaxies can host unusually large reservoirs of cold molecular gas given their star formation rates, implying long gas depletion times.

This legacy survey opens new avenues for research, including:

- **Molecular and ionised gas kinematics**: Joint modelling of CO and nebular emission lines, such as [O III], in the same galaxies will reveal whether different gas phases co-rotate or diverge (Rahmani et al. 2018a, Barfety et al. in prep.). In the pilot system Q2131-G1, the CO(3−2) and [O III] emission lines trace the same rotating disk, yet the H I absorption has an additional velocity component interpreted as infalling gas (Szakacs et al. 2021). More generally, detailed 3D kinematic modelling of ionised gas in 48 galaxies in MUSE-ALMA Haloes shows that roughly 10% of strong H I absorbers are consistent with accreting gas, and $\sim 30\%$ with outflowing gas, with the remainder tracing disk and intra-group material (Weng et al. 2023a). These coupled kinematic studies can quantify disk rotation, turbulence and inflow/outflow signatures in the multi-phase ISM.

- **JWST follow-up opportunities:** The galaxies in the MUSE-ALMA Haloes survey are good candidates for mid-infrared and near-infrared follow-up with JWST. MIRI and NIRCam observations can provide critical information on dust, molecular gas and stars, completing the baryon cycle picture. MIRI can map PAH emission features (3.3 μm, 7.7 μm, 11.3 μm) to trace dust properties, radiation fields and obscured star formations, while H_2 rotational lines (2.12 μm) can reveal diffuse molecular gas, complementing the CO maps from ALMA. Meanwhile, NIRCam will be able to constrain stellar masses, ages, and obscured star formation via broad-band imaging, closing the baryon budget when combined with the ionised and molecular gas phases.

- **Comparison to cosmological simulations:** The MUSE-ALMA Haloes survey observations can be directly compared to state-of-the-art hydrodynamical models to test key predictions about gas in galaxies. The TNG50 simulation predicts that galaxy haloes contain vast numbers of cold gas "cloudlets" ($\sim 10^4 - 10^5\,M_\odot$) formed by thermal instabilities (Nelson et al. 2020), which can be tested by measuring the covering fraction and kinematics of cold gas in the survey data. The COLDSim simulation (Maio et al. 2022) includes detailed non-equilibrium chemistry, providing insights into the CO-dark molecular gas that may be present in the studied systems. Upcoming COLIBRE simulations (Schaye et al. in prep., Chaikin et al. in prep.) will further advance these comparisons by modelling dust formation, destruction, and its role in enhancing H_2 formation in diffuse gas (Trayford et al. 2025). Meanwhile, the Magneticum simulation (Dolag et al. 2016) incorporates magnetic fields, offering predictions about the stability of CGM filaments seen in H I-selected systems. The MillenniumTNG simulation (Hernández-Aguayo et al. 2023) provides a framework to test whether H I absorbers at $z \sim 0.5$ trace predicted cosmic web structure. Finally, the CIELO simulations (Tissera et al. 2025) predict how gas accretion and feedback create metallicity gradients in galaxies (Tapia-Contreras et al. 2025). The MUSE-ALMA survey's spatially resolved metallicity measurements will test for these predicted features (inner drops, outer breaks, or central enhancements), linking CGM processes to galactic chemical evolution.

- **Galaxy environment and group dynamics:** The MUSE ALMA Haloes survey has revealed that absorption systems are frequently found in rich environments, with the majority of absorbers being associated with other galaxies ranging from 3 to 11. Observational efforts have demonstrated that the group environment has a significant impact on the kinematics and distribution of circumgalactic gas (e.g., Pointon et al. 2017; Nielsen et al. 2018; Nateghi et al. 2024). Absorbing gas in group environments often exhibits broader velocity spreads and larger equivalent widths than in isolated systems, consistent with contributions from multiple galaxies through outflows, tidal stripping, or gas recycling. The MUSE–ALMA Haloes survey, combining MUSE kinematics with ALMA CO observations, offers a unique opportunity to study how the environment shapes gas flows in different phases, probing how tidal interactions, mergers, and group potential wells influence gas accretion and removal.

- **Full census of the baryon cycle:** By design, the sample is H I-selected, so these galaxies are expected to host large amounts of atomic gas. Using the MUSE photometry, it is possible to derive accurate rest-frame B-band magnitudes (M_B), which have been empirically linked to the H I mass ($M_{\rm H\,I}$) by Dénes et al. (2014). This relation is based on an H I-selected sample closely matching our survey selection and has been shown to remain valid out to $z \sim 1$ via stacking studies (Chowdhury et al. 2024). By inserting each galaxy's M_B into this relation, it is possible to predict its H I mass (O'Beirne et al. in prep.). All these components tie into a coherent view of the galactic baryon cycle: how gas inflows from the CGM feed star formation, and how feedback returns gas to the halo. The MUSE-ALMA Haloes results highlight that the baryon cycle directly affects the dense neutral gas required for star formation. In the Q2131–G1 example, infalling neutral gas provides a potential fuel source, while outflowing components hint at recycling. Overall, this legacy dataset gives an unprecedented panoramic view of stars, H I, H_2 and ionised gas in galaxies at $z \sim 0.5$, establishing a foundation for many studies of galaxy evolution beyond the initial CO detections.

5.2.3 Big Picture and Concluding Remarks

To understand how galaxies evolve, we need to combine observations from multiple telescopes covering different wavelengths and state-of-the-art simulations. This section summarises current and future efforts connected to the results of this thesis, showing how they contribute to the broader picture of the galaxy evolution research field. Thinking carefully about current gaps can guide the upgrading and development of future facilities, which is essential to push forward our understanding of the cosmic gas census and the processes driving galaxy growth.

Ionised gas in galaxies has been extensively studied from the nearby Universe to the epoch of reionisation ($z > 6$). At low and intermediate redshifts, large surveys like SDSS, CALIFA, SAMI, MaNGA, and GAMA have mapped ionised gas in thousands of galaxies, revealing trends in metallicity, ionisation, and feedback (Sánchez et al. 2012; Bryant et al. 2015; Bundy et al. 2014; Hopkins et al. 2013). MUSE on the VLT has pushed this further by producing spatially resolved maps up to redshift $z \sim 6$, particularly through Lyα emission (Wisotzki et al. 2016; Leclercq et al. 2020; Puga et al. 2025). At higher redshifts

($z \sim 6-10$), JWST has detected rest-frame optical emission lines providing detailed insights into star formation and chemical enrichment (Curti et al. 2023; Cameron et al. 2023). New observational campaigns will shed light on how the first galaxies enriched their gas, whether stars or AGN dominated reionisation, and how black holes grew so rapidly (Hayes et al. 2023). In parallel, DESI is collecting high-resolution spectra of millions of galaxies and quasars across $0.4 < z < 3.5$, allowing absorption-line studies of ionised and neutral gas in the CGM and IGM (Napolitano et al. 2023; Chen et al. 2025; Dehghanian et al. 2025; Sharma et al. 2025). The hot ionised medium has been revealed by X-ray surveys from eROSITA in galaxy clusters and filaments across $z \sim 0-1.3$ (Merloni et al. 2024; Comparat et al. 2023; Yeung et al. 2023, 2024). In the far-infrared, fine-structure lines such as [N II], [O I], and [O III], observed with ALMA and Herschel, provide complementary diagnostics at high redshift, especially for dusty or optically faint galaxies (Looze et al. 2014; Breuck et al. 2019; Hodge & da Cunha 2020).

Atomic hydrogen has been detected across different redshifts through H I emission and Lyα absorption. At low redshift ($z \lesssim 0.1$), surveys like ALFALFA (Arecibo), HIPASS (Parkes), and WALLABY (ASKAP), MHONGOOSE (MeerKAT) have mapped the H I content of tens of thousands of galaxies (Jones et al. 2018; Zwaan et al. 2005; de Blok et al. 2024; Koribalski et al. 2024). WALLABY alone will detect over 500,000 galaxies across 14,000 deg^2 out to $z \lesssim 0.26$. At intermediate redshift ($z \lesssim 1$), direct H I detections become more difficult, but deep surveys such as CHILES (VLA), MIGHTEE-H I, DINGO (ASKAP) and LADUMA (MeerKAT) are now pushing the limits, using deep integrations and stacking to trace H I emission out to $z \sim 1.5$ (Fernández et al. 2016; Maddox et al. 2021; Baker et al. 2018; Rhee et al. 2023; Cook et al. 2024). At higher redshift ($z \gtrsim 2$), H I is primarily studied via Lyα absorption in DLAs, which dominate the neutral gas budget. Future surveys will also measure H I in 21-cm absorption against background radio sources, probing the cold neutral medium (Péroux et al. 2003; Prochaska & Wolfe 2009; Crighton et al. 2015).

Molecular gas has been mapped through CO and [C II] emission. At low redshift ($z \lesssim 0.3$), surveys like xCOLD GASS, PHANGS-ALMA, and ALMaQUEST use ALMA and NOEMA to measure gas masses and star formation timescales in thousands of galaxies (Leroy et al. 2021; Saintonge et al. 2017; Lin et al. 2020; Bolatto et al. 2017). At intermediate redshift ($0.5 \lesssim z \lesssim 3$) deep ALMA surveys like ASPECS, PHIBSS, and COLDz show that galaxies have higher molecular gas fractions (50–80%) and depletion times of

200–500 Myr (Decarli et al. 2020; Tacconi et al. 2018; Riechers et al. 2019). At $z > 4$, [C II] 158 μm becomes the main tracer. Surveys like ALPINE and REBELS have detected [C II] and dust emission, suggesting rapid early metal enrichment (Pozzi et al. 2024; Bowler et al. 2024). CRISTAL has recently resolved [C II] dynamics in $z \sim 4.5$ galaxies, showing signs of rotating disks (Villanueva et al. 2024; Herrera-Camus et al. 2025; Posses et al. 2025). However, detecting more typical galaxies at these early times remains challenging, as current instruments primarily observe only the most massive ($\gtrsim 10^{10}\,M_\odot$) or lensed systems (Boogaard et al. 2023; Catán et al. 2024).

Looking ahead, the next decade promises a revolution in how we map gas (atomic, molecular, and ionised) throughout cosmic history. New surveys across radio, millimetre, optical, and X-ray wavelengths are coming online, while powerful observatories will probe the CGM and IGM with unprecedented detail. These efforts will help pin down the cosmic gas census and trace its evolution, ultimately filling long-standing gaps in understanding how gas fuels galaxy growth.

Ionised gas studies are advancing quickly at all redshifts. At $z > 4$, JWST is routinely detecting rest-frame optical lines like Hα, [O III], and Hβ, as well as warm H_2 and PAH features (Duncan et al. 2025; Carniani et al. 2024; Shivaei & Boogaard 2024). At intermediate redshift ($z \sim 1-3$), wide-field spectroscopic surveys with Euclid and the Roman space telescopes will observe emission lines in millions of galaxies over wide areas, which allows statistical studies of ionised gas across environments (Laureijs et al. 2011; Yung et al. 2023; Collaboration et al. 2025). On the ground, instruments such as MOONS on the VLT will collect near-IR spectra for over a million galaxies in this same redshift range (Cirasuolo et al. 2020). Meanwhile, the new 30–40 m class telescopes will push ionised gas mapping to sub-kiloparsec resolution in high$-z$ galaxies. The Extremely Large Telescope (ELT, Padovani & Cirasuolo 2023; Cirasuolo et al. 2024) is expected to start its operations toward the end of this decade. The Giant Magellan Telescope (GMT, Fanson et al. 2022; Burgett et al. 2024) is projected to become operational during the early 2030s. The Thirty Meter Telescope (TMT, Sanders 2013; Kirshner 2024) remains uncertain; construction delays have pushed any possible first light to around 2033 or beyond.

Atomic gas will be transformed by new radio facilities. SKA Phase 1 (SKA1) will conduct blind H I surveys out to $z \sim 1.5$ in emission, detecting H I masses as low as $10^9\,M_\odot$, with sub-arc second resolution and μJy sensitivity (Bourke et al. 2018). SKA Phase 2 will further improve the sensitivity and angular resolution, mapping H I across

$0 < z < 2$, resolving gas disks, haloes, and cosmic web structures. In just 10 hours per pointing, it will detect H I column densities between $10^{19} - 10^{18}\,\mathrm{cm}^{-2}$ with $5 - 50''$ resolution. Over 100 hours, SKA will explore cold gas accretion and cosmic filaments, placing tight constraints on their role in sustaining star formation (Huynh & Lazio 2013; Hotan et al. 2021). In the Epoch of Reionisation ($z > 6$), SKA1-Low will statistically detect the redshifted 21-cm emission from the diffuse IGM (out to $z \sim 10$), mapping ionised bubbles with $\sim 600 - 1000$ h integrations (Koopmans et al. 2015; Mondal et al. 2020; Minoda et al. 2023). Complementary facilities like the Deep Synoptic Array-2000 (DSA-2000) with 2000 antennas and a dense short-baseline configuration, will survey the entire Northern Hemisphere and detect H I at column densities $N_{\mathrm{H\,I}} \gtrsim 10^{20}\,\mathrm{cm}^{-2}$ and 3" ($\sim 100\,\mathrm{pc}$) resolution in only 4 hours (Hallinan et al. 2019).

On the simulation side of molecular gas, recent advancements in modelling have been made, yet challenges persist in resolving CO emission and detailed ISM physics (Popping et al. 2019b). Large cosmological simulations, such as Magneticum (Dolag et al. 2016), IllustrisTNG (Pillepich et al. 2018), FIRE (Hopkins et al. 2014), and MilleniumTNG (Hernández-Aguayo et al. 2023) predict gas properties across cosmic time but often employ simplified sub-grid recipes for H_2 and CO. While IllustrisTNG-based models reproduce CO luminosity functions, they highlight significant CO-dark gas in dwarf galaxies. This limitation can be attributed to the reliance of these models on post-processing rather than direct ISM modelling (Inoue et al. 2020). To address this, newer high-resolution simulations incorporate non-equilibrium chemistry and local radiation effects (Richings & Schaye 2016; Lupi et al. 2018, 2020; Thompson et al. 2024), enabling more accurate CO–H_2 relations and star formation predictions. For example, ColdSIM shows coherence with observed molecular gas densities from $z \sim 7$ to 2 (Maio et al. 2022). The forthcoming COLIBRE simulations will explicitly model dust evolution (tracking carbon/silicate grains in multiple size bins) and its coupling to cold ISM phases, providing more realistic H_2 formation rates in diffuse gas (Schaye et al. in prep.). Radiative transfer tools like SÍGAME3 and SLICK (Olsen et al. 2021; Garcia & Narayanan 2023) enable line-emission predictions across gas phases, but full multi-phase modelling remains computationally demanding (Casavecchia et al. 2024).

On the observational side of molecular gas, ALMA and other sub-mm facilities continue to push faster and deeper cold-gas observations. The Wideband Sensitivity Upgrade (WSU), a keystone of the ALMA2030 roadmap, will expand the total bandwidth from

8 GHz to up to 32 GHz. It includes new Band 2, Band 6v2, Band 7v2, and Band 8v2 receivers, advanced digitisers, and a high-speed data system, along with a new correlator and GPU-based spectrometer (Carpenter et al. 2019, 2020, 2023). As a result, continuum imaging will be $\sim 3\times$ faster with doubled bandwidth, and spectral line imaging up to $\sim 2-3\times$ faster. At high spectral resolution (0.1−0.2 km/s), the achievable instantaneous bandwidth will increase by factors of 10–100, improving survey efficiency and coverage per tuning. This upgrade means ALMA will confidently detect and identify CO transitions in galaxies down to $L_{\rm CO} \sim 10^9$ K km s^{-1} pc^2 in a few-hours integration. Estimates of the molecular gas masses as low as $M_{\rm H_2} \sim 10^9$ $M_\odot$ at $z \sim 1$ will be possible, and down to $10^8 M_\odot$ at $z < 0.5$. This will enable robust statistical reconstructions of the H_2 mass density, $\rho_{\rm H_2}$, and better constraints on gas depletion timescales across time.

Meanwhile, the planned Next Generation Very Large Array (ngVLA), which is currently under design and pending funding (McKinnon et al. 2019), will achieve milliarcsecond resolution and a tenfold sensitivity improvement. The ngVLA is designed to operate from 1.2 to 116 GHz with baselines up to 30 times longer than the current VLA. It will be optimised to trace low− to mid−J CO transitions, dense gas tracers like HCN and HCO$^+$, and dust continuum, especially at $z \sim 2$−5 (Riechers et al. 2017; Di Francesco et al. 2019; Murphy & ngVLA Science Advisory Council 2020). Simulations show that the ngVLA could map CO(1−0) emission at ∼1 kpc scales in $z = 2-4$ galaxies in reasonable times, where ALMA would require more than 100 hours (Carilli et al. 2018).

Looking further ahead, the proposed 50-m Atacama Large-Aperture Submillimeter Telescope (AtLAST) is designed for wide-field CO and [C II] surveys at $0 < z < 3$ (Klaassen et al. 2020). With a huge collecting area, AtLAST could survey 1000 deg^2 at sub−μJy depths in $\sim 1,000$ hours. AtLAST could conduct blind line surveys (e.g., [C II], CO) across wide fields, detecting dusty galaxies up to $z \sim 5$. In practice, AtLAST will have $\sim 200\times$ the sky coverage of a 12 m dish in one pointing, and confusion limits thousands of times lower than current 6–12 m telescopes (van Kampen et al. 2024).

In parallel, ESO's Expanding Horizons initiative is now engaging the global community to envision the next transformational facility for the 2040s and beyond. This could take multiple forms: a new telescope (optical/IR/radio), a major upgrade to existing facilities (e.g., ALMA2040), or even a hybrid system combining ESO and partner capabilities. The selection process, driven by community input through workshops and proposals starting in 2026, will prioritise transformational science beyond the reach of the ELT.

Current CO surveys remain limited by bandwidth, sensitivity, and sky coverage. Deep ALMA state-of-the-art programs like ASPECS cover just a few arcmin^2 with cosmic variance dominating their uncertainties (Gkogkou et al. 2022). Surveys such as ALMACAL attempt to address this by exploiting serendipitous detections across the archival ALMA calibrator data to mitigate the effects of the field-to-field variance. The rising CMB temperature at high redshift reduces CO line contrast and alters excitation conditions (da Cunha et al. 2013; Bolatto et al. 2013; Vallini et al. 2015), biasing surveys toward IR–luminous, metal–rich systems and potentially missing large reservoirs of CO-dark gas. To overcome these challenges, alternative tracers like [C II] and polycyclic aromatic hydrocarbons (PAHs) have emerged as valuable tools. The [C II] 158 μm line, accessible to ALMA Band 3 out to $z \sim 9$, correlates well with molecular gas mass up to $z \sim 2$ (Zanella et al. 2018), though it requires careful calibration due to its mixed origin in ionised, atomic, and molecular phases (Casavecchia et al. 2025). ALMA programs such as ALPINE, REBELS, and CRISTAL are actively mapping [C II] in galaxies at $z = 4-7$. Meanwhile, PAH features in the mid-IR, traced by JWST/MIRI, show a tight relation to CO luminosities from $z \sim 0$ to 4 (Shivaei & Boogaard 2024), offering a stable proxy for H_2. Because PAHs trace shielded regions in molecular clouds, they also provide insight into cold gas conditions in systems where CO is faint or absent (Esparza-Arredondo et al. 2025). In combination, CO, [C II], [C I], and PAH diagnostics promise a more complete and unbiased picture of the cold ISM across cosmic time.

Despite the planned upgrade of ALMA, the collecting area (50 12-m antennas) and maximum baselines (16 km) remain unchanged. While WSU increases speed and spectral coverage, it does not significantly enhance the sensitivity needed to detect faint lines at high redshift. As a result, even in the 2030s, deep blind ALMA surveys will primarily detect bright CO emitters, while typical star-forming galaxies at $z \gtrsim 5$ will remain below detection thresholds. Observing these systems still requires tens of hours per target, and survey volumes remain small. In contrast, upcoming ionised and H I surveys will cover larger cosmic volumes, detecting low-mass and faint galaxies, and enabling statistical studies that will remain out of reach for molecular gas. To bring molecular gas studies into parity with those of atomic and ionised gas, a major investment in an ALMA2040 vision is needed (Baryshev 2024). Key steps include: increasing the collecting area by adding 12-m antennas to boost the sensitivity, implementing a next-generation GPU-based correlator capable of handling more than 70 antennas and large bandwidths (64 GHz or more), and

extending maximum baselines to 30–50 km to resolve galaxies on sub-kiloparsec scales at high redshift. In addition, modernising ALMA's data processing pipeline is critical, moving beyond the `CASA` software (McMullin et al. 2007), which struggles with scaling, reproducibility, and performance in big-data contexts.

To effectively probe the faint end of the CO and [C II] luminosity functions at high redshift ($z \gtrsim 3-6$), we need significantly deeper sensitivities than what is currently feasible in fairly short integration times. Low-mass, star-forming galaxies, which are most representative of the general population, emit weak molecular and atomic cooling lines that standard surveys cannot detect within 1−2 hour exposures. Deep ALMA surveys like ASPECS (Walter et al. 2016; Decarli et al. 2019) have shown that blind line search completeness exceeds 90% for lines with integrated fluxes $\gtrsim 0.2$ Jy km/s and peak flux densities > 1 mJy. These thresholds effectively set the detection floor for current surveys: below them, the probability of recovering a line drops rapidly, especially at the faint end of the luminosity function. As shown in both models (Popping et al. 2016) and observational data (e.g., Lagache et al. 2018), many typical star-forming galaxies at $z \gtrsim 3-6$ exhibit CO and [C II] line fluxes of 0.05-0.2 Jy km/s, with peak fluxes in the range of 0.5−2 mJy for line widths of ∼100 km/s.

For example, a galaxy at $z \sim 2$ with CO(1−0) luminosity of $L' \lesssim 10^9$ K km/s pc^2 produces a line flux on the order of 0.05-0.1 Jy km/s. Assuming a typical line width of 100 km/s, this translates into peak flux densities of ∼0.5–1 mJy. Similarly, at $z \sim 6-8$, [C II] line emission from galaxies with modest star formation rates ($\sim 10\,M_\odot\,\mathrm{yr}^{-1}$) results in integrated fluxes of ∼0.05–0.2 Jy km/s and peak fluxes of ∼0.5–2 mJy. To detect such galaxies individually, rather than statistically via stacking, observations have to reach an rms sensitivity of $\lesssim$ 0.2−0.3 mJy/beam per 50–100 km/s channel. Currently, ALMA achieves this sensitivity only with integration times exceeding 10–20 hours per source. To reduce this to a time frame of $\sim 1-2$ hours, it is necessary to enhance the sensitivity by a factor of $\sqrt{10} \sim 3.16$. This enhancement is required in order to follow the established relationship between the sensitivity of interferometers, as demonstrated in Eq. 1.4.2, which scales with the square root of the integration time and the number of antennas ($N(N-1)$). The ALMA2030 WSU will already improve the line sensitivity by a factor of 1.44, due to lower receiver noise and a more powerful correlator (Carpenter et al. 2020). To match the total required improvement, we need an additional factor of $3.16/1.44 = 2.2$, which must come from increased collecting area. To achieve this sensitivity gain from the collecting

area, we need $50 \times 2.2 \approx 110$ antennas. This would enhance the collecting area by 120%, which translates to a sensitivity boost of $\sqrt{(110 \cdot 109)/(50 \cdot 49)} \approx 2.2$. Combined with the $1.4\times$ factor, this yields the needed factor of ~ 3.16 improvement. Such a configuration would allow ALMA2040 to detect typical star-forming galaxies at $z \gtrsim 5-7$ with $M_{\mathrm{H_2}} \sim 10^9\,M_\odot$ in $\sim 1-2$ hours integration time, map transitions such as HCN, $\mathrm{HCO^+}$, and [C I] across large samples, and construct unbiased molecular gas scaling relations spanning $0 \lesssim z \lesssim 9$. Together with SKA and ELT, ALMA2040 would complete the picture of atomic, molecular, and ionised gas, offering a more complete view of galaxy evolution across cosmic time.

In conclusion, through the ALMACAL and MUSE-ALMA Haloes surveys, this thesis establishes a critical benchmark for future studies of galaxy evolution. It highlights that future research must be multiwavelength by design, combining optical, radio, submillimeter, and infrared data to trace the full lifecycle of baryons, from inflows and star formation to feedback and quenching. Over the next two decades, coordinated efforts across observatories are expected to provide a full census of baryons across cosmic time, clarifying the contributions of stars, dust, cold gas, and AGN activity to galaxy evolution. Such work will define how molecular gas connects to galaxy properties, trace the evolution of CO and H I emission, and test how feedback affects galaxies. Reaching these goals needs data from future facilities: SKA and ELT will target atomic and ionised gas, while ALMA2030 will improve molecular gas measurements. Detecting typical, faint galaxies at high redshift, however, requires ALMA2040 featuring more antennas, a broader bandwidth, and longer baselines. Ultimately, to fully understand these observations, we must pair them with theoretical models and cosmological simulations containing cold gas ($< 10^4\,\mathrm{K}$). Only by combining deep, multi-phase observations with theory will the physical processes driving galaxy evolution over cosmic time be uncovered.

Appendix A

Appendix to Chapter 2

We display the spectra and the moment map of the 87 detections found in the ALMACAL−22 survey, identified as CO emitters.

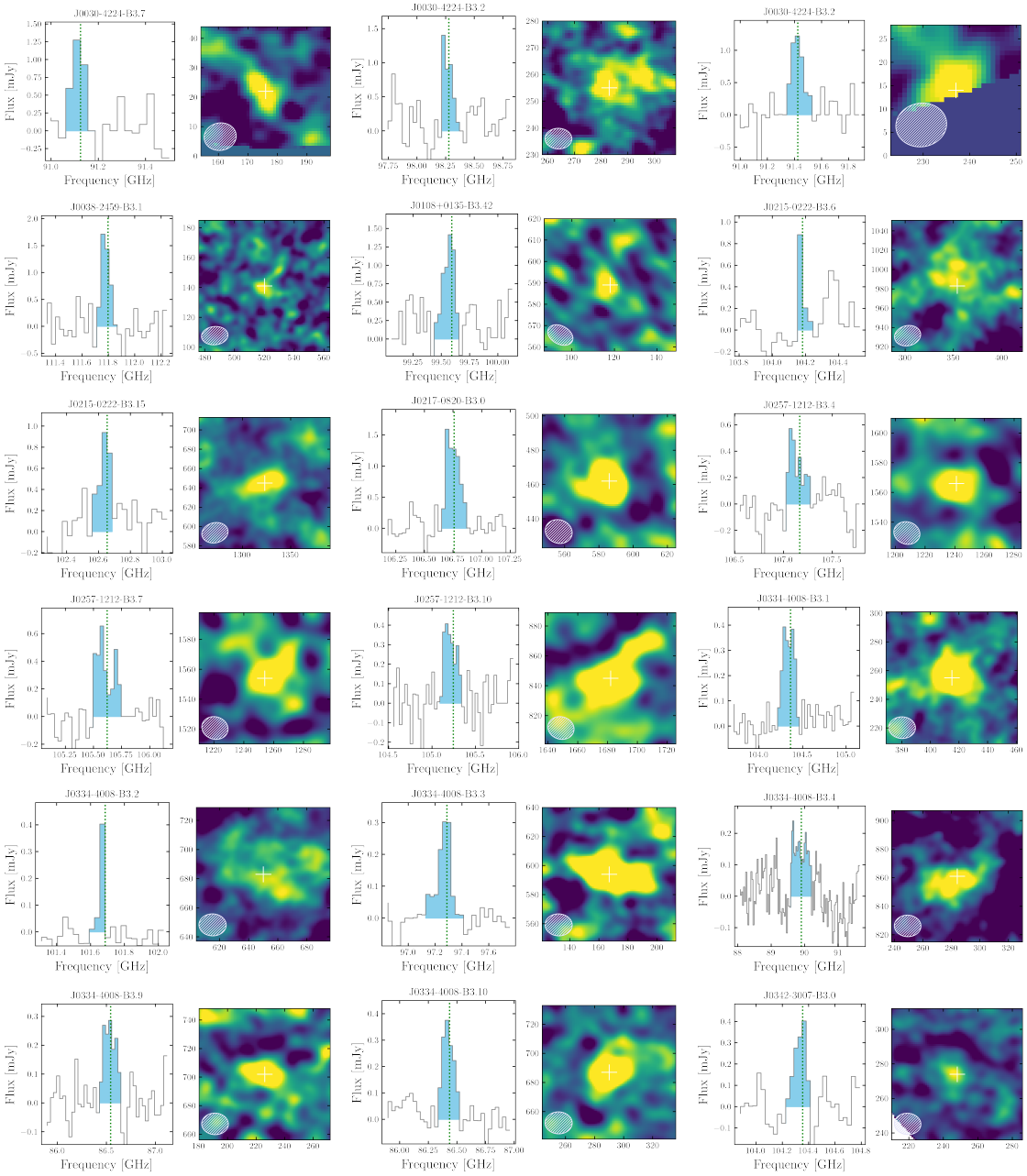

Figure A.1: ALMACAL−22 emission lines detections. Left panel: Spectrum of the emission lines at the brightest pixel. Right panel: The emission line moment map centred on the detection range. The parameters of all candidates are summarised in Table A.1.

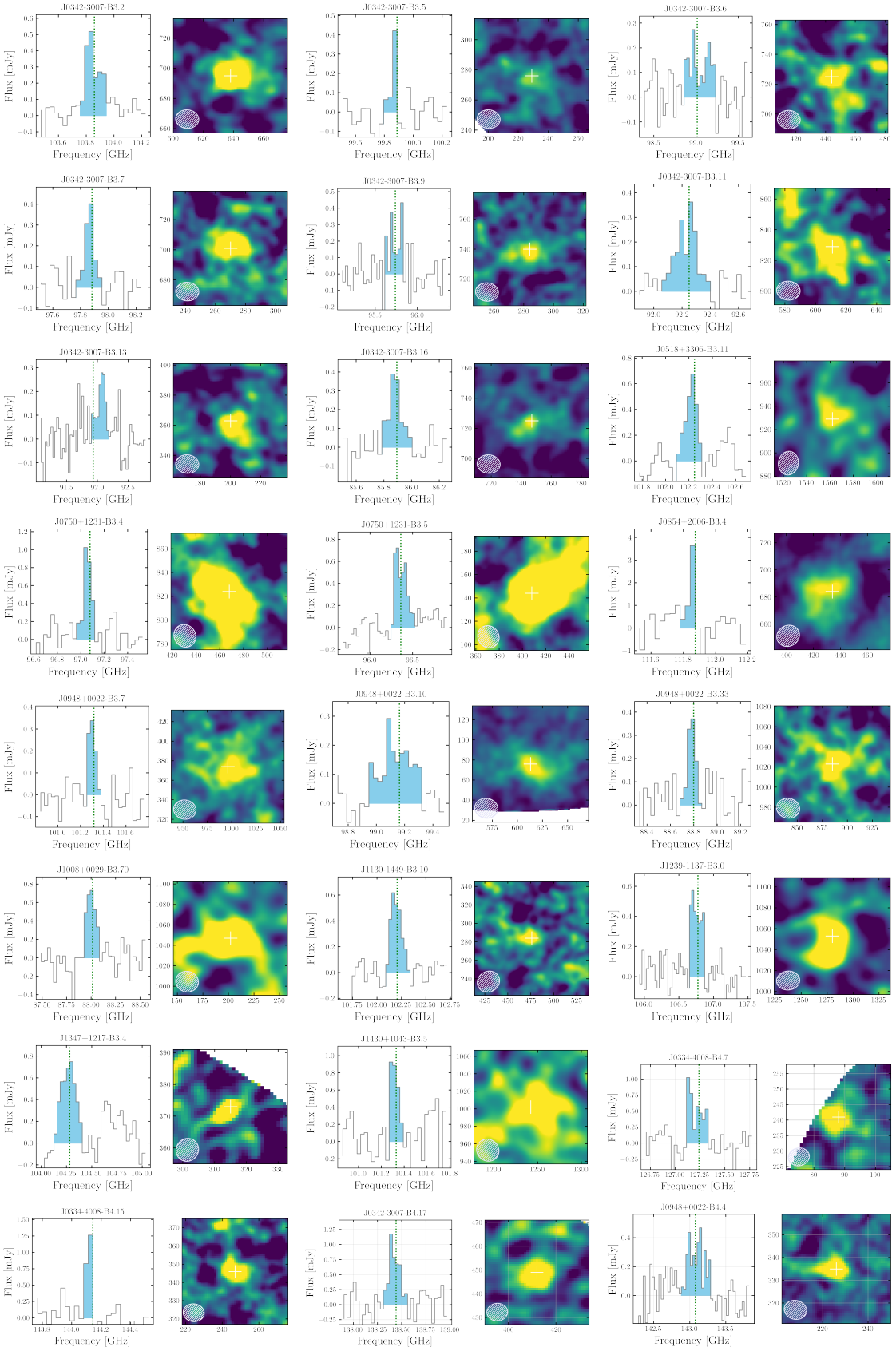

Figure A.2: continued.

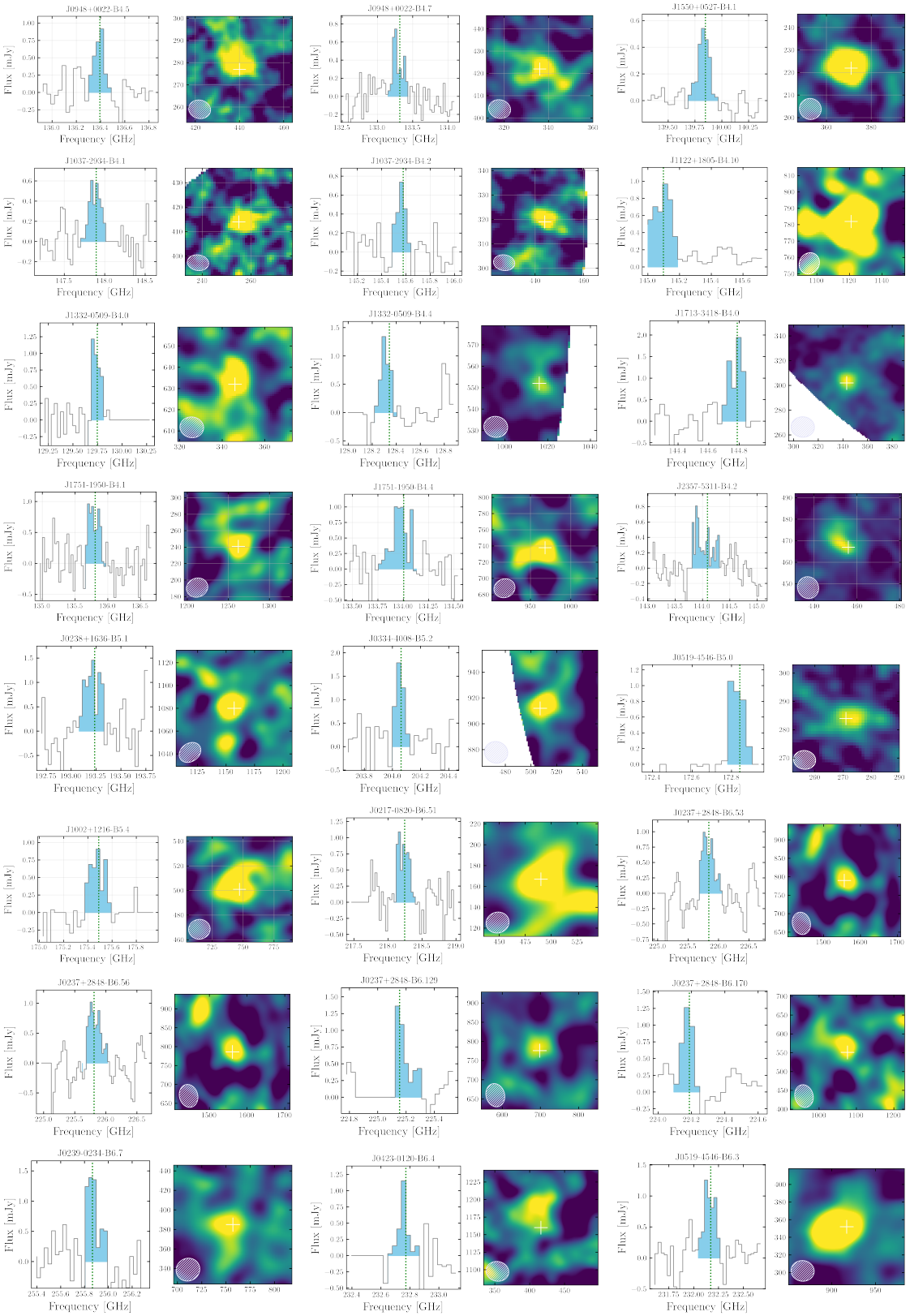

Figure A.3: continued.

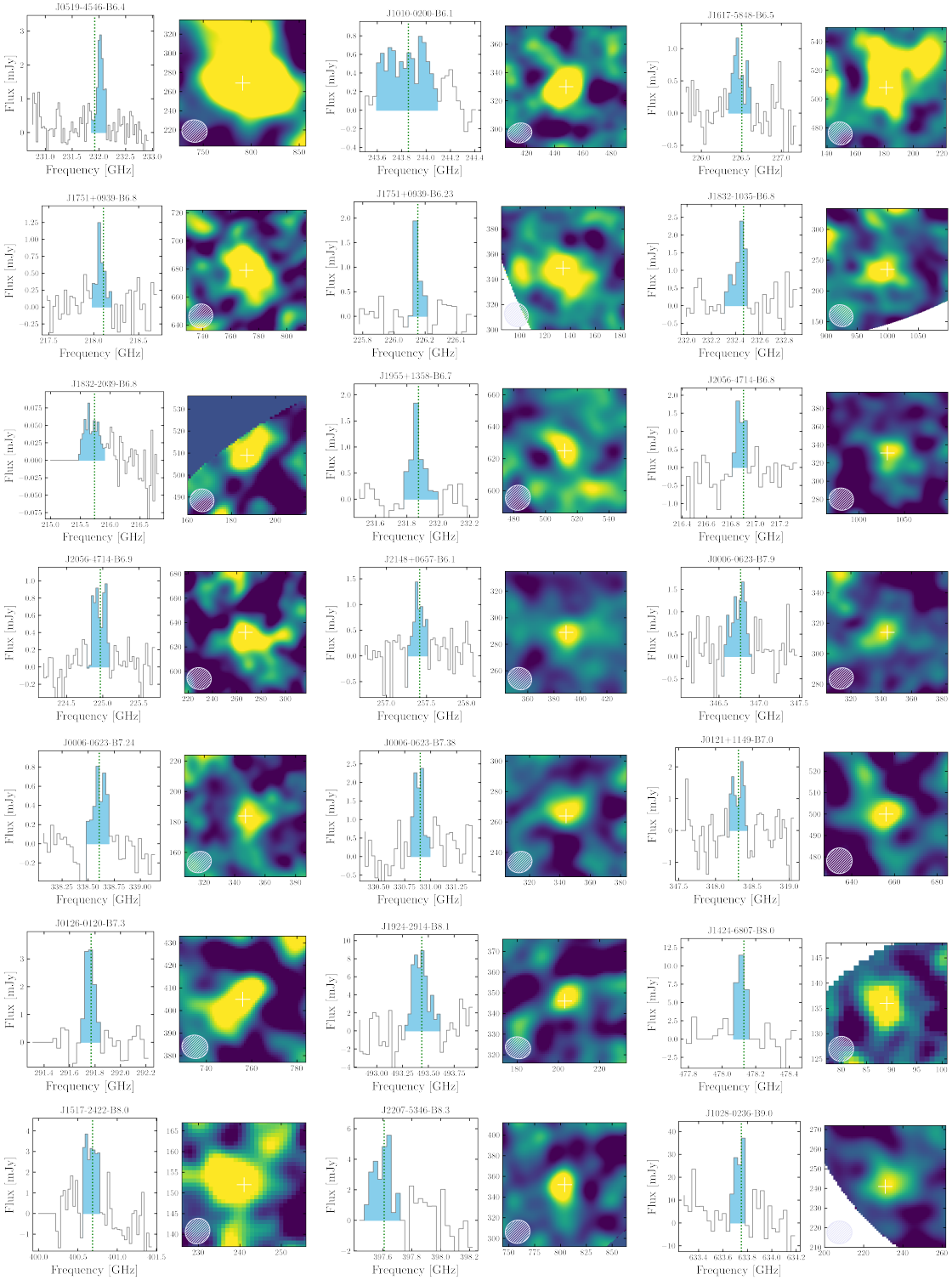

Figure A.4: continued.

Table A.1: Physical properties of the ALMACAL−22 CO detections

Name.ID	Frequency [GHz]	Flux [mJy km s^{-1}]	FWHM [km/s]	S/N	Completeness %	Reliability %
J0030-4224.2	91.4	643^{+260}_{-560}	318^{+34}_{-28}	7.24	0.78	0.27
J0030-4224.7	91.1	919^{+27}_{-37}	202^{+18}_{-18}	7.96	0.81	0.19
J0030-4224.2	98.3	942^{+19}_{-26}	287^{+42}_{-36}	6.43	0.80	0.27
J0038-2459.1	111.8	159^{+789}_{-159}	182^{+26}_{-28}	8.60	0.83	0.39
J0108+0135.42	99.6	946^{+22}_{-26}	338^{+33}_{-33}	8.55	0.82	0.52
J0215-0222.6	104.2	1000^{+0}_{-13}	53^{+51}_{-3}	10.48	0.81	0.25
J0215-0222.15	102.6	796^{+94}_{-252}	275^{+261}_{-69}	10.10	0.81	0.42
J0217-0820.0	106.7	854^{+62}_{-96}	384^{+69}_{-66}	12.43	0.79	0.71
J0257-1212.4	107.2	394^{+521}_{-394}	303^{+164}_{-125}	7.42	0.78	0.53
J0257-1212.7	105.6	929^{+56}_{-260}	300^{+253}_{-72}	5.05	0.81	0.37
J0257-1212.10	105.2	900^{+27}_{-38}	473^{+60}_{-50}	6.73	0.81	0.57
J0334-4008.1	104.4	929^{+23}_{-36}	442^{+39}_{-32}	13.67	0.82	0.92
J0334-4008.2	101.7	991^{+9}_{-30}	53^{+62}_{-14}	10.62	0.81	1.00
J0334-4008.3	97.3	935^{+36}_{-77}	285^{+63}_{-40}	9.10	0.81	1.00
J0334-4008.4	89.8	860^{+23}_{-36}	1756^{+92}_{-96}	6.37	0.75	1.00
J0334-4008.9	86.5	839^{+90}_{-316}	461^{+49}_{-41}	6.46	0.83	0.40
J0334-4008.10	86.4	907^{+54}_{-116}	378^{+59}_{-57}	9.93	0.79	0.66
J0342-3007.0	104.3	732^{+212}_{-538}	255^{+16}_{-19}	9.39	0.80	0.49
J0342-3007.2	103.8	627^{+208}_{-118}	419^{+100}_{-171}	12.68	0.80	0.40
J0342-3007.5	99.9	993^{+3}_{-4}	54^{+2}_{-2}	12.24	0.87	0.17
J0342-3007.6	99.0	349^{+651}_{-345}	1009^{+1214}_{-863}	5.43	0.76	0.34
J0342-3007.7	97.9	911^{+62}_{-196}	203^{+38}_{-32}	13.00	0.86	0.42
J0342-3007.9	95.7	1000^{+0}_{-148}	61^{+476}_{-4}	4.31	0.78	0.33
J0342-3007.11	92.3	427^{+167}_{-256}	549^{+67}_{-63}	9.76	0.75	0.39
J0342-3007.13	91.9	983^{+12}_{-67}	443^{+220}_{-88}	4.52	0.77	0.53
J0342-3007.16	85.9	275^{+488}_{-273}	328^{+78}_{-80}	9.48	0.80	0.36
J0518+3306.11	102.2	879^{+80}_{-87}	295^{+58}_{-65}	8.52	0.82	0.38
J0750+1231.4	97.1	907^{+85}_{-321}	221^{+74}_{-67}	8.55	0.83	0.39
J0750+1231.5	96.3	888^{+23}_{-21}	485^{+36}_{-35}	8.94	0.76	1.00
J0854+2006.4	111.9	973^{+20}_{-74}	81^{+35}_{-33}	8.46	0.78	0.05
J0948+0022.7	101.3	983^{+7}_{-9}	192^{+19}_{-16}	9.48	0.81	0.50
J0948+0022.10	99.1	348^{+48}_{-50}	917^{+103}_{-100}	8.23	0.74	0.22

Table A.2: Continuation.

Name.NID	Frequency [GHz]	Flux [mJy km s^{-1}]	FWHM [km/s]	S/N	Completeness %	Reliability %
J1239-1137.0	106.7	921^{+50}_{-74}	471^{+95}_{-146}	9.91	0.76	0.86
J1347+1217.4	104.3	763^{+173}_{-503}	464^{+27}_{-26}	6.92	0.81	0.72
J1430+1043.5	101.3	975^{+22}_{-56}	213^{+73}_{-90}	7.62	0.79	0.30
J0334-4008.7	127.2	8^{+555}_{-8}	111^{+300}_{-64}	4.38	0.80	0.48
J0334-4008.15	144.1	1000^{+0}_{-4}	39^{+54}_{-2}	9.29	0.75	0.34
J0342-3007.17	138.4	530^{+347}_{-492}	272^{+35}_{-28}	5.18	0.81	0.35
J0948+0022.4	143.1	479^{+214}_{-388}	634^{+114}_{-98}	5.04	0.81	0.38
J0948+0022.5	136.4	959^{+32}_{-52}	192^{+37}_{-46}	5.40	0.81	0.34
J0948+0022.7	133.3	24^{+776}_{-24}	318^{+177}_{-191}	3.98	0.82	0.15
J1550+0527.1	139.8	874^{+80}_{-211}	238^{+37}_{-34}	9.91	0.79	1.00
J1037-2934.1	147.9	374^{+480}_{-364}	346^{+33}_{-28}	4.09	0.82	0.12
J1037-2934.2	145.6	965^{+15}_{-18}	159^{+13}_{-12}	7.96	0.81	0.32
J1122+1805.10	145.1	830^{+58}_{-83}	375^{+72}_{-49}	12.52	0.79	0.77
J1332-0509.0	129.7	952^{+21}_{-24}	214^{+34}_{-44}	6.75	0.79	0.31
J1332-0509.4	128.3	978^{+5}_{-5}	201^{+11}_{-9}	6.25	0.80	0.25
J1713-3418.0	144.8	136^{+725}_{-136}	222^{+17}_{-14}	3.86	0.80	0.54
J1751-1950.1	135.8	819^{+108}_{-342}	428^{+46}_{-34}	4.18	0.80	0.48
J1751-1950.4	134.0	776^{+161}_{-137}	372^{+122}_{-163}	6.59	0.81	0.29
J2357-5311.2	144.1	34^{+803}_{-34}	302^{+495}_{-69}	4.56	0.78	0.94
J0238+1636.1	193.2	743^{+109}_{-113}	277^{+80}_{-66}	6.67	0.79	0.30
J0334-4008.2	204.0	920^{+15}_{-15}	92^{+5}_{-6}	7.78	0.78	0.20
J0519-4546.0	172.8	959^{+36}_{-82}	138^{+51}_{-55}	5.44	0.81	0.33
J1002+1216.4	175.5	744^{+130}_{-162}	228^{+54}_{-56}	5.97	0.80	0.20
J0217-0820.51	218.2	899^{+26}_{-28}	266^{+23}_{-23}	6.68	0.79	0.69
J0237+2848.53	225.8	932^{+28}_{-35}	264^{+33}_{-35}	4.45	0.81	0.52
J0237+2848.56	225.8	925^{+31}_{-37}	268^{+35}_{-35}	6.08	0.80	0.76
J0237+2848.129	225.1	950^{+48}_{-350}	90^{+121}_{-62}	7.38	0.79	0.22
J0237+2848.170	224.2	979^{+19}_{-371}	97^{+108}_{-25}	6.39	0.79	0.22
J0239-0234.7	255.9	973^{+15}_{-42}	93^{+25}_{-12}	10.14	0.81	0.57
J0423-0120.4	232.8	924^{+36}_{-69}	62^{+16}_{-10}	9.46	0.78	0.50
J0519-4546.3	232.1	830^{+69}_{-100}	174^{+22}_{-17}	10.86	0.82	0.22
J0519-4546.4	231.9	998^{+1}_{-1}	170^{+5}_{-6}	12.11	0.82	0.71

Table A.3: Continuation.

Name.NID	Frequency [GHz]	Flux [mJy km s^{-1}]	FWHM [km/s]	S/N	Completeness %	Reliability %
J0948+0022.33	88.8	931^{+53}_{-300}	245^{+29}_{-29}	8.45	0.83	0.37
J1008+0029.70	88.0	935^{+27}_{-60}	367^{+25}_{-26}	9.90	0.80	0.66
J1130-1449.10	102.2	814^{+115}_{-369}	348^{+44}_{-49}	9.77	0.81	0.65
J1010-0200.1	244.0	252^{+147}_{-167}	574^{+942}_{-219}	4.44	0.78	0.99
J1617-5848.5	226.5	760^{+106}_{-311}	257^{+34}_{-40}	6.06	0.80	0.47
J1751+0939.8	218.1	987^{+13}_{-51}	141^{+46}_{-44}	7.08	0.79	0.49
J1751+0939.23	226.1	997^{+3}_{-17}	43^{+12}_{-19}	11.57	0.84	0.18
J1832-1035.8	232.4	981^{+13}_{-60}	92^{+29}_{-13}	10.60	0.82	0.57
J1832-2039.8	215.7	454^{+291}_{-418}	513^{+39}_{-30}	8.71	0.77	1.00
J1955+1358.7	231.9	783^{+116}_{-101}	107^{+26}_{-36}	11.02	0.82	0.21
J2056-4714.8	216.9	935^{+19}_{-23}	111^{+11}_{-10}	7.26	0.78	0.27
J2056-4714.9	225.0	849^{+52}_{-77}	297^{+58}_{-39}	5.09	0.81	0.39
J2148+0657.1	257.4	883^{+85}_{-242}	161^{+42}_{-36}	8.99	0.83	1.00
J0006-0623.9	346.8	267^{+449}_{-263}	181^{+117}_{-56}	6.38	0.78	1.00
J0006-0623.24	338.6	834^{+53}_{-45}	121^{+14}_{-14}	7.30	0.79	0.37
J0006-0623.38	330.9	907^{+81}_{-408}	115^{+154}_{-40}	5.49	0.81	0.30
J0121+1149.0	348.3	925^{+42}_{-48}	153^{+23}_{-24}	3.12	0.75	0.15
J0126-0120.3	291.8	837^{+126}_{-463}	83^{+7}_{-7}	7.65	0.78	0.30
J1924-2914.1	493.4	635^{+231}_{-497}	95^{+11}_{-10}	6.75	0.77	0.44
J1424-6807.0	478.1	715^{+97}_{-160}	42^{+16}_{-9}	4.65	0.73	0.70
J1517-2422.0	400.7	910^{+47}_{-133}	139^{+55}_{-23}	4.60	0.78	0.29
J2207-5346.3	397.6	809^{+93}_{-285}	112^{+10}_{-8}	6.89	0.78	0.46
J1028-0236.0	633.7	939^{+20}_{-29}	42^{+4}_{-3}	5.15	0.74	0.41

Bibliography

Accurso, G., Saintonge, A., Catinella, B., et al. 2017, MNRAS, 470, 4750

Adams, E., Adebahr, B., de Blok, W. J. G., et al. 2018, in AAS Meeting, Vol. 231 (AAS Meeting), 354.04

Adams, E. A. K. & van Leeuwen, J. 2019, Nat Astron, 3, 188

Adams, N. J., Bowler, R. A. A., Jarvis, M. J., Häußler, B., & Lagos, C. D. P. 2021, MNRAS, 506, 4933

Akins, H. B., Fujimoto, S., Finlator, K., et al. 2022, ApJ, 934, 64

Allison, J. R., Sadler, E. M., Bellstedt, S., et al. 2020, MNRAS, 494, 3627

Andreani, P., Miyamoto, Y., Kaneko, H., et al. 2020, A&A, 643, L11

Anglés-Alcázar, D., Faucher-Giguère, C.-A., Kereš, D., et al. 2017, MNRAS, 470, 4698

Appleton, P. N., Guillard, P., Emonts, B., et al. 2023, ApJ, 951, 104

Aravena, M., Boogaard, L., Gónzalez-López, J., et al. 2020, ApJ, 901, 79

Aravena, M., Decarli, R., Gónzalez-López, J., et al. 2019, ApJ, 882, 136

Aravena, M., Heintz, K., Dessauges-Zavadsky, M., et al. 2024, A&A, 682, A24

Armus, L., Lai, T., U, V., et al. 2023, ApJL, 942, L37

Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, MNRAS, 310, 540

Audibert, A., Dasyra, K. M., Papachristou, M., et al. 2022, A&A, 668, A67

Augustin, R., Péroux, C., Karki, A., et al. 2024, MNRAS

Aung, H., Mandelker, N., Dekel, A., et al. 2024, MNRAS, 532, 2965

Avila-Reese, V., Zavala, J., & Lacerna, I. 2014, MNRAS, 441, 417

Bacon, R., Accardo, M., Adjali, L., et al. 2010, in Ground-Based and Airborne Instrumentation for Astronomy III, Vol. 7735, eprint: arXiv:2211.16795, 773508

Bacon, R., Brinchmann, J., Richard, J., et al. 2015, A&A, 575, A75

Baker, A. J., Blyth, S., Holwerda, B. W., & LADUMA Team. 2018, in American Astronomical Society Meeting Abstracts #231, Vol. 231, 231.07

Balashev, S. A., Klimenko, V. V., Noterdaeme, P., et al. 2019, MNRAS, 490, 2668

Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5

Bandura, K., Addison, G. E., Amiri, M., et al. 2014, in Ground-Based and Airborne Telescopes V, Vol. 9145, eprint: arXiv:1406.2288, 914522

Barbani, F., Pascale, R., Marinacci, F., et al. 2023, MNRAS, 524, 4091

Barfety, C., Jolly, J.-B., Schreiber, N. M. F., et al. 2025, ApJ, 988, 55

Baryshev, A. 2024, in The Promises and Challenges of the ALMA Wideband Sensitivity Upgrade, 32

Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2012, ApJL, 762, L31

Bennett, J. S. & Sijacki, D. 2020, MNRAS, 499, 597

Bera, A., Kanekar, N., Chengalur, J. N., & Bagla, J. S. 2019, ApJL, 882, L7

Berta, S., Lutz, D., Nordon, R., et al. 2013, A&A, 555, L8

Béthermin, M., Gkogkou, A., Cuyck, M. V., et al. 2022, A&A, 667, A156

Bicknell, G. V., Sutherland, R. S., van Breugel, W. J. M., et al. 2000, ApJ, 540, 678

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

Bigiel, F., Leroy, A., Walter, F., et al. 2008, AJ, 136, 2846

Birnboim, Y. & Dekel, A. 2003, MNRAS, 345, 349

Bisigello, L., Vallini, L., Gruppioni, C., et al. 2022, A&A, 666, A193

Blitz, L., Fukui, Y., Kawamura, A., et al. 2007, Giant Molecular Clouds in Local Group Galaxies

Blumenthal, G. R., Faber, S. M., Primack, J. R., & Rees, M. J. 1984, Nat, 311, 517

Blyth, S. L., Baker, A. J., Holwerda, B. W., et al. 2016, Conf.

Boettcher, E., Chen, H.-W., Zahedy, F. S., et al. 2021, ApJ, 913, 18

Bogdán, Á., Forman, W. R., Vogelsberger, M., et al. 2013, ApJ, 772, 97

Bogdan, A., Khabibullin, I., Kovacs, O., et al. 2023, Circumgalactic Medium on the Largest Scales: Detecting X-ray Absorption Lines with Large-Area Microcalorimeters

Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, Annu. Rev. Astron. Astrophys., 51, 207

Bolatto, A. D., Wong, T., Utomo, D., et al. 2017, ApJ, 846, 159

Bollo, V., González, V., Stefanon, M., et al. 2023, ApJ, 946, 117

Bollo, V., Péroux, C., Zwaan, M., et al. 2025, A&A, 695, A163

Bollo, V., Zwaan, M., Péroux, C., et al. 2024, A&A, 690, A258

Bonato, M., Liuzzo, E., Giannetti, A., et al. 2018, MNRAS, 478, 1512

Bonavera, L., Cueli, M. M., González-Nuevo, J., et al. 2021, A&A, 656, A99

Bond, J. R., Kofman, L., & Pogosyan, D. 1996, Nat Astron, 380, 603

Boogaard, L. A., Decarli, R., Walter, F., et al. 2023, A NOEMA Molecular Line Scan of the Hubble Deep Field North: Improved Constraints on the CO Luminosity Functions and Cosmic Density of Molecular Gas

Boogaard, L. A., van der Werf, P., Weiss, A., et al. 2020, ApJ, 902, 109

Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, A&A, 622, A103

Bordoloi, R., Lilly, S. J., Hardmeier, E., et al. 2014, ApJ, 794, 130

Bothwell, M. S., Smail, I., Chapman, S. C., et al. 2013, MNRAS, 429, 3047

Bouché, N., Bouché, M. D., & Péroux, C. 2005, MNRAS, 364, 319

Bouché, N., Dekel, A., Genzel, R., et al. 2010, ApJ, 718, 1001

Bouché, N., Finley, H., Schroetter, I., et al. 2016, ApJ, 820, 121

Bouché, N., Hohensee, W., Vargas, R., et al. 2012, MNRAS, 426, 801

Bouché, N., Lehnert, M. D., Aguirre, A., Péroux, C., & Bergeron, J. 2007, MNRAS, 378, 525

Bouché, N., Lehnert, M. D., & Péroux, C. 2006, MNRAS, 367, L16

Bourke, T., Braun, R., Bonaldi, A., et al. 2018, in American Astronomical Society Meeting Abstracts #231, Vol. 231, 152.07

Bouwens, R., González-López, J., Aravena, M., et al. 2020, ApJ, 902, 112

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, ApJ, 803, 34

Bouwens, R. J., Oesch, P. A., Stefanon, M., et al. 2021, AJ, 162, 47

Bouwens, R. J., Smit, R., Schouws, S., et al. 2022, ApJ, 931, 160

Bowen, D. V., Chelouche, D., Jenkins, E. B., et al. 2016, ApJ, 826, 50

Bowler, R. A. A., Inami, H., Sommovigo, L., et al. 2024, MNRAS, 527, 5808

Boylan-Kolchin, M. 2023, Nat Astron, 7, 731

Boylan-Kolchin, M., Springel, V., White, S. D. M., Jenkins, A., & Lemson, G. 2009, MNRAS, 398, 1150

Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503

Breuck, C. D., Weiß, A., Béthermin, M., et al. 2019, A&A, 631, A167

Brienza, M., Shimwell, T. W., de Gasperin, F., et al. 2021, Nat Astron, 5, 1261

Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151

Bromm, V. & Larson, R. B. 2004, Annu. Rev. Astron. Astrophys., 42, 79

Bromm, V., Yoshida, N., Hernquist, L., & McKee, C. F. 2009, Nat Astron, 459, 49

Brook, C. B., Stinson, G., Gibson, B. K., et al. 2014, MNRAS, 443, 3809

Brunner, H., Liu, T., Lamer, G., et al. 2022, A&A, 661, A1

Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000

Bryant, J. J., Owers, M. S., Robotham, A. S. G., et al. 2015, MNRAS, 447, 2857

Buck, T., Pfrommer, C., Pakmor, R., Grand, R. J. J., & Springel, V. 2020, MNRAS, 497, 1712

Bundy, K., Bershady, M. A., Law, D. R., et al. 2014, ApJ, 798, 7

Burgett, W., Bernstein, R., Ashby, D., et al. 2024, in Ground-Based and Airborne Telescopes X, Vol. 13094, 1309417

Burke, B. F., Graham, & Wilkinson, P. N. 2019, An Introduction to Radio Astronomy, 4th Edition

Butsky, I. S., Fielding, D. B., Hayward, C. C., et al. 2020, ApJ, 903, 77

Butsky, I. S., Werk, J. K., Tchernyshyov, K., et al. 2022, ApJ, 935, 69

Callingham, T. M. & Helmi, A. 2025, Group Accretion in Milky Way-like Stellar Haloes

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682

Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, ApJ, 429, 582

Cameron, A. J., Saxena, A., Bunker, A. J., et al. 2023, A&A, 677, A115

Cantalupo, S., Arrigoni-Battaia, F., Prochaska, J. X., Hennawi, J. F., & Madau, P. 2014, Nat Astron, 506, 63

Cappellari, M. 2023, MNRAS, 526, 3273

Caputi, K. I., Cirasuolo, M., Dunlop, J. S., et al. 2011, MNRAS, 413, 162

Caputi, K. I., Rinaldi, P., Iani, E., et al. 2024, ApJ, 969, 159

Carilli, C. L., Butler, B., Golap, K., Carilli, M. T., & White, S. M. 2018, Imaging Stellar Radio Photospheres with the Next Generation Very Large Array

Carilli, C. L. & Walter, F. 2013, Annu. Rev. Astron. Astrophys., 51, 105

Carnall, A. C., McLure, R. J., Dunlop, J. S., & Davé, R. 2018, MNRAS, 480, 4379

Carniani, S., D'Eugenio, F., Ji, X., et al. 2024, The Eventful Life of a Luminous Galaxy at z = 14: Metal Enrichment, Feedback, and Low Gas Fraction?

Carpenter, J., Brogan, C., Iono, D., & Mroczkowski, T. 2023, The ALMA2030 Wideband Sensitivity Upgrade

Carpenter, J., Iono, D., Kemper, F., & Wootten, A. 2020, The ALMA Development Program: Roadmap to 2030

Carpenter, J., Iono, D., Testi, L., et al. 2019, The ALMA Development Roadmap

CASA Team, Bean, B., Bhatnagar, S., et al. 2022, PASP, 134, 114501

Casavecchia, B., Maio, U., Péroux, C., & Ciardi, B. 2024, A&A, 689, A106

Casavecchia, B., Maio, U., Péroux, C., & Ciardi, B. 2025, A&A, 693, A119

Casey, C. M., Kartaltepe, J. S., Drakos, N. E., et al. 2023, ApJ, 954, 31

Catán, V., González-López, J., Solimano, M., et al. 2024, A&A, 692, A215

Catinella, B., Saintonge, A., Janowiecki, S., et al. 2018, MNRAS, 476, 875

Cedrés, B., Bongiovanni, Á., Cepa, J., et al. 2025, A&A, 696, A85

Cerardi, N., Pierre, M., Valageas, P., Garrel, C., & Pacaud, F. 2024, A&A, 682, A138

Chabrier, G. 2003, PASP, 115, 763

Chen, C.-C., Hodge, J. A., Smail, I., et al. 2017, ApJ, 846, 108

Chen, J., Ivison, R. J., Zwaan, M. A., et al. 2023, A&A, 675, L10

Chen, J., Ivison, R. J., Zwaan, M. A., et al. 2022, ALMACAL IX: Multi-Band ALMA Survey for Dusty Star-Forming Galaxies and the Resolved Fractions of the Cosmic Infrared Background

Chen, Z., Wang, E., Zou, H., et al. 2025, The Circumgalactic Medium Traced by Mg II Absorption with DESI: Dependence on Galaxy Stellar Mass, Star Formation Rate and Azimuthal Angle

Chevance, M., Kruijssen, J. M. D., Krumholz, M. R., et al. 2021, MNRAS, 509, 272

Chisholm, J., Tremonti, C. A., Leitherer, C., Chen, Y., & Wofford, A. 2016, MNRAS, 457, 3133

Chisholm, J., Tremonti, C. A., Leitherer, C., et al. 2015, ApJ, 811, 149

Chowdhury, A., Kanekar, N., & Chengalur, J. N. 2024, The HI Mass Function of Star-forming Galaxies at $z\approx1$

Chowdhury, A., Kanekar, N., Chengalur, J. N., Sethi, S., & Dwarakanath, K. S. 2020, Nat Astron, 586, 369

Christensen, C. R., Davé, R., Governato, F., et al. 2016, ApJ, 824, 57

Cirasuolo, M., Fairley, A., Rees, P., et al. 2020, The Messenger, 180, 10

Cirasuolo, M., Tamai, R., Koehler, B., et al. 2024, The Messenger, 192, 3

Collaboration, DESI., Adame, A. G., Aguilar, J., et al. 2024, AJ, 168, 58

Collaboration, DESI., Aghamousa, A., Aguilar, J., et al. 2016, The DESI Experiment Part I: Science,Targeting, and Survey Design

Collaboration, E., Mellier, Y., Abdurro'uf, et al. 2025, A&A, 697, A1

Collaboration, P., Ade, P. a. R., Aghanim, N., et al. 2011

Comparat, J., Luo, W., Merloni, A., et al. 2023, A&A, 673, A122

Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, AJ, 115, 1693

Conselice, C. J. 2014, Annu. Rev. Astron. Astrophys., 52, 291

Cook, R. H. W., Davies, L. J. M., Rhee, J., et al. 2024, Monthly Notices of the Royal Astronomical Society, 531, 708

Cook, R. J. 2024, ApJ, 965, 127

Cooper, T. J., Rudie, G. C., Chen, H.-W., et al. 2021, MNRAS, 508, 4359

Cornwell, T. & Fomalont, E. B. 1989, in Synthesis Imaging in Radio Astronomy, Vol. 6, 185

Cortes, P., Carpenter, J., Kameno, S., et al. 2025

Covelo-Paz, A., Giovinazzo, E., Oesch, P. A., et al. 2024, An H{\alpha} View of Galaxy Build-up in the First 2 Gyr: Luminosity Functions at Z~4-6.5 from NIRCam/Grism Spectroscopy

Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839

Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, MNRAS, 450, 1937

Crighton, N. H. M., Murphy, M. T., Prochaska, J. X., et al. 2015, MNRAS, 452, 217

Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11

Curti, M., Cresci, G., Mannucci, F., et al. 2017, MNRAS, 465, 1384

Curti, M., Maiolino, R., Curtis-Lake, E., et al. 2023, JADES: Insights on the Low-Mass End of the Mass–Metallicity–Star-Formation Rate Relation at $3 < z < 10$ from Deep JWST/NIRSpec Spectroscopy

Curti, M., Mannucci, F., Cresci, G., & Maiolino, R. 2020, MNRAS, 491, 944

da Cunha, E., Groves, B., Walter, F., et al. 2013, ApJ, 766, 13

Dabrowski, I. 1984, CaJP, 62, 1639

Daddi, E., Bournaud, F., Walter, F., et al. 2010, ApJ, 713, 686

Daddi, E., Dannerbauer, H., Liu, D., et al. 2015, A&A, 577, A46

Daddi, E., Dickinson, M., Morrison, G., et al. 2007, ApJ, 670, 156

Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, ApJ, 547, 792

Dashyan, G. & Dubois, Y. 2020, A&A, 638, A123

Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, MNRAS, 486, 2827

Davé, R., Finlator, K., & Oppenheimer, B. D. 2012, MNRAS, 421, 98

Davé, R., Oppenheimer, B. D., & Finlator, K. 2011, MNRAS, 415, 11

Davé, R., Thompson, R., & Hopkins, P. F. 2016, Mon Not R Astron Soc, 462, 3265

de Arriba, L. P., Alonso-Herrero, A., García-Burillo, S., et al. 2023, A&A, 675, A58

de Blok, W. J. G., Healy, J., Maccagni, F. M., et al. 2024, A&A, 688, A109

de los Reyes, M. A. C. & Kennicutt, R. C. 2019, ApJ, 872, 16

De Lucia, G., Springel, V., White, S. D. M., Croton, D., & Kauffmann, G. 2006, MNRAS, 366, 499

Decarli, R., Aravena, M., Boogaard, L., et al. 2020, ApJ, 902, 110

Decarli, R., Walter, F., Aravena, M., et al. 2016, ApJ, 833, 69

Decarli, R., Walter, F., Gónzalez-López, J., et al. 2019, ApJ, 882, 138

Dehghanian, M., Arav, N., Sharma, M., et al. 2025, Quasar Absorption Outflows on Galactic Scales: Insights from DESI

Deka, P. P., Gupta, N., Jagannathan, P., et al. 2023, The MeerKAT Absorption Line Survey (MALS) Data Release I: Stokes I Image Catalogs at 1-1.4 GHz

Dekel, A. & Birnboim, Y. 2006, MNRAS, 368, 2

Dekel, A., Birnboim, Y., Engel, G., et al. 2009, Nat Astron, 457, 451

Dénes, H., Kilborn, V. A., & Koribalski, B. S. 2014, MNRAS, 444, 667

Dessauges-Zavadsky, M., Calura, F., Prochaska, J. X., D'Odorico, S., & Matteucci, F. 2007, A&A, 470, 431

Dessauges-Zavadsky, M., Ginolfi, M., Pozzi, F., et al. 2020, A&A, 643, A5

Di Francesco, J., Chalmers, D., Denman, N., et al. 2019, in Canadian Long Range Plan for Astronomy and Astrophysics White Papers, Vol. 2020, eprint: arXiv:1911.01517, 32

Diemer, B., Stevens, A. R. H., Lagos, C. d. P., et al. 2019, MNRAS, 487, 1529

Dolag, K., Komatsu, E., & Sunyaev, R. 2016, MNRAS, 463, 1797

Downes, D. & Solomon, P. M. 1998, ApJ, 507, 615

Draine, B. T. 2011, Physics of the Interstellar and Intergalactic Medium

Draine, B. T., Dale, D. A., Bendo, G., et al. 2007, ApJ, 663, 866

Driver, S. P. & Robotham, A. S. G. 2010, MNRAS, 407, 2131

Duchesne, S. W., Grundy, J. A., Heald, G. H., et al. 2024, Publ. Astron. Soc. Aust., 41, e003

Duncan, K. J., McLeod, D. J., Best, P. N., et al. 2025, The JWST Emission Line Survey (JELS): Extending Rest-Optical Narrow-Band Emission Line Selection into the Epoch of Reionization

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

Dutta, S., Muzahid, S., Schaye, J., et al. 2025, ApJ, 980, 264

Dutta, S., Muzahid, S., Schaye, J., et al. 2024, MUSEQuBES: Mapping the Distribution of Neutral Hydrogen around Low-Redshift Galaxies

Efstathiou, G. 2025, Philosophical Transactions of the Royal Society of London Series A, 383, 20240022

Eisenstein, D. J., Willott, C., Alberts, S., et al. 2023, Overview of the JWST Advanced Deep Extragalactic Survey (JADES)

Elahi, P. J., Welker, C., Power, C., et al. 2018, MNRAS, 475, 5338

Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, A&A, 533, A119

Emonts, B. H. C., Breuck, C. D., Lehnert, M. D., et al. 2015, A&A, 584, A99

Erb, D. K. 2008, ApJ, 674, 151

Erb, D. K., Steidel, C. C., Shapley, A. E., et al. 2006, ApJ, 647, 128

Esparza-Arredondo, D., Almeida, C. R., Audibert, A., et al. 2025, A&A, 693, A174

Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019, The Astrophysical Journal, 875, L1

Fabian, A. C. 2012, Annu. Rev. Astron. Astrophys., 50, 455

Faisst, A. L., Schaerer, D., Lemaux, B. C., et al. 2020, ApJS, 247, 61

Fanson, J., Bernstein, R., Ashby, D., et al. 2022, in Ground-Based and Airborne Telescopes IX, Vol. 12182, 121821C

Farcy, M., Rosdahl, J., Dubois, Y., Blaizot, J., & Martin-Alvarez, S. 2022, MNRAS, 513, 5000

Faucher-Giguère, C.-A. & Oh, S. P. 2023, Annu. Rev. Astron. Astrophys., 61, 131

Ferkinhoff, C., Hailey-Dunsheath, S., Nikola, T., et al. 2010, ApJL, 714, L147

Fernández, X., Gim, H. B., van Gorkom, J. H., et al. 2016, ApJL, 824, L1

Ferrara, A., Scannapieco, E., & Bergeron, J. 2005, ApJ, 634, L37

Ferrara, A., Vallini, L., Pallottini, A., et al. 2019, MNRAS, 489, 1

Ferriere, K. M. 2001, Rev. Mod. Phys., 73, 1031

Feruglio, C., Fiore, F., Carniani, S., et al. 2015, A&A, 583, A99

Field, G. B., Goldsmith, D. W., & Habing, H. J. 1969, ApJ, 155, L149

Fielding, D. B., Ostriker, E. C., Bryan, G. L., & Jermyn, A. S. 2020, ApJL, 894, L24

Fillingham, S. P., Cooper, M. C., Boylan-Kolchin, M., et al. 2018, MNRAS, 477, 4491

Finkelstein, S. L., Leung, G. C. K., Bagley, M. B., et al. 2023, The Complete CEERS Early Universe Galaxy Sample: A Surprisingly Slow Evolution of the Space Density of Bright Galaxies at z ∼ 8.5-14.5

Fixsen, D. J., Bennett, C. L., & Mather, J. C. 1999, ApJ, 526, 207

Fletcher, T. J., Saintonge, A., Soares, P. S., & Pontzen, A. 2021, MNRAS, 501, 411

Flores, H., Hammer, F., Thuan, T. X., et al. 1999, ApJ, 517, 148

Fluetsch, A., Maiolino, R., Carniani, S., et al. 2019, MNRAS, 483, 4586

Forbes, J. C., Krumholz, M. R., Burkert, A., & Dekel, A. 2014, MNRAS, 443, 168

Ford, A. B., Davé, R., Oppenheimer, B. D., et al. 2014, MNRAS, 444, 1260

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

Fox, A. & Davé, R. 2017, Gas Accretion onto Galaxies, 430

Fox, A. J., Wakker, B. P., Barger, K. A., et al. 2014, ApJ, 787, 147

Franco, M., Elbaz, D., Béthermin, M., et al. 2018, A&A, 620, A152

Frank, S., Rasera, Y., Vibert, D., et al. 2012, MNRAS, 420, 1731

Freundlich, J., Combes, F., Tacconi, L. J., et al. 2019, A&A, 622, A105

Fu, H., Boco, L., Shankar, F., et al. 2025a, A&A, 695, A252

Fu, S., Sun, F., Jiang, L., et al. 2025b, ApJ, 987, 186

Fudamoto, Y., Smit, R., Bowler, R. A. A., et al. 2022, ApJ, 934, 144

Fumagalli, M., Labbé, I., Patel, S. G., et al. 2014, ApJ, 796, 35

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

Fynbo, J. P. U., Heintz, K. E., Neeleman, M., et al. 2018, MNRAS, 479, 2126

Fynbo, J. P. U., Laursen, P., Ledoux, C., et al. 2010, MNRAS, 408, 2128

Galliano, F., Galametz, M., & Jones, A. P. 2018, Annu. Rev. Astron. Astrophys., 56, 673

Garcia, K. & Narayanan, D. 2023, 241, 210.10

Garcia, K., Narayanan, D., Popping, G., et al. 2023, $\texttt{slick}$: Modeling a Universe of Molecular Line Luminosities in Hydrodynamical Simulations

Gehrels, N. 1986, ApJ, 303, 336

Genzel, R., Tacconi, L. J., Combes, F., et al. 2012, ApJ, 746, 69

Genzel, R., Tacconi, L. J., Lutz, D., et al. 2015, ApJ, 800, 20

Gkogkou, A., Béthermin, M., Lagache, G., et al. 2022, CONCERTO: Simulating the CO, [CII], and [CI] Line Emission of Galaxies in a 117 $\rm Deg^2$ Field and the Impact of Field-to-Field Variance

Glover, S. C. O. & Mac Low, M.-M. 2007, ApJ, 659, 1317

Glowacki, M., Collier, J. D., Kazemi-Moridani, A., et al. 2022, ApJL, 931, L7

Goldsmith, P. F., Heyer, M., Narayanan, G., et al. 2008, ApJ, 680, 428

Gómez-Guijarro, C., Elbaz, D., Xiao, M., et al. 2022, A&A, 658, A43

González, V., Labbé, I., Bouwens, R. J., et al. 2011, ApJ, 735, L34

Gozaliasl, G., Yang, L., Kartaltepe, J., et al. 2025, COSMOS Web: Morphological Quenching and Size-Mass Evolution of Brightest Group Galaxies from z = 3.7

Grazian, A., Fontana, A., Santini, P., et al. 2015, A&A, 575, A96

Grenier, I. A., Casandjian, J.-M., & Terrier, R. 2005, Science, 307, 1292

Groves, B., Brinchmann, J., & Walcher, C. J. 2012, MNRAS, 419, 1402

Gruen, D., McCullough, J., Amon, A., et al. 2023, The Messenger, 190, 28

Gruen, D., Seitz, S., Becker, M. R., Friedrich, O., & Mana, A. 2015, MNRAS, 449, 4264

Gruen, D., Zhang, Y., Palmese, A., et al. 2019, MNRAS, 488, 4389

Guilloteau, S., Delannoy, J., Downes, D., et al. 1992, A&A, 262, 624

Guo, H., Wang, J., Jones, M. G., & Behroozi, P. 2023a, ApJ, 955, 57

Guo, Y., Bacon, R., Bouché, N. F., et al. 2023b, Nat Astron, 624, 53

Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, ApJS, 207, 24

Haardt, F. & Madau, P. 1996, ApJ, 461, 20

Haehnelt, M. G., Steinmetz, M., & Rauch, M. 1998, ApJ, 495, 647

Hafen, Z., Faucher-Giguère, C.-A., Anglés-Alcázar, D., et al. 2019, MNRAS, 488, 1248

Hallinan, G., Ravi, V., Weinreb, S., et al. 2019, The DSA-2000 – A Radio Survey Camera

Hamadouche, M. L., McLure, R. J., Carnall, A., et al. 2024, JWST PRIMER: Strong Evidence for the Environmental Quenching of Low-Mass Galaxies out to $\mathbf{\textit{z} \simeq 2}$

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

Hamanowicz, A., Zwaan, M. A., Péroux, C., et al. 2022, ALMACAL VIII: A Pilot Survey for Untargeted Extragalactic CO Emission Lines in Deep ALMA Calibration Data

Harrison, C. M., Alexander, D. M., Rosario, D. J., Scholtz, J., & Stanley, F. 2019, Proc. IAU, 15, 199

Harwood, J. J., Hardcastle, M. J., Croston, J. H., & Goodger, J. L. 2013, MNRAS, 435, 3353

Hatfield, P. W. & Jarvis, M. J. 2017, MNRAS, 472, 3570

Hayes, M. J., Scarlata, C., Berg, D., et al. 2023, JWST Proposal. Cycle 2, 3290

Haynes, M. P. & Giovanelli, R. 1984, AJ, 89, 758

Haynes, M. P., Giovanelli, R., Martin, A. M., et al. 2011, AJ, 142, 170

Heckman, T. M. & Borthakur, S. 2016, ApJ, 822, 9

Heiderman, A., Evans, N. J., Allen, L. E., Huard, T., & Heyer, M. 2010, ApJ, 723, 1019

Hernández-Aguayo, C., Springel, V., Pakmor, R., et al. 2023, MNRAS, 524, 2556

Herrera-Camus, R., González-López, J., Schreiber, N. F., et al. 2025, The ALMA-CRISTAL Survey: Gas, Dust, and Stars in Star-Forming Galaxies When the Universe Was ~1 Gyr Old I. Survey Overview and Case Studies

Herrera-Camus, R., Schreiber, N. M. F., Price, S. H., et al. 2022, A&A, 665, L8

Ho, S. H., Martin, C. L., Kacprzak, G. G., & Churchill, C. W. 2017, ApJ, 835, 267

Hodge, J. A. & da Cunha, E. 2020, R. Soc. open sci., 7, 200556

Hodge, J. A., Riechers, D., Decarli, R., et al. 2014, ApJL, 798, L18

Högbom, J. A. 1974, Astron. Astrophys. Suppl. Ser., 15, 417

Hopkins, A. M. & Beacom, J. F. 2006, ApJ, 651, 142

Hopkins, A. M., Driver, S. P., Brough, S., et al. 2013, MNRAS, 430, 2047

Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, MNRAS, 445, 581

Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, MNRAS, 480, 800

Hopkins, P. F., Wetzel, A., Wheeler, C., et al. 2023, MNRAS, 519, 3154

Hotan, A. W., Bunton, J. D., Chippendale, A. P., et al. 2021, Publ. Astron. Soc. Aust., 38, e009

HSC Collaboration, Aguena, M., Aiola, S., et al. 2025, The Atacama Cosmology Telescope: DR6 Sunyaev-Zel'dovich Selected Galaxy Clusters Catalog

Hsiao, T. Y.-Y., Álvarez-Márquez, J., Coe, D., et al. 2024, ApJ, 973, 81

Hummels, C. B., Smith, B. D., Hopkins, P. F., et al. 2019, ApJ, 882, 156

Hunter, T. R., Indebetouw, R., Brogan, C. L., et al. 2023, PASP, 135, 074501

Husemann, B., Bennert, V. N., Jahnke, K., et al. 2019, ApJ, 879, 75

Huynh, M. & Lazio, J. 2013, An Overview of the Square Kilometre Array

Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841

Ilbert, O., McCracken, H. J., Fèvre, O. L., et al. 2013, A&A, 556, A55

Inami, H., Bacon, R., Brinchmann, J., et al. 2017, A&A, 608, A2

Inoue, S., Yoshida, N., & Yajima, H. 2020, MNRAS, 498, 5960

Ji, S., Chan, T. K., Hummels, C. B., et al. 2020, MNRAS, 496, 4221

Jin, S., Dannerbauer, H., Emonts, B., et al. 2021, A&A, 652, A11

Jones, M. G., Haynes, M. P., Giovanelli, R., & Moorman, C. 2018, MNRAS, 477, 2

Jurlin, N., Brienza, M., Morganti, R., et al. 2021, A&A, 653, A110

Kacprzak, G. G., Churchill, C. W., Evans, J. L., Murphy, M. T., & Steidel, C. C. 2011, MNRAS, 416, 3118

Kanekar, N., Prochaska, J. X., Christensen, L., et al. 2018, ApJL, 856, L23

Kanekar, N., Prochaska, J. X., Neeleman, M., et al. 2020, ApJL, 901, L5

Kanekar, N., Prochaska, J. X., Smette, A., et al. 2014, MNRAS, 438, 2131

Karki, A., Kulkarni, V. P., Weng, S., et al. 2023, MNRAS, 524, 5524

Katsianis, A., Blanc, G., Lagos, C. P., et al. 2017, MNRAS, 472, 919

Katsianis, A., Xu, H., Yang, X., et al. 2021, MNRAS, 500, 2036

Katz, H., Kimm, T., Sijacki, D., & Haehnelt, M. G. 2017, MNRAS, 468, 4831

Kaur, B., Kanekar, N., Neeleman, M., et al. 2024, An HI-absorption-selected Cold Rotating Disk Galaxy at $z\approx2.193$

Kaur, B., Kanekar, N., Neeleman, M., et al. 2025, ApJ, 982, L26

Kaur, B., Kanekar, N., Rafelski, M., et al. 2021, ApJ, 921, 68

Kaur, B., Kanekar, N., Revalski, M., et al. 2022, ApJ, 934, 87

Kaviraj, S., Devriendt, J., Dubois, Y., et al. 2015, MNRAS, 452, 2845

Keating, L. C., Richings, A. J., Murray, N., et al. 2020, MNRAS, 499, 837

Keenan, R. P., Marrone, D. P., & Keating, G. K. 2020, ApJ, 904, 127

Kennicutt, R. C. & De Los Reyes, M. A. C. 2021, ApJ, 908, 61

Kennicutt, R. C. & Evans, N. J. 2012, Annu. Rev. Astron. Astrophys., 50, 531

Kennicutt, Jr., R. C. 1989, ApJ, 344, 685

Kennicutt, Jr., R. C. 1998, ApJ, 498, 541

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

Kewley, L. J. & Ellison, S. L. 2008, ApJ, 681, 1183

Kim, J.-h., Ma, X., Grudić, M. Y., et al. 2018, MNRAS, 474, 4232

Kirshner, R. 2024, in American Astronomical Society Meeting Abstracts, Vol. 244, 122.04

Kitzbichler, M. G. & White, S. D. M. 2007, MNRAS, 376, 2

Klaassen, P. D., Mroczkowski, T. K., Cicone, C., et al. 2020, in Ground-Based and Airborne Telescopes VIII, Vol. 11445, eprint: arXiv:2011.07974, 114452F

Klessen, R. S. & Glover, S. C. O. 2016, in Saas-Fee Advanced Course 43. Swiss Society for Astrophysics and Astronomy, Saas-Fee (Berlin, Heidelberg: Springer), 85–249

Klessen, R. S. & Glover, S. C. O. 2023, The First Stars: Formation, Properties, and Impact

Klimenko, V. V., Kulkarni, V., Wake, D. A., et al. 2023, The Baryonic Content of Galaxies Mapped by MaNGA and the Gas Around Them

Klitsch, A., Christensen, L., Valentino, F., et al. 2022, MNRAS, 514, 2346

Klitsch, A., Davis, T. A., Hamanowicz, A., et al. 2023, ALMACAL X: Constraints on Molecular Gas in the Low-Redshift Circumgalactic Medium

Klitsch, A., Péroux, C., Zwaan, M. A., et al. 2021, MNRAS, 506, 514

Klitsch, A., Péroux, C., Zwaan, M. A., et al. 2019a, MNRAS, 490, 1220

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

Klitsch, A., Zwaan, M. A., Péroux, C., et al. 2019b, MNRAS, 482, L65

Klitsch, A., Zwaan, M. A., Smail, I., et al. 2020, MNRAS, 495, 2332

Komugi, S., Toba, Y., Matsuoka, Y., Saito, T., & Yamashita, T. 2022, ApJ, 930, 3

Koopmans, L. V. E., Pritchard, J., Mellema, G., et al. 2015

Koribalski, B. S., Duchesne, S. W., Lenc, E., et al. 2024, MNRAS, 533, 608

Koribalski, B. S., Staveley-Smith, L., Westmeier, T., et al. 2020, Astro. Space Sci, 365, 118

Krumholz, M. R., McKee, C. F., & Tumlinson, J. 2009, ApJ, 693, 216

Kulkarni, V. P., Bowen, D. V., Straka, L. A., et al. 2022, ApJ, 929, 150

Kulkarni, V. P., Fall, S. M., Lauroesch, J. T., et al. 2005, ApJ, 618, 68

Kulkarni, V. P., Hill, J. M., Schneider, G., et al. 2000, ApJ, 536, 36

Kulkarni, V. P., Som, D., Morrison, S., et al. 2015, ApJ, 815, 24

Lacy, M., Baum, S. A., Chandler, C. J., et al. 2020, PASP, 132, 035001

Lagache, G., Cousin, M., & Chatzikos, M. 2018, A&A, 609, A130

Lagos, C. D. P., Bravo, M., Tobar, R., et al. 2023, Quenching Massive Galaxies across Cosmic Time with the Semi-Analytic Model SHARK v2.0

Lagos, C. d. P., Bravo, M., Tobar, R., et al. 2024, MNRAS

Lagos, C. d. P., Crain, R. A., Schaye, J., et al. 2015, MNRAS, 452, 3815

Lagos, C. d. P., Lacey, C. G., & Baugh, C. M. 2012, Simulations and Modelling of the ISM in Galaxies

Lagos, C. d. P., Tobar, R. J., Robotham, A. S. G., et al. 2018, MNRAS, 481, 3573

Lammers, C., Iyer, K. G., Ibarra-Medel, H., et al. 2023, ApJ, 953, 26

Langan, I., Zabl, J., Bouché, N. F., et al. 2023, MNRAS, 521, 546

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, Euclid Definition Study Report

Le Fèvre, O., Béthermin, M., Faisst, A., et al. 2020, A&A, 643, A1

Leclercq, F., Bacon, R., Verhamme, A., et al. 2020, A&A, 635, A82

Lee, B., Wang, J., Chung, A., et al. 2022, ApJS, 262, 31

Lehner, N., Howk, J. C., Tripp, T. M., et al. 2013, ApJ, 770, 138

Lenkić, L., Bolatto, A. D., Fisher, D. B., et al. 2023, CO Excitation in High-z Main Sequence Analogues: Resolved CO(4-3)/CO(3-2) Line Ratios in DYNAMO Galaxies

Lenkić, L., Bolatto, A. D., Förster Schreiber, N. M., et al. 2020, AJ, 159, 190

Leroy, A. K., Schinnerer, E., Hughes, A., et al. 2021, ApJS, 257, 43

Leroy, A. K., Walter, F., Bigiel, F., et al. 2009, AJ, 137, 4670

Leroy, A. K., Walter, F., Brinks, E., et al. 2008, AJ, 136, 2782

Leslie, S. K., Schinnerer, E., Liu, D., et al. 2020, ApJ, 899, 58

Li, C. 2013, in Molecular Gas, Dust, and Star Formation in Galaxies, Vol. 292, 149–152

Li, Q., Narayanan, D., Davè, R., & Krumholz, M. R. 2018, ApJ, 869, 73

Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y. 2013, ApJ, 772, 119

Lilly, S. J., Le Fevre, O., Hammer, F., & Crampton, D. 1996, ApJL, 460, L1

Lin, L., Ellison, S. L., Pan, H.-A., et al. 2020, ApJ, 903, 145

Lin, L., Zhang, Z.-Y., Wang, J., et al. 2025, Nat Astron, 1

Lister, M. L., Homan, D. C., Hovatta, T., et al. 2019, ApJ, 874, 43

Liu, A., Bulbul, E., Ghirardini, V., et al. 2022, A&A, 661, A2

Liu, D., Schinnerer, E., Groves, B., et al. 2019, ApJ, 887, 235

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

Lofthouse, E. K., Kaviraj, S., Conselice, C. J., Mortlock, A., & Hartley, W. 2017, MNRAS, 465, 2895

Looze, I. D., Cormier, D., Lebouteiller, V., et al. 2014, A&A, 568, A62

Lu, S., Fang, G., Gu, Y., et al. 2021, ApJ, 913, 81

Luo, B., Brandt, W. N., Xue, Y. Q., et al. 2017, ApJS, 228, 2

Lupi, A., Bovino, S., Capelo, P. R., Volonteri, M., & Silk, J. 2018, MNRAS, 474, 2884

Lupi, A., Pallottini, A., Ferrara, A., et al. 2020, MNRAS, 496, 5160

Madau, P. & Dickinson, M. 2014, Annual Review of Astronomy and Astrophysics, 52, 415

Madau, P., Ferguson, H. C., Dickinson, M. E., et al. 1996, MNRAS, 283, 1388

Madden, S. C., Cormier, D., Hony, S., et al. 2020, A&A, 643, A141

Madden, S. C., Galliano, F., Jones, A. P., & Sauvage, M. 2006, A&A, 446, 877

Maddox, N., Frank, B. S., Ponomareva, A. A., et al. 2021, A&A, 646, A35

Magnelli, B., Boogaard, L., Decarli, R., et al. 2020, ApJ, 892, 66

Mahatma, V. H., Basu, A., Hardcastle, M. J., Morabito, L. K., & van Weeren, R. J. 2023, MNRAS, 520, 4427

Mahony, E. K., Morganti, R., Emonts, B. H. C., Oosterloo, T. A., & Tadhunter, C. 2013, MNRAS, 435, L58

Mahony, E. K., Sadler, E. M., Croom, S. M., et al. 2011, MNRAS, 417, 2651

Maio, U., Péroux, C., & Ciardi, B. 2022, A&A, 657, A47

Maiolino, R., Carniani, S., Fontana, A., et al. 2015, MNRAS, 452, 54

Maiolino, R., Nagao, T., Grazian, A., et al. 2008, A&A, 488, 463

Marasco, A., Fraternali, F., Lehner, N., & Howk, J. C. 2022, Monthly Notices of the Royal Astronomical Society, 515, 4176

Marinacci, F., Vogelsberger, M., Pakmor, R., et al. 2018, MNRAS, 480, 5113

Martig, M., Bournaud, F., Teyssier, R., & Dekel, A. 2009, ApJ, 707, 250

Martin, C. L. 2005, ApJ, 621, 227

Martin, C. L., Shapley, A. E., Coil, A. L., et al. 2012, ApJ, 760, 127

Martin, D. C., O'Sullivan, D., Matuszewski, M., et al. 2019, Nat Astron, 3, 822

Martin, D. C., Wyder, T. K., Schiminovich, D., et al. 2007, ApJS, 173, 342

Masui, K. W., Switzer, E. R., Banavar, N., et al. 2013, ApJL, 763, L20

Mathes, N. L., Churchill, C. W., Kacprzak, G. G., et al. 2014, ApJ, 792, 128

Mauch, T., Murphy, T., Buttery, H. J., et al. 2013, VizieR Online Data Catalog, 8081, VIII/81B

McClymont, W., Tacchella, S., Smith, A., et al. 2025, The THESAN-ZOOM Project: Burst, Quench, Repeat – Unveiling the Evolution of High-Redshift Galaxies along the Star-Forming Main Sequence

McConnell, D., Hale, C. L., Lenc, E., et al. 2020, Publ. Astron. Soc. Aust., 37, e048

McCourt, M., Oh, S. P., O'Leary, R., & Madigan, A.-M. 2018, MNRAS, 473, 5407

McKee, C. F. 1995, in The Physics of the Interstellar Medium and Intergalactic Medium, Vol. 80, 292

McKee, C. F. & Ostriker, J. P. 1977, ApJ, 218, 148

McKinnon, M., Beasley, A., Murphy, E., et al. 2019, in Bulletin of the American Astronomical Society, Vol. 51, 81

McLeod, A. F., Ali, A. A., Chevance, M., et al. 2021a, MNRAS, 508, 5425

McLeod, D. J., McLure, R. J., Dunlop, J. S., et al. 2021b, MNRAS, 503, 4413

McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, 376, 127

Merloni, A., Lamer, G., Liu, T., et al. 2024, A&A, 682, A34

Merloni, A., Predehl, P., Becker, W., et al. 2012, eROSITA Science Book: Mapping the Structure of the Energetic Universe

Messias, H., Guerrero, A., Nagar, N., et al. 2024, MNRAS, 533, 3937

Meyer, E. T., Georganopoulos, M., Sparks, W. B., et al. 2015, ApJ, 805, 154

Michael Shull, J., Danforth, C. W., & Tilton, E. M. 2014, ApJ, 796, 49

Michałowski, M. J., Gall, C., Hjorth, J., et al. 2024, The Fate of the Interstellar Medium in Early-Type Galaxies. III. The Mechanism of ISM Removal and Quenching of Star Formation

Minoda, T., Saga, S., Takahashi, T., et al. 2023, Publ. Astron. Soc. Japan, 75, S154

Mitchell, P. D. & Schaye, J. 2022, MNRAS, 511, 2948

Mitra, S., Davé, R., & Finlator, K. 2015, MNRAS, 452, 1184

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

Molyneux, S. J., Smit, R., Schaerer, D., et al. 2022, MNRAS, 512, 535

Mondal, R., Shaw, A. K., Iliev, I. T., et al. 2020, MNRAS, 494, 4043

Morganti, R., Oosterloo, T., Oonk, J. B. R., Frieswijk, W., & Tadhunter, C. 2015a, A&A, 580, A1

Morganti, R., Sadler, E. M., Curran, S. J., & Members, t. S. H. S. 2015b, Cool Outflows and HI Absorbers with SKA

Moster, B. P., Somerville, R. S., Newman, J. A., & Rix, H.-W. 2011, ApJ, 731, 113

Mouri, H. & Taniguchi, Y. 2006, A&A, 459, 371

Mukherjee, D., Bicknell, G. V., & Wagner, A. Y. 2021, AsNa, 342, 1140

Mukherjee, D., Bicknell, G. V., Wagner, A. Y., Sutherland, R. S., & Silk, J. 2018, MNRAS, 479, 5544

Murphy, E. & ngVLA Science Advisory Council. 2020, in American Astronomical Society Meeting Abstracts #235, Vol. 235, 364.01

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

Muzahid, S., Srianand, R., & Charlton, J. 2015, MNRAS, 448, 2840

Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, ApJ, 777, 18

Naab, T. & Ostriker, J. P. 2017, Annu. Rev. Astron. Astrophys., 55, 59

Naiman, J. P., Pillepich, A., Springel, V., et al. 2018, MNRAS, 477, 1206

Napolitano, L., Pandey, A., Myers, A. D., et al. 2023, AJ, 166, 99

Narayanan, D., Krumholz, M. R., Ostriker, E. C., & Hernquist, L. 2012, MNRAS, 421, 3127

Nateghi, H., Kacprzak, G. G., Nielsen, N. M., et al. 2024, Signatures of Gas Flows-II: Connecting the Kinematics of the Multiphase Circumgalactic Medium to Galaxy Rotation

Neeleman, M., Kanekar, N., Prochaska, J. X., et al. 2018, ApJL, 856, L12

Neeleman, M., Kanekar, N., Prochaska, J. X., et al. 2017, Science, 355, 1285

Neeleman, M., Kanekar, N., Prochaska, J. X., Rafelski, M. A., & Kahinga, L. A. 2025, ApJ, 983, 26

Neeleman, M., Kanekar, N., Xavier Prochaska, J., Rafelski, M. A., & Carilli, C. L. 2019, ApJL, 870, L19

Neeleman, M., Prochaska, J. X., Kanekar, N., & Rafelski, M. 2020, Nat Astron, 581, 269

Neeleman, M., Prochaska, J. X., Zwaan, M. A., et al. 2016, ApJL, 820, L39

Nelson, D., Pillepich, A., Springel, V., et al. 2018, MNRAS, 475, 624

Nelson, D., Sharma, P., Pillepich, A., et al. 2020, MNRAS, 498, 2391

Nelson, D., Springel, V., Pillepich, A., et al. 2019, Comput. Astrophys., 6, 2

Nelson, D., Vogelsberger, M., Genel, S., et al. 2013, MNRAS, 429, 3353

Newman, S. F., Genzel, R., Förster-Schreiber, N. M., et al. 2012, ApJ, 761, 43

Nicastro, F., Kaastra, J., Krongold, Y., et al. 2018, Nat Astron, 558, 406

Nielsen, N. M., Kacprzak, G. G., Pointon, S. K., Churchill, C. W., & Murphy, M. T. 2018, ApJ, 869, 153

Nielsen, N. M., Kacprzak, G. G., Pointon, S. K., et al. 2020, ApJ, 904, 164

Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJL, 660, L43

Noterdaeme, P., Ledoux, C., Petitjean, P., & Srianand, R. 2008, A&A, 481, 327

Obreschkow, D., Croton, D., Lucia, G. D., Khochfar, S., & Rawlings, S. 2009, ApJ, 698, 1467

Olsen, K. P., Burkhart, B., Mac Low, M.-M., et al. 2021, ApJ, 922, 88

Oosterloo, T., Morganti, R., Tadhunter, C., et al. 2019, A&A, 632, A66

Oppenheimer, B. D., Babul, A., Bahé, Y., Butsky, I. S., & McCarthy, I. G. 2021, Universe, 7, 209

Oppenheimer, B. D. & Davé, R. 2006, MNRAS, 373, 1265

Oppenheimer, B. D. & Davé, R. 2008, MNRAS, 387, 577

Oppenheimer, B. D., Davé, R., Kereš, D., et al. 2010, MNRAS, 406, 2325

Oteo, I., Zwaan, M., Ivison, R., Smail, I., & Biggs, A. 2016a, The Messenger, 164, 41

Oteo, I., Zwaan, M. A., Ivison, R. J., Smail, I., & Biggs, A. D. 2016b, ApJ, 822, 36

Oteo, I., Zwaan, M. A., Ivison, R. J., Smail, I., & Biggs, A. D. 2017, ApJ, 837, 182

Oxland, M., Parker, L. C., de Carvalho, R. R., & Sampaio, V. M. 2024, MNRAS, 529, 3651

Oyarzún, G. A., Rafelski, M., Christensen, L., et al. 2025, The Qz5 Survey (I): How the HI Mass Density of the Universe Evolves With Cosmic Time

Padovani, P. 2017, Nat Astron, 1, 0194

Padovani, P. & Cirasuolo, M. 2023, Cont. Phys., 64, 47

Pagel, B. E. J. 1999, Early Chemical Evolution of Galaxies

Pakmor, R., Springel, V., Coles, J. P., et al. 2023, MNRAS, 524, 2539

Pallottini, A., Ferrara, A., Gallerani, S., et al. 2022, MNRAS [[arXiv]2201.02636]

Pallottini, A., Ferrara, A., Gallerani, S., et al. 2017, MNRAS, 465, 2540

Pavesi, R., Sharon, C. E., Riechers, D. A., et al. 2018, ApJ, 864, 49

Pearson, T. J. & Readhead, A. C. S. 1984, Annu. Rev. Astron. Astrophys., 22, 97

Peebles, P. J. E. 1982, ApJ, 263, L1

Peebles, P. J. E. 2025, Philosophical Transactions of the Royal Society of London Series A, 383, 20240021

Peeples, M. S., Werk, J. K., Tumlinson, J., et al. 2014, ApJ, 786, 54

Pelupessy, F. I. & Papadopoulos, P. P. 2009, ApJ, 707, 954

Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010, ApJ, 721, 193

Peng, Y.-j. & Maiolino, R. 2014, MNRAS, 443, 3643

Pensabene, A., Cantalupo, S., Cicone, C., et al. 2024, A&A, 684, A119

Penzias, A. A. & Wilson, R. W. 1965, ApJ, 142, 419

Pereira-Santaella, M., Colina, L., García-Burillo, S., et al. 2016, A&A, 594, A81

Pereira-Santaella, M., Colina, L., García-Burillo, S., et al. 2018, A&A, 616, A171

Perivolaropoulos, L. & Skara, F. 2022, Challenges for ΛCDM: An Update

Péroux, C., Bouché, N., Kulkarni, V. P., York, D. G., & Vladilo, G. 2011, MNRAS, 410, 2237

Péroux, C. & Howk, J. C. 2020, Annu. Rev. Astron. Astrophys., 58, 363

Péroux, C., McMahon, R. G., Storrie-Lombardi, L. J., & Irwin, M. J. 2003, MNRAS, 346, 1103

Péroux, C., Meiring, J. D., Kulkarni, V. P., et al. 2006, MNRAS, 372, 369

Peroux, C. & Nelson, D. 2024, The Multi-Scale Multi-Phase Circumgalactic Medium: Observed and Simulated

Péroux, C., Nelson, D., van de Voort, F., et al. 2020, MNRAS, 499, 2462

Péroux, C., Quiret, S., Rahmani, H., et al. 2016, MNRAS, 457, 903

Péroux, C., Rahmani, H., Quiret, S., et al. 2017, MNRAS, 464, 2053

Péroux, C., Weng, S., Karki, A., et al. 2022, MNRAS, 516, 5618

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

Pettini, M., Ellison, S. L., Steidel, C. C., & Bowen, D. V. 1999, ApJ, 510, 576

Pettini, M. & Pagel, B. E. J. 2004, MNRAS, 348, L59

Pillepich, A., Nelson, D., Hernquist, L., et al. 2018, MNRAS, 475, 648

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6

Pointon, S. K., Nielsen, N. M., Kacprzak, G. G., et al. 2017, ApJ, 844, 23

Ponnada, S. B., Panopoulou, G. V., Butsky, I. S., et al. 2022, MNRAS, 516, 4417

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

Popping, G., Narayanan, D., Somerville, R. S., Faisst, A. L., & Krumholz, M. R. 2019a, MNRAS, 482, 4906

Popping, G., Pillepich, A., Somerville, R. S., et al. 2019b, ApJ, 882, 137

Popping, G., Somerville, R. S., & Trager, S. C. 2014, MNRAS, 442, 2398

Popping, G., van Kampen, E., Decarli, R., et al. 2016, MNRAS, 461, 93

Popping, G., Walter, F., Behroozi, P., et al. 2020, ApJ, 891, 135

Posses, A., Aravena, M., González-López, J., et al. 2025, A&A, 699, A256

Posses, A. C., Aravena, M., González-López, J., et al. 2023, A&A, 669, A46

Pozzetti, L., Bolzonella, M., Zucca, E., et al. 2010, A&A, 523, A13

Pozzi, F., Calura, F., D'Amato, Q., et al. 2024, The ALPINE-ALMA [CII] Survey: Dust Emission Effective Radius up to 3 Kpc in the Early Universe

Predehl, P., Andritschke, R., Arefiev, V., et al. 2021, A&A, 647, A1

Prochaska, J. X. & Neeleman, M. 2018, MNRAS, 474, 318

Prochaska, J. X., Werk, J. K., Worseck, G., et al. 2017, ApJ, 837, 169

Prochaska, J. X. & Wolfe, A. M. 1998, ApJ, 507, 113

Prochaska, J. X. & Wolfe, A. M. 2009, ApJ, 696, 1543

Puga, M. C., Balmaverde, B., Capetti, A., Massaro, F., & Gilli, R. 2025, A&A, 698, A241

Puskás, D., Tacchella, S., Simmonds, C., et al. 2025, Constraining the Major Merger History of $z \sim 3-9$ Galaxies Using JADES: Dominant in-Situ Star Formation

Putman, M. E., Peek, J. E. G., & Joung, M. R. 2012, Annu. Rev. Astron. Astrophys., 50, 491

Qu, Z., Chen, H.-W., Johnson, S. D., et al. 2024, ApJ, 968, 8

Quiret, S., Péroux, C., Zafar, T., et al. 2016, MNRAS, 458, 4074

Rafelski, M., Wolfe, A. M., Prochaska, J. X., Neeleman, M., & Mendez, A. J. 2012, ApJ, 755, 89

Ragone-Figueroa, C., Granato, G. L., Parente, M., et al. 2024, A&A, 691, A200

Rahmani, H., Péroux, C., Augustin, R., et al. 2018a, MNRAS, 474, 254

Rahmani, H., Péroux, C., Schroetter, I., et al. 2018b, MNRAS, 480, 5046

Rangwala, N., Maloney, P. R., Glenn, J., et al. 2011, ApJ, 743, 94

Rees, M. J. & Ostriker, J. P. 1977, MNRAS, 179, 541

Rhee, J., Lah, P., Briggs, F. H., et al. 2018, MNRAS, 473, 1879

Rhee, J., Meyer, M., Popping, A., et al. 2023, Monthly Notices of the Royal Astronomical Society, 518, 4646

Riaz, S., Hartwig, T., & Latif, M. A. 2022, ApJL, 937, L6

Richings, A. J. & Faucher-Giguère, C.-A. 2018, MNRAS, 474, 3673

Richings, A. J., Faucher-Giguere, C.-A., Gurvich, A. B., Schaye, J., & Hayward, C. C. 2022, MNRAS, 517, 1557

Richings, A. J. & Schaye, J. 2016, MNRAS, 458, 270

Richter, P., Nuza, S. E., Fox, A. J., et al. 2017, A&A, 607, A48

Riechers, D. A., Boogaard, L. A., Decarli, R., et al. 2020a, ApJL, 896, L21

Riechers, D. A., Carilli, C. L., Casey, C., et al. 2017, in American Astronomical Society Meeting Abstracts #229, Vol. 229, 348.12

Riechers, D. A., Hodge, J. A., Pavesi, R., et al. 2020b, ApJ, 895, 81

Riechers, D. A., Pavesi, R., Sharon, C. E., et al. 2019, ApJ, 872, 7

Rosenberg, M. J. F., van der Werf, P. P., Aalto, S., et al. 2015, ApJ, 801, 72

Rowan-Robinson, M. 2001, New Astr. Rev., 45, 631

Rowson, B. 1962, MNRAS, 125, 177

Rubin, K. H. R., Xavier Prochaska, J., Koo, D. C., & Phillips, A. C. 2012, ApJL, 747, L26

Ruffa, I., Prandoni, I., Davis, T. A., et al. 2022, MNRAS, 510, 4485

Ruffa, I., Spavone, M., Iodice, E., et al. 2025, The Link between Galaxy Merger, Radio Jet Expansion and Molecular Outflow in the ULIRG IRAS 00183-7111

Rupke, D. S., Veilleux, S., & Sanders, D. B. 2005, ApJS, 160, 87

Saintonge, A. & Catinella, B. 2022, Annu. Rev. Astron. Astrophys., 60, 319

Saintonge, A., Catinella, B., Tacconi, L. J., et al. 2017, ApJS, 233, 22

Salomé, Q., Salomé, P., Miville-Deschênes, M.-A., Combes, F., & Hamer, S. 2018, A&A, 617, C3

Salpeter, E. E. 1955, ApJ, 121, 161

Salvadori, S. & Ferrara, A. 2012, MNRAS, 421, L29

Sánchez, S. F., Kennicutt, R. C., de Paz, A. G., et al. 2012, A&A, 538, A8

Sánchez Almeida, J., Elmegreen, B. G., Muñoz-Tuñón, C., & Elmegreen, D. M. 2014, Astron Astrophys Rev, 22, 71

Sanders, D. B. & Mirabel, I. F. 1996, Annu. Rev. Astron. Astrophys., 34, 749

Sanders, G. H. 2013, J Astrophys Astron, 34, 81

Santini, P., Fontana, A., Grazian, A., et al. 2012, A&A, 538, A33

Santos, M. G., Bull, P., Alonso, D., et al. 2015, Cosmology with a SKA HI Intensity Mapping Survey

Schaerer, D., Ginolfi, M., Bethermin, M., et al. 2020, A&A, 643, A3

Scharré, L., Sorini, D., & Davé, R. 2024, MNRAS, 534, 361

Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521

Schechter, P. 1976, ApJ, 203, 297

Schinnerer, E. & Leroy, A. K. 2024

Schmidt, M. 1959, ApJ, 129, 243

Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, A&A, 575, A74

Schreiber, N. M. F., Übler, H., Davies, R. L., et al. 2019, ApJ, 875, 21

Schroetter, I., Bouché, N., Péroux, C., et al. 2015, ApJ, 804, 83

Schroetter, I., Bouché, N., Wendt, M., et al. 2016, ApJ, 833, 39

Schruba, A., Leroy, A. K., Walter, F., et al. 2012, AJ, 143, 138

Scoville, N., Lee, N., Bout, P. V., et al. 2017, ApJ, 837, 150

Scoville, N. Z., Yun, M. S., Clemens, D. P., Sanders, D. B., & Waller, W. H. 1987, ApJS, 63, 821

Serra, P., Oosterloo, T., Morganti, R., et al. 2012, MNRAS, 422, 1835

Serra, P., Westmeier, T., Giese, N., et al. 2015, MNRAS, 448, 1922

Shangguan, J., Ho, L. C., Bauer, F. E., Wang, R., & Treister, E. 2020, ApJ, 899, 112

Shapiro, P. R. & Field, G. B. 1976, ApJ, 205, 762

Sharma, M., Arav, N., Dehghanian, M., et al. 2025, A&A

Sharma, S. S., Roy, J., Kudale, S., et al. 2023, ApJ, 947, 88

Shen, S., Madau, P., Aguirre, A., et al. 2012, ApJ, 760, 50

Shen, S., Madau, P., Conroy, C., Governato, F., & Mayer, L. 2014, ApJ, 792, 99

Shivaei, I. & Boogaard, L. 2024, The Tight Correlation of PAH and CO Emission from Z~0-4

Sibony, Y., Liu, B., Simmonds, C., Meynet, G., & Bromm, V. 2022, A&A, 666, A199

Silk, J. & Mamon, G. A. 2012, RAA, 12, 917

Smethurst, R. J., Lintott, C. J., Simmons, B. D., et al. 2015, MNRAS, 450, 435

Smit, R., Bouwens, R. J., Labbé, I., et al. 2016, ApJ, 833, 254

Smit, R., Bouwens, R. J., Labbé, I., et al. 2014, ApJ, 784, 58

Smoot, G. F., Bennett, C. L., Kogut, A., et al. 1992, ApJ, 396, L1

Solimano, M., González-López, J., Aravena, M., et al. 2024, The ALMA-CRISTAL Survey. Discovery of a 15 Kpc-Long Gas Plume in a $z=4.54$ Lyman-α Blob

Solomon, P. M., Downes, D., & Radford, S. J. E. 1992, ApJ, 387, L55

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

Solomon, P. M., Rivolo, A. R., Barrett, J., & Yahil, A. 1987, ApJ, 319, 730

Somerville, R. S. & Davé, R. 2015, Annu. Rev. Astron. Astrophys., 53, 51

Somerville, R. S., Primack, J. R., & Faber, S. M. 2001, MNRAS, 320, 504

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

Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, ApJS, 148, 175

Spitoni, E., Calura, F., Mignoli, M., et al. 2020, A&A, 642, A113

Spitzer, L. 1978, Physical Processes in the Interstellar Medium

Springel, V. & Hernquist, L. 2003, MNRAS, 339, 289

Springel, V., Pakmor, R., Pillepich, A., et al. 2018, MNRAS, 475, 676

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629

Stach, S. M., Dudzevičiūtė, U., Smail, I., et al. 2019, MNRAS, 487, 4648

Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M. 1999, ApJ, 519, 1

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

Stern, J., Hennawi, J. F., Prochaska, J. X., & Werk, J. K. 2016, ApJ, 830, 87

Sternberg, A., Le Petit, F., Roueff, E., & Le Bourlot, J. 2014, ApJ, 790, 10

Stocke, J. T., Keeney, B. A., Danforth, C. W., et al. 2013, ApJ, 763, 148

Strickland, D. K., Heckman, T. M., Colbert, E. J. M., Hoopes, C. G., & Weaver, K. A. 2004, ApJ, 606, 829

Strong, A. W. & Mattox, J. R. 1996, A&A, 308, L21

Su, R., Gu, M., Curran, S. J., et al. 2023, ApJ, 956, L28

Su, Y., Fang, G., Lu, S., & Lin, Z. 2025, A&A, 699, A184

Sun, F., Wang, F., Yang, J., et al. 2025, ApJ, 980, 12

Sushch, I. & Böttcher, M. 2015, A&A, 573, A47

Switzer, E. R., Masui, K. W., Bandura, K., et al. 2013, MNRAS, 434, L46

Szakacs, R., Péroux, C., Zwaan, M., et al. 2021, MNRAS, 505, 4746

Tacchella, S., Dekel, A., Carollo, C. M., et al. 2016, MNRAS, 458, 242

Tacchella, S., Smith, A., Kannan, R., et al. 2021, H-Alpha Emission in Local Galaxies: Star Formation, Time Variability and the Diffuse Ionized Gas

Tacconi, L. J., Genzel, R., Neri, R., et al. 2010, Nature, 463, 781

Tacconi, L. J., Genzel, R., Saintonge, A., et al. 2018, ApJ, 853, 179

Tacconi, L. J., Genzel, R., & Sternberg, A. 2020, Annu. Rev. Astron. Astrophys., 58, 157

Tacconi, L. J., Neri, R., Genzel, R., et al. 2013, ApJ, 768, 74

Tadhunter, C. 2016, A&AR, 24, 10

Tanimura, H., Aghanim, N., Douspis, M., & Malavasi, N. 2022, A&A, 667, A161

Tapia-Contreras, B., Tissera, P. B., Sillero, E., et al. 2025, Insight into the Physical Processes That Shape the Metallicity Profiles in Galaxies

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

Thompson, A. R., Moran, J. M., & Swenson, Jr., G. W. 2017, Interferometry and Synthesis in Radio Astronomy, 3rd Edition

Thompson, O. A., Richings, A. J., Gibson, B. K., et al. 2024, Predictions for CO Emission and the CO-to-H$_2$ Conversion Factor in Galaxy Simulations with Non-Equilibrium Chemistry

Thorne, J. E., Robotham, A. S. G., Davies, L. J. M., et al. 2021, MNRAS, 505, 540

Tielens, A. G. G. M. 2005, The Physics and Chemistry of the Interstellar Medium (Cambridge University Press)

Tissera, P. B., Bignone, L., Gonzalez-Jara, J., et al. 2025, A&A, 697, A134

Tollet, É., Cattaneo, A., Macciò, A. V., Dutton, A. A., & Kang, X. 2019, MNRAS, 485, 2511

Tollet, É., Cattaneo, A., Macciò, A. V., & Kang, X. 2022, MNRAS, 515, 3453

Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2014, ApJ, 783, 85

Trayford, J. W., Schaye, J., Correa, C., et al. 2025, Modelling the Evolution and Influence of Dust in Cosmological Simulations That Include the Cold Phase of the Interstellar Medium

Tremblay, G. R., Combes, F., Oonk, J. B. R., et al. 2018, ApJ, 865, 13

Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898

Tumlinson, J., Peeples, M. S., & Werk, J. K. 2017, Annu. Rev. Astron. Astrophys., 55, 389

Tumlinson, J., Thom, C., Werk, J. K., et al. 2011, Science, 334, 948

Urry, C. M. & Padovani, P. 1995, PASP, 107, 803

Vallini, L., Ferrara, A., Pallottini, A., Carniani, S., & Gallerani, S. 2021, MNRAS, 505, 5543

Vallini, L., Gallerani, S., Ferrara, A., Pallottini, A., & Yue, B. 2015, ApJ, 813, 36

Vallini, L., Pallottini, A., Ferrara, A., et al. 2018, MNRAS, 473, 271

Vallini, L., Witstok, J., Sommovigo, L., et al. 2024, MNRAS, 527, 10

van Cappellen, W. A., Oosterloo, T. A., Verheijen, M. a. W., et al. 2022, A&A, 658, A146

van Cittert, P. H. 1934, Physica, 1, 201

van de Voort, F., Bahé, Y. M., Bower, R. G., et al. 2017, MNRAS, 466, 3460

van de Voort, F., Bieri, R., Pakmor, R., et al. 2021, MNRAS, 501, 4888

van de Voort, F., Schaye, J., Booth, C. M., Haas, M. R., & Dalla Vecchia, C. 2011, MNRAS, 414, 2458

van de Voort, F., Springel, V., Mandelker, N., van den Bosch, F. C., & Pakmor, R. 2019, MNRAS, 482, L85

van Kampen, E., Bakx, T., De Breuck, C., et al. 2024, Open Res Eur, 4, 122

Veilleux, S., Cecil, G., & Bland-Hawthorn, J. 2005, Annu. Rev. Astron. Astrophys., 43, 769

Veilleux, S., Maiolino, R., Bolatto, A. D., & Aalto, S. 2020, Astron Astrophys Rev, 28, 2

Villanueva, V., Herrera-Camus, R., Gonzalez-Lopez, J., et al. 2024, The ALMA-CRISTAL Survey: Dust Temperature and Physical Conditions of the Interstellar Medium in a Typical Galaxy at Z=5.66

Voit, G. M. 2005, Rev. Mod. Phys., 77, 207

Walter, F., Bañados, E., Carilli, C., et al. 2025, ApJ, 983, L8

Walter, F., Brinks, E., de Blok, W. J. G., et al. 2008, AJ, 136, 2563

Walter, F., Carilli, C., Neeleman, M., et al. 2020, ApJ, 902, 111

Walter, F., Decarli, R., Aravena, M., et al. 2016, ApJ, 833, 67

Walter, F., Weiß, A., Downes, D., Decarli, R., & Henkel, C. 2011, ApJ, 730, 18

Wang, J., Yang, D., Lin, X., et al. 2025, ApJ, 980, 25

Wang, J., Yang, D., Oh, S. H., et al. 2023, ApJ, 944, 102

Wang, L., Dutton, A. A., Stinson, G. S., et al. 2015, MNRAS, 454, 83

Waterval, S., Cannarozzo, C., & Macciò, A. V. 2025, Gas Accretion at High Redshift: Cold Flows All the Way

Waterval, S., Macciò, A. V., Buck, T., et al. 2024, MNRAS, 533, 1463

Weaver, J. R., Davidzon, I., Toft, S., et al. 2023, A&A, 677, A184

Weinberger, R., Springel, V., & Pakmor, R. 2020, ApJS, 248, 32

Weiner, B. J., Coil, A. L., Prochaska, J. X., et al. 2009, ApJ, 692, 187

Weiß, A., Walter, F., & Scoville, N. Z. 2005, A&A, 438, 533

Weng, S., Péroux, C., Karki, A., et al. 2023a, MNRAS, 523, 676

Weng, S., Péroux, C., Karki, A., et al. 2022, MUSE-ALMA Haloes VIII: Statistical Study of Circumgalactic Medium Gas

Weng, S., Péroux, C., Karki, A., et al. 2023b, MNRAS, 519, 931

Weng, S., Péroux, C., Ramesh, R., et al. 2024, MNRAS, 527, 3494

Weng, S., Sadler, E. M., Kerrison, E., et al. 2025, MNRAS, 539, 1977

Wenger, M., Ochsenbein, F., Egret, D., et al. 2000, Astron. Astrophys. Suppl. Ser., 143, 9

Werk, J. K., Prochaska, J. X., Tumlinson, J., et al. 2014, ApJ, 792, 8

Westmeier, T., Kitaeff, S., Pallot, D., et al. 2021, MNRAS, 506, 3962

Whitaker, K. E., Franx, M., Leja, J., et al. 2014, ApJ, 795, 104

Whitaker, K. E., Williams, C. C., Mowla, L., et al. 2021, Nature, 597, 485

White, S. D. M. & Frenk, C. S. 1991, ApJ, 379, 52

White, S. D. M. & Rees, M. J. 1978, MNRAS, 183, 341

Wilson, T. L., Rohlfs, K., & Hüttemeister, S. 2013, Tools of Radio Astronomy

Wisotzki, L., Bacon, R., Blaizot, J., et al. 2016, A&A, 587, A98

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

Wolfe, A. M., Gawiser, E., & Prochaska, J. X. 2005, Annu. Rev. Astron. Astrophys., 43, 861

Wolfire, M. G., Hollenbach, D., & McKee, C. F. 2010, ApJ, 716, 1191

Wolfire, M. G., Hollenbach, D., McKee, C. F., Tielens, A. G. G. M., & Bakes, E. L. O. 1995, ApJ, 443, 152

Wolfire, M. G., McKee, C. F., Hollenbach, D., & Tielens, A. G. G. M. 2003, ApJ, 587, 278

Wright, A. H., Driver, S. P., & Robotham, A. S. G. 2018, MNRAS, 480, 3491

Wu, X., Cai, Z., Lan, T.-W., et al. 2024, Tracing the Evolution of the Cool Gas in CGM and IGM Environments through Mg II Absorption from Redshift Z=0.75 to Z=1.65 Using DESI-Y1 Data

Wuyts, E., Kurk, J., Förster Schreiber, N. M., et al. 2014, ApJL, 789, L40

Xi, H., Peng, B., Staveley-Smith, L., et al. 2024, ApJL, 966, L36

Yan, L., Sajina, A., Loiacono, F., et al. 2020, ApJ, 905, 147

Yeung, M. C. H., Freyberg, M. J., Ponti, G., et al. 2023, A&A, 676, A3

Yeung, M. C. H., Ponti, G., Freyberg, M. J., et al. 2024, A&A, 690, A399

Yoon, H., Sadler, E. M., Mahony, E. K., et al. 2024, The First Large Absorption Survey in HI (FLASH): II. Pilot Survey Data Release and First Results

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

York, D. G., Dopita, M., Green, R., & Bechtold, J. 1986, ApJ, 311, 610

Yu, N., Ho, L. C., & Wang, J. 2022, ApJ, 930, 85

Yu, N.-P., Zhu, M., Xu, J.-L., et al. 2024, MNRAS, 532, 1744

Yung, L. Y. A., Somerville, R. S., Finkelstein, S. L., et al. 2023, MNRAS, 519, 1578

Zahedy, F. S., Chen, H.-W., Cooper, T. M., et al. 2021, MNRAS, 506, 877

Zahid, H. J., Kashino, D., Silverman, J. D., et al. 2014, ApJ, 792, 75

Zanella, A., Daddi, E., Magdis, G., et al. 2018, MNRAS, 481, 1976

Zavala, J. A., Casey, C. M., Manning, S. M., et al. 2021, ApJ, 909, 165

Zernike, F. 1938, Physica, 5, 785

Zhang, Y., Comparat, J., Ponti, G., et al. 2025, A&A, 693, A197

Zheng, H., Tegmark, M., Dillon, J. S., et al. 2017, MNRAS, 464, 3486

Zu, Y. & Mandelbaum, R. 2016, MNRAS, 457, 4360

Zwaan, M., Ivison, R., Peroux, C., et al. 2022, ALMACAL: Surveying the Universe with ALMA Calibrator Observations, https://doi.eso.org/10.18727/0722-6691/5256

Zwaan, M. A., Meyer, M. J., Staveley-Smith, L., & Webster, R. L. 2005, MNRAS, 359, L30

List of Publications

First Author Publications

During the duration of the doctoral degree, the author has undertaken the following works as the primary author, which were included in this thesis as Chapters 2, 3 & 4.

- **MUSE-ALMA Haloes XIII: Molecular gas in $z \sim 0.5$ H I–selected galaxies**
  Bollo, V.; Péroux, C.; Zwaan, M.; et al. submitted to A&A on July 21, 2025
- **ALMACAL XIII. Evolution of the CO luminosity function and the molecular gas density out to $z \sim 6$**
  Bollo, V.; Péroux, C.; Zwaan, M.; et al. (2025) A&A, 695, A163.
- **ALMACAL XII. Data Characterisation and Products**
  Bollo, V.; Zwaan, M.; Péroux, C.; et al. (2024) A&A, 690, 258.

Contributed Author Publications

These include works related to the thesis, as well as other research activities the author engaged in during the PhD. This also includes publications from the author's master's research completed during this period.

- **ALMACAL XIV: X-Shooter Spectroscopy, Infrared Properties and Radio SEDs of Calibrators**
  Weng, S; Sadler, E.; Kerrison, E.; Bollo, V.; et al., 2025, MNRAS, 539, 1977.
- **An Hα view of galaxy build-up in the first 2 Gyr: luminosity functions at $z \sim 4 - 6.5$ from NIRCam/grism spectroscopy**
  Covelo-Paz, A.; Giovinazzo; E; et al. (incl. Bollo, V.) 2025, A&A, 694, A178.

- **The physical origins of gas in the circumgalactic medium using observationally motivated TNG50 mocks**
  Weng, S.; Péroux, C.; et al. (incl. Bollo, V.) 2024, MNRAS, 527, 3494.

- **ALMACAL. XI. Over-densities as signposts for proto-clusters? A cautionary tale**
  Chen, J., Ivison, R. J.; et al. (incl. Bollo, V.) 2023, A&A, 675, L10.

- **MUSE-ALMA Haloes XI: gas flows in the circumgalactic medium**
  Weng, S.; Péroux, C.; et al. (incl. Bollo, V.) 2023, MNRAS, 523, 676.

- **The BarYon Cycle project (ByCycle): Identifying and localizing Mg II metal absorbers with machine learning**
  Szakacs, R.; Péroux, C.; et al. (incl Bollo, V.) 2023, MNRAS, 526, 3744.

- **The Hα Luminosity Function of Galaxies at** $z \sim 4.5$
  Bollo, V.; González, V.; Stefanon, M.; et al. (2023), ApJ, 946, 117.

Acknowledgements

Exactly three years ago, I set foot for the first time in a distant and unfamiliar land. Carrying a suitcase full of emotions, I still vividly recall the mingling of excitement and fear. I survived this journey, while trying to enjoy it, thanks to the amazing people who surrounded me.

First of all, my deepest thanks to my supervisors: Céline and Martin. For your constant support, your time, your patience, and for caring deeply while pushing me to reach places I never imagined I could. Thank you for giving me this opportunity and choosing me to carry out such exciting projects. I will keep all the lessons I have learned from you with me always. You are both a true source of inspiration, and I can only hope to be like you as I continue my career. To my colleagues and team: Alex, Jianhang, Simon, Rob, Varsha, Benedetta, Capucine, Tamsyn, Nico; thank you for the warm and supportive environment you create. I feel incredibly lucky to do science along with such kind and brilliant people.

Thanks to ESO and its warming environment. It is the people who transform this labyrinthine building into a pleasant place to do science. I feel so lucky to have spent three years here. I have met incredible people, who have inspired me deeply. Sharing this experience with folks from all around the world has been such an eye-opening experience, both professionally and personally. From casual conversations to words of encouragement, I have treasured every moment. I greatly appreciate the respect and consideration I have always received, and I am grateful for the many opportunities to learn and grow during my time here. In particular, I wish to thank my mentors: Paula, Martina, and Amelia. I have relied on you to share the weight of my thoughts, anxieties and insecurities related to work, and I always felt supported and heard. I hold a deep admiration for each of you.

Thanks to all my friends, and especially to Veronica, having your company has been a like a warm hug for my heart from the very beginning, and I am so grateful that our paths crossed. Gracias a mis compañeras de NiUnaMenosMunich: en este espacio encontré un

lugar cercano a mis raíces y a aquello que me mueve en el mundo. Ejercer el feminismo junto a mujeres guerreras ha sido una de las experiencias más hermosas que me ha dado migrar por primera vez. Gracias a mis amigos de Chile, por mantenerse en contacto y seguir creyendo en mí. En especial a Alex, por tu lealtad y estar siempre presente en mis tristezas y alegrías.

Gracias a mi familia. Los extraño tanto que no puedo contener las lágrimas al escribir estas palabras. Siempre los llevo conmigo, en todo lo que hago y a cada lugar al que voy. Gracias a las familia Avendaño, Santibáñez, Orellana, Doizi, por estar siempre presentes, por sus gestos de amistad y cariño. Gracias a mi segunda familia, Farías Espinoza, por recibirme con los brazos abiertos desde el principio, y por su apoyo incondicional. Gracias a mis hermanos. Andrea, gracias por visitarme, por ser constante; apoyarnos en nuestros procesos de sanación y crecimiento ha sido un regalo precioso de la vida, y agradezco la profunda conexión que estamos construyendo. Mariano, gracias por cuidarme, por tus detalles de cariño, por apoyarme siempre que te he necesitado y animarme a perseguir mis sueños. Gracias a Don Gato y Lola, por enseñarme una de las maneras más hermosas de amar; perderlos a la distancia es una herida que sigue sanando. A mis espíritus, papá y abuela, por todas las enseñanzas; me gusta pensar que siguen cuidando de mí. Gracias a mi mamá, la lucecita más brillante de mi vida. Gracias por dar tanto de ti para que pudiera cumplir mi sueños. Todos mis logros son tuyos. Siempre lo han sido y siempre lo serán. Gracias a mi Frani, mi compañera de vida. La vida es infinitamente más bella contigo a mi lado. Gracias por tu paciencia, tu apoyo incondicional, y por tanto amor. Siempre recuerdo la inmensa emoción que sentí cuando llegaste a vivir a Alemania conmigo, y ahora me siento increíblemente afortunada de que quieras seguir recorriendo el camino de la vida a mi lado. Contigo, mi hogar, hasta el fin del mundo.

"Gracias a la vida que me ha dado tanto
me dio dos luceros que, cuando los abro
perfecto distingo, lo negro del blanco
y en el alto cielo su fondo estrellado"
– Violeta Parra