



PHYSICS AREA - PH.D. COURSE IN
ASTROPHYSICS AND COSMOLOGY

**VISIBILITY STACKING OF
ALMA ARCHIVAL DATA:
UNVEILING THE MOLECULAR ISM
OF DUSTY STAR-FORMING
GALAXIES AT COSMIC NOON**

Candidate:
MARTINA TORSSELLO

Supervisors:
ANDREA LAPI
MARCELLA MASSARDI
FRANCESCA PERROTTA

ACADEMIC YEAR 2025/2026

*To Arte,
who illuminated the darkest steps of my path, with the vow to be a light
unto her life in return.*

"Dang it, Jim! I'm an astronomer, not a doctor! I mean, I am a doctor, but I'm not that kind of doctor. I have a doctorate, it's not the same thing. You can't help people with a doctorate, you just sit there, and you're useless!"

— Dr. Delbert Doppler, *Treasure Planet* (Disney, 2002)

ACKNOWLEDGMENTS

This page is an opportunity for me to thank everyone who, even with just a small crumb of care and attention, helped me achieve this result and navigate these past four years.

The biggest thank you of all goes to Giacomo, my constant. Everything I have become, the goals I have reached, and the parts of myself that I like the most are all thanks to him. Thank you for all the strength, fun, and love you have given me over these years, and that I know you will continue to give me. If I stand here today, proud of myself, it is thanks to you, and to your *pesce-lesso* face that is crying while reading these words.

Thank you to my precious colleagues, official and "adopted": Carlos, Emma, Claudio, Sara, Florie, Lorenzo and Franceschis, for always reminding me that I was never alone and that, even in the heaviest moments, we were all in the same boat. A special thanks to Carlos, my almost-boyfriend (or as he would say, "eh magari"), because from day zero when I rescued him, he has been an incredible friend. I know that with him, and with all the others, these four years won't just be a temporary friendship, but merely the years we spent living in the same city. All the rest is yet to come.

Thank you to Gayathri: a friend, companion, teacher, and student, everything I sought and needed during this PhD. Life might take us on different paths, but the impact we have had on each other's journey has been an immense gift. I will always be grateful to you, both from a scientific perspective and, above all, on a human level. I am so happy to have had the privilege of knowing a wonderful person and researcher like you.

Thank you to everyone at SISSA, to APCs, APGs, TPPs, and all the other silly acronyms. To all the students and postdocs who, even with just a brief chat or some company, lightened this journey and gave each other strength. To those who were there until the very last year and to all those we have said goodbye to, you made this environment a wonderful place to work. A special thank you to Debora, Barbara, and Patrizia for the human and sweet touch you brought to my days, even if only for a couple of minutes a day.

Thank you to all my HPC Master's classmates, especially Andrea and Francesco, who, like a true military regiment, defended and ac-

accompanied me in the trenches. Together, we survived the endless sieges of sleepless nights, fought side-by-side on the front lines of impossible deadlines, and held our ground when giving up seemed like the easiest way out. After all, we won the war together.

Thank you to all my supervisors and to all the professors, researchers, PhD students, and postdocs who have known, helped, and supported me, shaping me into the researcher I am today and allowing me to forge my unique academic path.

Thanks to my nutcrackers: Celeste, Mate, Martina, Fede, and Angi, my chosen family. Thank you for the strength you give me every day through your stories and passions, and for making me feel perfect just being myself.

Thanks to all of Ciketo, the Girls group, and all my friends in Trieste. I cannot name you all because I would surely forget someone and feel terrible about it, so please forgive me. To the ones who stayed, the ones who left, and the ones who return: every single one of you, for a specific time, gave me a part of yourselves and your company, making my life shine in a unique way. You will always be a home for me to return to, no matter where any of you are in the world.

Thanks to my sister and her new family, Antonio and Arte, for making my problems lighter. And to my mother, who loves and supports me, even if she sometimes drives me crazy (but you can tell her only the first part, because she doesn't understand English ♡).

Finally, I have to thank the city of Trieste. For 11 years, it was the cradle of my adulthood, a father who taught me to fly and granted me the freedom to carve my own path.

ABSTRACT

Dusty star-forming galaxies (DSFGs) host some of the most intense episodes of star formation in cosmic history, dominating the cosmic star formation rate density at $z \sim 2-3$, during the cosmic noon. Characterising the state of their molecular gas is an essential step towards understanding how galaxies assembled their stellar mass across cosmic time. Submillimetre spectral lines observed with interferometers such as ALMA provide the ideal diagnostic of these physical conditions. However, individual transitions are often faint, making stacking techniques essential to recover weak signals buried in noise and to build reliable statistical studies using public archives. Because these galaxies have been targeted across numerous independent programs, archival datasets are deeply heterogeneous in angular resolution, spectral setup, sensitivity, and redshift. Combining such data via post-processed images introduces severe artifacts and beam discrepancies, highlighting the need for a novel methodology capable of handling disparate datasets across broad redshift ranges.

In this thesis, we address this challenge by presenting ViSTA, a stacking framework that operates natively in the Fourier domain. By combining data directly on a common rest-frame uv plane and spectral grid, ViSTA preserves native noise statistics and eliminates image-plane artifacts. Validated on simulated and real observations, ViSTA reliably recovers faint line signals buried in noise. Notably, the noisier the data, the greater the gain over traditional image-plane methods. Although an initial version tailored for ALMA demonstrated the validity of this approach, poor computational scaling limited its application to large archival samples. We therefore re-engineered the framework using high-performance computing, achieving over an order-of-magnitude speedup and enabling joint stacking across multiple interferometers. We then apply ViSTA to a sample of 104 DSFGs observed with ALMA to reconstruct their average CO Spectral Line Energy Distribution, constrain excitation conditions, and evaluate gravitational lensing biases. The L_{IR} -normalised ladder peaks at mid- J transitions, tracing warm, moderately dense gas. Confirmed lensed and unlensed systems yield consistent ladders and excitation parameters with no evidence for differential lensing, supporting strongly lensed galaxies as proxies for the broader population. Finally, non-LTE modelling confirms that a two-component ISM combining warm and cool gas best reproduces the observed emission, yielding physical conditions aligned with previous studies. These findings establish a benchmark for high-redshift ISM studies and demonstrate the power of uv -stacking for archival science in the SKAO era.

PUBLICATIONS

The core results and contents of this thesis are based on the following main publications:

1. Martina Torsello et al. “The ViSta Method for Stacking in the Fourier Domain and its Application to the Dusty Star-forming Galaxies in the ALMA Science Archive.” In: *PASP* 137.12, 124501 (2025). doi: [10.1088/1538-3873/ae1faf](https://doi.org/10.1088/1538-3873/ae1faf).
2. Martina Torsello et al. “The ViSta method for optimized stacking of broadband interferometric data in the Fourier domain.” In: *Proceedings of SPIE*. Vol. 14153. SPIE, 2026. doi: [10.1117/12.3103792](https://doi.org/10.1117/12.3103792).
3. Martina Torsello et al. “The average CO spectral line energy distribution of dusty star-forming galaxies at cosmic noon from ALMA uv-plane stacking in the ALMA Science Archive.” In: *Astronomy & Astrophysics* (2026). Submitted.

The following co-authored publications directly enriched and supported the research presented in this thesis:

1. Francesca Perrotta, Martina Torsello, Marika Giulietti, and Andrea Lapi. “Astrochemistry of the Molecular Gas in Dusty Star-Forming Galaxies at the Cosmic Noon.” In: *Galaxies* 12.2, 18 (2024). doi: [10.3390/galaxies12020018](https://doi.org/10.3390/galaxies12020018).
2. T. Ronconi et al. “GalaPy: A highly optimised C++/Python spectral modelling tool for galaxies. I. Library presentation and photometric fitting.” In: *A&A* 685, A161 (2024). doi: [10.1051/0004-6361/202346978](https://doi.org/10.1051/0004-6361/202346978).
3. Francesca Perrotta et al. “The Way of Water: ALMA Resolves H₂O Emission Lines in a Strongly Lensed Dusty Star-forming Galaxy at $z \sim 3.1$.” In: *ApJ* 952.1, 90 (2023). doi: [10.3847/1538-4357/accd72](https://doi.org/10.3847/1538-4357/accd72).
4. Marika Giulietti et al. “ALMA Resolves the First Strongly Lensed Optical/Near-IR-dark Galaxy.” In: *ApJ* 943.2, 151 (2023). doi: [10.3847/1538-4357/aca53f](https://doi.org/10.3847/1538-4357/aca53f).

CONTENTS

1	INTRODUCTION	1
1.1	Dusty Star-forming Galaxies	2
1.1.1	The obscured cosmic star formation history	2
1.1.2	Discovery, selection and physical properties	4
1.1.3	Strong gravitational lensing as a natural telescope	8
1.1.4	The role of DSFGs in galaxy evolution	10
1.2	The interstellar medium and its tracers	12
1.2.1	Cold gas tracers: Carbon Monoxide and Atomic Carbon	15
1.2.2	High- J CO transitions and the Spectral Line Energy Distribution	17
1.2.3	Other gas tracers	21
1.3	The need for a statistical approach	22
1.3.1	The Stacking technique	23
1.3.2	Exploiting public archives and their heterogeneity	25
1.4	Aim and structure of this thesis	26
2	THE SAMPLE	29
2.1	The selection criteria	29
2.1.1	Parent sample and flux density selection	29
2.1.2	Spectroscopic redshifts	32
2.2	ALMA: The ideal tool for studying high- z DSFGs	33
2.2.1	Arrays and configurations	33
2.2.2	Receivers, spectral setup, and signal calibration	33
2.2.3	The ALMA archive as a scientific resource	35
2.3	The final ALMA ensemble	37

2.4	Wrap-up	43
3	VISTA: A NEW STACKING METHOD IN THE FOURIER DOMAIN	45
3.1	Interferometry and the visibility domain	45
3.1.1	Principles of interferometry	46
3.1.2	The uv -plane and the Visibility function	49
3.1.3	Visibility noise and baseline sensitivity	50
3.1.4	Calibration errors	51
3.1.5	Spatial scales and response to source structures	52
3.2	From the Visibilities to the Image	55
3.2.1	The dirty image and the sampling function	55
3.2.2	Weighting and tapering	55
3.2.3	Deconvolution: CLEAN and data combination	56
3.2.4	Stacking before or after image reconstruction	57
3.3	The ViSta Method	59
3.3.1	A Common Rest-Frame uv Domain	59
3.3.2	Phase Centre Alignment and Primary Beam Correction	60
3.3.3	Spectral Alignment onto a Common Grid	62
3.4	Statistical combination of visibilities	63
3.4.1	Natural and democratic weighting	64
3.4.2	Flux normalisation	65
3.4.3	Continuum subtraction	66
3.5	Signal extraction	67
3.5.1	Extraction in the image domain	67
3.5.2	Extraction in the visibility domain	68
3.6	Wrap-up	69
4	VISTA IMPLEMENTATION FOR STACKING ALMA DATA	71
4.1	ALMA data format and the CASA software	71
4.1.1	The Measurement Set format	72

4.2	ViSta CASA-based pipeline	73
4.2.1	Pre-processing	74
4.2.2	Stacking	77
4.2.3	Post-processing	80
4.3	Code validation on simulated data	84
4.3.1	The simulated sample	84
4.3.2	Unresolved source	88
4.3.3	Resolved source	91
4.4	Application to a real case	98
4.5	Wrap-up	105
5	VISTA 2.0: THE OPTIMISED FRAME- WORK	107
5.1	Limitations of the CASA-based implementa- tion	107
5.2	A self-contained pipeline	108
5.2.1	Building the stacked dataset with <code>dask-ms</code>	110
5.2.2	The compute core	111
5.2.3	Post-processing	113
5.3	Flux density extraction	115
5.3.1	Compression into sufficient statistics	116
5.3.2	Line profile and uv -plane flux fitting	117
5.3.3	Uncertainties and null tests	118
5.4	Performance validation	120
5.5	Wrap-up	120
6	THE AVERAGE CO SLED OF DSFGS AT COS- MIC NOON	123
6.1	The ALMA sample	123
6.1.1	Lensing classification	127
6.1.2	Sample distributions	128
6.2	Stacking with ViSta	132
6.2.1	Data retrieval and spectral window selec- tion	132
6.2.2	Weighting scheme: Democratic	134

6.2.3	Flux normalisation: rescaling by infrared luminosities	134
6.2.4	Continuum subtraction and treatment of multiple sources	139
6.2.5	Signal extraction	140
6.3	Average CO SLED	143
6.3.1	Lensed vs unlensed sample	152
6.4	Average physical properties of the molecular gas	153
6.4.1	Dust continuum of the stack and gas mass cross-check	154
6.4.2	Single-component LVG modelling	154
6.4.3	Double component LVG modelling	158
6.5	Discussion	163
6.5.1	Stacked CO SLED	163
6.5.2	Gas derived properties	165
6.6	Wrap-up	167
7	CONCLUSIONS AND PROSPECTS	169
7.1	Future prospects	170
7.1.1	Expanding the CO ladder	170
7.1.2	Beyond CO: Additional tracers	172
7.1.3	Redshift evolution and subsamples	172
7.1.4	Beyond Dusty Star-Forming Galaxies (DSFGs): A general framework for visibility stacking	174
7.1.5	ViSTA in the SKAO era	175
	Appendix	179
A	SUPPORTING CODE FOR THE ANALYSIS	181
A.1	The mock observation simulator	181
A.2	Building the archival line coverage catalogue	184
A.3	Image-plane stacking	186
B	BENCHMARKING OF THE HPC PIPELINE	189

b.1	Benchmark setup	189
b.2	Results	190
C	PHOTOMETRIC FLUXES AND ALMA ARCHIVAL DATA	195
c.1	100–2100 μm observed fluxes	195
c.2	ALMA archival observations with a possible CO emission line detection	201
	BIBLIOGRAPHY	211

LIST OF FIGURES

Figure 1	The cosmic star formation rate density	3
Figure 2	Spectral energy distribution of a dusty star-forming galaxy	5
Figure 3	The negative K -correction	6
Figure 4	Schematic representation of strong gravitational lensing	9
Figure 5	Evolutionary link between DSFGs and local early-type galaxies	11
Figure 6	Evolution of the CO excitation ladder as a function of density and temperature	18
Figure 7	CO SLEDs of local galaxies and high-redshift SMGs	19
Figure 8	Average rest-frame spectrum of a sample of DSFGs	24
Figure 9	Cumulative submillimetre number counts at $500\,\mu\text{m}$	30
Figure 10	ALMA antenna arrays	34
Figure 11	ALMA bands and atmospheric transmission	35
Figure 12	Distribution of refereed publications across ALMA project codes	36
Figure 13	Single-antenna response	47
Figure 14	Schematic geometry of a two-element radio interferometer	48
Figure 15	uv coverage as a function of the number of antennas	50
Figure 16	Fourier-transform pairs for representative source morphologies	53
Figure 17	Rest-framing of the uv -plane for a simulated source	60
Figure 18	Main table of a Measurement Set	72
Figure 19	ViSta schematic pipeline	74
Figure 20	Rebinning and regridding of the spectral axis	78
Figure 21	Image and uv distribution of a simulated source	86
Figure 22	uv coverage of stacked simulated sources	87
Figure 23	Peak channel image of single vs. stacked source	88
Figure 24	ViSta-ImFit performance scaling for unresolved sources	89

Figure 25	Peak SNR as a function of the number of stacked sources	89
Figure 26	Stacking method comparison for unresolved sources	90
Figure 27	ViSta-ImFit performance scaling for resolved sources	92
Figure 28	Spectral line recovery using known source size	93
Figure 29	Spectral line recovery using estimated source size	94
Figure 30	Spectral line recovery at lowest common resolution	95
Figure 31	Stacked central channel images at lowest resolution	96
Figure 32	Stacked continuum emission maps	97
Figure 33	Stacked CO(3–2) spectra comparison	101
Figure 34	Stacked continuum image of the observational sample	102
Figure 35	Stacked CO(3–2) spectra at degraded sensitivity	103
Figure 36	Stacking method comparison across sensitivity regimes	104
Figure 37	Threading scheme of the streaming pipeline	110
Figure 38	Visibility-domain flux extraction	119
Figure 39	Redshift distribution of the sample	129
Figure 40	Angular resolution distribution of the sample	129
Figure 41	Observing year distribution of the sample	130
Figure 42	Distribution of unique CO observations per transition	131
Figure 43	Angular resolution of the observations per transition	132
Figure 44	Comparison of weighting methods in the uv -plane	135
Figure 45	Image-plane comparison of weighting schemes	135
Figure 46	Examples of modified-blackbody SED fits	140
Figure 47	Example uv fit for CO(6 – 5)	142
Figure 48	L_{IR} -normalised CO SLED	144
Figure 49	Non-normalised CO SLED	145
Figure 50	Stacked continuum SED	146
Figure 51	Apparent line source sizes	147
Figure 52	Apparent continuum source sizes	148

Figure 53	Posterior distributions for the single-component LVG fit 155
Figure 54	Lensed vs. unlensed posteriors for the single-component fit 156
Figure 55	CO SLED fits across sample splits 159
Figure 56	Posterior distributions for the two-component LVG fit 160
Figure 57	Lensed vs. unlensed posteriors for the two-component fit 161
Figure 58	Comparison of average CO SLED with literature values 166
Figure 59	Multi-facility interferometric coverage enabled by ViSta 171
Figure 60	Serendipitous line coverage for future stacking 173
Figure 61	Wall-clock scaling of the HPC pipeline 190

LIST OF TABLES

Table 1	Sources of the parent sample with public ALMA observations (updated with multiple components). 37
Table 2	Peak SNR Results for ViSTA-ImFit 90
Table 3	Properties of the sources and of their ALMA observations 99
Table 4	Best-fit Gaussian parameters of the stacked CO(3–2) line 100
Table 5	ALMA sample of <i>Herschel</i> -selected DS-FGs. 124
Table 6	MBB temperature (T) and apparent infrared luminosity ($\log(\mu L_{\text{IR}})$) for the sample. 137
Table 7	L_{IR} -normalised line SLED measurements 149
Table 8	L_{IR} -normalised continuum measurements 151
Table 9	LVG excitation parameters of the stacked CO ladder. 164
Table 10	Pipeline timing breakdown as a function of N_{MS} . 191
Table 11	Pipeline timing breakdown as a function of N_{chan} . 191
Table 12	Pipeline timing breakdown as a function of N_{readers} . 191

Table 13	Pipeline timing breakdown as a function of OMP_NUM_THREADS (CPU). 191
Table 14	Pipeline timing breakdown as a function of N_{batch} (GPU). 191
Table 15	100–2100 μm observed flux densities for the sources in the sample. 196
Table 16	ALMA archival observations with a possible CO emission line detection by source ID. 202

ACRONYMS

ACA	Atacama Compact Array
AGN	Active Galactic Nucleus
ALMA	Atacama Large Millimeter/submillimeter Array
ASDM	ALMA Science Data Model
CASA	Common Astronomy Software Applications
CMB	Cosmic Microwave Background
CO	carbon monoxide
CSFH	Cosmic Star Formation History
DSFG	Dusty Star-Forming Galaxy
FIR	far-infrared
FUV	far-ultraviolet
FWHM	Full Width at Half Maximum
GPU	Graphics Processing Unit
ISM	Interstellar Medium
LTE	Local Thermodynamic Equilibrium
LVG	Large Velocity Gradient
MBB	Modified Blackbody
MS	Measurement Set
PDR	Photodissociation Region
PSF	Point Spread Function
RMS	Root Mean Square
SED	Spectral Energy Distribution
SFRD	Star Formation Rate Density

SLED	Spectral Line Energy Distribution
SMG	Submillimetre Galaxy
SNR	Signal-to-Noise Ratio
ULIRG	Ultra-Luminous Infrared Galaxy
UV	ultraviolet
XDR	X-ray Dominated Region

INTRODUCTION

Galaxies are the fundamental structural units of the observable Universe, functioning as dynamic ecosystems whose growth and transformation continuously reshape the cosmic landscape. The primary driver of galaxy growth is star formation, in which cold and dense molecular gas collapses under gravity to form new stars. This process converts baryonic gas into stellar mass and enriches the Interstellar Medium (ISM) with heavy elements through stellar evolution. Mapped across cosmic time, this collective activity outlines the Cosmic Star Formation History (CSFH), which traces how the Universe assembled its stellar mass over time. Unveiling this history is a fundamental step toward understanding the origin and evolution of galaxies.

Tracing this history empirically presents major observational challenges. Although young massive stars emit strong ultraviolet (UV) radiation, interstellar dust absorbs a large fraction of this radiation and re-emits it as thermal emission at far-infrared (FIR) wavelengths. This obscuration peaks precisely during the epoch of maximum cosmic star formation. Consequently, an uncorrected UV census remains incomplete and introduces a systematic, epoch-dependent bias across cosmic time.

DSFGs, the central population studied in this thesis, represent the most extreme manifestation of obscured star formation. Identified through their intense FIR and (sub-)millimetre emission, they contribute a substantial fraction of the cosmic star-formation activity at its peak ($z \sim 2-3$), an epoch commonly known as *cosmic noon*, which makes them the primary candidate progenitors of massive quiescent galaxies. Understanding how they sustain and eventually quench such rapid growth is therefore essential to galaxy evolution models and their extreme interstellar environments serve as unique laboratories to test whether standard empirical scaling relations remain valid at physical limits.

Characterising these extreme environments requires probing the physical state of the ISM, rather than measuring total star-formation output alone. Key properties such as gas density, temperature, and excitation dictate how molecular clouds collapse, how long starburst activity lasts, and how feedback mechanisms inject energy back into the gas. Because these physical state parameters cannot be measured directly, they are usually inferred from the emission lines of the cold gas. Both the thermal continuum from dust and these line diagnostics

fall within the FIR and (sub-)millimetre window, which makes this spectral region indispensable for understanding this population.

Measuring cold gas line emission across high-redshift galaxies remains observationally expensive. Individual emission lines are faint and require substantial integration times, leaving direct detections restricted to a small sample of bright or gravitationally lensed sources. To obtain a population-wide view, astronomers rely on stacking to combine many targets into a single average. Although public archives of interferometric observations contain vast amounts of suitable data, these observations were taken by unrelated projects and are highly heterogeneous in redshift, spectral configuration and spatial resolution. Attempting to combine reconstructed images introduces severe artifacts, compromising the accuracy of the stacked measurement.

This thesis directly addresses these observational and methodological challenges through ViSTA, a new stacking framework that combines heterogeneous interferometric observations directly in the visibility domain, before image reconstruction. To demonstrate the capabilities of this novel approach, its first scientific application focuses on constraining the average excitation conditions of the molecular gas in a sample of DSFGs.

The remainder of this chapter establishes the scientific and methodological background for this research. Section 1.1 provides an overview of the DSFG population, their physical properties, and their evolutionary context. Section 1.2 then explores the ISM of these galaxies, focusing on the line diagnostics used to infer the excitation state of the molecular gas. The methodological framework is presented in Section 1.3, which outlines the rationale and challenges of stacking heterogeneous interferometric data. Finally, Section 1.4 defines the main goals and structure of the thesis.

1.1 DUSTY STAR-FORMING GALAXIES

1.1.1 *The obscured cosmic star formation history*

The evolution of star formation across cosmic time is quantified by the Star Formation Rate Density (SFRD), defined as the stellar mass formed per unit time and per unit comoving volume. Rather than tracing the history of an individual galaxy, this quantity measures the integrated activity of the entire galaxy population as a function of redshift, and its reconstruction provides one of the tightest benchmarks available for models of galaxy formation and cosmological simulations.

One of the most popular ways to trace star formation is by recovering the light of young massive stars through two complementary channels, the emitted rest-frame UV continuum [1] and the obscured FIR continuum absorbed and thermally re-emitted by dust [2, 3], both

necessary to recover the total star-formation activity. The balance between the two shifts across cosmic time, from comparable contributions in the local Universe to a growing dominance of the obscured term at higher redshifts.

The evolution of the SFRD as a function of redshift has been well determined in many studies [4–6]. In Fig. 1 an example is reported from the compilation of Madau and Dickinson [6]. It rises steeply from early epochs, reaches a broad peak between $z \sim 1$ and $z \sim 3$, and then declines by an order of magnitude toward the present day. This maximum occurred when the Universe was about 3.5 Gyr old, in an era known as *cosmic noon*. This period is critical for galaxy growth, because more than half of the stellar mass observed today was formed during this time, offering a prime laboratory to probe the mechanisms driving star formation.

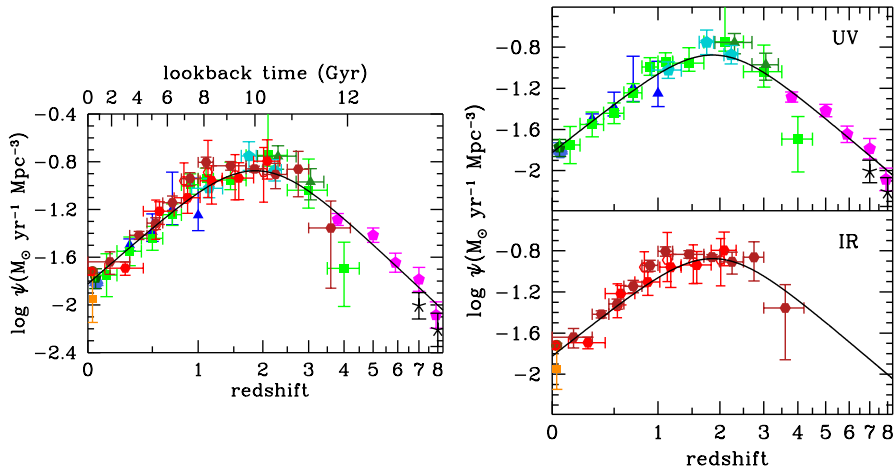


Figure 1: Cosmic SFRD as a function of redshift (and lookback time), from rest-frame UV + IR measurements (left) and from the UV (top right) and IR (bottom right) data separately. The different symbols and colours correspond to determinations from different surveys, blue-to-magenta for the UV-based ones and brown-to-red for the IR-based ones. The solid curve in all panels is the best-fit SFRD, peaking at $z \simeq 1.9$ and declining towards the present epoch. Adapted from [6].

The dominance of the obscured star formation component at the peak carries a direct methodological consequence. Since most of the star formation at cosmic noon is heavily dust-attenuated in the UV and accessible only through its reprocessed FIR emission [7, 8], a census built on the UV alone must rely on substantial extinction corrections at the epoch that matters most. Consequently, the CSFH became reliable at these redshifts only once wide infrared and (sub-)millimetre surveys measured the obscured component directly [6]. Beyond the peak, at $z \gtrsim 3$, the obscured budget rests on far fewer

observational constraints and remains poorly characterised, even though recent facilities have shown that it remains significant, driven in part by a population of optically dark galaxies [9–12].

The galaxy population responsible for this heavily obscured star formation is collectively referred to as Dusty Star-Forming Galaxies (DSFGs). Although rare, these systems host such high star formation rates that they contribute roughly 20% of the SFRD at $z \sim 2$ –3 and are thought to dominate the assembly of the most massive galaxies at those epochs [13, 14]. As the most extreme and heavily dust-enshrouded environments at cosmic noon, they not only dominate the dust-reprocessed emission of this era, but also reach star formation rates roughly an order of magnitude above those of typical star-forming galaxies at the present epoch.

1.1.2 *Discovery, selection and physical properties*

DSFGs are the most vigorous star-forming systems known but, since their activity is almost entirely reprocessed by dust, they were invisible to the optical surveys that first mapped the galaxy population [8, 15]. Their discovery followed directly from the opening of the FIR and sub-millimetre window. The grains that absorb the UV radiation of young massive stars are heated to a few tens of Kelvin and re-emit thermally, producing a broad continuum that peaks around 60–100 μm in the rest frame. At the wavelengths where this emission is observed, the same grains are optically thin, so the measured flux density is a nearly unbiased tracer of the reprocessed stellar radiation [8, 16]. The resulting Spectral Energy Distribution (SED), which describes how the emitted energy of a source is distributed as a function of wavelength or frequency, encodes the dust temperature, mass and luminosity of the system. An example is shown in Fig. 2.

Another remarkable property of this emission is the negative K -correction. For a source of fixed intrinsic luminosity, increasing the redshift shifts progressively brighter parts of the dust spectrum into a fixed observing band. Along the Rayleigh-Jeans tail (the low-frequency side of the dust peak where $S_\nu \propto \nu^{2+\beta}$) this effect nearly cancels the dimming produced by the growing luminosity distance. As a result, the observed flux density at sub-millimetre wavelengths remains almost constant over a wide redshift interval, from $z \sim 1$ to $z \sim 6$, as illustrated in Fig. 3. A sub-millimetre survey is therefore sensitive to dusty galaxies over a broad range of cosmic epochs with only a mild luminosity bias, a property that no other wavelength regime shares [8, 15].

Space-based missions initially opened this spectral window, starting in the infrared with *IRAS* and *Spitzer* and extending to the sub-millimetre with *COBE* and *Herschel*. These space observatories were complemented by ground-based facilities, such as the JCMT

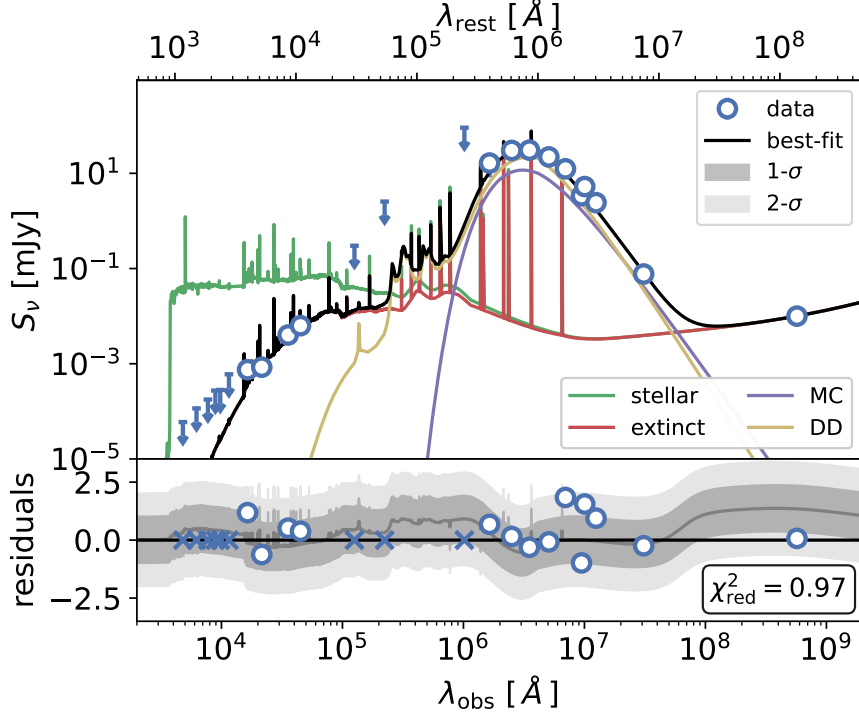


Figure 2: SED of a lensed DSFG at $z \sim 3.1$, modelled with GALAPY [17] from the photometry of Giulietti et al. [18]. Circles are photometric detections and downward arrows are upper limits, while the black line is the best-fitting model and the shaded areas its 1σ and 2σ confidence regions. The coloured lines separate the components that build up the emission, namely the intrinsic stellar continuum produced by the composite stellar population (green) and the fraction of it that survives absorption by dust (red), together with the thermal re-emission of that absorbed energy by the warm dust of the molecular clouds (purple) and by the colder diffuse dust distributed on galactic scales (yellow). The stellar light is strongly suppressed in the rest-frame UV and optical, and the energy it loses reappears as the FIR peak, where the two dust components blend into the broad continuum that characterises these galaxies.

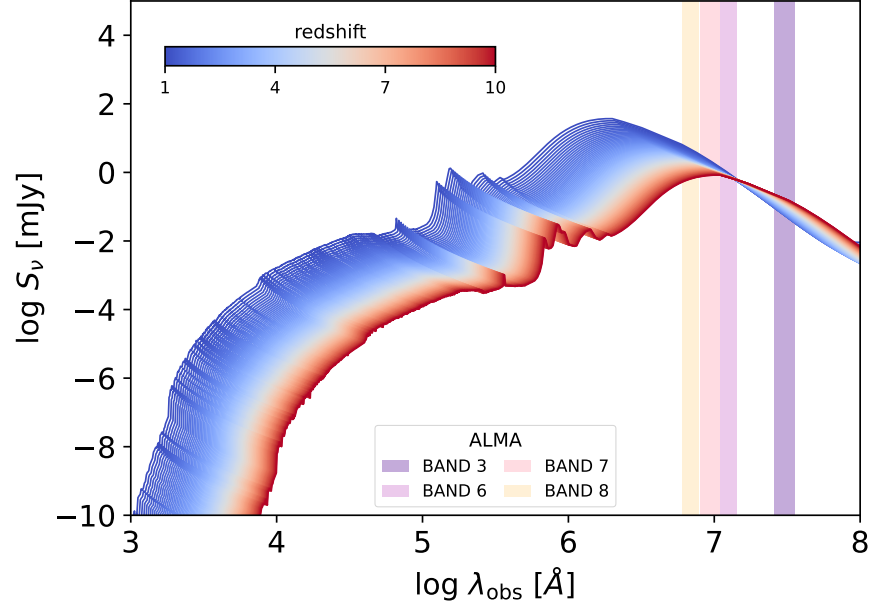


Figure 3: Simulated SED of a DSFG generated with GALAPY [17] and redshifted, at fixed intrinsic luminosity, from $z = 1$ to $z = 10$. Shaded vertical stripes mark ALMA bands 3, 6, 7 and 8. At rest-frame UV and optical wavelengths the observed flux density drops by several orders of magnitude as the source is moved to higher redshift, whereas in the (sub-)mm the curves converge. The steeply rising Rayleigh-Jeans tail of the dust emission is shifted into the observing band and almost exactly compensates the cosmological dimming.

equipped with SCUBA in the sub-millimetre, as well as APEX and the South Pole Telescope in the millimetre. Through these early efforts, *IRAS* first uncovered local luminous and ultraluminous infrared galaxies [19], while *COBE* showed that the cosmic infrared background carries as much energy as the optical background, revealing widespread dust-obscured star formation [20, 21]. Shortly after, pioneering blank-field surveys at $850\ \mu\text{m}$ with SCUBA resolved this background into individual high-redshift sources, historically known as Submillimetre Galaxies (SMGs) [22–24].

Building the large samples needed for statistical studies later relied on wide-area surveys such as H-ATLAS [25, 26], HerMES [27], and HerS [28] with *Herschel* [29], together with millimetre surveys from the South Pole Telescope [30] and all-sky selection with *Planck* [31]. Each survey named its sources after its own selection criteria, producing several overlapping terms that are not interchangeable—from ultraluminous and hyperluminous infrared galaxies defined by luminosity thresholds of 10^{12} and $10^{13}\ L_{\odot}$ to SMGs selected at $850\ \mu\text{m}$. To avoid tying the population to any single threshold or filter, I adopt throughout this thesis the broader term DSFG, referring to galaxies beyond the local Universe whose star formation is largely obscured and whose energy emerges predominantly in the FIR. This class is internally diverse, and the objects considered here occupy its most luminous end [8, 16].

The physical properties of these galaxies directly reflect the intense FIR emission that defines them, with infrared luminosities exceeding $10^{12}\ L_{\odot}$ and reaching beyond $10^{13}\ L_{\odot}$, corresponding to star formation rates of 10^2 to $10^3\ M_{\odot}\ \text{yr}^{-1}$. This energy output arises from dust masses of 10^8 to $10^9\ M_{\odot}$ at temperatures of a few tens of Kelvin, powered by massive molecular gas reservoirs of 10^{10} to $10^{11}\ M_{\odot}$ [32, 33]. These vast gas supplies are consumed rapidly, indicating a brief starburst phase with depletion timescales as short as a few hundred million years. Across cosmic time, these galaxies peak in space density around $z \sim 2$ to 3 and reside preferentially in overdense environments, with a tail extending to higher redshifts populated mainly by millimetre-selected sources [34–36].

Their characterisation was initially limited by the low angular resolution of single-dish telescopes, which made counterpart identification ambiguous. As discussed in detail in the following chapters, this obstacle was overcome through high-resolution interferometric follow-up that characterised these sources and measured their redshifts directly from molecular emission lines [37]. These capabilities turned ambiguous single-dish detections into well-characterised populations. Among these facilities are the Submillimeter Array (SMA), the NOthern Extended Millimeter Array (NOEMA), and Atacama Large Millimeter/submillimeter Array (ALMA), with the latter providing most of the data used in this project. High-resolution inter-

ferometric imaging, particularly with ALMA, has further shown that the activity is spatially concentrated, with the dust continuum typically confined to compact regions of a few kiloparsecs or less while the molecular gas and the stellar component are more extended, and with morphologies ranging from ordered rotating disks to clearly interacting systems [38–41]. This morphological and structural diversity implies that a fixed angular scale corresponds to different physical regions in different sources, and that the emission of different tracers does not necessarily arise from the same volume. Overcoming these observational limits to resolve such compact, multi-component structures and detect faint emission lines requires superior sensitivity and angular detail, which can be effectively achieved with the help of strong gravitational lensing.

1.1.3 Strong gravitational lensing as a natural telescope

Strong gravitational lensing occurs when light emitted by a distant background source is deflected by the gravitational potential of a massive foreground object, such as an elliptical galaxy or a galaxy cluster, stretching the image into arcs, multiple images or a complete Einstein ring, as depicted in Fig. 4. The mapping between the source plane and the observer plane is governed by the so-called *caustics*, which are curves in the source plane that mark where the lensing distortion is the strongest. Since the magnification formally diverges along the caustic, a source far from a caustic is only mildly magnified, while one that lies close to or crosses it is stretched into long and highly amplified arcs. In reality, this divergence is never reached, since a real source has a finite size and the magnification is averaged over it, but it explains why the strongest amplifications and the most spectacular arcs and rings are seen for sources that sit close to a caustic.

This physical phenomenon produces two main observational effects, both quantified by the *magnification factor* μ . First, spatial stretching expands the apparent angular size of the source, magnifying linear scales by a factor proportional to $\sqrt{\mu}$ and improving the effective spatial resolution down to sub-kiloparsec scales. Second, since lensing conserves surface brightness while enlarging the apparent solid angle, the total flux is amplified by μ , typically ranging from a factor of a few up to several tens. This flux boost is the critical advantage for the study of DSFGs, since it raises the Signal-to-Noise Ratio (SNR) and brings faint molecular and atomic lines within reach of feasible integration times. A further advantage in the (sub-)millimetre is that the foreground galaxies acting as lenses are predominantly passive ellipticals with negligible cool dust, so they are virtually dark at these wavelengths and their emission does not contaminate the observed flux of the background source.

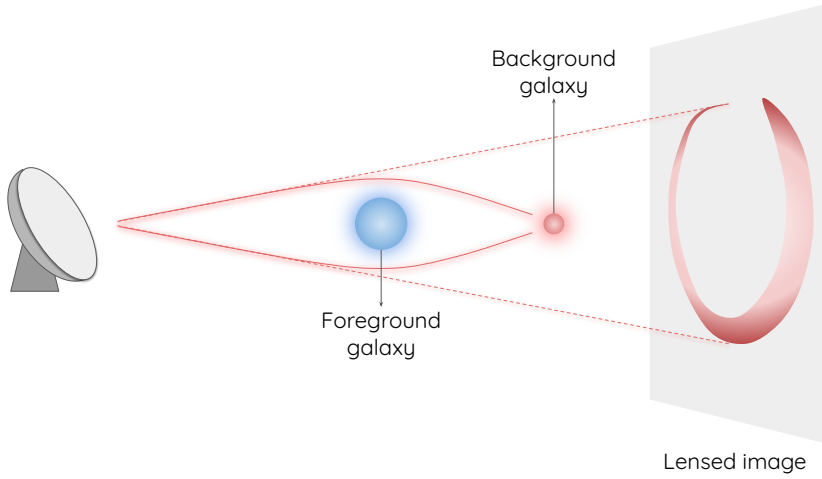


Figure 4: Schematic representation of strong gravitational lensing, where light emitted by a distant background galaxy is deflected by the gravitational potential of a massive foreground lens, producing a distorted and magnified image in the observer plane.

Before using lensed systems as reliable proxies for the unlensed population, it is essential to verify that lensed and unlensed sources trace the same underlying physical population. Ideally, the two subsets should differ only by an overall magnification factor, without intrinsic physical or spectral biases being introduced by the lensing itself. In particular, one must check whether magnification acts uniformly across the source or introduces inhomogeneous effects that could reshape the emission and bias derived quantities. A primary concern of this nature is spatially *differential magnification*, which occurs when a magnification gradient across the source plane amplifies different regions of the same galaxy by unequal factors. For example, compact and high-density structures, such as localised star-forming cores or a central Active Galactic Nucleus (AGN), occupy small regions that can lie close to high-magnification caustics. In contrast, extended and diffuse gas and dust are spread over much larger areas and can be amplified more modestly. As a result, line fluxes from distinct spatial regions undergo uneven magnifications, distorting observed line ratios and integrated profiles based on the source geometry and caustic alignment. Quantifying this effect is therefore essential to prevent systematic biases when comparing lensed sources with the unlensed populations.

1.1.4 The role of DSFGs in galaxy evolution

DSFGs play a central role in modern galaxy formation paradigms, serving as an empirical link between the early collapse of gas and the assembly of the most massive stellar structures in the Universe. The physical properties reviewed in the previous sections are consistent with this evolutionary picture, with star formation rates of 10^2 to $10^3 \text{ M}_\odot \text{ yr}^{-1}$ sustained over depletion timescales of about 100 Myr within radii of one or two kiloparsecs. Together, these properties describe a rapid and self-limiting phase that builds a compact stellar core, matching the conditions required to form the dense centres of local spheroids. Clustering measurements add that these galaxies live in dark matter halos of 10^{12} to $10^{13} \text{ M}_\odot$, comparable to those of quasars at $z \sim 2$ [35, 42], and their number density peaks at $z \sim 2.5$ [35, 43], essentially where the CSFH and the black hole accretion history reach their own maxima [6, 44]. These galaxies host both intense star formation and, in many cases, an active AGN. A common cold gas reservoir can feed both processes, but their triggering is not driven by the gas supply alone, since physical mechanisms such as major mergers and disk instabilities can channel gas towards the centre and set off both the starburst and the AGN activity [45, 46]. This population therefore lies at the intersection of star formation and nuclear activity, and it provides a natural laboratory for the stellar and black hole co-evolution that models must reconcile with local scaling relations.

The discovery of this high-redshift population has brought major theoretical challenges. Reproducing the observed abundance and rapid assembly of the brightest DSFGs at early epochs remains difficult for cosmological models, which often require ad hoc adjustments to the star formation efficiency or the initial mass function [47–51]. Observationally, massive galaxies follow an evolutionary trend known as *archaeological downsizing*, forming their stars earlier and over shorter timescales than lower-mass systems [52, 53]. Early- and late-type galaxies follow markedly different tracks [54], as the former assemble their stars in an early episode lasting no more than about 1 Gyr before feedback dissipates the gas, whereas the latter form stars almost continuously over several billion years. Because their star formation is concentrated in an intense initial burst, the most massive star-forming galaxies, DSFGs among them, are observed almost exclusively at $z \gtrsim 2$. Halo growth models and downsizing considerations thus converge, identifying DSFGs as the candidate progenitors of the massive quiescent galaxies seen at $z \sim 1.5$ to 2 and of local giant ellipticals [8, 14, 42, 55–59], as illustrated in Fig. 5.

While downsizing explains when massive galaxies form their stars, it does not specify what triggers this activity or what ultimately quenches it. Evolutionary frameworks diverge in how galaxies

populate the $\text{SFR}-M_*$ plane, where most star-forming galaxies at any epoch trace a tight correlation known as the *main sequence*, while starbursts such as DSFGs lie well above it [60–62]. In merger-driven scenarios, major mergers boost the cold gas supply and drive starbursts that are eventually quenched by AGN feedback [45, 46, 63]. A second framework describes a smoother process fed by cold gas streams along cosmic web filaments [64, 65]. Finally, in-situ scenarios instead attribute the burst to the early collapse and compaction of halo gas, driving rapid core assembly alongside nuclear accretion until combined feedback quenches the galaxy [58, 59]. Although these scenarios may describe complementary stages of galaxy growth [66], major mergers account for only about ten per cent of the SFRD at cosmic noon [67], leaving the primary trigger uncertain.

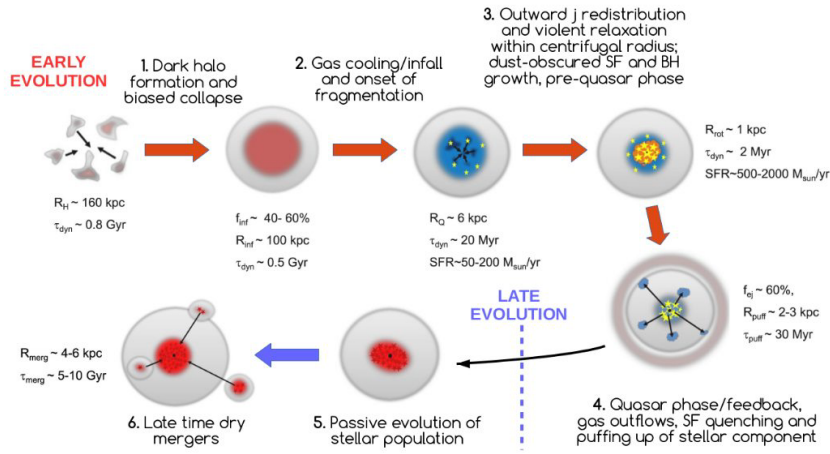


Figure 5: An example evolutionary pathway connecting high-redshift DSFGs to the massive quiescent galaxies of the local Universe. Starting from the collapse of gas within the host halo, the galaxy undergoes a heavily dust-obscured starburst phase during which most of the stellar mass is assembled alongside central black hole growth, followed by feedback-driven quenching and passive evolution into an early-type galaxy. The spatial scales and timescales follow the in situ scenario of Lapi et al. [58] and Pantoni et al. [59], illustrating the overall evolutionary connection independently of the specific triggering mechanism. Adapted from Pantoni et al. [59].

DSFGs offer a valuable testbed to help discriminate among these evolutionary scenarios, since the three frameworks predict different gas distributions, triggering mechanisms, and quenching timescales for a given star formation rate. Integrated observations alone, however, cannot break the degeneracy between an elevated star formation efficiency and a large, replenished gas reservoir, because all three can

produce comparable total gas masses and differ instead in how that gas is structured and converted into stars, whether through heightened efficiency from merger-induced compression [49, 65, 68, 69], an exceptionally massive reservoir converted at standard efficiencies and replenished by cold streams [64, 70], or a combination of both [71, 72].

These theoretical pictures are distinguished instead by the physical state of the cold molecular ISM, and in particular by the volume densities, kinetic temperatures, thermal pressures, and radiation field intensities that govern cloud collapse [73, 74]. Characterising these local properties determines whether the molecular gas in high-redshift DSFGs differs fundamentally from that of local luminous and ultraluminous infrared galaxies, testing the universality of star-formation laws across galaxy populations. Achieving this requires resolving both the diffuse and dense gas phases through the line emission of their primary molecular tracers.

Obtaining these detailed ISM measurements across the galaxy population, however, is hampered by two major observational challenges related to sample representativeness. First, theoretical models predict ensemble properties, whereas high-resolution line observations typically target individual sources. A single DSFG provides only a single orientation and a brief snapshot of a phase lasting a few hundred million years within a population that displays broad diversity in dust temperature, morphology, and compactness (Sec. 1.1.2). Physical conditions inferred from an individual object cannot be generalised to the entire population, meaning excitation diagnostics must be averaged over homogeneously selected samples observed in identical transitions, a task that requires observationally expensive and scarce datasets. Second, as introduced in Section 1.1.3, most DSFGs with detailed line detections are gravitationally lensed. This introduces a strong selection bias toward the most compact, high-surface-brightness sources, while differential magnification across the source plane can distort the intrinsic line ratios used to constrain physical conditions [75, 76].

The remainder of this chapter addresses these observational limitations. Section 1.2 reviews the physics of the cold ISM and the line diagnostics used to probe its excitation, while Section 1.3 details the statistical framework used to combine heterogeneous archival interferometric data and recover the average excitation properties of the DSFG population.

1.2 THE INTERSTELLAR MEDIUM AND ITS TRACERS

Composed of gas, dust grains, magnetic fields, and cosmic rays filling the space between stars, the ISM serves as the primary bary-

onic reservoir within galaxies, mediating the cycle of matter and continuously exchanging material with the surrounding intergalactic medium through cosmological accretion streams and energetic stellar outflows. It is characterised by a complex multiphase structure maintained in rough thermal and pressure equilibrium, spanning several orders of magnitude in temperature, volume density, and ionisation state. The relative proportions of gas and dust, and their total contribution to the baryonic budget of the galaxy, vary substantially depending on metallicity, galaxy morphology, and redshift. Despite their lower mass contribution, irregular silicate and aromatic hydrocarbon dust particles are fundamental components of the ISM, absorbing optical and UV radiation, growing via icy mantles in cold clouds, and re-emitting energy thermally in the infrared [77, 78].

The gaseous component of the ISM is dominated by hydrogen and is distributed across distinct physical regimes defined by temperature, density, and ionisation state [78, 79]. Rather than representing sharp thermodynamic boundaries, these multiphase regimes coexist dynamically in a global equilibrium governed by heating and cooling processes, pressure balance, radiation fields, turbulence, and magnetic forces. Supernova blast waves carve out a tenuous, hot ionised medium ($T \sim 10^6$ K, $n < 10^{-3} \text{ cm}^{-3}$), while massive stars and ambient radiation sustain warm ionised and warm neutral phases ($T \sim 10^4$ K, $n \sim 0.1\text{--}1 \text{ cm}^{-3}$). Where gas cools and condenses ($T \sim 100$ K, $n \sim 10 \text{ cm}^{-3}$), it forms the *cold neutral medium* (CNM), which accounts for roughly half of the local interstellar gas mass and is most reliably traced through the 21 cm hyperfine transition. At even higher densities ($n > 10^2 \text{ cm}^{-3}$) and lower temperatures ($T \sim 10$ K), hydrogen atoms coalesce into molecular hydrogen (H_2) via catalytic surface reactions on dust grains [66, 80], forming the cold, dense molecular clouds where star formation occurs. Newly formed massive stars then photoionise the surrounding gas with intense UV radiation to form emission nebulae known as *HII regions*.

H_2 represents the primary fuel for star formation, establishing a direct physical link between molecular reservoirs and galactic growth. This empirical connection is formalised by the *Kennicutt-Schmidt relation* between the surface density of star formation rate and gas mass surface density,

$$\Sigma_{\text{SFR}} \propto \Sigma_{\text{gas}}^n, \quad (1)$$

originally proposed with $n \sim 1\text{--}3$ [81] and later recalibrated to $n \sim 1.4$ on local spirals and infrared-luminous galaxies [82]. Across cosmic time, the space density of molecular gas tightly tracks the overall SFRD, rising toward $z \sim 1.5\text{--}2$ before declining by nearly a factor of six toward the present day [83]. This evolution implies that starburst activity at cosmic noon is sustained by vast molecular reservoirs replenished by cosmic web accretion streams [64, 70, 84–87].

In local environments, the transition to predominantly molecular gas is not determined by a universal density or temperature threshold, but depends on total column density, metallicity, dust attenuation, and H_2 self-shielding, typically resulting in cold, dense molecular clouds ($T \lesssim 10\text{--}20\text{ K}$, $n \gtrsim 10^2\text{ cm}^{-3}$). In high-redshift DSFGs, instead, exceptionally high gas surface densities and strong gravitational instabilities allow this cold molecular phase to dominate the total baryonic gas mass budget, with self-gravity easily overcoming thermal, magnetic, and turbulent support.

Interpreting this environment requires accounting for several physical conditions that set high-redshift DSFGs apart from local galaxies. First, the Cosmic Microwave Background (CMB) temperature scales as $T_{\text{CMB}}(z) = T_{\text{CMB}}^0(1+z)$, reaching approximately 13.6 K at $z \sim 4$. At high redshift, the increasing CMB energy density establishes a lower limit for the dust temperature, modifies the excitation of molecular species via enhanced radiative pumping, and reduces the observed contrast of both spectral line and dust continuum emission against the background [88, 89]. Second, intense starburst activity drives a far-UV radiation field that can exceed local interstellar values by three to five orders of magnitude [90, 91]. This field heats Photodissociation Regions (PDRs) [92], while heavy dust attenuation simultaneously shields the inner cloud cores, preserving vast molecular gas reservoirs. Third, high supernova rates produce elevated cosmic-ray fluxes that penetrate deep into these cores. These cosmic rays regulate ion-neutral chemistry and heating throughout molecular clouds, increasing the abundance of neutral or ionised carbon [93–95]. Finally, dust and metals regulate all of these processes together. The high dust-to-stellar mass ratios of DSFGs point to large gas reservoirs, and dust grains provide the catalytic sites for H_2 formation and mantle chemistry while driving UV attenuation and the overall atomic-to-molecular balance [96]. Despite their high redshift, these galaxies are already chemically enriched to near-solar values [97, 98], and the interplay between metallicity and the dust-to-gas ratio sets the dust opacity that governs the atomic-to-molecular transition.

These physical properties make DSFGs ideal targets for study in the FIR and (sub-)millimetre window, where both the thermal dust continuum and the cold gas emission lines fall. A natural starting point is the cold dense molecular phase, which directly fuels star formation. Although H_2 is its most abundant constituent, it remains essentially invisible under these conditions. Lacking a permanent electric dipole moment, the homonuclear H_2 molecule only emits via weak quadrupolar rotational transitions, whose lowest excitation levels require temperatures of several hundred kelvin. This is far above the 10 to 20 K typical of cold star-forming gas, so most of the cold gas mass produces no detectable H_2 emission, and its quantitative determination must rely on indirect methods [74, 99]. The main tracers avail-

able in this spectral window begin with the principal indirect probes of cold molecular gas, specifically carbon monoxide and atomic carbon, and extend to complementary molecular and atomic diagnostics that probe denser and more excited phases, although their interpretation is complicated by the blending of regions with distinct physical conditions within these unresolved high-redshift sources.

1.2.1 Cold gas tracers: Carbon Monoxide and Atomic Carbon

Carbon monoxide (^{12}CO , hereafter CO) is the second most abundant molecule in the ISM, and it has traditionally been used to trace the molecular H_2 content of a galaxy, since its weak permanent dipole moment enables allowed rotational transitions [73, 74]. The ground transition $\text{CO}(1-0)$ has a rest frequency of 115.27 GHz, an upper-level energy equivalent to 5.5 K, and a formal critical density of $\approx 10^3 \text{ cm}^{-3}$, which is further reduced to $\sim 10^2 \text{ cm}^{-3}$ by radiative trapping in optically thick gas [100]. Being the least sensitive to excitation variations, it is uniquely suited to reveal extended, subthermally excited gas reservoirs, whereas higher rotational lines require higher temperatures and densities, thereby isolating the warmer regions associated with ongoing starbursts [101].

Observed line fluxes are expressed as velocity-integrated brightness temperature luminosities through the standard relation

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

where $S_\nu \Delta v$ is the integrated line flux, D_L is the luminosity distance, and ν_{obs} is the observed frequency [100]. For optically thick emission in Local Thermodynamic Equilibrium (LTE), and assuming a common emitting area and beam filling factor, the ratio of two such luminosities equals the ratio of their corresponding brightness temperatures, making it the natural quantity for constructing excitation ladders [100]. Converting a line luminosity into total molecular gas mass relies on the conversion factor α_{CO} , defined by $M_{\text{H}_2} = \alpha_{\text{CO}} L'_{\text{CO}(1-0)}$.

This conversion factor is well calibrated in the nearby Universe, with local Galactic disk clouds typically exhibiting $\alpha_{\text{CO}} \sim 4.3 M_\odot (\text{K km s}^{-1} \text{ pc}^2)^{-1}$ (inclusive of heavy elements, primarily helium) [73, 102, 103]. In high-redshift DSFGs, the appropriate choice of conversion factor remains actively debated, as dynamical constraints reveal a broad continuum of values spanning from the Galactic-disk baseline down to the $\alpha_{\text{CO}} \simeq 0.8 M_\odot (\text{K km s}^{-1} \text{ pc}^2)^{-1}$ characteristic of local ultraluminous infrared galaxies [104], rather than a strict bimodality [73, 74, 105].

Major limitations affect CO as a quantitative tracer of molecular mass, driven by several physical processes that induce spatial and excitation variations. Strong UV radiation fields can dissociate CO

in molecular cloud envelopes while H_2 remains self-shielded, producing a CO-dark molecular component whose mass fraction depends strongly on metallicity, shielding, and radiation field strength [106, 107]. At low metallicities, reduced dust shielding contracts the CO emitting core, and this sharply increases the conversion factor α_{CO} [73]. In extreme star-forming environments, cosmic ray destruction further suppresses CO abundance [94], while heightened local gas turbulence ($\langle \Delta V_{\text{turb}} \rangle \sim 100 \text{ km/s}$ excluding galactic rotation and large-scale bulk motion) alters line opacities and broadens line profiles [105].

To bypass the optical depth saturation of CO and the uncertainties inherent to excitation-derived mass estimates, atomic carbon ([CI]) provides an alternative and independent mass proxy through its fine structure transitions [CI](1 – 0) and [CI](2 – 1), whose upper levels lie at 23.6 K and 62.5 K above the ground state, respectively. At $z > 1$ it becomes particularly advantageous, since its ground transition is redshifted into millimetre wavelengths and is therefore far easier to detect than faint CO(1 – 0) emission. Once thought to exist only in thin transition layers of PDRs, its widespread detection spatially coexistent with CO across Galactic and high-redshift starbursts established [CI] as an extended tracer of bulk molecular gas [105, 108, 109]. Under typical gas conditions, the [CI](1 – 0) transition is often assumed to be optically thin, and it originates from a volume similar to CO(1 – 0) with a low excitation temperature ($\sim 24 \text{ K}$), although its actual optical depth depends on column density and local physical conditions [105, 108]. In environments with elevated cosmic ray fluxes where CO is efficiently dissociated, carbon is redistributed between neutral carbon ([CI]) and ionised carbon ([CII]) according to local chemical balance, which allows [CI] to trace molecular gas provided that the ionisation fraction remains moderate [94, 105]. Its abundance X_{CI} is often less sensitive to metallicity than α_{CO} [110, 111], but it is not a universal constant and varies significantly with local environmental conditions.

Observing both fine structure lines provides an independent constraint on the excitation state of atomic carbon, as the luminosity ratio $R_{\text{CI}} = L'_{[\text{CI}](2-1)} / L'_{[\text{CI}](1-0)}$ yields the excitation temperature under the assumptions of LTE, a single excitation temperature, and optically thin emission via

$$T_{\text{ex}} = \frac{38.8}{\ln(2.11/R_{\text{CI}})} \quad [\text{K}], \quad (3)$$

[112, 113]. Once T_{ex} is determined and fixes the excitation factor Q_{10} , the molecular gas mass can be estimated from the [CI](1 – 0) line luminosity as

$$M_{\text{H}_2}^{[\text{CI}]} = 1375.8 \frac{D_L^2 S_\nu \Delta \nu}{(1+z) A_{10} Q_{10} X_{\text{CI}}} \quad [\text{M}_\odot], \quad (4)$$

[114], where $A_{10} = 0.793 \times 10^{-7} \text{ s}^{-1}$ is the Einstein coefficient of the transition. Here the neutral carbon abundance ratio $X_{\text{CI}} = [\text{CI}]/\text{H}_2$ enters as the primary free parameter, and it represents the main source of systematic uncertainty in the mass conversion. Comparing $[\text{CI}]$ and $\text{CO}(1-0)$ luminosities thus provides an independent way to calibrate α_{CO} across high-redshift populations, where joint analyses reveal that DSFGs tend towards ULIRG-like values [115, 116].

1.2.2 High- J CO transitions and the Spectral Line Energy Distribution

As anticipated in Section 1.2.1, $\text{CO}(1-0)$ provides the most direct measure of the total cold molecular gas reservoir but is intrinsically faint and shifted to low observing frequencies at high redshift, so that brighter transitions higher up the rotational ladder are frequently adopted as proxies. Each transition is labelled by the rotational quantum number of the two levels involved, so that $\text{CO}(J_{\text{up}} - J_{\text{low}})$ denotes the decay from the upper level J_{up} to $J_{\text{low}} = J_{\text{up}} - 1$, with the upper-level energy scaling as $E_J \propto J_{\text{up}}(J_{\text{up}} + 1)$. Because these energies and the corresponding critical densities grow rapidly with J_{up} , mid- and high- J lines trace progressively warmer and denser phases of the ISM rather than the diffuse bulk gas [74, 100]. The excitation state of the molecular gas, however, varies widely across distant star-forming galaxies. Dense starbursts can remain close to thermalisation up to $J_{\text{up}} \sim 6$, while more typical galactic environments become sub-thermal as early as $J_{\text{up}} = 3$ [74, 101, 117]. Line flux ratios are therefore far from universal, which prevents the use of a single conversion factor to extrapolate high- J observations back to $\text{CO}(1-0)$ [32, 118]. Reconstructing the complete rotational excitation spectrum is thus essential to establish the physical state and mass of molecular gas across cosmic epochs [117, 118].

The distribution of line fluxes or luminosities across this ladder, plotted against J_{up} , constitutes the CO Spectral Line Energy Distribution (SLED). Its shape is determined by the balance between collisional excitation and radiative processes [74, 101, 119]. Collisional excitation depends primarily on the gas volume density n_{H_2} and kinetic temperature T_{kin} . As illustrated in Fig. 6, increasing either quantity systematically elevates high- J line emission relative to $\text{CO}(1-0)$ [74]. At a fixed kinetic temperature, higher density pushes the excitation peak from $J_{\text{up}} \sim 3$ up to $J_{\text{up}} \sim 7$, while at a fixed volume density, elevating the kinetic temperature boosts overall excitation and shifts the peak toward higher J_{up} . Radiative processes, in turn, depend on line optical depths τ_J , CMB heating, and ambient radiation fields. High star formation rate surface densities Σ_{SFR} and intense interstellar radiation fields U elevate both gas temperatures and the dense gas fraction, systematically shifting the excitation peak toward higher rotational transitions [119, 120].

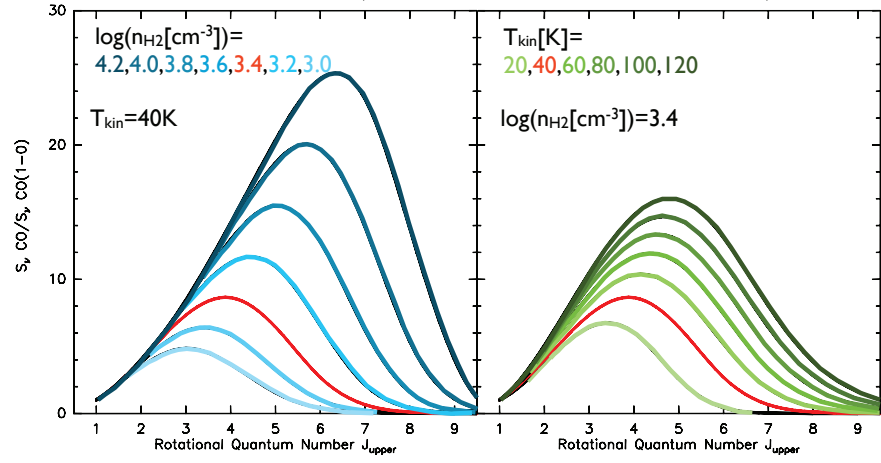


Figure 6: Evolution of the CO SLEDs as a function of physical gas properties. The left panel shows the impact of varying H_2 gas volume density $\log(n_{H_2} [\text{cm}^{-3}])$ from 3.0 to 4.2 at a fixed kinetic temperature of $T_{kin} = 40 \text{ K}$. The right panel displays the effect of varying kinetic temperature T_{kin} from 20 K to 120 K at a fixed volume density of $\log(n_{H_2} [\text{cm}^{-3}]) = 3.4$. Increasing either volume density or kinetic temperature enhances collisional excitation, shifting the excitation peak toward higher rotational quantum numbers and boosting flux ratios normalised to $\text{CO}(1-0)$. Adapted from Carilli and Walter [74].

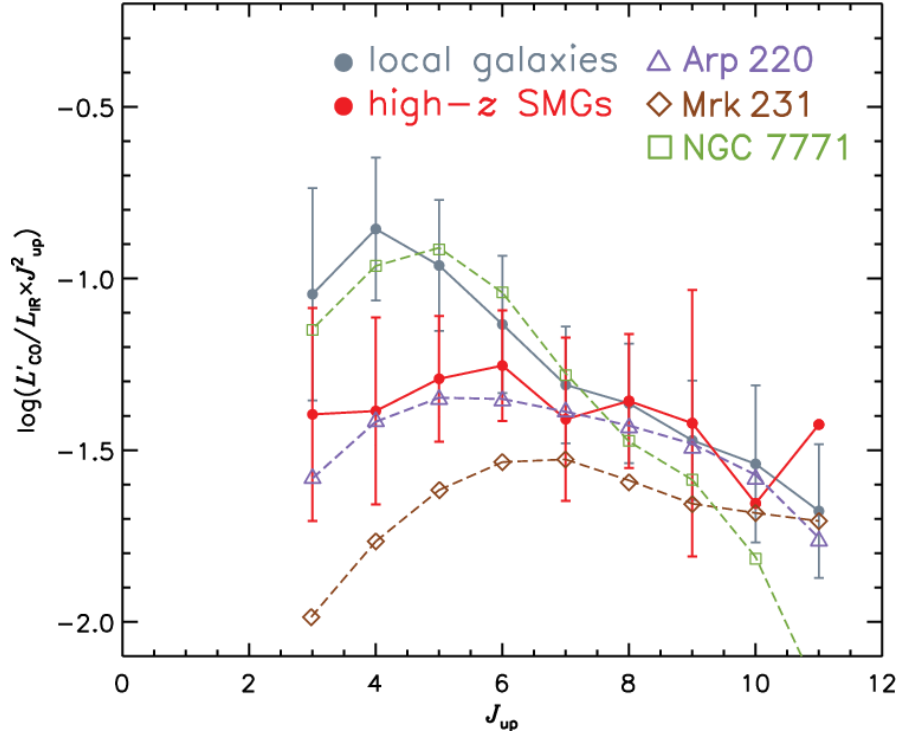


Figure 7: CO SLEDs of local galaxies (grey) and high-redshift SMGs (red), together with the local ULIRG Arp 220, the AGN-dominated source Mrk 231, and the LIRG NGC 7771, each normalised to the infrared luminosity and scaled by J_{up}^2 . For local star-forming galaxies, the normalised ladder declines toward higher rotational numbers, whereas the high-redshift SMGs and Arp 220 remain flat out to high J_{up} , indicating a larger contribution of dense, warm and highly excited gas. Adapted from Yang et al. [121].

This excitation diagnostic is clearly demonstrated by comparing local galaxy samples with distant submillimetre populations, as shown in Fig. 7 [121]. In local normal disk galaxies, the CO SLED declines rapidly beyond $J_{\text{up}} \sim 3\text{--}4$, reflecting subthermal excitation in gas where volume densities fall below the critical density of high- J levels ($n_{\text{crit}} \gtrsim 10^4\text{--}10^6 \text{ cm}^{-3}$). Conversely, high-redshift DSFGs, extreme starbursts, and local Ultra-Luminous Infrared Galaxies (ULIRGs) display broad, flat, or high-peaking excitation ladders ($J_{\text{up}} \sim 5\text{--}7$ or higher). This profile demonstrates that a large fraction of their molecular reservoir resides in dense, warm environments driven by intense far-ultraviolet (FUV) heating in PDRs, mechanical dissipation from turbulence, or X-ray heating from AGNs (X-ray Dominated Regions (XDRs)) [101, 121, 122].

Inferring physical ISM conditions from an observed CO SLED requires non-LTE radiative transfer modelling. In optically thick interstellar environments, this is routinely performed using Large Velocity Gradient (LVG) escape probability models [123–125]. These tools simplify the interaction between matter and radiation by treating line photon escape as a local process rather than solving the full three-dimensional radiation field. Under the LVG approximation, large-scale velocity gradients are assumed to dominate over local thermal broadening, so that photons shift out of resonance as they travel and escape locally with a probability determined by the optical depth. By simultaneously solving statistical equilibrium and radiative transport, these algorithms constrain the H_2 volume density, the kinetic gas temperature, the CO column density per velocity gradient and the effective beam filling factor [124, 126].

When applied to emission integrated over an entire galaxy, single-component models represent a simplified approximation. Analyses across multiple transitions consistently reveal that the ISM in high-redshift starbursts and DSFGs exists in several physical phases, so that the integrated CO SLED is usually decomposed into a cool and extended component, which dominates the low- J emission and contains most of the cold gas mass, and a warmer and denser one associated with compact nuclear starburst regions, which drives the emission at higher J_{up} [101, 117, 118, 121, 127]. Systems hosting energetic AGNs may require a further, more excited component heated by X-rays [128, 129], while continuous models incorporating supersonic turbulence distributions can reproduce the excitation spectrum without enforcing discrete phases [105].

Despite these modelling frameworks, systematic population analyses remain limited by selection effects, small sample sizes, and non-uniform normalisation choices across literature studies [32, 33, 36, 130, 131]. Homogeneous observations across multiple transitions for representative galaxy samples remain imperative to accurately char-

acterise CO ladder evolution and calibrate high- J line ratios as total molecular mass estimators across cosmic time.

1.2.3 Other gas tracers

Beyond CO and CI, several alternative gas and dust tracers provide complementary insights into the physical conditions, dense star-forming fuel, and radiative processes governing interstellar reservoirs. Optically thin dust continuum emission on the Rayleigh-Jeans tail offers a third independent estimate of gas mass [132–134]. Described by a Modified Blackbody (MBB) relation,

$$S_\nu \propto \frac{M_{\text{dust}} \kappa_\nu B_\nu(T_{\text{dust}})}{D_L^2}, \quad \kappa_\nu = \kappa_0 \left(\frac{\nu}{\nu_0} \right)^\beta, \quad (5)$$

where B_ν is the Planck function and β is the emissivity index taking values between 1.5 and 2 [77, 135], the continuum yields a dust mass that converts to total gas mass via an assumed dust-to-gas ratio.

Water is among the most abundant oxygen-bearing molecules and can become highly abundant in warm molecular gas. Its gas-phase abundance varies by orders of magnitude across interstellar phases, dropping below 10^{-8} in cold clouds where water freezes onto icy dust grains and rising above 10^{-4} in warm environments where ices evaporate and gas-phase reactions activate [136]. The molecule possesses a large dipole moment of 1.84 D, which couples it strongly to the radiation field and pushes its critical densities up to 10^8 – 10^9 cm^{-3} . Since such densities are rarely reached, excitation above the ground levels is driven by the absorption of FIR photons emitted by warm dust rather than by collisions, and the molecule then cascades down to emit intense (sub-)millimetre lines [136]. Because this emission is ultimately powered by FIR radiation, water line luminosity correlates strongly with infrared luminosity across four orders of magnitude, making water a direct tracer of obscured star formation, and of galactic outflows when detected in absorption [137–140]. Radiative pumping therefore allows water to trace warm and dense gas ($n_{\text{H}} \sim 10^5$ – 10^6 cm^{-3} , $T \sim 100$ – 200 K) associated with starbursts, shocks and AGNs, well below its formal critical densities, although collisions still govern the lowest levels and coupled radiative transfer modelling of gas and dust is required to interpret line ratios [136, 141].

Dense gas molecular species with high dipole moments and ground transition critical densities around 10^5 cm^{-3} , such as hydrogen cyanide (HCN), hydrogen isocyanide (HNC), the formyl cation (HCO^+), and carbon monosulfide (CS), probe a different physical regime [142]. While low- J CO traces the diffuse bulk gas, high critical density molecules isolate the dense star-forming cores [74]. Assuming that the $L'_{\text{HCN}(1-0)}/L'_{\text{CO}(1-0)}$ line luminosity ratio can be used as a proxy for the dense molecular gas fraction, this

ratio is systematically higher in local luminous star-forming galaxies than in normal spirals [142]. However, high-redshift observations of $\text{HCN}(1-0)$ remain limited and suggest a sublinear relation between dense gas luminosity and FIR luminosity, while local studies indicate that $\text{HCN}(1-0)$ emission can also originate from gas of moderate density [143, 144].

These high-density tracers also serve as diagnostics for the primary excitation mechanism. FUV photons from starbursts create PDRs, whereas hard X-rays from active nuclei penetrate deeper to form XDRs. Although an elevated $\text{HCN}(1-0)/\text{HCO}^+(1-0)$ ratio was long considered a signature of nuclear activity, the ratio is affected by opacity, temperature, and starburst contamination, rendering its physical interpretation degenerate [145]. Mid and high rotational lines such as $\text{HCN}(4-3)$, $\text{HCO}^+(4-3)$, and $\text{CS}(7-6)$ provide cleaner discriminants of the neutral region, though the physical origin of high- J HCN enhancement remains actively debated [146]. Additionally, faint isotopologues such as ^{13}CO and C^{18}O constrain the optical depth of the primary species, breaking degeneracies between column density and excitation [147, 148].

Because these dense gas tracers are intrinsically faint, individual detections at high redshift remain restricted to a few extreme, strongly lensed, or highly peculiar sources that are not representative of the broader galaxy population. Consequently, systematic constraints across main-sequence systems rely primarily on spectral stacking techniques to recover faint emission from species such as HCN, HCO^+ , CH, and CCH [36, 90, 130]. Just as with the CO rotational ladders, systematic population analyses of these alternative tracers remain severely limited by selection effects, small sample sizes, and non-uniform observational strategies across literature studies. Overcoming these observational barriers requires uniform multi-transition surveys targeting unbiased galaxy populations, which is essential to build a self-consistent picture of molecular gas physics across cosmic epochs.

1.3 THE NEED FOR A STATISTICAL APPROACH

The previous section has shown that characterising the ISM of DSFGs requires measuring several transitions of the same species, and that the transitions carrying the most information are also the faintest. This combination defines the observational problem that this thesis addresses. Sampling a full excitation ladder in a single galaxy demands long integration times on each target and each line, so that well-sampled ladders exist only for a handful of exceptionally bright or strongly lensed objects, and rarer tracers have been detected in only a few sources. Consequently, the average behaviour of the population is usually reconstructed by combining measurements drawn

from different studies, each based on its own selection criteria, spectral setup, and normalisation scheme. Averaging quantities obtained under such different assumptions yields a mean, or a median, that does not correspond to any well-defined population, and the resulting scatter mixes genuine physical diversity with the inhomogeneity of the measurements themselves.

This limitation is further compounded by the transient nature of the starburst phase, so that any single system provides only an isolated snapshot in time, and by the intrinsic diversity of the population, which prevents one object from representing the whole. Furthermore, individual detections are affected by selection effects such as Malmquist and amplification biases, which naturally favour the brightest and most magnified objects that may not represent the typical population.

To overcome these limitations and extract robust physical insights from current and future surveys, a statistical framework is required.

1.3.1 The Stacking technique

The standard answer to this problem is stacking, a statistical signal-processing technique designed to extract faint emission from astronomical populations. By averaging the aligned measurements across an ensemble of N sources, the uncorrelated background fluctuations tend to cancel out, while the coherent astronomical signal adds up constructively, so that the mean physical properties of the underlying sample emerge from the noise.

Let S_i denote the flux measured for the i -th source, and assume that it is affected by an independent, zero-mean Gaussian noise of variance σ_i^2 . Under these assumptions, the overall SNR of the combined measurement is maximised by an inverse-variance weighted mean,

$$\langle S \rangle = \frac{\sum_{i=1}^N w_i S_i}{\sum_{i=1}^N w_i}, \quad \text{where } w_i = \frac{1}{\sigma_i^2}. \quad (6)$$

Standard error propagation then yields the uncertainty on the stacked flux,

$$\sigma_{\text{stacked}} = \frac{1}{\sqrt{\sum_{i=1}^N \frac{1}{\sigma_i^2}}}, \quad (7)$$

which reduces to the classical scaling relation $\sigma_{\text{stacked}} = \sigma_0 / \sqrt{N}$ in the special case of uniform single-target noise. For a sample sharing a characteristic flux density $\langle S \rangle$, the resulting SNR then scales as

$$\text{SNR}_{\text{stacked}} = \frac{\langle S \rangle}{\sigma_{\text{stacked}}} \approx \sqrt{N} \text{SNR}_{\text{single}}. \quad (8)$$

Under the assumption of a homogeneous population with comparable intrinsic emission, stacking recovers average signals that would

otherwise remain hidden beneath individual noise levels. Achieving the theoretical $\sqrt{N}$ gain requires precise astrometry to prevent signal attenuation from misalignment alongside statistically uncorrelated noise across all combined measurements. Two limitations apply because the result represents an average that loses population dispersion and yields physical meaning only when the ensemble stems from a controlled selection, making sample construction as critical as the combination process itself.

The effectiveness of this stacking approach is well established across diverse galactic environments ranging from local normal and main-sequence star-forming galaxies to extreme high-redshift populations. In the local Universe, analyses of high-order molecular transitions and average CO SLEDs across normal and starburst systems provide vital baseline constraints [127], and spectral stacking techniques have been extensively deployed on large representative samples from legacy surveys like GASS and xCOLD GASS to measure average atomic and molecular gas fractions and scale relations along the star-formation main sequence [149, 150]. Similarly at cosmic noon, most systematic detections of faint molecular gas, dense gas tracers, and rare species originate from spectral stacking applied to DSFGs [36, 90, 130], successfully recovering emission that remains entirely inaccessible in any single unlensed source as illustrated in Fig. 8.

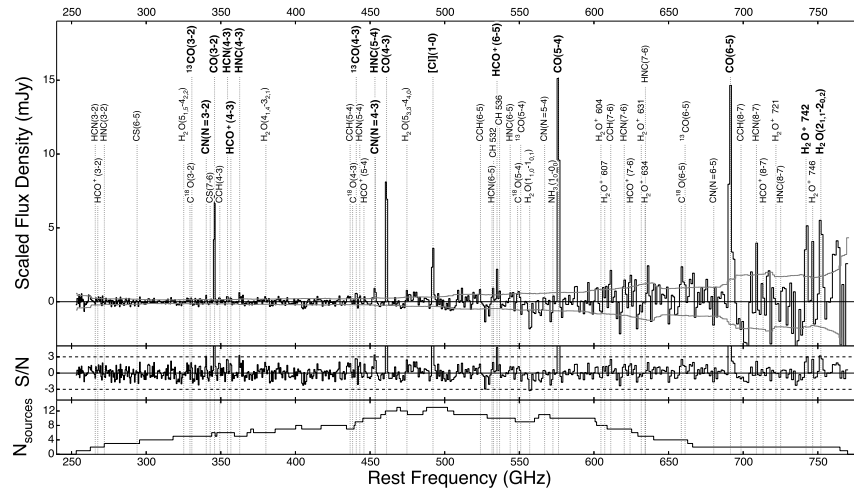


Figure 8: Average rest-frame spectrum obtained by stacking the spectra of a sample of DSFGs. Besides the bright CO and atomic carbon transitions, the stacked spectrum reveals the emission of dense gas tracers and of rarer species that remain undetected in the individual sources, showing how the combination of many observations gives access to diagnostics that are otherwise out of reach. Adapted from Spilker et al. [130].

Despite these substantial advances, the vast majority of existing stacking analyses in the literature have been performed using measurements originating from the same specific observational setup, instrumental configuration, or narrow redshift intervals. What would truly be powerful and scientifically compelling is to use this methodology as a robust tool to fully exploit the vast content of public archives, leveraging the wealth of observations and heterogeneous data already executed across disparate programmes.

1.3.2 *Exploiting public archives and their heterogeneity*

Building the large samples required for statistical measurements does not necessarily demand new observing campaigns, as astronomical archives now hold the accumulated output of decades of operations. Following a standard proprietary period across most major facilities, raw and calibrated data eventually enter the public domain, where they can be retrieved and repurposed for scientific goals beyond their original scope. This transforms isolated programmes into shared resources that address questions unforeseen by their original proposers and significantly enhance the legacy value of these facilities. A large fraction of public observations have never appeared in refereed publications, making astronomical archives an immense, unexploited reservoir of data. Because telescope time, operations, and calibration are already funded, reusing this data requires no additional observational resources. When the goal shifts toward characterising entire populations rather than individual objects, archive mining and dedicated programmes operate as complementary strategies. While a new proposal offers complete control over setup parameters and a homogeneous dataset, it comes at the cost of intense competition and long scheduling delays on heavily oversubscribed facilities. Archive mining, by contrast, offers three distinct advantages. The first is statistical: archives contain extensive observations where lines of interest were either not the primary target or remain hidden beneath the noise floor of individual pointings. Stacking these datasets boosts collective SNRs, enabling the recovery of faint signals from otherwise inaccessible sources. The second concerns sample fairness: dedicated follow-up programmes inherently favour the brightest or strongly magnified systems, whereas retrieving archival data for targets drawn from uniform flux-limited surveys captures lensed and unlensed populations far more impartially. The third is sustainability: storing and transferring massive datasets carries financial and environmental costs that yield no return if the data remain unanalysed. Public reuse aligns directly with open science principles (requiring data to be findable, accessible, interoperable, and reusable (FAIR) [151]) and helps archives evolve into dynamic discovery platforms [152].

The population studied here is particularly well represented in these repositories because wide-area (sub-)millimetre and millimetre surveys described in Section 1.1.2 delivered large catalogues of bright candidates. Due to their high apparent brightness and compact angular scales, these sources have been observed repeatedly across independent programmes by major facilities such as ALMA, NOEMA and APEX, so that they now populate public archives densely and across a wide range of rotational transitions. In particular, the ALMA Science Archive offers a large number of observations of DSFGs spanning a broad range of observing settings, redshifts and rotational transitions, as will be detailed in the following chapter.

This richness, however, comes with severe instrumental and observational challenges. Archival observations are collected across different projects, cycles, and instrumental configurations such that angular resolutions span orders of magnitude while spectral setups, spatial frequency sampling, and integration times vary widely. Such heterogeneity requires extreme caution in statistical studies, because fixed angular scales correspond to varying physical sizes across different redshifts and instrumental responses probe distinct spatial scales from one observation to another, running the risk of mixing intrinsically different physical regions within galaxies and biasing derived averages if combined uncritically. Furthermore, because the vast majority of observations in these bands are provided by interferometers, working directly with image data is precisely where this observational heterogeneity manifests its most critical challenges. Unlocking the full potential of these rich archives therefore relies on developing a specialised stacking methodology capable of properly accounting for data heterogeneity, thereby combining diverse observations into robust measurements. Developing this methodology and applying it to probe the average cold gas properties of DSFGs forms the core of this thesis, as detailed in the following section.

1.4 AIM AND STRUCTURE OF THIS THESIS

The primary challenge addressed in this thesis is unveiling the properties of cold molecular gas in DSFGs at cosmic noon, a critical step for characterising this galaxy population. To overcome the faintness of individual lines and exploit the wealth of observations present in public archives, the goal is to perform a statistical stacking analysis over a sample representative of the whole population, using the ALMA Science Archive as the primary data source. However, as discussed previously, these archival interferometric datasets exhibit severe instrumental and observational heterogeneity, spanning a broad range of angular resolutions, spectral setups, and redshifts, which prevents any standard combination at the image level.

To unlock this potential, I present the development of ViSTA, a novel stacking framework designed to operate natively in the visibility domain rather than on reconstructed images. By transforming every observation into a common rest-frame reference, ViSTA brings sources at different redshifts onto the same physical spatial and spectral scales, and preserves noise statistics and spatial scale coverage while combining diverse interferometric datasets. The framework is provided in two implementations, a reference version tailored to ALMA and built almost entirely on its native software environment, and an optimised high-performance version capable of handling data from a larger number of facilities, establishing a versatile tool that extends well beyond this specific study.

Building upon this methodological foundation, the thesis will present the first scientific application of ViSTA, targeting CO lines as the primary and most extensively observed molecular tracers. Once the average CO SLED has been constructed from a sample of bright DSFGs observed with ALMA, radiative transfer modelling will be used to constrain the average physical properties of the gas. Moreover, by comparing lensed and unlensed systems within a well-defined statistical sample, I will assess whether they genuinely represent the same underlying population.

- **Chapter 2** presents the sample by describing the parent catalogues, the selection criteria and the archival queries assembling the final dataset along with its properties and heterogeneity.
- **Chapter 3** develops the method, reviewing interferometry fundamentals and the visibility domain, explaining why statistical combination must precede image reconstruction, and detailing the components of ViSTA.
- **Chapter 4** describes the reference implementation and its validation on simulated and real data.
- **Chapter 5** presents the high-performance implementation that removes computational bottlenecks.
- **Chapter 6** reports the first scientific application, presenting the average CO SLED, deriving physical conditions via radiative transfer modelling, and comparing the lensed and unlensed subsets.
- **Chapter 7** summarises the main findings and outlines future developments enabled by this work.

THE SAMPLE

The previous chapter outlined the physical and chemical properties of DSFGs, their significance in cosmic star formation and the diagnostic power of the submillimetre regime. Given their intrinsic diversity, individual systems offer an incomplete view, necessitating a statistically representative sample to constrain their average properties. Building such an ensemble demands more than sheer numbers, as each target must also satisfy key observational prerequisites, notably secure redshift measurements and accessible interferometric data.

In this chapter, I turn to the practical assembly of this sample, explaining how these criteria shape the selection strategy, how they yield the final dataset used in subsequent analyses, and why its intrinsic heterogeneity ultimately calls for a new stacking method.

Specifically, Section 2.1 details the sample construction, explaining how the parent catalogue is derived from wide-area *Herschel* surveys via a flux threshold, outlining the role of gravitational lensing, and introducing the necessary observational requirements. Section 2.2 presents ALMA, describing its array configurations, receivers, and spectral setups, with an emphasis on its scientific archive as the primary data source for this study. Section 2.3 then introduces the final collection resulting from the intersection of the parent sample with the ALMA archive. Finally, Section 2.4 summarises the chapter and motivates the need for the dedicated *uv*-plane stacking framework detailed in the following chapters.

2.1 THE SELECTION CRITERIA

The diversity discussed in Section 1.1.4 makes a carefully defined selection essential to constrain the average properties of the population rather than those of individual objects. This section details the origin of the candidates, the criteria governing the selection and the observational requirements each target must satisfy.

2.1.1 Parent sample and flux density selection

Exploiting the amplification effect of gravitational lensing, the selection strategy follows the observational framework pioneered by Negrello et al. [153] and Negrello et al. [154], and further refined by Bakx et al. [155] and Borsato et al. [156]. Within this framework, the candi-

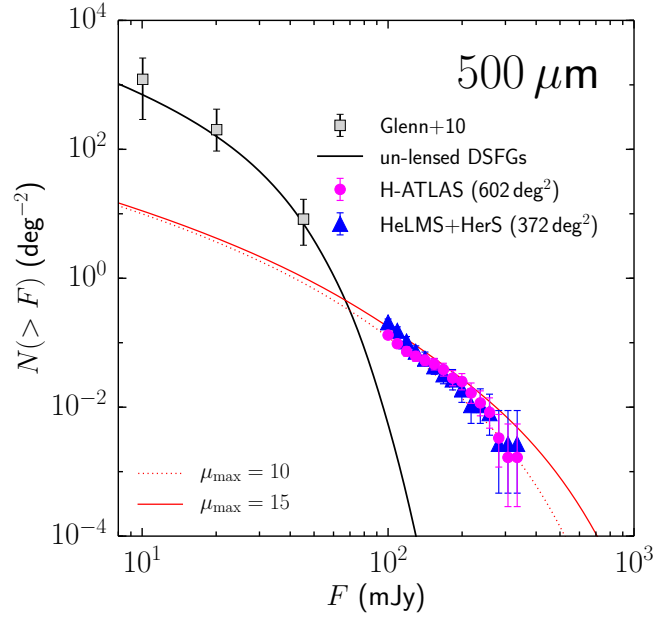


Figure 9: Cumulative submillimetre number counts at $500\ \mu\text{m}$ adapted from Negrello et al. [153], illustrating the steep decline in the unlensed DSFG population (solid black line) at bright fluxes and the clear dominance of strongly lensed sources above $S_{500} \approx 80\text{--}100\ \text{mJy}$ predicted by theoretical models with varying maximum magnification values.

date parent sample is drawn from the wide-area photometric surveys conducted with the *Herschel Space Observatory* [29] introduced in Section 1.1.2, namely H-ATLAS [25, 26], HerMES/HeLMS [27], and HerS [28], which together mapped several hundred square degrees at 250, 350, and 500 μm and, thanks to the negative K -correction discussed there, recover sources of comparable intrinsic luminosity with almost equal efficiency across $z \sim 1\text{--}4$.

As demonstrated by Negrello et al. [153] (see Fig. 9), applying a flux density threshold at 500 μm ($S_{500} \geq 100 \text{ mJy}$) drastically suppresses the unlensed population, isolating a sample overwhelmingly dominated by high-redshift lensed starbursts. Bakx et al. [155] later showed that lowering this limit to $S_{500} \geq 80 \text{ mJy}$, combined with a photometric redshift cut $z_{\text{phot}} > 2$ to filter out low-redshift interlopers, recovers additional lensed targets with moderate magnifications while capturing a significant portion of the high-flux tail of the unlensed DSFG distribution, and defined on this basis the *Herschel* Bright Sources (HerBS) sample. The resulting selection combines both strongly lensed systems and intrinsically hyper-luminous unlensed sources.

However, a submillimetre flux density threshold alone does not guarantee a pure DSFG sample, as unrelated sources can still exceed 500 μm . These interlopers consist mainly of nearby spiral galaxies that appear bright due to their proximity, and radio-loud AGNs or blazars whose emission stems from synchrotron radiation rather than dust. Matching submillimetre detections with optical and radio catalogues easily identifies and removes these contaminants from the candidate list. To this end, I assemble the parent sample by collecting the submillimetre sources with $S_{500} \geq 80 \text{ mJy}$ after they have been systematically purged of such contaminants in earlier works. Specifically, this selection relies on the cleaned candidate lists from Negrello et al. [153] and Bakx et al. [155] across the H-ATLAS fields, from Wardlow et al. [157] over HerMES, and on the filtered HeLMS and HerS catalogues from Nayyeri et al. [158]. Overall, the selection criteria follow the approach of Borsato et al. [156], except that this parent sample is not restricted to the targets observed by their specific *HST* snapshot programme, thereby retaining well-studied systems such as the SDP sources.

This framework provides a robust sample of both lensed and bright unlensed DSFGs, which is essential to test whether the two populations share similar physical properties or represent different stages of extreme star formation, and to control for the biases discussed in Section 1.1.3. The lensing nature of most of these objects has already been established in the literature. The majority were classified by Borsato et al. [156] through *HST* snapshot imaging, in which the foreground deflector is modelled and subtracted so that the residual emission reveals the arcs, rings, or counter-images that identify a confirmed strong lens. For a subset of the Southern sky sources, an

independent confirmation comes from Bakx et al. [159], whose high-resolution ALMA observations resolve the characteristic morphological signatures of lensing and provide a ready lens model for some of the targets.

2.1.2 Spectroscopic redshifts

A key prerequisite for the sample is the availability of secure spectroscopic redshifts z_{spec} . While photometric estimates from *Herschel* colours suffice for initial target selection, their precision is too coarse for detailed ISM characterisation. Accurate redshifts are necessary to unambiguously identify (sub-)millimetre transitions within dense molecular spectra, preventing misidentifications that would distort excitation models and chemical diagnostics. They also enable reliable calculations of intrinsic line and dust luminosities via the luminosity distance, directly impacting derived gas masses, gas fractions, and depletion timescales. From a kinematic perspective, they establish the rest frame required to model line profiles, measure velocity dispersions, and identify broad outflow or merger signatures.

This precision is particularly critical for the stacking methodology central to this thesis (see Section 1.3.1), where any redshift uncertainty causes spectral misalignment prior to combination. Jolly, Knudsen, and Stanley [160] quantified this effect through simulated stacking experiments, demonstrating that line reconstruction deteriorates rapidly once the uncertainty exceeds $\Delta z \sim 0.01$. In particular, the amplitude of a 400 km s^{-1} -wide line drops to approximately 20 per cent for $\Delta z = 0.05$. Moreover, this sensitivity depends strongly on intrinsic linewidth, rendering narrower features significantly more vulnerable than broader ones. Complex spectral signatures, such as double-peaked profiles, degrade even faster and become indistinguishable already for $\Delta z \gtrsim 0.001$. Given that typical photometric uncertainties for these targets ($\Delta z \sim 0.2\text{--}0.5$) exceed this threshold by orders of magnitude, stacking without secure spectroscopic redshifts would severely attenuate or entirely wash out the coherent carbon monoxide (CO) signal. Consequently, only sources with confirmed z_{spec} are retained.

Most spectroscopic redshifts for the HerBS sources in the South Galactic Pole were retrieved from the BEARS survey [159, 161, 162], based on ALMA spectral scans. For the HeLMS and HerS fields, I rely primarily on the z-GAL NOEMA programme [163, 164], complemented by literature values. Bright, strongly lensed sources originally identified in H-ATLAS [153, 154], such as the SDP targets, have dedicated follow-up studies from which their values are drawn [e.g. 165–171]. Finally, for HerBS sources [155] resolved into multiple components, the specific redshift adopted for each component is addressed in the following sections.

High-sensitivity (sub-)millimetre observations are required to detect these molecular transitions across both lensed and unlensed populations. ALMA, through its public archive, provides high-quality data for these survey-selected targets, enabling a robust statistical characterisation of high-redshift DSFGs.

2.2 ALMA: THE IDEAL TOOL FOR STUDYING HIGH-Z DSFGs

The interferometric data with which the sample is analysed are obtained with Atacama Large Millimeter/submillimeter Array (ALMA). This section describes the observatory and, in particular, its scientific archive, which is not only the medium through which the data are retrieved but also a defining element of the strategy adopted in this work.

ALMA is an interferometer operating at millimetre and submillimetre wavelengths, located on the Chajnantor plateau in the Chilean Andes at an altitude of about 5000 m. The site is one of the driest on Earth, minimising atmospheric water vapour absorption that would otherwise limit observations in these bands, and providing the stable and transparent sky required for high-frequency interferometry.

2.2.1 *Arrays and configurations*

ALMA comprises 66 high-precision antennas, including fifty-four 12 m dishes and twelve 7 m dishes. Fifty of the 12 m antennas form the main array, with baselines reconfigurable from about 150 m in the most compact configuration to about 16 km in the most extended one. Depending on the observing frequency and configuration, this setup yields an angular resolution ranging from several arcseconds down to about 0.005 arcseconds at 950 GHz, making the main array well suited to probe the smallest angular scales.

The ACA comprises twelve 7 m antennas in a fixed configuration of short baselines from 9 to 30 m, together with the four remaining 12 m antennas used for total power measurements, yielding a coarser resolution on the order of several to a few tens of arcseconds. The ACA recovers extended and large scale emission that the main array resolves out, offering high efficiency for bright and extended sources. Combining these configurations samples the uv plane over a broad range of spatial scales.

2.2.2 *Receivers, spectral setup, and signal calibration*

To cover a broad frequency range spanning roughly 84 to 950 GHz, each antenna houses receivers corresponding to ten planned

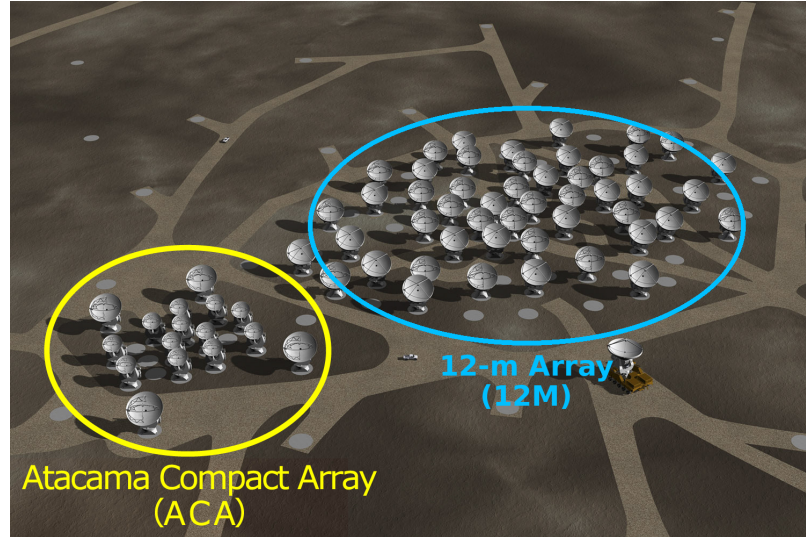


Figure 10: Artist’s impression of ALMA. It shows the two components of the array, including the 12 m Array, shown in the blue circle, consisting of fifty 12 m antennas arranged in reconfigurable baselines, and the Atacama Compact Array (ACA), in the yellow circle, which comprises twelve 7 m antennas plus four 12 m antennas, closely packed to provide sensitivity to extended emission. Adapted from ALMA (ESO/NAOJ/NRAO).

frequency bands. Band 1 has been available since Cycle 10, and Band 2 has been available since Cycle 13, so the system now offers the full complement of ten bands. Because incoming high-frequency signals cannot be processed directly, they are first downconverted by mixing them with a narrow local oscillator tone to generate an intermediate frequency signal. This signal is split into an upper and a lower sideband, with a bandwidth of 4–8 GHz for most bands and 4–12 GHz for Bands 1, 9 and 10. A second downconversion places up to four basebands within these sidebands, with each baseband hosting up to four spectral windows targeted at specific science goals, spectral lines, or line-free continuum. The exact setups used in this work are provided alongside the observations [see the ALMA Observing Tool and Technical Handbook, 172].

Each spectral window is independently digitised and channelised at the antenna level. The resulting frequency-resolved signal is sent to the correlator, which cross-multiplies signals between antenna pairs within each spectral window to generate complex visibilities, which are the fundamental interferometric observable encoding sky brightness spatial coherence on the scale of each baseline. Cross correlation is performed continuously and averaged over short integration times, typically a few seconds, resolving each visibility in time. Longer ef-

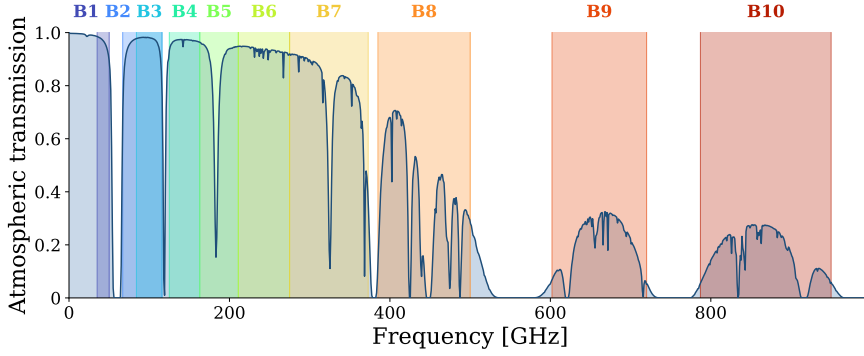


Figure 11: ALMA receiver bands overlaid on the atmospheric transmission curve, highlighting how the observable bands align with atmospheric millimetre and submillimetre windows.

fective target integrations are built up by accumulating many such dumps over an observation.

However, the raw signals recorded by each antenna are corrupted by instrumental response and fluctuating atmospheric conditions, which distort both amplitude and phase. Calibration resolves this by comparing measured visibilities of well-characterised reference sources against their expected signals to derive correction factors for the science target. Consequently, target observations are interleaved with short scans of nearby calibrators. The resulting dataset across all baselines, spectral channels, integrations, and polarisations constitutes the raw visibilities delivered directly to the archive.

2.2.3 *The ALMA archive as a scientific resource*

The data used in this work are retrieved entirely from the [ALMA Science Archive](#), for the reasons developed in Section 1.3.2. The ALMA Science Archive is the central repository for all scientific observations executed by the array, and over more than a decade of operations it has matured into a resource exceptionally well suited for statistical studies of high-redshift DSFGs. It grows continuously through standard operations: once a proposal submitted by a Principal Investigator is accepted, it is partitioned into science goals, scheduled, observed, and evaluated through quality assurance before deposition, remaining proprietary to the original team for one year before entering the public domain.

This repository is expanding rapidly and currently holds on the order of two petabytes of data, yet its scientific exploitation remains markedly uneven. This imbalance is quantified in Fig. 12, which shows how refereed publications are distributed across the 5200 project codes observed from Cycle 0 to Cycle 11, i.e. up to the 2024

call. Counting unique project codes (left panel), only 60.9% have produced at least one paper, while the 39.1% that remain unpublished account for 40.0% of the 742 599 science observations once the distribution is weighted by the number of observations per project (right panel). The unexploited fraction is therefore not confined to small or marginal programmes but is essentially proportional to the observing effort invested. The distribution is also strongly skewed at the opposite end, with a handful of legacy programmes reaching up to 109 associated papers, so that the scientific return is concentrated in a small minority of projects [173].

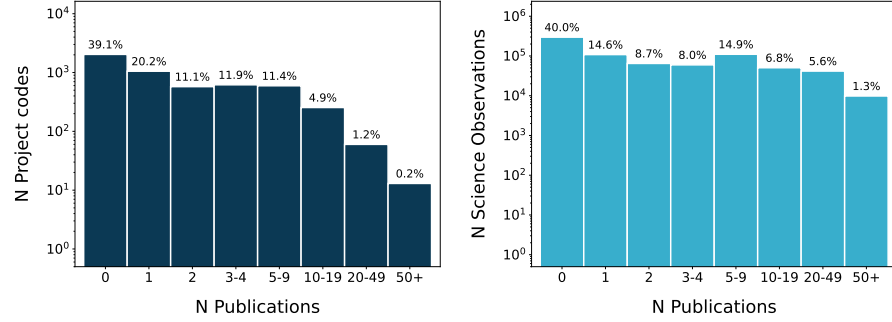


Figure 12: Number of refereed publications associated with ALMA project codes, for all project codes from Cycle 0 to Cycle 11, i.e. up to the 2024 call ($N = 5\,200$ projects, 742 599 science observations). Publications are counted per project code, since ALMA bibliographic links are assigned at that level. *Left:* unique project codes per bin. *Right:* the same distribution weighted by the number of science observations belonging to each project. Percentages above each bar give the fraction of the corresponding sample.

Beyond the sustainability argument, ALMA observations represent an ideal resource for this study, as the target population of high-redshift DSFGs is exceptionally well represented within the archive. Indeed, the wide-area submillimetre surveys used to build the parent sample (Section 2.1.1), most notably those carried out with *Herschel*, delivered large catalogues of bright candidates, many of which are strongly lensed [25, 154]. Owing to their high apparent brightness and compact scales, these sources have been observed extensively by ALMA across many independent programmes, so that they now populate the archive densely across a wide range of molecular transitions. I therefore choose to exploit this rich repository for this analysis, using public visibilities to build a large sample that includes both lensed and unlensed DSFGs. Nevertheless, the brightest lensed candidates are naturally the most frequently targeted in the literature, since their high flux densities make them the most secure lensed identifications and provide the high SNR required for efficient, high-resolution follow-up. As a result, coverage is expected to be most complete pre-

cisely for this brightest subset, so that a residual statistical bias toward these sources remains even within an archival sample.

However, this observational richness comes with the intrinsic heterogeneity anticipated in Section 1.3.2, as synthesised beams range from 0.01 to 10 arcsec, while uv -plane coverage, spectral resolution, and integration times vary by over an order of magnitude. Combining such disparate observations uncritically risks probing mismatched spatial scales and biasing average population properties, so fully exploiting the potential of the archive requires a dedicated methodology designed to handle this heterogeneity consistently.

2.3 THE FINAL ALMA ENSEMBLE

I construct the final sample by combining the criteria established in the previous sections. Starting from the wide-area *Herschel* candidate catalogues, I retain only sources with a flux density $S_{500} \geq 80$ mJy and a secure spectroscopic redshift, and I further restrict the selection to targets with observations available in the ALMA Science Archive.

The parent sample assembled in this way comprises 207 *Herschel* sources with a secure spectroscopic redshift, of which 192 possess public observations in the ALMA Science Archive and enter the final dataset presented in Table 1. Ten of these sources are resolved by high-resolution ALMA follow-up into multiple components, each with an individual spectroscopic redshift, bringing the total number of entries in the table to 202. The sample spans redshifts from $z = 1.03$ to 5.40 , with three-quarters of the targets lying between $z = 2$ and $z = 3.5$, and is drawn predominantly from H-ATLAS (133 sources), followed by HeLMS (30), HerMES (17), and HerS (12).

Where high-resolution ALMA follow-up resolved a single *Herschel* source into multiple components, each component is reported separately and the appropriate coordinates are adopted from these observations [174]. Components are labelled by their HerBS identifier and a component letter, and only those with a measured spectroscopic redshift are retained.

While the final ensemble encompasses both candidate strongly lensed systems and unlensed sources, the detailed classification of each target is deferred to Chapter 6 for the subset driving the primary scientific analysis of this thesis.

Table 1: Sources of the parent sample with public ALMA observations (updated with multiple components).

Source	z_{spec}	RA (J2000)	Dec (J2000)	Ref	HerBS_ID
HATLASJ000007.4-334059	3.307	00:00:07.5	-33:41:03.1	Ba18	42A

continued on next page

Table 1 – continued

Source	z_{spec}	RA (J2000)	Dec (J2000)	Ref	HerBS_ID
HATLASJ000007.4-334059	3.314	00:00:07.4	-33:40:55.9	Ba18	42B
HATLASJ000124.9-354212	4.098	00:01:24.8	-35:42:11.0	Ba18	41A
HELMSJ000215.9-012829	1.4202	+00:02:15.9	-01:28:29.0	Na16	
HATLASJ000330.6-321136	3.077	00:03:30.6	-32:11:35.1	Ba18	155A
HATLASJ000745.8-342014	3.140	00:07:46.2	-34:20:03.0	Ba18	163A
HATLASJ000806.8-351205	4.526	00:08:07.2	-35:12:05.0	Ba18	117A
HELMSJ001325.7+042509	3.3896	+00:13:25.7	+04:25:09.0	Na16	
HERMESJ001353.5-060200	1.9494	+00:13:53.5	-06:02:00.0	Na16	
HERMESJ001615.8+032433	2.765	00:16:15.9	+03:24:38.0	Na16	
HERMESJ001626.0+042613	2.5093	00:16:26.0	+04:26:12.9	Na16	
HELMSJ001800.1-060235	2.7710	+00:18:00.1	-06:02:35.0	Na16	
HATLASJ001802.2-313505	2.369	00:18:02.5	-31:35:05.1	Ba18	106A
HATLASJ001802.2-313505	2.588	00:18:01.1	-31:35:08.0	Ba18	106B
HATLASJ002054.6-312752	3.160	00:20:54.2	-31:27:57.4	Ba18	81A
HATLASJ002054.6-312752	2.588	00:20:54.7	-31:27:50.8	Ba18	81B
HELMSJ002208.1+034044	2.1898	+00:22:08.1	+03:40:44.0	Na16	
HELMSJ002220.9-015524	5.1614	00:22:20.8	-01:55:20.8	Na16	
HATLASJ002533.6-333826	2.059	00:25:33.7	-33:38:26.3	Ba18	87
HATLASJ002624.8-341737	3.050	00:26:24.9	-34:17:38.1	Ba18	22A
HELMSJ002719.5+001204	2.2714	+00:27:19.5	+00:12:04.0	Na16	
HATLASJ003207.7-303724	2.561	00:32:07.6	-30:37:35.2	Ba18	56C
HELMSJ003519.7+072806	1.9817	+00:35:19.7	+07:28:06.0	Na16	
HELMSJ003619.8+002420	1.6169	+00:36:19.8	+00:24:20.0	Na16	
HATLASJ003717.0-323307	2.883	00:37:16.7	-32:32:57.4	Ba18	122A
HELMSJ003814.1-002252	4.984	00:38:14.0	-00:22:52.5	Na16	
HELMSJ003929.6+002426	2.848	00:39:29.5	+00:24:25.8	Na16	
HERMESJ004410.2+011821	4.1625	00:44:10.3	+01:18:22.2	Na16	
HELMSJ004622.3+073509	2.5765	00:46:22.3	+07:35:09.0	Na16	
HERMESJ004714.1+032453	2.2919	00:47:14.1	+03:24:56.2	Na16	
HERMESJ004723.3+015749	1.441	00:47:23.4	+01:57:51.5	Na16	
HATLASJ004736.0-272951	2.198	00:47:36.1	-27:29:52.0	Ba18	24
HELMSJ004747.1+061444	2.6875	+00:47:47.1	+06:14:44.0	Na16	
HATLASJ004853.3-303110	3.265	00:48:53.4	-30:31:09.9	Ba18	57
HATLASJ005132.0-302011	2.432	00:51:31.7	-30:20:20.6	Ba18	63A
HATLASJ005132.8-301848	2.434	00:51:32.9	-30:18:49.6	Ba18	45A
HERMESJ005159.4+062240	2.392	+00:51:59.5	+06:22:41.0	Na16	
HATLASJ005506.5-300027	1.569	00:55:06.5	-30:00:28.3	Ba18	207
HATLASJ005629.6-311206	2.228	00:56:29.2	-31:12:07.5	Ba18	77A
HATLASJ005659.4-295039	3.992	00:56:59.2	-29:50:39.3	Ba18	90A

continued on next page

Table 1 – continued

Source	z_{spec}	RA (J2000)	Dec (J2000)	Ref	HerBS_ID
HATLASJ005724.2-273122	3.261	00:57:24.3	-27:31:23.3	Ba18	60
HELMSJ005841.2-011149	1.498	+00:58:41.2	-01:11:49.0	Na16	
HATLASJ011014.5-314814	3.955	01:10:14.4	-31:48:15.9	Ba18	160
HERSJ011640.1-000454	3.791	+01:16:40.1	-00:04:54.0	Na16	
HERSJ011722.3+005624	2.4688	01:17:22.3	+00:56:22.7	Na16	
HATLASJ011730.3-320719	1.407	01:17:30.5	-32:07:20.9	Ba18	138B
HATLASJ011850.1-283642	2.658	01:18:50.2	-28:36:44.0	Ba18	178A
HATLASJ011850.1-283642	2.655	01:18:50.1	-28:36:40.5	Ba18	178B
HATLASJ011850.1-283642	2.656	01:18:49.9	-28:36:43.2	Ba18	178C
HERSJ012041.6-002705	2.015	+01:20:41.6	-00:27:05.0	Na16	
HATLASJ012222.3-274456	3.125	01:22:22.4	-27:44:53.7	Ba18	120A
HATLASJ012222.3-274456	3.124	01:22:22.1	-27:44:59.0	Ba18	120B
HATLASJ012335.1-314619	2.730	01:23:34.6	-31:46:23.6	Ba18	145A
HATLASJ012407.3-281434	2.631	01:24:07.5	-28:14:34.7	Ba18	11
HATLASJ012415.9-310500	2.075	01:24:16.1	-31:04:59.5	Ba18	69A
HATLASJ012415.9-310500	2.073	01:24:15.8	-31:05:05.1	Ba18	69B
HERSJ012521.0+011724	2.4759	01:25:21.0	+01:17:24.5	Na16	
HERSJ012546.3-001143	2.2706	+01:25:46.3	-00:11:43.0	Na16	
HERSJ012620.5+012950	1.449	+01:26:20.5	+01:29:50.0	Na16	
HERSJ012754.1+004940	3.0607	+01:27:54.1	+00:49:40.0	Na16	
HATLASJ012853.0-332719	3.026	01:28:53.0	-33:27:19.1	Ba18	73
HERSJ013212.2+001754	1.6926	01:32:12.3	+00:17:55.5	Na16	
HATLASJ013239.9-330906	1.971	01:32:40.2	-33:09:08.0	Ba18	40
HATLASJ013840.4-281855	3.782	01:38:40.4	-28:18:57.5	Ba18	14
HATLASJ013951.9-321446	2.656	01:39:52.0	-32:14:45.5	Ba18	55
HERSJ014057.3-010547	3.3441	+01:40:57.3	-01:05:47.0	Na16	
HATLASJ014313.2-332633	2.151	01:43:13.3	-33:26:33.1	Ba18	200
HATLASJ014520.0-313835	2.553	01:45:20.0	-31:38:32.5	Ba18	107
HERSJ020941.2+001558	2.5534	02:09:41.2	+00:15:58.2	Na16	
HERSJ021402.6-004612	3.0182	+02:14:02.6	-00:46:12.0	Na16	
HERMESJ021831-053131	3.3950	02:18:30.7	-05:31:31.8	W13	
HERSJ021434.4+005926	2.1981	+02:14:34.4	+00:59:26.0	Na16	
HERMESJ022017-060143	2.307	02:20:16.6	-06:01:43.3	W13	
HERMESJ022135-062617	2.72	+02:21:35.0	-06:26:17.0	W13	
HERMESJ022548-041750	2.985	+02:25:48.0	-04:17:50.0	W13	
HATLASJ083051.0+013225	3.634	08:30:51.2	+01:32:24.4	Ba18	
HATLASJ083817-004134	3.7168	+08:38:17.4	-00:41:34.4	Ba18	
HATLASJ083859.3+021325	2.6977	+08:38:59.3	+02:13:26.0	Ba18	
HATLASJ083932-011760	2.669	+08:39:32.2	-01:17:58.2	Ba18	

continued on next page

Table 1 – continued

Source	z_{spec}	RA (J2000)	Dec (J2000)	Ref	HerBS_ID
HATLASJ083945+021023	2.4793	+08:39:45.1	+02:10:22.1	Ba18	
HATLASJ084958+010713	1.896	+08:49:57.7	+01:07:13.4	Ba18	
HATLASJ085309-005727	3.2884	+08:53:08.6	-00:57:28.1	Ba18	
HATLASJ085358.9+015537	2.0925	08:53:58.9	+01:55:37.0	Na17	
HATLASJ090302.9-014127	2.3049	09:03:03.0	-01:41:27.1	Na17	
HATLASJ090311.6+003907	3.042	09:03:11.6	+00:39:06.7	Ba18	
HATLASJ090453.2+022017	1.8910	09:04:53.0	+02:20:16.8	Ba18	
HATLASJ090740.0-004200	1.577	09:07:40.1	-00:42:00.0	Na17	
HATLASJ091043.0-000322	1.784	09:10:43.1	-00:03:22.8	Na17	
HATLASJ091304.9-005344	2.6256	09:13:05.1	-00:53:43.0	Ba18	
HATLASJ084259.9+024959	2.7675	+08:43:00.0	+02:49:59.2	Ba18	
HATLASJ091840.8+023048	2.581	+09:18:40.9	+02:30:48.6	Ba18	
HATLASJ092136+000132	2.4047	+09:21:35.7	+00:01:31.8	Ba18	
HATLASJ092409-005018	4.3238	+09:24:08.8	-00:50:17.5	Ba18	
HATLASJ113243.0-005108	2.58	+11:32:43.1	-00:51:09.0	Ba18	
HATLASJ113526.2-014606	3.1276	11:35:26.3	-01:46:06.6	Ba18	
HATLASJ113803.6-011737	3.15	11:38:03.2	-01:17:38.0	Ba18	
HATLASJ115112-012638	1.4219	+11:51:12.2	-01:26:37.3	Ba18	
HATLASJ114637.9-001132	3.259	+11:46:38.0	-00:11:32.3	Ba18	
HATLASJ114753-005832	2.8169	+11:47:52.8	-00:58:31.4	Ba18	
HATLASJ115433.6+005042	3.9625	+11:54:33.6	+00:50:42.4	Ba18	
HATLASJ115820.1-013752	2.1911	+11:58:20.2	-01:37:52.7	Ba18	
HATLASJ120127.6-014043	3.7293	12:01:27.6	-01:40:45.6	Ba18	
HATLASJ120319.1-011253	2.9283	+12:03:19.1	-01:12:54.0	Ba18	
HATLASJ121301.5-004922	3.1438	+12:13:01.6	-00:49:22.4	Ba18	
HATLASJ121348.0+010812	3.1546	+12:13:48.1	+01:08:12.5	Ba18	
HATLASJ121416.3-013704	2.0126	+12:14:16.4	-01:37:04.4	Ba18	
HATLASJ121542.7-005220	2.5738	+12:15:42.7	-00:52:20.3	Ba18	
HATLASJ122034.2-003805	2.5890	+12:20:34.3	-00:38:06.0	Ba18	
HATLASJ120600.7+003459	2.5596	+12:06:00.7	+00:34:59.5	Ba18	
HATLASJ120709.2-014702	2.1827	+12:07:09.3	-01:47:02.8	Ba18	
HATLASJ125125.8+254929	3.4419	+12:51:25.9	+25:49:29.6	Ba18	
HATLASJ125135.3+261457	3.675	+12:51:35.3	+26:14:57.8	Ba18	
HATLASJ125632.4+233627	3.565	12:56:32.5	+23:36:27.5	Ba18	
HATLASJ125652.5+275903	2.79	+12:56:52.5	+27:59:00.2	Ba18	
HATLASJ130054+260303	3.3074	13:00:53.8	+26:03:00.5	Ba18	
HATLASJ130118.0+253708	4.04	+13:01:18.1	+25:37:08.4	Ba18	
HATLASJ131611+281220	2.9497	+13:16:11.5	+28:12:19.1	Ba18	
HATLASJ132128+282023	3.1958	+13:21:28.6	+28:20:20.0	Ba18	

continued on next page

Table 1 – continued

Source	z_{spec}	RA (J2000)	Dec (J2000)	Ref	HerBS_ID
HATLASJ132301.7+341649	2.19	+13:23:01.7	+34:16:49.4	Ba18	
HATLASJ132419+320754	3.2121	+13:24:19.0	+32:07:52.0	Ba18	
HATLASJ132427.0+284449	1.676	+13:24:27.0	+28:44:49.9	Ba18	
HATLASJ132630.1+334407	2.951	13:26:30.1	+33:44:08.8	Ba18	
HATLASJ132859.2+292326	2.778	13:28:59.3	+29:23:26.3	Ba18	
HATLASJ132909+300958	3.4935	+13:29:09.5	+30:09:56.9	Ba18	
HATLASJ133008.5+245900	3.1112	13:30:08.5	+24:58:59.2	Ba18	
HATLASJ133256+342210	2.9268	+13:32:55.8	+34:22:08.0	Ba18	
HATLASJ133440+353140	3.1725	+13:34:40.4	+35:31:41.2	Ba18	
HATLASJ133542.9+300401	2.685	+13:35:43.0	+30:04:01.6	Ba18	
HATLASJ133534+341837	2.3310	+13:35:34.1	+34:18:34.9	Ba18	
HATLASJ133649.9+291800	2.2024	+13:36:49.9	+29:18:01.0	Na17	
HATLASJ134429+303036	2.3010	+13:44:29.5	+30:30:34.2	Ba18	
HATLASJ140422-001218	2.8122	+14:04:21.8	-00:12:17.3	Ba18	
HATLASJ141118-010655	4.1408	+14:11:17.9	-01:06:55.8	Ba18	
HATLASJ141810.0-003747	2.2406	14:18:10.1	-00:37:46.1	Ba18	
HATLASJ141351.9-000026	2.4782	14:13:52.1	-00:00:24.4	Ba18	
HATLASJ141833+010212	2.6810	+14:18:33.0	+01:02:12.5	Ba18	
HATLASJ142140+000448	2.781	+14:21:40.4	+00:04:48.0	Ba18	
HATLASJ142413.9+022303	4.243	14:24:14.0	+02:23:03.7	Ba18	
HERMESJ142824+352620	1.325	14:28:24.1	+35:26:19.5	W13	
HATLASJ142935.3-002836	1.027	+14:29:35.2	-00:28:35.7	Na17	
HATLASJ143403.5+000234	3.1150	+14:34:03.6	+00:02:34.4	Ba18	
HATLASJ144243+015506	3.1501	+14:42:43.4	+01:55:04.4	Ba18	
HATLASJ144556.1-004853	1.8349	+14:45:56.1	-00:48:53.3	Ba18	
HATLASJ145135.2-011418	2.5875	14:51:35.2	-01:14:10.3	Ba18	
HATLASJ145337.2+000407	3.0408	+14:53:37.2	+00:04:08.0	Ba18	
HATLASJ223615.2-343301	3.741	22:36:15.3	-34:33:02.3	Ba18	121A
HATLASJ223753.8-305828	2.719	22:37:53.8	-30:58:27.6	Ba18	68
HATLASJ223942.4-333304	2.371	22:39:42.3	-33:33:04.1	Ba18	111
HATLASJ224759.7-310135	2.085	22:47:59.7	-31:01:36.0	Ba18	141
HATLASJ224920.6-332940	2.272	22:49:21.0	-33:29:41.5	Ba18	209A
HATLASJ225045.5-304719	2.583	22:50:45.4	-30:47:20.3	Ba18	168A
HATLASJ225250.7-313657	2.433	22:52:50.7	-31:36:59.9	Ba18	47
HATLASJ225600.7-313232	3.300	22:56:00.7	-31:32:33.0	Ba18	189
HATLASJ225611.7-325653	2.401	22:56:11.7	-32:56:52.0	Ba18	135A
HATLASJ225744.6-324231	2.478	22:57:44.5	-32:42:33.0	Ba18	208A
HATLASJ225744.6-324231	2.483	22:57:44.8	-32:42:32.9	Ba18	208B
HATLASJ230002.6-315005	2.231	23:00:02.5	-31:50:08.9	Ba18	80A

continued on next page

Table 1 – continued

Source	z_{spec}	RA (J2000)	Dec (J2000)	Ref	HerBS_ID
HATLASJ230002.6-315005	1.968	23:00:02.8	-31:50:08.0	Ba18	80B
HATLASJ230538.5-312204	2.227	23:05:38.8	-31:22:05.6	Ba18	182
HATLASJ230546.2-331038	2.724	23:05:46.4	-33:10:38.1	Ba18	49A
HATLASJ230546.2-331038	2.730	23:05:46.5	-33:10:43.1	Ba18	49B
HATLASJ230815.5-343801	3.925	23:08:15.7	-34:38:00.5	Ba18	28
HELMSJ231857.2-053035	2.8187	+23:18:57.2	-05:30:35.0	Na16	
HELMSJ232210.3-033559	4.688	+23:22:10.3	-03:35:59.0	Na16	
HATLASJ232210.9-333749	2.003	23:22:10.6	-33:37:58.4	Ba18	146B
HATLASJ232419.8-323926	2.182	23:24:19.8	-32:39:26.5	Ba18	18
HERMESJ232439.4-043934	2.4726	23:24:39.4	-04:39:34.2	Na16	
HELMSJ001226.9+020810	5.3994	+00:12:26.9	+02:08:10.0	Na16	
HELMSJ232558.3-044525	2.2978	+23:25:58.3	-04:45:25.0	Na16	
HELMSJ232617.5-025319	2.1567	+23:26:17.5	-02:53:19.0	Na16	
HATLASJ232623.0-342642	2.619	23:26:23.1	-34:26:44.0	Ba18	37
HELMSJ232831.8-004035	2.2834	23:28:31.8	-00:40:35.6	Na16	
HELMSJ232833.6-031416	3.3514	+23:28:33.6	-03:14:16.0	Na16	
HATLASJ232900.6-321744	3.229	23:29:00.8	-32:17:45.0	Ba18	39
HATLASJ233024.1-325032	3.287	23:30:24.4	-32:50:32.3	Ba18	102A
HATLASJ233037.2-331217	2.170	23:30:37.4	-33:12:16.8	Ba18	123
HELMSJ231447.5-045658	1.37	+23:14:47.5	-04:56:58.0	Na16	
HERMESJ233255.5-031134	2.6895	23:32:55.5	-03:11:36.4	Na16	
HERMESJ233255.7-053424	2.4022	23:32:55.5	-05:34:26.6	Na16	
HELMSJ233441.0-065220	1.905	+23:34:41.0	-06:52:20.0	Na16	
HELMSJ233633.5-032119	2.335	+23:36:33.5	-03:21:19.0	Na16	
HERMESJ233620.7-060826	3.4346	23:36:20.9	-06:08:26.0	Na16	
HELMSJ234014.6-070738	1.9558	+23:40:14.6	-07:07:38.0	Na16	
HERMESJ234051.3-041937	3.5027	23:40:51.6	-04:19:40.3	Na16	
HELMSJ233721.9-064740	2.2154	+23:37:21.9	-06:47:40.0	Na16	
HELMSJ233728.8-045106	2.1947	+23:37:28.8	-04:51:06.0	Na16	
HELMSJ234314.0+012152	3.9802	+23:43:14.0	+01:21:52.0	Na16	
HATLASJ234418.1-303936	3.323	23:44:18.1	-30:39:38.9	Ba18	21A
HATLASJ234750.5-352931	2.400	23:47:50.4	-35:29:30.2	Ba18	93
HELMSJ234951.6-030019	2.2232	+23:49:51.6	-03:00:19.0	Na16	
HATLASJ234955.7-330833	2.507	23:49:55.6	-33:08:34.4	Ba18	184
HELMSJ235101.7-024425	2.0531	+23:51:01.7	-02:44:26.0	Na16	
HATLASJ235122.0-332902	2.236	23:51:21.7	-33:29:00.4	Ba18	159A
HATLASJ235122.0-332902	2.235	23:51:22.3	-33:29:08.0	Ba18	159B
HATLASJ235324.7-331111	2.564	23:53:24.5	-33:11:11.8	Ba18	86
HELMSJ235331.7+031717	3.1420	+23:53:32.0	+03:17:18.0	Na16	

continued on next page

Table 1 – continued

Source	z_{spec}	RA (J2000)	Dec (J2000)	Ref	HerBS_ID
HATLASJ235623.1-354119	3.095	23:56:23.0	-35:41:19.5	Ba18	36
HATLASJ235827.6-323244	2.912	23:58:27.5	-32:32:44.8	Ba18	25
HATLASJ001838.7-354133	1.536	00:18:38.8	-35:41:33.1	Ba18	104B
HATLASJ000455.4-330812	4.182	00:04:55.4	-33:08:12.8	Ba18	170
HATLASJ225324.2-323504	2.942	22:53:24.2	-32:35:04.2	Ba18	103
HATLASJ225339.1-325550	2.197	22:53:39.5	-32:55:52.3	Ba18	131B
HATLASJ231205.2-295027	2.473	23:12:05.3	-29:50:26.5	Ba18	132

Columns: (1) *Source*: IAU name from the discovery survey. (2) z_{spec} : spectroscopic redshift of the background source. (3) RA (J2000) and (4) Dec (J2000): sexagesimal coordinates, given with one decimal on the seconds. Coordinates preceded by a dagger (†) correspond to the original *Herschel* discovery positions. For sources observed at high resolution with ALMA, coordinates have been updated accordingly. (5) *Ref*: discovery reference – Ba18 = Bakx et al. (2018); Na17 = Negrello et al. (2017); Wa22 = Ward et al. (2022); Na16 = Nayyeri et al. (2016); W13 = Wardlow et al. (2013). Sources present in more than one catalogue are assigned to Bakx et al. (2018). (6) *HERBS_ID*: HerBS identifier for sources resolved into components by Bakx et al. [159] and Bendo et al. [174]. Only components with a measured spectroscopic redshift are listed.

2.4 WRAP-UP

This chapter has defined the sample upon which the remainder of this thesis is built, highlighting why ALMA and its scientific archive represent the ideal tools for this investigation. By applying a flux threshold to wide-area *Herschel* surveys, requiring a secure spectroscopic redshift, and selecting sources with public data in the ALMA archive, I assembled a final ensemble that encompasses both strongly lensed and unlensed populations.

Starting from this sample, one can search for the emission lines needed to characterise the properties of the interstellar gas. Many of these diagnostic transitions, however, are intrinsically faint. Despite the exceptional sensitivity of ALMA, their signal frequently lies below the noise floor of individual pointings, rendering single-object detections challenging. Recovering the average properties of the population therefore calls for statistical approaches such as stacking (Sec. 1.3.1), which combines the signal from numerous targets into a single high-fidelity measurement.

However, as I have emphasised throughout this chapter, these archival data are remarkably heterogeneous, having been collected across different projects, cycles, and array configurations, with disparate angular resolutions, spectral setups, and sensitivities. Stacking reconstructed ALMA images under such mismatched conditions introduces severe systematic uncertainties, since combining synthesised maps with different beam sizes produces an ill-defined effective

resolution and impairs flux accuracy. Overcoming these limitations demands a stacking methodology that circumvents image-plane combination by operating directly in the uv plane, where individual observations can be properly weighted and combined coherently, preserving the intrinsic spatial frequencies and the Gaussian noise statistics. To address this, the following chapter presents a dedicated methodology developed to tackle these challenges.

VISTA: A NEW STACKING METHOD IN THE FOURIER DOMAIN

Understanding the physical properties of DSFGs requires observations in the (sub-)millimetre regime (Sec. 1.2), where achieving the necessary spatial resolution relies on interferometers. These instruments sample the Fourier domain rather than directly retrieving sky images, which complicates the combination of independent datasets, especially when they are as heterogeneous in redshift and resolution as the archival sample assembled in Chapter 2. Recovering the average properties of these faint sources therefore requires a stacking approach able to circumvent the systematic uncertainties that arise when reconstructed images with mismatched properties and resolutions are combined. To this end, I developed VISTA, a methodology designed to operate natively in the visibility domain and to stack diverse interferometric datasets seamlessly. Before detailing the method itself, this chapter establishes the observational and theoretical background required to merge such datasets coherently.

Section 3.1 reviews the fundamentals of submillimetre interferometry, describing how an array samples the sky in the Fourier domain through the visibilities, and how baseline length, source structure, and noise shape the measured signal. Section 3.2 follows the data from the visibilities to the image domain, discusses the limitations inherent to this reconstruction, and outlines the advantages of performing the statistical combination prior to imaging. Section 3.3 details the core of the method, namely how it aligns the spatial and spectral properties of heterogeneous datasets onto a shared rest frame. Section 3.4 formalises the visibility averaging, together with the redshift and luminosity normalisations needed to homogenise the sample and with the treatment of the dust continuum. Section 3.5 presents the signal-extraction methods, both in the image and in the visibility domain.

3.1 INTERFEROMETRY AND THE VISIBILITY DOMAIN

Achieving the angular resolution required to resolve compact, high-redshift sources at (sub-)millimetre wavelengths exceeds the physical

limits of any single-dish telescope, a constraint that directly motivates the use of interferometers [e.g., 175] (Sec. 1.1.2). To illustrate this, consider the response of a single antenna. The power per unit frequency it collects from a source of intensity I_ν over the beam solid angle Ω , given an effective aperture A_e , is

$$P_\nu = \frac{1}{2} A_e I_\nu \Omega, \quad (9)$$

where the factor $1/2$ arises because an antenna feed responds to a single polarisation state, while the emission from an incoherent astronomical source is unpolarised and carries half of its power in each state. The angular response of the antenna, its *primary beam*, has a bright central lobe of Full Width at Half Maximum (FWHM)

$$\theta_{\text{PB}} \propto \frac{\lambda}{D}, \quad (10)$$

together with sidelobes inherent to the aperture's diffraction pattern (Fig. 13), whose amplitude is modulated by the non-uniform illumination. From equation (10), a 2.4 m optical telescope observing at $0.5 \mu\text{m}$ reaches $\sim 0.04''$, whereas matching that resolution at $\lambda = 1 \text{ mm}$ would require a dish some 5 km across, which cannot be built as a single structure. The primary beam determines the *field of view*, which is defined as the region of the sky from which emission is collected during a single pointing. For an array of identical antennas, the angular extent of this field of view corresponds simply to θ_{PB} , the beamwidth of a single element.

The way to get around this limitation is to replace the single dish with an array of smaller antennas whose signals are correlated in pairs. In the radio and (sub-)millimetre regime, the signal frequency is low enough to preserve the phase of the incoming radiation, and this allows all the correlated signals to be combined coherently into a single measurement. Through this technique, known as *aperture synthesis*, it is possible to reconstruct the angular resolution of a dish as large as the maximum separation between the antennas while retaining the collecting area and practicality of the individual smaller elements.

3.1.1 Principles of interferometry

The concept at the core of an interferometer is interference [175]: when two coherent wavefronts reach spatially separated receivers, the way they combine depends on their relative phase, a phenomenon first illustrated by Young's double-slit experiment. Here, a single wavefront illuminates two parallel, narrow slits, dividing the incoming light into two coherent secondary sources. As these waves propagate toward a downstream screen, they overlap and interfere with one another. Because the distances from each slit to

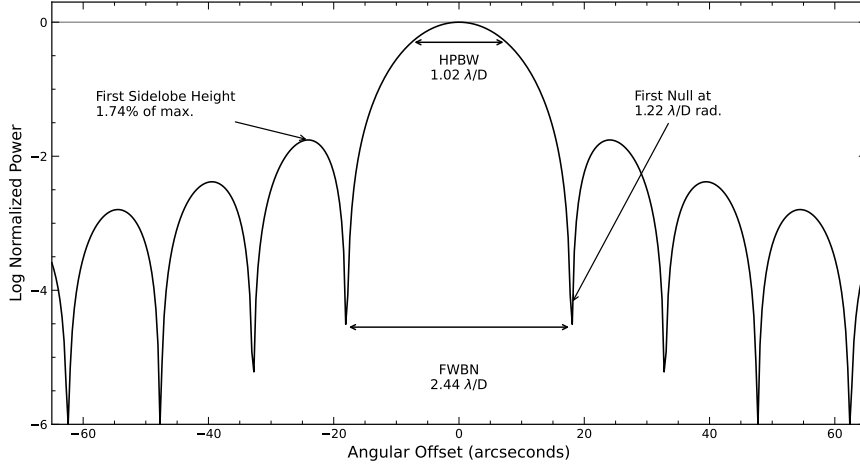


Figure 13: Response of a single antenna: the bright central *primary beam* and the surrounding sidelobes (e.g. a 12 m dish at a representative frequency). Adapted from Remijan et al. [172].

any given point on the screen differ slightly, the waves arrive with a relative phase shift. Where the path difference equals an integer number of wavelengths, they reinforce each other to create a bright fringe, whereas a half-integer path difference leads to destructive interference and a dark fringe.

A two-element radio interferometer reproduces this scheme in reception rather than in transmission. Instead of two slits illuminated by a distant coherent source and observed on a screen, there are two antennas A_1 and A_2 that *receive* the wavefront coming from a distant, incoherent astronomical source, and the "screen" is replaced by an electronic correlator that multiplies the two received signals. The separation between the slits becomes the *baseline* vector $\vec{B}$ connecting the two antennas. Just as the path difference between the two slits sets the local phase of the fringes on the screen, the path difference between the two antennas sets the phase difference between the two received signals.

Consider a source in the direction of the unit vector $\hat{s}$, making an incidence angle α with the baseline (Fig. 14). The wavefront reaches the two antennas at slightly different times because of the extra path $\vec{B} \cdot \hat{s}$ that one of the two must travel; this translates into a time lag, the *geometric delay*,

$$\tau_g = \frac{\vec{B} \cdot \hat{s}}{c}. \quad (11)$$

If the field at A_2 is $E_2(\nu) = E(\nu)$, the field at A_1 is the same signal, but phase-shifted by the amount corresponding to the geometric delay, $E_1(\nu) = E(\nu) e^{ik(\nu) \vec{B} \cdot \hat{s}}$, where $k(\nu) = 2\pi\nu/c$.

For a single frequency, the correlator forms the time-averaged product $R(\nu) = \langle E_1(\nu) E_2^*(\nu) \rangle$. For an extended source with brightness distribution $I(\vec{\sigma}, \nu)$ observed over a finite receiver bandwidth $\Delta\nu$, each sky element at an angular offset $\vec{\sigma}$ contributes to the fringe response across the passband. The total correlator response is obtained by integrating over both the celestial surface and the frequency bandpass $G(\nu)$:

$$R = \int_{\Delta\nu} G(\nu) \left[e^{i \frac{2\pi\nu}{c} \vec{B} \cdot \hat{s}} \int I(\vec{\sigma}, \nu) e^{i \frac{2\pi\nu}{c} \vec{B} \cdot \vec{\sigma}} d\vec{\sigma} \right] d\nu. \quad (12)$$

Under the *quasi-monochromatic approximation* (i.e., when $\Delta\nu \ll \nu_0$, where ν_0 is the central frequency), the source intensity $I(\vec{\sigma}, \nu) \approx I(\vec{\sigma}, \nu_0)$ can be assumed constant across the bandpass. Nevertheless, the frequency dependence in the exponent must be properly tracked, as bandwidth integration over ν introduces a delay-dependent attenuation factor (often represented by a sinc function for a square passband), leading to the well-known *bandwidth smearing* effect at large radial offsets $\vec{\sigma}$.

The leading factor $e^{ik\vec{B} \cdot \hat{s}}$ is called the *fringe-tracking* phase and it is simply the phase corresponding to the reference direction, compensated by the correlator and refined on calibrators (Sec. 3.1.4). It does not carry information about the source structure, unlike the integral where the physics of the source lives.

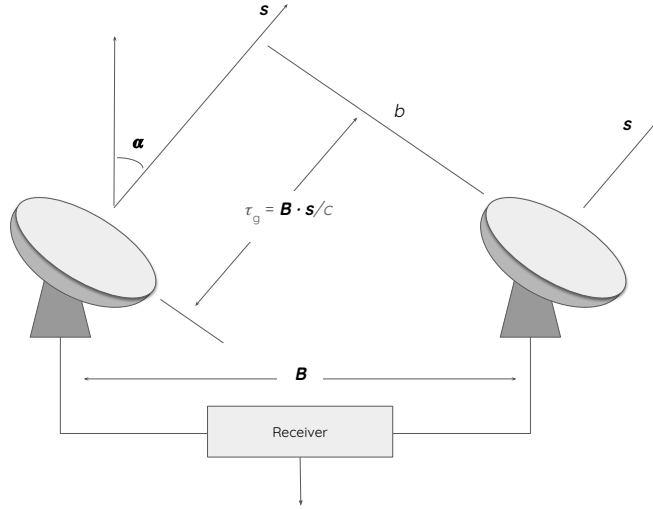


Figure 14: Geometric layout of a basic two-element radio interferometer observing a distant source along the direction unit vector $\mathbf{s}$. The antennas are separated by a baseline vector $\mathbf{B}$. Because the incoming wavefront arrives at an angle α relative to the zenith, it travels an additional path length b to reach the second dish. This path difference introduces a geometric time delay $\tau_g = \mathbf{B} \cdot \mathbf{s} / c$ between the received signals before they are combined at the central receiver.

3.1.2 The uv -plane and the Visibility function

In order to recover the brightness distribution I from Eq. (12), both the sky and the baseline are referred to a common Cartesian frame, decomposing each into a pair of orthogonal components. The source offset vector $\vec{\sigma}$ introduced above is decomposed as

$$\vec{\sigma} = x \hat{i} + y \hat{j}, \quad (13)$$

where $\hat{i}$ and $\hat{j}$ are unit vectors along the East–West and North–South directions, so that x and y are the two angular offsets of the sky element in that plane.

Regarding the baseline, the only component that determines the geometric delay is its projection onto the plane orthogonal to the line of sight towards the source. Equivalently, an observer located at the source would see the baseline foreshortened onto that plane. Expressed in units of the observing wavelength, this projected baseline splits into two orthogonal components: u , directed parallel to the celestial equator (East–West), and v , directed towards the North Celestial Pole (North–South). Both coordinates are therefore dimensionless, and the plane they span defines the uv -plane.

With these definitions the argument of the exponential in Eq. (12) becomes $k \vec{B} \cdot \vec{\sigma} = 2\pi(ux + vy)$. Dropping the fringe-tracking factor, which the correlator has already compensated, the response can be written as

$$V(u, v) = \iint I(x, y) e^{2\pi i(ux + vy)} dx dy \quad (14)$$

where $V(u, v)$ is the *visibility function*. It can also be expressed as a complex quantity defined by its modulus A and argument ϕ ,

$$V(u, v) = A e^{i\phi}. \quad (15)$$

The amplitude A measures the flux density contributed by the source on a given baseline, while the phase ϕ encodes the spatial position of the emission relative to the phase centre.

Thus, a pair of antennas does not measure the sky brightness $I(x, y)$ directly, but its two-dimensional Fourier transform, sampled at the single point (u, v) fixed by the projected baseline from the pointing position. Since $I(x, y)$ is a real quantity, each antenna pair actually delivers two samples at once per polarisation observed: the point (u, v) and its mirror $(-u, -v)$, related by $V(-u, -v) = V^*(u, v)$.

For an array composed of N elements, the $N(N - 1)/2$ distinct baselines instantaneously sample the uv -plane at $N(N - 1)$ visibility points. As the Earth rotates, the orientation of each projected baseline relative to the source drifts continuously, tracing out elliptical arcs over the course of the observation (Fig. 15). The number of antennas thus sets how many Fourier components are sampled at any given instant, while the duration of the observation governs how densely

the uv -plane is filled, as illustrated by the comparison between the instantaneous and the 1 h coverage in Fig. 15. The receiver capabilities and their combination with the correlator set the number of polarisations that can be instantaneously measured. The larger the number of antennas and the longer the integration time, the more faithfully the sky brightness can subsequently be reconstructed.

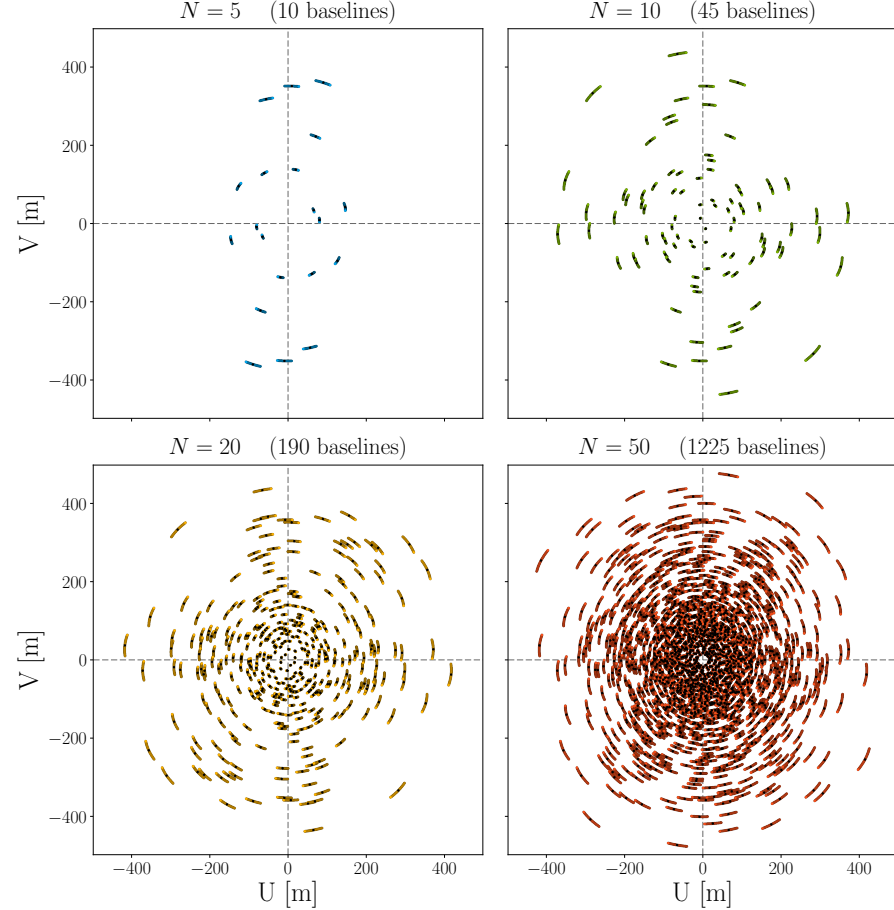


Figure 15: uv coverage of a mock source observed for 1 h around transit, with nested subsets of $N = 5, 10, 20$, and 50 antennas randomly drawn from the ALMA configuration C-3 of Cycle 11. Coloured points show the elliptical arcs traced by each baseline as the Earth rotates, sampled every 30 s, while black points mark the instantaneous sampling of a single integration at transit.

3.1.3 Visibility noise and baseline sensitivity

Every measured visibility $V(u, v)$ is contaminated by thermal noise originating from receiver electronics, atmospheric emission, and ground spillover. This noise is modelled as a zero-mean, complex Gaussian random variable that corrupts the real and imaginary

components independently. The precision of a single visibility measurement, namely the baseline *sensitivity*, is expressed in units of flux density (Jy) and quantified by its Root Mean Square (RMS) noise σ_V . It is governed by the single-baseline radiometer equation:

$$\sigma_V \approx \frac{2k_B T_{\text{sys}}}{\eta A \sqrt{n_p \Delta\nu \tau_{\text{acc}}}}, \quad (16)$$

where k_B is the Boltzmann constant, T_{sys} the system temperature, η the overall efficiency, A the dish collecting area, n_p the number of polarisations correlated, $\Delta\nu$ the channel bandwidth, and τ_{acc} the integration time per visibility record. Recording both polarisation states recovers the factor lost in Eq. (9), and improves the sensitivity by $\sqrt{2}$.

The resulting variance σ_V^2 defines the intrinsic statistical weight $w_i \propto 1/\sigma_{V,i}^2$ assigned to each sample in the uv plane, accounting for variations in system performance and atmospheric conditions across different baselines and time intervals.

3.1.4 Calibration errors

Unlike the random, additive thermal noise described by Eq. (16), calibration procedures (Section 2.2.2) introduce residual errors that are systematic and multiplicative. Each antenna i introduces a time- and frequency-dependent complex gain g_i , meaning the visibility measured on baseline ij is $V_{ij}^{\text{obs}} = g_i g_j^* V_{ij} + n_{ij}$. Calibration estimates these gains using reference sources, but the correction is imperfect due to atmospheric and instrumental fluctuations between calibrator and target scans. The resulting residual phase and amplitude errors are antenna-based, meaning they correlate across time, frequency channels, and baselines sharing a common antenna.

At submillimetre wavelengths, these phase errors originate primarily from tropospheric water vapour inhomogeneities. They are mitigated by water vapour radiometers and phase referencing, but residual fluctuations persist and they cause *decorrelation*. For an RMS σ_ϕ in radians, this effect reduces the coherently averaged visibility by the factor

$$\langle V \rangle = V e^{-\sigma_\phi^2/2}. \quad (17)$$

The signal loss affects the vector-averaged visibilities and their real part, whereas the scalar amplitude is not reduced but is positively biased by noise at low SNR [175]. Before calibration, the phase fluctuations grow with baseline length, since longer baselines sample larger tropospheric structures that vary on longer timescales. Phase referencing, however, corrects the fluctuations on timescales longer than the calibration cycle, so that the residual phase RMS increases with baseline length only up to a scale set by the cycle time, and remains approximately constant on longer baselines [176]. The residual decorrelation is therefore larger on the longer baselines and can make re-

solved sources appear artificially extended, while the shortest baselines are the least affected and preserve the total flux. Phase errors also scatter flux to incorrect positions, which raises the local noise and limits the dynamic range of bright sources.

Amplitude errors originate mainly from imperfect corrections of the atmospheric opacity and antenna response, and from uncertainties in the absolute flux scale. This scale is transferred to the target from calibrators with approximate and time-varying flux densities, and the resulting nominal accuracy ranges from a few to ten per cent depending on the observing band. Flux scale errors act as a common factor on all baselines so they directly affect the total flux, and they become critical whenever fluxes from independently calibrated observations or bands are compared.

Residual phase errors are commonly mitigated through *self-calibration*, where the science target itself is used as a calibrator. A model of the source is obtained by imaging the data, and antenna gains are iteratively re-derived by comparing observed visibilities with model predictions [177]. Since Cycle 10, the ALMA pipeline attempts self-calibration automatically on the continuum of science targets, and calibrated measurement sets contain these self-calibrated visibilities whenever the procedure succeeds. However, self-calibration requires a high SNR within each solution interval, and it is unviable for the intrinsically faint targets typically used for stacking.

In this work, calibration errors are therefore not modelled explicitly, and the visibilities are used as delivered by the ALMA archive (Sec. 2.2.3). For the faint targets that require stacking, thermal noise dominates over residual calibration errors. The total flux, the main quantity of interest, is set by the short baselines, which are the least affected by decorrelation. Moreover, since the datasets are calibrated independently, random calibration errors tend to average out during stacking rather than accumulate. In the case of the CO SLED of Chapter 6, however, brighter sources are also stacked together and calibration errors may become comparable to the thermal noise. For this reason, a calibration floor on the absolute flux scale is added in quadrature to the uncertainties of the stacked fluxes.

3.1.5 *Spatial scales and response to source structures*

Because the visibility is the Fourier transform of the sky brightness, each baseline corresponds to a distinct spatial frequency, and the range of baseline lengths present in an observation sets the range of angular scales to which the array responds. The longest baseline $B_{\max}$ samples the highest spatial frequency and fixes the angular resolution,

$$\theta_{\min} \simeq \frac{\lambda}{B_{\max}}, \quad (18)$$

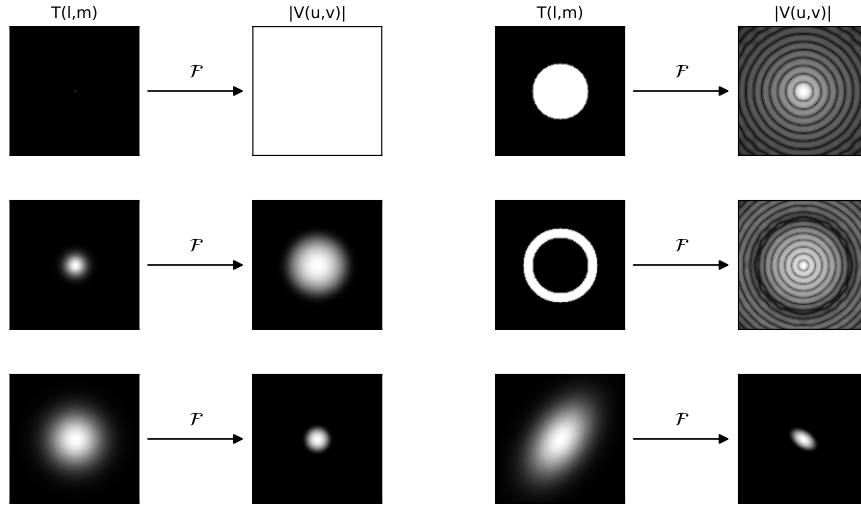


Figure 16: Model sky brightness distributions $T(l,m)$ (left of each pair) and the modulus of their visibilities $|V(u,v)|$ (right), related by the Fourier transform $\mathcal{F}$. *Left column, top to bottom:* a point source transforms into a flat, constant amplitude; a compact Gaussian into a broad one, and a broad Gaussian into a compact one, illustrating the reciprocity between source size and visibility width. *Right column, top to bottom:* a uniform disk into a Bessel function; a thin ring into oscillating Bessel-like patterns; an inclined elliptical Gaussian maps into an ellipse rotated by 90° .

the finest detail the array can distinguish. The shortest baseline $B_{\min}$, conversely, samples the lowest measured spatial frequency and sets the largest structure that can be recovered, the *maximum recoverable scale*,

$$\theta_{\max} \simeq \frac{\lambda}{B_{\min}}. \quad (19)$$

Emission on scales finer than $\theta_{\min}$ remains unresolved, whereas emission on scales coarser than $\theta_{\max}$ is progressively *resolved out*, because the corresponding low spatial frequencies fall within the unsampled central hole of the uv coverage. Extended structures are therefore attenuated and the total flux is not conserved. Only zero-spacing baselines would recover the total flux by capturing smooth, large-scale structures, whereas longer baselines increase the fringe density across the source to resolve fine details, at the cost of lower signal per fringe.

The visibility function encodes the response of the interferometer to the structure of the source, so that within this window of spatial frequencies the source morphology is directly imprinted on the functional form of $V(u, v)$. A few representative cases are shown in Fig. 16.

A point source is modelled as a Dirac delta distribution on the sky, yielding a flat Fourier transform. Its visibility modulus remains constant and equal to the total flux density across every baseline, while the phase simply records the offset from the phase centre. In contrast, an extended source behaves differently because emission originating from distinct regions arrives at a given baseline with different phases. The resulting constructive and destructive interference causes both amplitude and phase to vary from baseline to baseline. A Gaussian brightness distribution transforms into another Gaussian, causing its visibility amplitude to decay smoothly with baseline length. A uniform circular disk transforms into a first-order Bessel function, also known as a jinc profile, whose characteristic oscillations and nulls reveal the sharp edges of the disk. More intricate structures, such as gravitationally lensed arcs, generate multi-scale interference patterns across the uv -plane.

The main challenge in reconstructing the true sky brightness distribution of a source is that, despite combining baselines across a wide range of lengths, the coverage of the uv -plane remains inherently discrete and incomplete. Because the visibility function is sampled at only a finite set of locations, this discrete sampling fundamentally constrains which spatial structures can be recovered. Reconstructing a continuous image from such sparse Fourier measurements requires specialised inversion algorithms and image processing techniques. These reconstruction methods, along with their practical caveats and limitations, are discussed in detail in the next section.

3.2 FROM THE VISIBILITIES TO THE IMAGE

3.2.1 The dirty image and the sampling function

In an ideal scenario, the true sky brightness distribution $I(x, y)$ could be recovered directly through an inverse Fourier transform of the visibility function $V(u, v)$. As anticipated above, however, an interferometric array samples the Fourier domain only at a discrete and incomplete set of points. This limited baseline coverage is represented mathematically by a two-dimensional sampling function $S(u, v)$, which equals unity where a visibility measurement exists and zero elsewhere, so that the acquired data correspond to the product $S(u, v) V(u, v)$.

Applying an inverse Fourier transform to these sampled visibilities yields the observed map, commonly referred to as the dirty image $I_D(x, y)$. By the convolution theorem, this map equals the true sky brightness convolved with the Point Spread Function (PSF) of the array, known as the dirty beam $B(x, y)$:

$$I_D(x, y) = I(x, y) * B(x, y), \quad B(x, y) = \mathcal{F}^{-1}\{S(u, v)\}. \quad (20)$$

The dirty beam describes the response of the system to a point source. Because $S(u, v)$ contains unobserved gaps in uv -coverage, its Fourier transform exhibits prominent sidelobes that spread flux from bright individual sources across the surrounding field. Consequently, the dirty image is a distorted representation of the celestial sky and cannot be interpreted directly.

3.2.2 Weighting and tapering

To optimise the synthesised beam shape and map sensitivity, the binary sampling function $S(u, v)$ is generalised by assigning a weighting function $W(u, v)$ to the visibilities, modifying the dirty beam to $B(x, y) = \mathcal{F}^{-1}\{W(u, v)S(u, v)\}$. The choice of $W(u, v)$ establishes the fundamental trade-off between angular resolution and realised noise level.

Under *natural weighting*, the imaging weight $W(u, v)$ is set directly equal to the intrinsic statistical weight $w_i \propto 1/\sigma_{V,i}^2$ derived from the radiometer equation (16). This scheme allows densely sampled short baselines to dominate the transform, minimising overall image noise and maximising sensitivity to faint or extended emission, at the cost of a broader synthesised beam and higher sidelobes. Conversely, *uniform weighting* compensates for sampling density by down-weighting crowded regions of the uv -plane. This yields the sharpest angular resolution, but at the expense of significantly elevated noise levels. The *Briggs (robust)* scheme [178] provides a continuous interpolation

between these two limits via a tunable parameter, serving as the standard practical compromise in most imaging workflows.

A complementary operation is *tapering*, which applies an additional spatial frequency weighting envelope, typically Gaussian, in the uv -plane. Most commonly, an *outer taper* down-weights long baselines, suppressing high spatial frequencies to deliberately sacrifice angular resolution in favour of enhanced surface-brightness sensitivity to smooth, extended emission. Conversely, an *inverse* (or *inner*) taper can be applied to down-weight short baselines, filtering out large-scale diffuse structure and sharpening the angular resolution at the expense of sensitivity. Beyond filtering spatial scales, tapering also provides a versatile mechanism to match observations from different array configurations to a common synthesised beam size, a key requirement when combining heterogeneous images.

3.2.3 Deconvolution: CLEAN and data combination

While visibility weighting optimises the beam shape, it cannot eliminate sidelobes caused by unsampled uv regions. The Fourier inversion is mathematically underdetermined, as infinitely many sky brightness distributions can reproduce the measured visibilities while differing at unmeasured spatial frequencies. Unambiguous image reconstruction therefore requires a deconvolution step that incorporates prior physical assumptions about source morphology.

The standard method for this task is the CLEAN algorithm [179], which models the sky as a collection of point sources. Operating iteratively on the dirty image, it identifies the peak intensity, subtracts a scaled dirty beam centred at that location, and records the feature as a "clean" component. This cycle repeats on the residual image $R(x, y)$ until the remaining peaks merge into the noise floor σ . The accumulated components are then convolved with a Gaussian *restoring beam* $B_{\text{clean}}(x, y)$ and added back to the final residual map:

$$I_{\text{restored}}(x, y) = \sum_k A_k B_{\text{clean}}(x - x_k, y - y_k) + R(x, y), \quad (21)$$

where A_k and (x_k, y_k) denote the amplitudes and positions of the clean components.

The noise in the imaged data, namely the *point-source sensitivity*, follows from the radiometer equation Eq. (16) applied to an interferometer,

$$\sigma_S = \frac{w_r \sqrt{2} k_B T_{\text{sys}}}{A_{\text{eff}} (1 - f_s) \eta \sqrt{n_b n_p t_{\text{int}} \Delta \nu}}, \quad (22)$$

where $n_b = N(N - 1)/2$ is the number of baselines, t_{int} the total integration time, w_r a factor accounting for the adopted visibility weighting, f_s the fraction of data lost to flagging, and the remaining symbols carry the definitions of Eq. (16). The sensitivity thus improves as the

square root of the number of baselines, the integration time and the bandwidth.

The restored map represents a model-dependent, non-linear reconstruction rather than an exact inverse transform of the physical data. When faint astronomical targets fall below the thermal noise floor of the observation, individual detections become impossible. Given that telescope time is highly competitive and expensive, arbitrarily increasing integration times on a single target to reduce noise is rarely feasible. Recovering such sub-threshold emissions thus necessitates alternative strategies to systematically boost the SNR, paving the way for specialised signal-recovery techniques.

3.2.4 *Stacking before or after image reconstruction*

As noted in Section 3.2.3, faint sources frequently fall below the sensitivity threshold of an individual interferometric observation, while extending the integration time is practically constrained by telescope time allocations. This limitation is particularly severe for the cold gas emission lines discussed in Section 1.3, where the diagnostic transitions carrying the most physical information are often buried beneath the channel noise. Stacking offers a statistical way around this barrier by aggregating weak line or continuum signals across an ensemble of targets, yielding the $\sqrt{N}$ gain in sensitivity of Eq. (8).

Traditionally, this combination is performed in the image plane, on data that have already gone through the full reconstruction chain. The target fields are first imaged and deconvolved with non-linear algorithms such as CLEAN (Sec. 3.2.3), then spatial cutouts centred on the target coordinates are extracted and either coadded via pixel-wise statistics or photometrically measured over predefined apertures [e.g., 130, 180, 181]. Early image-stacking studies were explicitly tailored to homogeneous datasets, where observational setups, synthesised beam shapes, and noise properties remained consistent across the entire sample [e.g., 182].

Severe complications arise, however, when the same approach is applied to heterogeneous datasets collected across different projects or instrument configurations. Because different baseline distributions sample distinct spatial scales in the uv plane (Sec. 3.1.5), combining the resulting synthesised images at their native mismatched resolutions produces a stacked map governed by an ill-defined hybrid PSF, which prevents reliable flux recovery and physical interpretation. To avoid such distortions, all cutouts are typically convolved to a common angular resolution matching the largest synthesised beam in the sample [e.g., 183]. As discussed in Section 3.2.2, forcing a uniform resolution through spatial smoothing or tapering severely degrades the high-resolution data, increases the spatial noise correlations, and risks diluting or erasing compact signal components.

In addition, image-plane stacking inherits all the caveats of the reconstruction process described above. While the visibilities carry uncorrelated thermal noise (Sec. 3.1.3), the noise in deconvolved images (Sec. 3.2.1) becomes spatially correlated over the synthesised beam area and non-Gaussian, because of uncleaned sidelobes and incomplete uv coverage. Non-linear deconvolution algorithms can also introduce a systematic clean bias or suppress faint emission below the threshold (Sec. 3.2.3). Furthermore, centring the target cutouts requires subpixel shifting and regridding, which introduces interpolation noise and additional smoothing artefacts.

Performing the combination directly in the visibility domain circumvents these limitations. The thermal component of the visibility noise is additive, independent and complex Gaussian, as described by the radiometer equation (Sec. 3.1.3), whereas calibration errors are treated separately (Sec. 3.1.4). Working directly on the visibilities keeps the thermal noise in this simple form, which ensures unbiased error propagation and reliable statistical estimates. Visibilities from different baseline configurations are combined according to their intrinsic spatial frequencies (Sec. 3.1.5), preserving the high spatial resolution without any destructive smoothing. By bypassing non-linear image reconstruction entirely, this approach avoids clean bias, side-lobe contamination, and pixel interpolation errors (Sec. 3.2.3).

In practice, stacking in the uv domain requires that the phase centre of each field be shifted first, so that the target coordinates are aligned precisely to the origin of the plane (Sec. 3.1.2). Once all targets share a common spatial centre, the corresponding complex visibilities can be coadded through a weighted average across the overlapping regions of the uv plane.

Combining visibility datasets becomes far more delicate, however, when the targets span different cosmic epochs or are observed with different spectral configurations. Because redshift shifts the emission lines across observed frequencies and alters the angular diameter distance, sources at different redshifts map identical physical scales onto different uv distances in units of wavelength. Variations in the native channel widths introduce further mismatches in velocity resolution that complicate the spectral alignment of the lines.

Consequently, existing uv -plane stacking implementations have been predominantly restricted to local galaxy samples or to targets confined within a narrow redshift slice, where these baseline and frequency variations remain negligible [e.g., 183–185]. Extending the approach to heterogeneous multi-redshift samples instead requires a dynamic rescaling of the uv distances, so that identical physical dimensions are aligned across the sample, together with a dynamic channel regridding and velocity alignment to handle shifted lines and heterogeneous spectral resolutions. Conventional pipelines do

not account for these corrections, which motivates the framework presented in the following section.

3.3 THE VISTA METHOD

ViSTA (**Vi**sibility **St**acking) has been developed precisely to address these limitations. By mapping every observation onto a common, rest-frame, spectrally coherent uv plane, the framework enables the coherent combination of datasets spanning diverse cosmic epochs, native spectrometer configurations and array setups. The underlying principle relies on rescaling the physical spatial baselines by $(1+z)^{-1}$ and normalising the spatial frequencies by the rest-frame frequency ν_{rest} of the targeted transition.

The remainder of this section develops the theoretical foundations of the method, namely the construction of a common rest-frame uv domain, the alignment of each source to a shared phase centre with the associated primary-beam corrections, and the projection of heterogeneous spectra onto a unified frequency grid.

3.3.1 A Common Rest-Frame uv Domain

Consider a baseline vector $\mathbf{b} = (X, Y, Z)$ (Sec. 3.1.1), expressed in physical units of metres on the Earth’s surface. At an observed frequency ν_{obs} , the dimensionless baseline coordinates (u, v, w) , which measure spatial scales in units of the observing wavelength $\lambda_{\text{obs}} = c/\nu_{\text{obs}}$, are given by

$$(u, v, w) = \frac{\nu_{\text{obs}}}{c}(X, Y, Z). \quad (23)$$

To align an emission line with rest-frame frequency ν_{rest} , the corresponding observed frequency relates to the source redshift z via the standard relation $\nu_{\text{obs}} = \nu_{\text{rest}}/(1+z)$. Substituting this expression into the baseline coordinate equation reveals the fundamental geometric transformation defined in the ViSTA framework,

$$(u, v, w) = \frac{\nu_{\text{rest}}}{c(1+z)}(X, Y, Z) = \frac{1}{1+z}(u_0, v_0, w_0), \quad (24)$$

where $(u_0, v_0, w_0) = (\nu_{\text{rest}}/c)(X, Y, Z)$ represents the intrinsic spatial baseline coordinates in the rest frame of the target. Equation (24) establishes that cosmic expansion compresses the effective spatial frequency sampling by a factor of $(1+z)$.

Figure 17 illustrates this transformation for a single simulated source, comparing the uv -plane coverage in the observed frame with its rest-framed counterpart.

This transformation extends systematically across all frequency and geometric scales. Individual spectral channels of observed

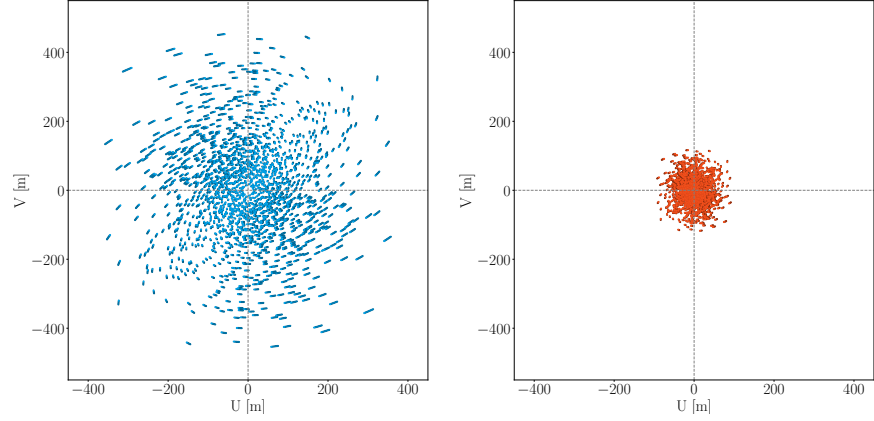


Figure 17: Visual example of the rest-framing of the uv -plane for a simulated source at redshift $z \simeq 2.9$, according to the rescaling of Equation (24) that contracts the sampled baselines toward the origin.

width $\Delta\nu_{\text{obs}}$ and total bandwidth B_{obs} map to rest-frame equivalents scaled by cosmic expansion, so that $\Delta\nu_{\text{rest}} = \Delta\nu_{\text{obs}}(1+z)$ and $B_{\text{rest}} = B_{\text{obs}}(1+z)$. In terms of Doppler velocity v , the native velocity channel width Δv is preserved identically between frames, since $\Delta v/c = \Delta\nu_{\text{obs}}/\nu_{\text{obs}} = \Delta\nu_{\text{rest}}/\nu_{\text{rest}}$. The same logic applies to the geometry of the array. To preserve the scale invariance of the synthesised beam profile across redshifts, the physical diameter D of an individual array dish undergoes an identical spatial contraction in the rest frame, $D_{\text{rest}} = D/(1+z)$. The native primary beam response, governed by the diffraction scale $\theta_{\text{PB}} \propto \lambda/D = c/(\nu_{\text{obs}}D)$, remains intrinsically invariant under this transformation, as $c/[\nu_{\text{rest}}(1+z)^{-1}D] = c/(\nu_{\text{rest}}D_{\text{rest}})$. Consequently, angular resolution elements scale consistently with physical baseline dimensions in spatial wavelength space across all sample redshifts.

3.3.2 Phase Centre Alignment and Primary Beam Correction

Before visibility datasets from heterogeneous fields can be coherently combined within the ViSTA framework, every target source must be brought to a common spatial reference. As shown in Section 3.1.2, the uv plane is centred on the pointing direction of the observation, known as the phase centre, and the position of a source is encoded in the phase ϕ of its visibilities (Eq. 15). Combining sources observed at different sky positions therefore requires that each target be relocated to the same point of the field before the datasets can be treated as repeated observations of a single object.

Once accurate coordinates of the target are available, the phase centre of every observation is shifted so that the source lies at the centre of the field, and its coordinates are set to zero, so that all sources in the

sample are treated as if they had been observed at the same position. Because this operation acts on the phase of the complex visibilities rather than on image-plane pixels, it avoids the interpolation noise and spatial covariance that regridding in the image domain would introduce, leaving the original amplitude information untouched.

Relocating the source to the field centre, however, does not by itself restore its true flux. A target that lay off-axis in its original observation had its emission attenuated by the primary beam response of the array elements, which falls off with the angular offset from the pointing centre. Recovering the intrinsic flux density therefore requires a *primary beam correction*, in which the visibilities of each dataset are multiplied by a single factor $1/\text{PB}(\theta_0)$, where θ_0 is the angular separation between the target and the original pointing centre, stored before the phase-centre shift. This places all sources on a consistent flux scale regardless of where they fell in their respective fields. The primary beam response of an individual array element can be approximated by a normalised symmetric Gaussian depending on the observing frequency ν and on the angular offset θ from the pointing centre,

$$\text{PB}(\theta, \nu) = \exp \left[-4 \ln 2 \left(\frac{\theta}{1.13 \lambda/D} \right)^2 \right], \quad (25)$$

where $\lambda = c/\nu$ is the observing wavelength and D the antenna diameter. Since the attenuation is a physical property of the observation, occurring before any rest-framing, the correction must be computed at the observed frequency of the line, $\nu_{\text{obs}} = \nu_{\text{rest}}/(1+z)$, with the physical dish diameter, and not at the rest-frame frequency. Because the attenuation must reflect the position the source actually occupied, the correction is applied at the original offset θ and not at the field centre, where the response would be unity and the flux would be underestimated for off-axis targets.

It is important to note that this is an image-plane prescription applied to the visibilities. The measured visibilities are the Fourier transform of the product $\text{PB}(x, y) I(x, y)$, so the attenuation acts in the uv plane as a convolution of the source visibilities with the Fourier transform of the primary beam, and a single multiplicative factor is a simplification. The scalar correction is evaluated at one position and one frequency, and it neglects the variation of the primary beam across the spectral window, so it is exact only for a point source. For an extended source, the response varies across the emission. For a circular Gaussian source of dispersion σ_s centred at θ_0 and a Gaussian primary beam of dispersion σ_{PB} , the ratio between the recovered and the true flux is

$$\mathcal{R} = \frac{\sigma_{\text{PB}}^2}{\sigma_{\text{PB}}^2 + \sigma_s^2} \exp \left[\frac{\theta_0^2 \sigma_s^2}{2 \sigma_{\text{PB}}^2 (\sigma_{\text{PB}}^2 + \sigma_s^2)} \right] \simeq 1 - \left(\frac{\sigma_s}{\sigma_{\text{PB}}} \right)^2 \left[1 - \frac{\theta_0^2}{2 \sigma_{\text{PB}}^2} \right], \quad (26)$$

and the error is of second order in $\sigma_s/\sigma_{\text{PB}}$ because the linear gradient of the primary beam across a symmetric source cancels out. The attenuated source also appears slightly smaller and it is displaced towards the pointing centre. This approximation degrades when the source extent becomes comparable with its offset. For a source centred where $\text{PB} = 0.8$, for instance, the scalar correction recovers $\sim 94\%$ of the flux when the source FWHM equals its offset, and $\sim 80\%$ when it is twice as large. In the regime analysed in this thesis, targets are close to the phase centre and compact with respect to the primary beam, so the error remains below 1% for simulated sources and reaches at most a few per cent for compact Band 7 sources, staying below the absolute flux calibration uncertainty (Sec. 3.1.4). Larger offsets and asymmetric or lensed morphologies, for which the gradients of the primary beam across the source do not cancel, require a more precise treatment, in which the primary beam is included in the forward modelling of the visibilities through its convolution kernel. This will be implemented in future versions of the code.

3.3.3 Spectral Alignment onto a Common Grid

Once each source has been centred and rest-framed, the visibilities of a multi-redshift sample still carry heterogeneous native channel widths, spectral resolutions, and observed bandpasses, which must be reconciled onto a single spectral axis before the datasets can be merged. Cosmic expansion contributes to this mismatch, since the rest-frame channel spacing of each dataset scales as $\Delta\nu_{i,\text{rest}} = \Delta\nu_{i,\text{obs}}(1 + z_i)$ and therefore differs from source to source. Directly combining visibilities across unaligned channels would misregister the spectral features and corrupt the kinematics of the stacked emission line. The alignment proceeds in two stages: a rebinning of each dataset to a common resolution followed by a regridding onto a common channelisation.

First, the rebinning operation brings every dataset to a shared spectral resolution and channel width. The resolution of each observation is adjusted to match the widest, coarsest resolution present in the sample, which prevents the finer datasets from imposing a channelisation that the coarser ones cannot support and so avoids introducing artifacts when the data are later combined.

Regridding then standardises the datasets so that they share the same frequency range and the same number of channels, which is a prerequisite for concatenating them into a single stackable cube. A common frequency range can be defined from the lowest and highest frequencies observed across the sample, and a reference channelisation is taken from the dataset with the largest number of channels. Each dataset is then interpolated onto the target channelisation. A

target frequency ν_j falling between two native channel centres ν_k and ν_{k+1} is assigned the linearly interpolated visibility

$$V(\nu_j) = (1 - f) V(\nu_k) + f V(\nu_{k+1}), \quad f = \frac{\nu_j - \nu_k}{\nu_{k+1} - \nu_k}, \quad (27)$$

where $f \in [0, 1]$ is the fractional distance of ν_j between the two bracketing channels. The interpolated visibility is thus a weighted average of the two neighbours, reducing to $V(\nu_k)$ when $f = 0$ and to $V(\nu_{k+1})$ when $f = 1$.

The noise associated with each visibility must be rescaled consistently with this operation. Following the radiometer equation (16), the per-channel sensitivity scales as $\sigma_V \propto (\Delta\nu)^{-1/2}$, so that widening the spectral window onto coarser channels lowers the noise per channel, while the reverse holds when the grid is finer than the native one. Averaging channels reduces the noise of the data themselves, but the uncertainties stored with each visibility must be updated accordingly, so that the statistical weights used in the combination remain consistent with the resampled data. In practice, when n native channels are averaged into one, the noise becomes $\sigma_V / \sqrt{n}$ and the weight $w \propto \sigma_V^{-2}$ is multiplied by n . The interpolation of Eq. (27), moreover, introduces correlations between adjacent output channels, since each native channel contributes to more than one of them.

3.4 STATISTICAL COMBINATION OF VISIBILITIES

Once every dataset has been brought onto the common rest-frame uv plane and the shared spectral grid, the individual observations can be treated as repeated measurements of a single underlying source and combined into a single statistical estimate. This combination relies on a weighted average of the aligned visibilities, but the choice of the weighting scheme is not purely technical, since it depends directly on the scientific objective, particularly when dealing with heterogeneous archive data. Depending on whether the goal is to maximise the formal SNR or to recover the unbiased average behaviour of a population, different strategies must be applied. Weighting alone, however, does not render the contributions of different sources physically comparable, because objects at different distances and with different intrinsic luminosities enter the average with amplitudes that are not directly commensurate. Addressing this requires proper flux normalisations to place all targets on a common scale, alongside a careful treatment of the dust continuum underlying the spectral lines of interest and a careful choice of where to remove it within the workflow.

3.4.1 Natural and democratic weighting

The natural weight of a visibility is defined as the inverse of its variance, because a measurement with RMS noise σ_V (Sec. 3.1.3) carries an information content proportional to σ_V^{-2} . Forming the weighted mean with $w_k = 1/\sigma_k^2$ (Eq. 6) yields the minimum-variance estimate of the combined signal, $\bar{V}$, maximising the overall SNR. This is known as *natural* weighting because it uses the intrinsic statistical weights directly, and it mirrors the imaging scheme described in Section 3.2.2, where setting the imaging weight $W(u, v)$ equal to the radiometric weight $w_i \propto 1/\sigma_{V,i}^2$ minimises the map noise. Although this choice is optimal for faint signals or for homogeneous samples, where all visibilities sample the same intrinsic emission, it becomes problematic for heterogeneous datasets. Indeed, since the radiometer equation (16) implies that $\sigma_V \propto \tau_{\text{acc}}^{-1/2}$, the individual weights scale linearly with integration time as $w \propto \tau_{\text{acc}}$. Consequently, a target observed a hundred times longer than the others accumulates a hundred times more weight, which biases the combined estimate toward its own emission and lets a few deeply observed sources dominate the stack. In an archive spanning broad ranges of integration times, angular resolutions, and sensitivities (Sec. 2.2.3), this unevenness produces a combined estimate that boasts the highest formal SNR but fails to represent the target population as a whole.

When the primary scientific objective shifts toward uncovering the average behaviour of the ensemble, sources must instead be weighted *democratically*, i.e. with equal total weight per source, so that every object contributes equally regardless of its observational depth. For any given source i , the internal visibility weights are rescaled to satisfy

$$\sum_{k \in i} w_k = 1, \quad (28)$$

which ensures that each object enters the average on equal footing. This is not a unity weighting of the individual visibilities. Within each source, the relative natural weights $w_k \propto \sigma_k^{-2}$ are preserved, and only their sum is normalised, so that each source, rather than each visibility, carries the same weight. This scheme achieves population representativeness at the expense of formal sensitivity, because deep observations are no longer permitted to drive the result. Such a trade-off is acceptable when the targeted lines are bright enough to remain well defined at lower SNR regimes, whereas natural weighting remains preferable for extremely faint emission that is undetectable in the individual sources.

Whether sources detected individually should be included in stacking depends entirely on the scientific objective. When the analysis aims to determine the average properties of the full sample, including them is essential, as their exclusion would bias the mean downward. In such cases, the democratic weighting of Eq. (28)

prevents bright objects from dominating the resulting measurement. Conversely, when the primary objective is to recover a faint signal from a noise-dominated population, excluding individual detections is preferred. The stack then accurately represents the undetected sources rather than being skewed by high-significance objects, which can instead be measured independently and combined a posteriori if required.

A further nuance arises when individual sources in the sample have been observed across multiple epochs or configurations. The equal-weight condition of Eq. (28) applies strictly among distinct physical sources, which are the independent units of the population average. When a single source has multiple observations, treating each dataset as an independent target would give that object a disproportionate weight in the ensemble average. To prevent this, repeated observations of the same source are first combined internally using natural weighting, which preserves the optimal intra-source sensitivity, and the resulting single dataset then enters the stack with its total contribution normalised to unity.

3.4.2 Flux normalisation

Weighting fixes how much each source contributes to the stack, but sometimes this is not enough to make the contributions physically comparable. Two further sources of heterogeneity affect the amplitudes themselves and must be corrected before the visibilities of different objects are averaged together.

The first correction addresses the redshift. A line of fixed intrinsic luminosity produces an observed flux that depends on the distance to the source, so that in a sample spanning a broad redshift range the same emission appears with amplitudes differing by nearly an order of magnitude, purely because of the different luminosity distances. Averaging the visibilities in this state would blend together fluxes that are not physically comparable, and the stacked amplitude would carry no well-defined luminosity. Each source must therefore be transported to a common reference redshift z_{ref} before the combination. The visibility amplitudes of a source at redshift z_i are multiplied by

$$\alpha_i = \left[\frac{D_L(z_i)}{D_L(z_{\text{ref}})} \right]^2 \frac{1 + z_{\text{ref}}}{1 + z_i}, \quad (29)$$

where $D_L(z)$ is the luminosity distance and the factor $(1 + z_{\text{ref}})/(1 + z_i)$ accounts for the bandwidth compression between the two frames. This converts the velocity-integrated flux of each source into the flux it would have if the source were placed at z_{ref} , so that the stacked amplitude corresponds to a well-defined line luminosity. Adopting a single common reference redshift is the natural choice for a popula-

tion study, since it places all objects at a representative distance for the sample, typically its median redshift.

A second correction accounts for intrinsic luminosity differences. Even at a common redshift, an intrinsically brighter object contributes proportionally more flux, allowing a bright minority to skew the stack. Rescaling the amplitudes by the ratio of a common reference luminosity to the intrinsic luminosity of each source removes this degeneracy and places all objects on an equivalent footing, so that the stacked profile reflects the average intrinsic behaviour of the population rather than its brightest members. Ideally, the reference quantity traces the stacked emission directly, but such a measurement is rarely available for an entire archival sample, and one must then rely on a proxy, such as a continuum flux density or an integrated luminosity known to correlate with the line under study.

Both corrections are optional, and whether to apply them depends on the scientific question being asked. The redshift rescaling has little effect over a narrow redshift range, but becomes essential over a broad one to obtain a well-defined stacked luminosity. The luminosity normalisation can instead be omitted when the goal is the total emission of the sample weighted by the true intrinsic luminosities rather than the average spectral shape, or when no reliable proxy is available for the whole sample. Comparing the normalised and non-normalised stacks is itself informative, since any discrepancy between them is a diagnostic of the underlying luminosity distribution and of possible selection effects.

3.4.3 *Continuum subtraction*

In the (sub-)millimetre regime, the spectral lines emitted by these objects often sit on top of dust continuum emission (Sec. 1.1.2). This continuum is a broadband signal present across all channels, so the stacking operation combines both components at once (Eq. 6). Deciding whether to fit the two together or to treat them separately depends on the specific scientific case. The continuum level could be significantly weaker than the expected line emission for all sources in the sample, lying well below the individual noise floor. In this case, the continuum can be preserved during combination. Full visibilities are stacked together first, and the continuum is fitted and subtracted from the combined dataset. This approach is not restricted to simulated data and applies equally to real samples, provided that the continuum is faint and homogeneous across the sample and that no source is bright enough to distort the stacked baseline. On the other hand, sometimes the continuum can dominate the line emission, or a few sources in the sample have a much brighter continuum than the rest. Such bright continua distort the baseline of the stacked line profile. In this situation, the continuum must be removed from each

source prior to line stacking, and the per-source continuum models can be stacked in a separate workflow. Both situations occur in the real datasets analysed in this thesis, and the strategy adopted in each case is specified in the following chapters.

The choice of the fitting function for removing the continuum depends on the spectral bandwidth. When fitting locally near the line profile, across narrow spectral windows, a low-order polynomial of degree 0 or 1 is sufficient. For broadband observations spanning the thermal dust emission, however, the continuum flux density follows a MBB spectrum (Eq. 5). In the Rayleigh–Jeans limit it scales as $\nu^{2+\beta}$, which behaves approximately as ν^4 for a typical dust emissivity index $\beta \approx 2$, and the fitting function must account for this steep dependence to prevent baseline residuals.

3.5 SIGNAL EXTRACTION

The stacking formalism described so far combines the visibilities of many sources into a single averaged dataset. The final step consists of recovering the stacked signal, which can be achieved either by imaging the stacked visibilities or by fitting a model directly in the uv plane. In both approaches, the morphology of the stacked source dictates how the flux is extracted.

3.5.1 *Extraction in the image domain*

One way to recover the stacked flux is to image the combined visibilities and deconvolve them once as a single dataset. As discussed in Section 3.2.4, imaging the stack rather than stacking images benefits from the denser and more uniform uv coverage built up by combining sources observed in different configurations, and this ensures that the reconstruction rests on a better-sampled Fourier plane than any individual observation provides. The procedure nonetheless inherits the general caveats of image reconstruction, and the flux measured from the stacked map remains tied to the choice of weighting and restoring beam.

The extraction of the stacked flux depends directly on the morphology of the source. In the simplest case, where the entire signal is contained within the synthesised beam and the stacked source remains unresolved, the flux density is given by the peak intensity because all emission falls within a single resolution element. When the angular extent approaches or exceeds the beam size, the peak intensity no longer captures the total flux and the emission must instead be integrated over an extended region, and this can be done either using a fixed aperture large enough to encompass it or by fitting a two-dimensional Gaussian to the restored image, which simultaneously yields an estimate of the source size. The restoring beam itself

is determined in the exact same manner by fitting a two-dimensional Gaussian to the central peak of the dirty beam, and this establishes the effective resolution against which any measured size must be compared.

As introduced in Section 3.2.2, a further complication arises when the sample contains strongly resolved sources whose emission is spread over many long baselines. Merging these directly with lower-resolution data combines incompatible spatial frequency content and distorts the stacked image. While tapering can bring such observations to a common effective resolution, deconvolving data spanning extreme resolution gaps remains challenging because of the large unsampled regions left in the Fourier plane. Moreover, an aggressive taper can suppress long baselines to the point where the contribution of high-resolution sources is almost entirely filtered out of the stack.

3.5.2 *Extraction in the visibility domain*

An alternative method is to bypass image reconstruction entirely and fit a parametric source model directly to the stacked visibilities. This approach is naturally suited for heterogeneous samples because it avoids the deconvolution steps and restoring beams that make image-domain extraction delicate.

As for the image-domain extraction described above, the source morphology enters through the choice of model, since different spatial regimes correspond to distinct visibility amplitude shapes as a function of baseline length (Sec. 3.1.5). When every source in the sample is unresolved, the visibility amplitude remains flat and independent of baseline length, so the stacked signal can be modelled as a constant, any baseline range provides an equally valid measurement of the same flux, and observations of different resolutions all sample the identical signal.

For a resolved source, a simple Gaussian model remains adequate even when the true brightness distribution is complex. This is because the total flux is governed by the shortest baselines, which sample the largest angular scales and thus capture the integrated emission. Substructure appearing at longer baselines, such as the Bessel rings of a sharp-edged disk or the fringes of a lensed arc, encodes the detailed morphology rather than the total flux. Since the goal is to recover the integrated emission, fitting a Gaussian profile to the visibility amplitudes yields both the total flux, through the zero-baseline amplitude, and an effective size describing the overall source extent, without needing to reproduce the internal structure.

Fitting models directly to the measured Fourier components therefore ensures a robust and unbiased extraction of the integrated signal. By bypassing image reconstruction entirely, this approach avoids

image-domain artefacts and restoring-beam dependencies, offering a reliable framework for analysing complex and heterogeneous datasets.

3.6 WRAP-UP

This chapter introduced the ViSTA method and outlined the core principles of visibility domain stacking. Submillimetre interferometers sample the Fourier transform of sky brightness, where baseline lengths correspond to specific spatial scales. Transforming these measurements into images via CLEAN introduces correlated noise and reconstruction artifacts. Stacking overcomes these detection limits, delivering a theoretical $\sqrt{N}$ sensitivity gain by combining multiple faint sources. Performing this combination directly in the visibility domain preserves the Gaussian statistics of the thermal noise while avoiding image plane biases. Ultimately, ViSTA solves the challenge of stacking heterogeneous datasets across different redshifts by remapping visibilities onto a shared rest frame uv plane, aligning phase centres, and unifying spectral channels before signal extraction. The combination itself can then be tuned to the scientific goal, weighting the sources either by their intrinsic sensitivity or democratically, normalising their amplitudes to a common redshift and luminosity, and recovering the final signal either by imaging the combined dataset or by fitting a model directly to the visibilities.

With these foundations in place, the next chapter moves from theory to practice by presenting a proof-of-concept implementation tailored to analyse ALMA data. Built upon the native software of the observatory and third-party astronomical libraries, it validates the core mechanics of the method using both simulated and real datasets.

ViSTA IMPLEMENTATION FOR STACKING ALMA DATA

While the theoretical formalism established in the preceding chapter defines the objectives and the mathematical framework of interferometric stacking, it does not address the practical execution required to process real observational data. In this chapter I present the first version of ViSTA, a reference implementation tailored to work with ALMA data and built on its standard software infrastructure.

Section 4.1 describes how ALMA data are formatted and calibrated, and introduces the data structure that constitutes the input to the pipeline.¹ Section 4.2 details the stacking pipeline itself, following each theoretical step of the method described in Chapter 3 as it is translated into a concrete sequence of operations on the data. Section 4.3 validates the implementation on simulated mock observations, while Section 4.4 applies it to a small sample of real ALMA data.

4.1 ALMA DATA FORMAT AND THE CASA SOFTWARE

As described in Section 2.2.2, the correlator delivers the complex visibilities resolved in time across all baselines, spectral channels and polarisations. Applying the calibration solutions derived from the interleaved calibrator scans converts them into calibrated visibilities. The archive stores the raw data together with the calibration products, from which the calibrated datasets can be restored. In this work, the calibrated datasets were obtained through the ALMA Helpdesk, which provides them upon request after processing with the ALMA pipeline, or, for the early cycles, from the reprocessing of the ARI-L project whenever available [186].

The reduction, calibration, and imaging of ALMA data are carried out within Common Astronomy Software Applications (CASA), the reference environment for ALMA and VLA data developed jointly by NRAO, ESO, and NAOJ [187, 188]. Its performance-critical routines are implemented in a C++ core, *casacore*, exposed through a Python

¹ Throughout this thesis, the term *pipeline* refers to the ViSTA processing workflow, while the automated calibration and imaging pipeline of the observatory is always referred to explicitly as the ALMA pipeline.

interface that provides both individual processing tasks and higher-level automated pipelines. This section describes this environment and the format in which the data are stored, both of which constitute the starting point of the pipeline.

4.1.1 The Measurement Set format

Raw data reach the archive in the ALMA Science Data Model (ASDM) format and are converted, through the `importasdm` task, into a Measurement Set (MS), the standard directory-based binary table format adopted by `casacore` [189]. The MS is not a single file but a structured collection of tables that jointly describe the visibilities and all the metadata required to interpret them, and its content can be inspected with the `casatasks.browsetable()` task (Fig. 18).

Table Browser <1>

File

Edit

View

Tools

Export

Help

About

uid_A002_Xc3ecf6_X126c.ms

table data

table keywords

field keywords

	UVW	FLAG	_AG_CATEGOF	WEIGHT	SIGMA	ANTENNA1	ANTENNA2	ARRAY_ID	DATA_DESCRIPTOR
0	[0, 0, 0]	[2, 1] Boolean	[0, 0, 0] Boolean	[3.2e+07, 3....	[0.0001767...	0	0	0	0
1	[0, 0, 0]	[2, 1] Boolean	[0, 0, 0] Boolean	[3.2e+07, 3....	[0.0001767...	1	1	0	0
2	[0, 0, 0]	[2, 1] Boolean	[0, 0, 0] Boolean	[3.2e+07, 3....	[0.0001767...	2	2	0	0
3	[0, 0, 0]	[2, 1] Boolean	[0, 0, 0] Boolean	[3.2e+07, 3....	[0.0001767...	3	3	0	0
4	[0, 0, 0]	[2, 1] Boolean	[0, 0, 0] Boolean	[3.2e+07, 3....	[0.0001767...	4	4	0	0
5	[0, 0, 0]	[2, 1] Boolean	[0, 0, 0] Boolean	[3.2e+07, 3....	[0.0001767...	5	5	0	0
6	[0, 0, 0]	[2, 1] Boolean	[0, 0, 0] Boolean	[3.2e+07, 3....	[0.0001767...	6	6	0	0
7	[0, 0, 0]	[2, 1] Boolean	[0, 0, 0] Boolean	[3.2e+07, 3....	[0.0001767...	7	7	0	0
8	[0, 0, 0]	[2, 1] Boolean	[0, 0, 0] Boolean	[3.2e+07, 3....	[0.0001767...	8	8	0	0
9	[0, 0, 0]	[2, 1] Boolean	[0, 0, 0] Boolean	[3.2e+07, 3....	[0.0001767...	9	9	0	0
10	[0, 0, 0]	[2, 1] Boolean	[0, 0, 0] Boolean	[3.2e+07, 3....	[0.0001767...	0	0	0	0
11	[0, 0, 0]	[2, 1] Boolean	[0, 0, 0] Boolean	[3.2e+07, 3....	[0.0001767...	1	1	0	0

Restore Columns

Resize Headers

PAGE NAVIGATION

First

<<

[1 / 1392]

>>

Last

1

Go

Loading

10000

rows

Figure 18: Main table of an ALMA MS as displayed by `casatasks.browsetable()`. Each row stores the visibilities of all channels and polarisations for a pair of antennas, identified by `ANTENNA1` and `ANTENNA2`, at a given time, together with the associated `UVW`, `FLAG`, `WEIGHT` and `SIGMA` columns. Adapted from the CASA documentation².

The visibility measurements themselves are stored in the main table, where each row is identified by the pair of antennas forming the baseline, stored in the separate columns `ANTENNA1` and `ANTENNA2`, by the field, the timestamp, and the spectral setup. Its principal columns separate the different processing states and the ancillary quantities of each visibility: `DATA` holds the raw visibilities, `CORRECTED_DATA` the calibrated ones produced once the calibration solutions have been

² <https://casadocs.readthedocs.io/en/v6.2.0/api/tt/casatasks.information.browsetable.html>

applied, and `MODEL_DATA` the Fourier transform of any sky model associated with the field. Alongside these, `WEIGHT` stores the statistical weight of each visibility and `SIGMA` its associated uncertainty, `FLAG` marks the samples to be excluded, `UVW` records the baseline coordinates in the Fourier plane, and `TIME` and `OBSERVATION_ID` identify the integration and the execution each sample belongs to. This separation between raw, calibrated, and model columns allows successive processing steps to be applied without destroying the original measurements, a property that ViSTA exploits throughout the pipeline.

The main table is complemented by a set of subtables that store the observational metadata, which can be accessed from the *table keywords* tab of the same browser. The `ANTENNA` subtable lists the array element positions and dish diameters, the `FIELD` subtable records the phase tracking centres of each pointing through its `PHASE_DIR`, `REFERENCE_DIR`, and `DELAY_DIR` entries, `SPECTRAL_WINDOW` describes the frequency setup of each window, `POLARIZATION` specifies the correlation products, and `DATA_DESCRIPTION` links each spectral window to its polarisation setup. Further subtables, including `SOURCE`, `FEED`, `POINTING`, and `OBSERVATION`, trace the properties of the observed sources, the receiver feeds, the antenna pointing directions, and the individual executions, so that the full observational context of every visibility can be reconstructed.

4.2 VISTA CASA-BASED PIPELINE

Taking the calibrated MSs described above as input, this section shows how each theoretical step of the ViSTA method described in Chapter 3 is implemented in practice. The pipeline is illustrated in Figure 19 and is structured into three main phases. It starts with the pre-processing operations that prepare the individual datasets to be combinable in the Fourier domain, continues with the core stacking procedure that combines the visibility measurements, and ends with an optional post-processing phase, covering operations such as normalisation and continuum subtraction, that leads to the extraction of the stacked flux.

The operational sequence combines standard CASA tasks with custom Python routines, the latter modifying the MS tables directly whenever CASA lacks a dedicated functionality. These routines rely on established scientific libraries, such as `Astropy`, `NumPy`, and `SciPy`, to perform the numerical calculations and the coordinate transformations. The workflow described in the following paragraphs is a general framework, applicable to data from other interferometric facilities by substituting the software tools while preserving the core processing logic.

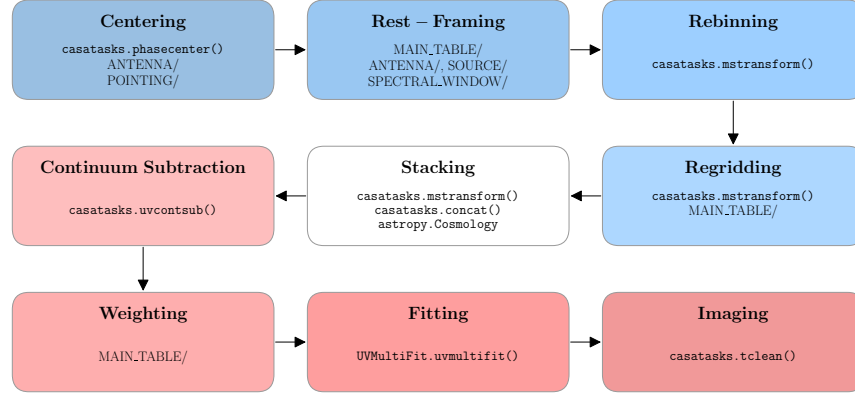


Figure 19: General workflow of the ViSTA method. Specific implementations for ALMA MSs are also provided, detailing the relevant functions from CASA or external astronomical libraries alongside manually modified data tables. Blue cells denote the input data preparation phase, whereas red cells represent the output extraction steps.

4.2.1 Pre-processing

4.2.1.1 Centering

To align all target sources to a common reference position and enable coherent visibility stacking (Sec. 3.3.2), each observation undergoes a phase centering process. Using the `casatasks.phaseshift()` function, a phase correction based on the target coordinates is applied to each input MS, hereafter referred to as *file.ms*. This operation outputs a new dataset, *file.ms.centre*, where the source is positioned at the field centre, preserving the original data file untouched.

The phase rotation is the only operation performed by `casatasks.phaseshift()`. Following the spatial shift, the observational metadata are modified by custom Python routines of the ViSTA workflow, acting directly within the FIELD, SOURCE, and POINTING subtables to set all source coordinates to zero. In particular, the PHASE_DIR, REFERENCE_DIR, and DELAY_DIR entries in the FIELD subtable, alongside the pointing directions in the POINTING subtable, are explicitly zeroed. Simulating a shared origin at coordinate zero ensures full structural consistency, allowing CASA to treat all targets as co-spatial during subsequent concatenation.

Relocating the source to the field centre does not by itself restore its intrinsic flux, since a target observed off-axis had its emission attenuated by the primary beam response of the array elements. Recovering the intrinsic flux density therefore requires multiplying the visibili-

ties of each dataset by a single factor, the inverse of the attenuation of Equation 25 evaluated at the original offset of the source, within the scalar approximation discussed in Section 3.3.2. This correction, however, acts on the visibility amplitudes rather than on the meta-data, and it is therefore applied at the stacking stage, as described in Section 4.2.2.

4.2.1.2 Restframing

As discussed in Section 3.3.1, in order to shift visibility measurements into the source rest frame, observed frequencies and their associated parameters must be transformed using the redshift factor $1 + z$. Within the MS, frequency dependent parameters are stored in the SPECTRAL_WINDOW subtable. To apply these corrections, a new dataset named *file.ms.centre.rest* is created, where channel central frequencies in CHAN_FREQUENCY, reference frequencies in REF_FREQUENCY, individual channel widths in CHAN_WIDTH, effective noise bandwidths in EFFECTIVE_BW, channel resolutions in RESOLUTION, and total bandwidths in TOTAL_BANDWIDTH are scaled accordingly.

Because spatial coordinates and physical antenna dimensions depend on wavelength, additional corrections are applied to maintain physical consistency in the rest frame. The spatial coordinates UVW in the main table, alongside antenna coordinates in the POSITION column of the ANTENNA subtable, are divided by $1 + z$. Furthermore, the antenna dish diameter DISH_DIAMETER is rescaled by the same $1 + z$ factor to preserve the primary beam, since $\theta_{PB} \propto \lambda/D$. Because the UVW coordinates in metres are divided by $1 + z$ while the frequencies are multiplied by the same factor, the uv coordinates in units of wavelength are unchanged. An image of a single dataset made before and after rest-framing is therefore identical in angular size and shape, and only the labelling of the frequency axis changes, now expressed in the rest frame of the source.

4.2.1.3 Rebinning

Prior to combining multiple observations, spectral consistency across all MSs must be ensured by matching their spectral resolution and channel widths, allowing spectral features to align accurately. The initial resolution adjustment is performed using `casatasks.mstransform()`, which rebins each observation to match the broadest channel width in the sample, outputting *file.ms.centre.rest.rebin* and avoiding spectral artifacts during data combination.

During this rebinning process, the noise properties of the dataset are updated directly in the MS. As described in Section 3.3.3, widening the channels by a given binning factor lowers the per-channel noise according to the radiometer equation, where sensitivity scales

inversely with the square root of the channel width. Consequently, `casatasks.mstransform()` automatically updates the `SIGMA` and `WEIGHT` columns in the output dataset, scaling the visibility uncertainties and multiplying the weights by the corresponding channel enlargement factor to reflect the modified noise variance. This update simply accounts for the larger amount of information collected in each wider channel, and it is not to be confused with the reweighting schemes of Section 4.2.3.2, which instead set the relative contribution of each source to the stack.

Optionally, to enhance spectral smoothness, Hanning smoothing can be applied prior to rebinning via `casatasks.hanningsmooth()`, which executes a weighted running average across adjacent channels, assigning weights of 0.25 to neighbouring channels and 0.5 to the central channel, while flagging the boundary channels. This operation generates `file.ms centre.rest.smooth`, which can then be rebinned using `casatasks.mstransform()` to produce `file.ms centre.rest.smooth.rebin`. In ALMA data a spectral smoothing is often already applied at the correlator stage or during calibration, in which case the task detects this status and bypasses the smoothing step for the affected spectral windows, so that this step is only needed for data that have not been smoothed before.

4.2.1.4 Regridding

To enable data merging through `casatasks.concat()`, all observations must share the same number of channels, since the task requires spectral windows of identical dimensions in order to merge them into a single window in the subsequent steps. This constraint derives from the CASA-based implementation rather than from the method itself, and two alternative pathways are available to satisfy it, illustrated in Figure 20.

In the first pathway, adopted in this work, all datasets are cut to the same number of channels, necessarily limited by the dataset with the fewest channels. Each dataset keeps this number of channels around the target line, selected so as to maximise the mutual overlap, and the rest of its band is discarded. Since the line does not necessarily lie at the centre of each spectral window, the retained ranges do not coincide exactly, and the coverage varies across the final spectrum, with some channels sampled by the full dataset and others by fewer observations. Rather than restricting the analysis strictly to the common overlap region, individual channels can then be weighted according to their observation count (Sec. 4.2.3.2). This pathway requires every dataset to overlap substantially with the retained range, which holds for the samples analysed here, since the observations were set up around the target line at a known redshift and the datasets in which the line falls too close to the edge of a window are discarded (Sec. 6.2.1).

In the second pathway, every dataset is instead expanded with flagged dummy channels up to a common frequency range, so that no part of the band is lost. The global minimum and maximum frequencies of the sample define this range, and the observation with the largest number of channels provides the reference grid. The expansion is performed with custom routines that generate copies of each MS, preserving its table structure and metadata. These copies are fully flagged with `casatasks.flagdata()`, so that their visibilities become invalid placeholders, and are merged with the original dataset via `casatasks.concat()`. Finally, `casatasks.mstransform()` with `combinespws = True` combines the resulting spectral windows into a single one, trims the excess outer channels and regrids the spectrum onto the reference grid, yielding a uniformly sampled dataset, *file.ms.centre.rest.rebin.regrid*. Because it multiplies the volume of data written to disk, this pathway is the less efficient of the two, and it is useful only when a wide common band is needed, for instance to fit several lines simultaneously.

4.2.2 Stacking

Once the data preparation phase is complete, the individual MSs are combined into a single dataset, *concat.ms*, using `casatasks.concat()`. Unless the input datasets match in frequency coverage and channelisation within the tolerance defined by the `frequ_tol` parameter, this task preserves their spectral windows as distinct entries in the `SPECTRAL_WINDOW` subtable, treating the concatenated file as a collection of separate spectral setups sharing a common field. Individual observations therefore remain identifiable after merging, and per-source modifications can in principle still be applied at this stage, by selecting the corresponding spectral window in the tasks that support it.

The combination stage is the point where all individual corrections are applied, because it represents the final step where the datasets remain distinguishable as independent files. Two distinct mechanisms are available for this purpose. The first acts on the visibility amplitudes through a multiplicative factor applied to the `DATA` column of each dataset before merging, and it carries the flux normalisations. The second mechanism acts on the statistical weights through the `visweightscale` parameter of `casatasks.concat()`, which accepts a rescaling factor for each dataset and sets the relative contribution of each observation to the stack. By default this parameter is left untouched so that the combination reduces to the natural weighting detailed in Section 3.4.1, and it is modified only when the democratic scheme is adopted as discussed in Section 4.2.3.2.

The amplitude normalisation is controlled through a set of independent flags in the stacking routine, where each enables a multiplicative

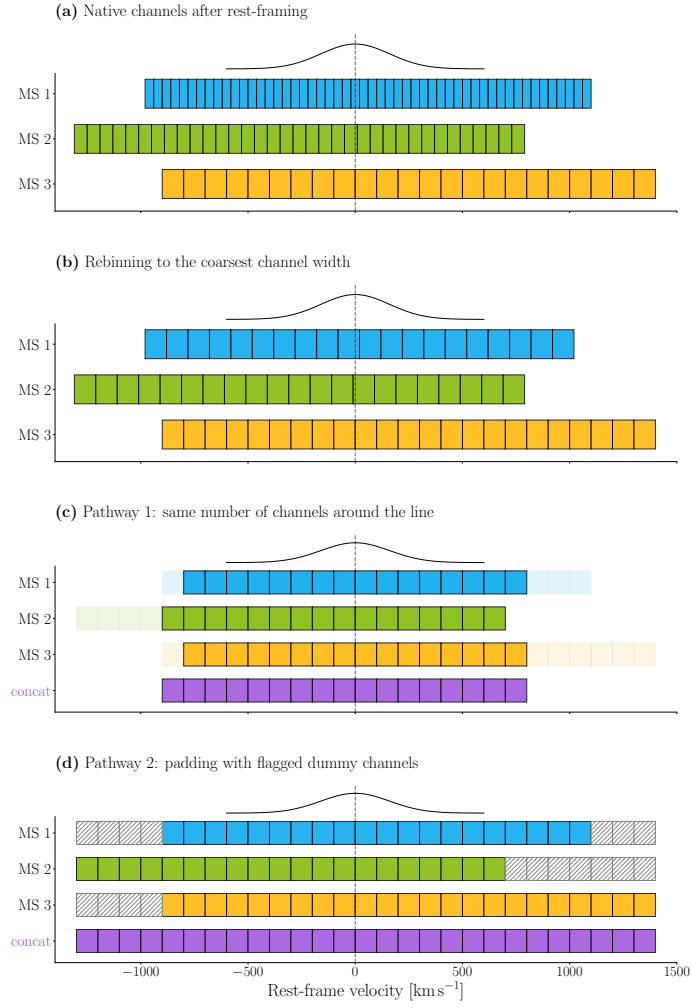


Figure 20: Schematic representation of the spectral alignment for three datasets, with the target line centred at zero. (a) Native channelisation after rest-framing, with different channel widths and frequency ranges. (b) After rebinning, all datasets share the coarsest channel width, but their channel edges are not yet aligned. (c) First pathway: the datasets are regridded onto a common grid and cut to the same number of channels around the line, while the discarded channels are shown in lighter colours. (d) Second pathway: each dataset is padded with flagged dummy channels, shown as hatched cells, up to the common frequency range of the sample. The purple row shows the spectral coverage of the final concatenated dataset.

factor that is accumulated per source and applied immediately before concatenation. Three factors are available, namely the primary beam correction, the redshift rescaling, and a user-defined normalisation.

The first factor is the primary beam correction defined in Equation 25, which is deferred from the centering step of Section 4.2.1.1. Relocating a target to the field centre restores its geometric position but not its intrinsic flux, since the signal remains attenuated by the off-axis response of the array elements. To correct for this effect, the angular separation θ between the original phase centre recorded before the shift and the target coordinates supplied in the input file is evaluated using *Astropy*. The attenuation of Equation 25 is then evaluated at this offset, at the observed frequency of the line, $\nu_{\text{obs}} = \nu_{\text{rest}}/(1+z)$, and with the physical dish diameter of the array. The visibilities of the dataset are finally multiplied by the inverse of this single value, whose validity is discussed in Section 3.3.2.

The second factor is the rescaling of the amplitudes to a common reference redshift. This adjustment accounts for the different luminosity distances across the sample through Equation 29 and is computed internally from the redshift of each source assuming a flat Λ CDM cosmology.

The third factor is left to the user and covers any further normalisation detailed in Section 3.4.2 that cannot be derived from the observation itself. Such adjustments include a rescaling by intrinsic luminosity or a correction for lensing magnification, and they are supplied through an optional column in the input file. These three factors are independent and are multiplied together before being applied to the visibilities.

Some operations require instead the data to be averaged across the sample, either in frequency or in time. This is the case for the time averaging needed to reduce the computational load before the model fitting, and for the visualisation of the stacked visibilities in the uv plane with tools such as *casaplotms*. Both require the spectral windows to be definitively merged into one, achieved by running `casatasks.mstransform()` with `combinespws = True`, and the observation entries to be unified as well, since each source remains registered as an independent entry of the `OBSERVATION` subtable, a structure that would otherwise prevent averaging across scans belonging to different executions. The unification is obtained by editing that subtable so that a single observation identifier spans the whole temporal range of the sample, and by extracting it into a unified dataset, *concat.ms.unified*.³

Until this step, each source remains a separate spectral window and observation within *concat.ms*, while all sources share the same field,

³ In practice, the global extremes of the `TIME_RANGE` column are assigned to the first row of the `OBSERVATION` subtable, all the entries of the `OBSERVATION_ID` column of the main table are set to zero, and `casatasks.mstransform()` extracts observation ID 0 into the new dataset.

since their coordinates were set to zero during the centering. Because this unification is irreversible, it also sets the boundary of the pipeline. Every operation that distinguishes one source from another must be carried out before it, which is precisely why the normalisations described above are bound to the concatenation step, while everything that treats the stack as a single object can only follow.

4.2.3 *Post-processing*

The operations described in this section are classified as post-processing because they serve a fundamentally different purpose compared to the preceding steps. While pre-processing is strictly dictated by the need to harmonise heterogeneous data into a combinable format, the procedures detailed here encode the scientific intent of the analysis. The treatment of the continuum, the relative weighting of the sources and the final flux extraction all depend on the specific question being asked, meaning a different scientific goal would naturally lead to an alternative configuration of the pipeline. The flux normalisations detailed in Section 4.2.2 conceptually belong to this same category, and they were discussed alongside the combination step solely because the technical implementation requires them to be applied at that stage.

4.2.3.1 *Continuum subtraction and analysis*

As discussed in Section 3.4.3, the continuum can be removed either from the individual sources before combination or from the stacked spectrum afterwards. The choice is dictated by the sample rather than by the software, since both orderings are available through the same CASA task `casatasks.uvcontsub()` which performs the subtraction natively in the *uv* domain. Applying this task to *concat.ms.unified* realises the post-stacking strategy, while applying it to each *file.ms* before concatenation realises the pre-stacking approach. This implementation adopts the post-stacking strategy by default, and this constitutes the appropriate choice for a sample where the individual continuum levels lie below the noise floor and no single source is bright enough to distort the stacked baseline.

The task estimates the continuum by fitting polynomials to the real and imaginary parts of the visibilities across the line free channels specified via the `fitspw` parameter. Typically a zero order polynomial representing a constant value is selected to prevent overfitting and preserve the line flux. Subtracting this fitted model outputs a new MS named *concat.ms.unified.contsub* containing exclusively the line emission. Since the task fits each visibility row independently, and each row belongs to a single source, applying it to the unified dataset is equivalent to a per-source subtraction with a common set of line-free channels, defined once on the shared rest-frame grid. A

continuum model of the stack as a whole is instead obtained by fitting the continuum as the constant baseline of the stacked spectrum, as done in Sections 4.3 and 4.4.

Setting the parameter `write_model = True` additionally stores the fitted continuum model in the `MODEL` column of the output dataset. Imaging this column yields a noiseless representation of the fitted continuum rather than an image of the measured emission. A stacked continuum map is instead obtained by imaging the line-free channels of the stacked visibilities with `casatasks.tclean()` in multifrequency synthesis mode, as done in Sections 4.3 and 4.4.

4.2.3.2 Spectral and statistical weighting

As described in Section 4.2.2, the pipeline defaults to natural weighting to maximise the overall SNR of the stacked signal. If sample representation is prioritised over sensitivity, the democratic weighting scheme of Equation 28 can be applied instead by supplying individual rescaling factors to the `visweightscale` parameter of `casatasks.concat()` during concatenation.

A spectral channel weighting can additionally be superimposed on either scheme. Unlike the statistical weighting choices, this procedure does not reflect a scientific preference but compensates for an inhomogeneity introduced by the pipeline itself. As detailed in Section 4.2.1.4, regridding datasets with different initial bandwidths onto a common frequency grid causes individual channels to be sampled by a varying number of observations. The effective number of datasets N_{eff} contributing to each channel of the common grid is calculated from the frequency coverage of the individual MSs, and the weights are rescaled by $\sqrt{N_{\text{eff}}/N_{\text{max}}}$. This downweights channels near the spectral boundaries, which are sampled by fewer observations, relative to central channels with full sample coverage. Because this correction relies on the common grid, it represents the only weighting step performed after concatenation rather than during it.

A final consideration concerns the interaction between the weighting and the flux normalisations. Any multiplicative factor applied to the visibilities inevitably scales the noise as well, meaning the variance of a normalised stack no longer reflects the true instrumental sensitivity. The recommended practice is therefore to carry two concatenated datasets through the pipeline. A normalised dataset named `concat.ms.norm` is used to measure the physical flux, while an unnormalised dataset named `concat.ms.unnorm` is retained to recompute the weights, ensuring they continue to represent the true instrumental sensitivity rather than the arbitrary scale imposed by the normalisation.

4.2.3.3 *uv-model fitting with UVMultiFit*

To model the stacked emission directly in the visibility plane without relying on image reconstruction, the `UVMultiFit` package is employed. Fitting models directly to visibilities bypasses the non linearities and spatial correlations introduced by deconvolution, enabling unbiased parameter estimation for unresolved or partially resolved sources. Within CASA, the `uvmultifit()` function fits analytical intensity distributions to complex visibilities across user defined frequency channels. Depending on source morphology, several functional forms can be selected, including a delta function for unresolved point sources, a two dimensional Gaussian profile for resolved emission, or circular and elliptical disk models.

When executing per channel fitting with the setting `OneFitPerChannel = True`, the function fits the selected model independently across the spectral axis, producing channel by channel estimates of flux density and associated statistical uncertainties. Additionally, by setting the parameter `write = 'residuals'`, the task writes the residual visibilities directly into a column of the MS, allowing fit quality and residual distributions to be evaluated.

Because `UVMultiFit` processes visibilities in memory, running the task on large concatenated datasets can lead to severe memory saturation and execution crashes. To prevent computational failure, the dataset must first undergo the time averaging described in Section 4.2.2, which requires the spectral windows and the observation entries to be unified beforehand. This step reduces the total visibility volume, but the averaging interval must be kept short enough that the *uv* position of each baseline changes little within it. Otherwise, visibilities that sample different spatial frequencies are merged together, an effect known as time-average smearing, which grows with baseline length and with the extent of the source and is negligible for a point source. Finally, plotting the extracted channel flux densities yields a spectrum in the visibility domain, which can be modelled with a Gaussian profile to determine the integrated line flux, line width, and central velocity offset directly from the interferometric data.

4.2.3.4 *Imaging with tclean*

The visibilities can also be transformed into reconstructed image cubes using the `casatasks.tclean()` task, which synthesises images from interferometric data by deconvolving the PSF from the observed visibilities. Beyond providing an alternative route to the flux extraction, imaging the stack serves to test the consistency of the two approaches against each other and to inspect the stacked source visually, which is a useful validation of the results obtained in the *uv* plane, where the emission is never seen directly.

To image stacked spectra, the deconvolution algorithm, the visibility weighting and the spatial grid dimensions are selected based on the scientific goals and on the heterogeneity of the sample, as discussed in Section 3.5.1. Deconvolution is performed using the Högbom algorithm, a standard iterative method suited for fields dominated by compact or point like emission. The imaging weights are generally set to natural weighting, which maximises the point source sensitivity by emphasising the densely sampled regions near the origin of the uv plane at the expense of angular resolution, a trade off that is optimal for recovering the total flux density of compact or faint stacked targets. Any other scheme remains available, although it is advisable to keep the imaging weights consistent with the weighting adopted for the combination, particularly when the democratic scheme has been used, so that the two do not act against each other. The spatial cell size is defined as one fifth of the major axis of the restoring beam, satisfying the Nyquist sampling requirements for the synthesised beam.

Initially, a dirty image cube is generated by setting `niter = 0` to verify data alignment, inspect potential stacking artifacts, and evaluate preliminary signal significance.

The cleaning threshold for the deconvolution is established from the theoretical point source sensitivity σ of Equation 22, introduced in Section 3.2.3. This threshold depends on the choice of the weighting scheme (Sec. 3.4.1), which enters through the factor w_r of that same equation. Under natural weighting it corresponds to the minimum variance estimate, whereas under democratic weighting it must be updated to account for the rescaled visibility weights. Deconvolution is then executed down to a stopping threshold defined as a multiple of this theoretical noise level, typically two or three times σ .

When dealing with a sample containing diverse source morphologies or heterogeneous spatial resolutions, specific parameters of `casatasks.tclean()` can be adjusted prior to the image extraction. For partially resolved sources whose size exceeds the synthesised beam, a common spatial resolution can be enforced across the sample using the `restoringbeam` parameter to set the beam size to that of the coarsest observation. For strongly resolved or lensed targets observed at high spatial resolution, the uv tapering introduced in Section 3.2.2 can be enabled within the task, downweighting the long baselines to smooth the high resolution visibilities to a beam compatible with the rest of the sample.

Running `casatasks.tclean()` generates final image products including the restored image cube, residual map, model image, PSF, and primary beam response. Since the primary beam correction has already been applied to the visibilities of each dataset, the primary beam image produced by the task must not be used for a further correction of the stacked image. To quantify the noise statistics, measure

peak intensities for unresolved targets, or extract spatial emission properties directly from these output maps, `casatasks.imstat()` or two dimensional Gaussian fitting via `casatasks.imfit()` can be applied on selected regions of the restored cube. Finally, the reconstructed images can be visualised in interactive software such as CARTA for further spatial and spectral examination.

4.3 CODE VALIDATION ON SIMULATED DATA

Before applying ViSTA to observational data, I test its accuracy and reliability on a controlled synthetic dataset, where the intrinsic properties of the sources are known *a priori* and the recovered signal can therefore be benchmarked directly against the input model. This section describes the simulated sample and presents the results of the validation, carried out separately for an unresolved and for a spatially resolved source, and for each of the extraction strategies introduced in Section 3.5.

4.3.1 The simulated sample

Two structural cases are considered, an unresolved point source and a spatially resolved Gaussian source with a FWHM of 5 arcseconds, and for each of them a separate sample of 50 mock galaxies is generated, so that the two configurations can be validated independently. Within each sample the 50 observations are distributed across redshifts between 2 and 4 and across the array configurations `alma.cycle.11.n.cfg`, with n ranging from 2, corresponding to a maximum baseline of 314 m, to 5, corresponding to 1400 m, which reproduces the spread in angular resolution of a real archival dataset. The observations are produced with the `simobserve` task of CASA and subsequently corrupted with custom routines that inject realistic per-antenna thermal noise and elevation-dependent atmospheric phase offsets, as detailed in Appendix A.1. The mocks do not include the turbulent tropospheric phase fluctuations, consistently with the use of calibrated data, in which these are removed by the calibration and only residual errors remain (Sec. 3.1.4). They therefore do not test the effect of such residuals on the stack.

Each input sky model is a spectral cube consisting of 100 channels with a 1 MHz width, containing a flat continuum of 0.6 mJy per channel and a Gaussian line profile with a 2 mJy peak flux confined to the 20 central channels. Both values refer to the flux density integrated over the whole source, and not to the peak surface brightness, since the spatial profile is normalised to unit sum. This spectral setup is identical for both samples, ensuring that their spatial extent remains the sole difference between them. The central channel is set to the rest frequency of the CO(3–2) transition at 345.80 GHz, with each cube

constructed so that the rest frame line frequency aligns across the entire sample (Sec. 3.3.3). Because the line lies at the centre of every input cube, which has the same number and width of channels for all sources, the rebinning and the partial overlap at the edges of the common grid are exercised only in a symmetric configuration. The trimming of lines close to the edge of a spectral window was instead tested on individual sources, validating each function of the pipeline against the corresponding CASA reference task. Selecting this transition anticipates the scientific application presented in the following section and in Chapter 6, where the average line emission of the population is reconstructed. Over the target redshift range, the typical thermal noise at 1 MHz resolution is of the order of 1 mJy, rendering neither the line nor the continuum detectable in individual observations. This establishes the low SNR regime specifically required to evaluate the recovery performance of the stacking pipeline. The behaviour of the code in the absence of noise was verified separately. Since the simulator keeps the noiseless model in the DATA column, running the pipeline on that column provides the exact response of ViSTA to the rest-framing, regridding and primary beam correction, and the stacked fluxes recovered in this way are consistent with the input model. The real sample of Section 4.4 provides a further test at higher SNR, since all its targets are detected in their moment zero maps.

Figure 21 illustrates the effect of the rest framing procedure of Section 4.2.1.2 on an individual source: the visibility distances scale down by a factor $(1+z)$, while the reconstructed rest frame image remains structurally identical to the input, since the uv coordinates in units of wavelength are unchanged (Sec. 4.2.1.2). No primary beam correction is applied to these images. Figure 22 shows the concatenated uv coverage of the stacked mock observations, and how combining multiple array configurations and redshifts yields a dense and uniform spatial sampling.

Two choices in the configuration of the pipeline follow directly from the properties of these samples. No continuum subtraction was applied to the individual sources: following the post-stacking strategy of Section 4.2.3.1, all visibilities were stacked together and the full line-plus-continuum profile was fitted at once, with the continuum recovered as the constant baseline of the fit. This ordering is justified here because all mock lines lie below the individual noise level and the injected continuum is faint, as expected for star-formation-dominated emission, so that no single source carries a continuum bright enough to distort the stacked baseline. For the same reason, natural weighting was adopted throughout (Section 4.2.3.2), since all simulated sources share identical intrinsic fluxes and no target dominates the sample.

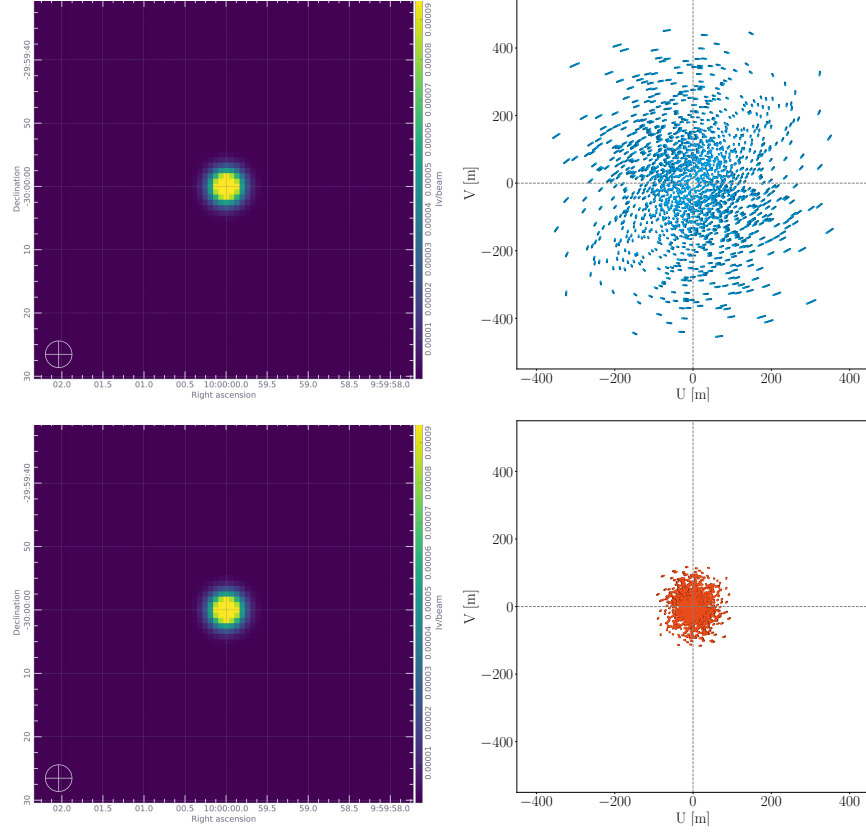


Figure 21: Example image and uv plane distribution in metres for an uncorrupted simulated source at redshift 2.9 observed in ALMA configuration 4. The upper panels showcase the image and uv plane distribution in the observed frame, while the bottom panels display the corresponding results following rest framing.

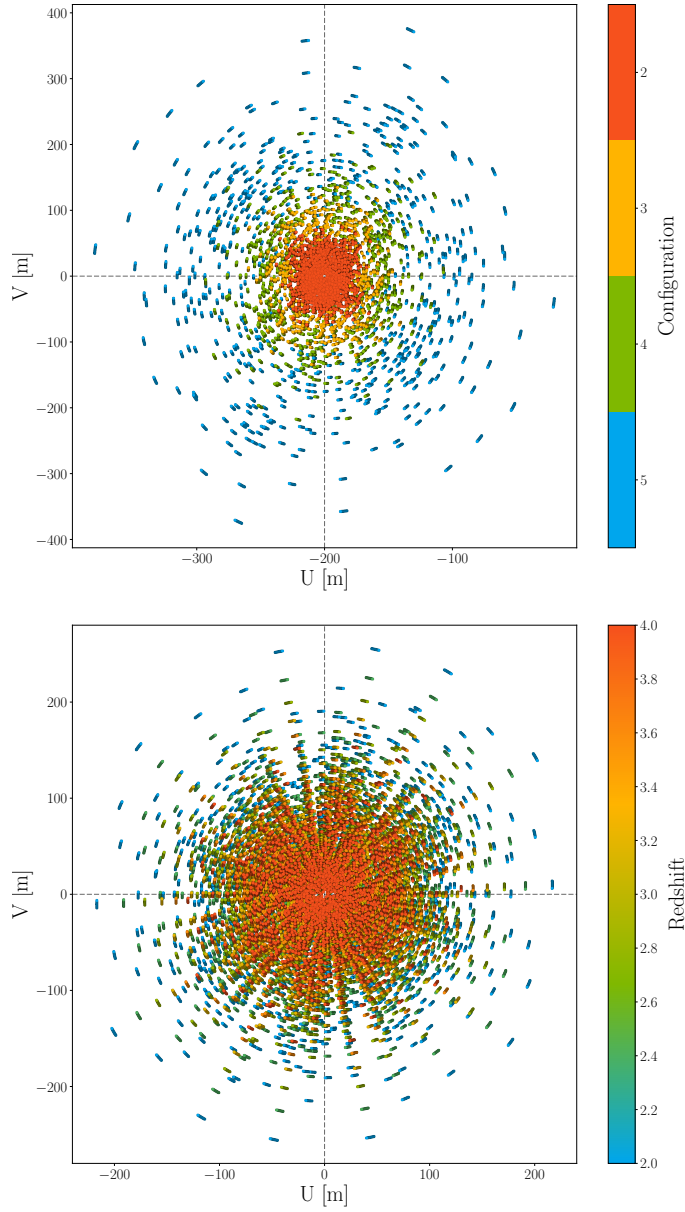


Figure 22: uv plane coverage of the stacked simulated sources. Visibilities from different array configurations have been centred on the phase centre and shifted to the rest frame using `VISTA`. The top panel shows rest framed visibility positions for all 50 mock observations colour coded by array configuration. The bottom panel illustrates, as an example, the variation in uv coverage for five sources observed in configuration 4 across different redshifts.

Following the centering, rest framing, rebinning and regridding onto a shared frequency axis, the new channel spacing is set by the observation with the largest intrinsic frequency dispersion, that is by the highest redshift source. The fits are then executed both in the uv domain, using `uvmultifit()` and hereafter denoted ViSTA-uvFit, and in the image plane after the reconstruction of the stack, hereafter ViSTA-ImFit. Both are compared against the traditional image plane stacking, denoted ImSTA-ImFit, in which the individual datasets are imaged separately with their native setups and the fluxes are averaged channel by channel, as detailed in Appendix A.3. For the two image domain methods, the deconvolution is performed with default parameters and a synthesised beam fitted to the central lobe of the concatenated PSF under natural weighting.

4.3.2 Unresolved source

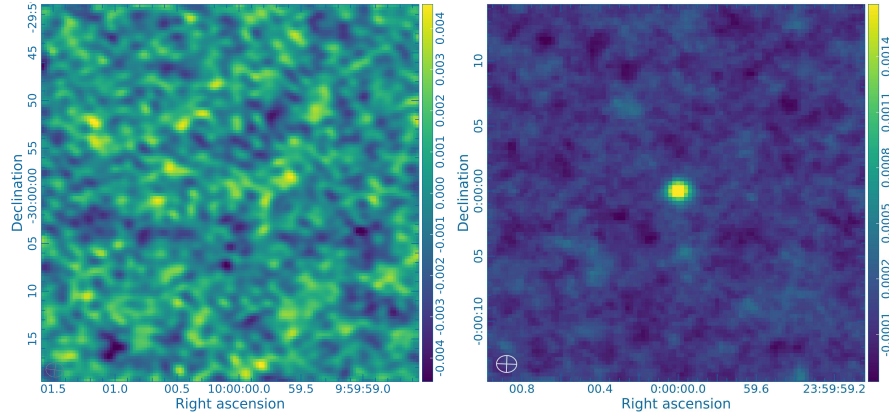


Figure 23: Comparison between the peak channel image of an individual mock source before stacking, where emission remains below noise level (left panel), and the resulting image after stacking 50 mock sources using ViSTA (right panel).

Figure 23 compares an individual unstacked mock image, representative of all the individual observations, against the stacked image derived from 50 sources. While individual targets remain hidden within thermal noise, stacking recovers a clear high significance detection. For the unresolved sample, the flux density is measured in the image plane directly from the central peak pixel, without spatial integration, and in the uv plane by fitting a point source model, in accordance with Section 3.5. Figure 24, Figure 25 and Table 2 document the SNR evolution as a function of stacked sample size N . The channel noise decreases according to $1/\sqrt{N}$ scaling, reducing the initial average channel noise of 1.2 mJy by a factor of 8 and yielding a final peak SNR of 17.3 for 50 stacked sources.

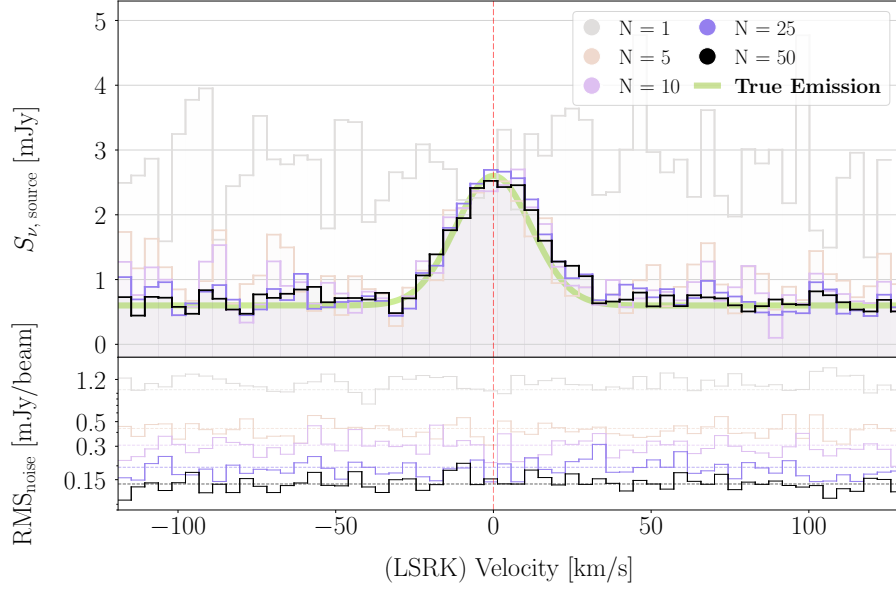


Figure 24: Top panel displays spectral line profiles recovered via ViSTA-ImFit across varying numbers of stacked point sources ($N = 1, 5, 10, 25, 50$) compared to the true input model shown in green. Bottom panel indicates measured RMS noise as a function of N , where the dashed line denotes theoretical $1/\sqrt{N}$ noise scaling.

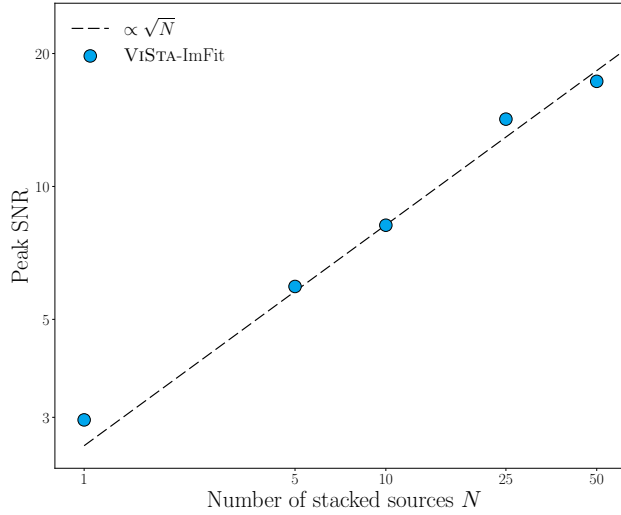


Figure 25: Peak SNR of the central channel recovered with ViSTA-ImFit as a function of the number N of stacked unresolved mock sources (Table 2). The dashed line shows the $\sqrt{N}$ scaling expected for noise-limited stacking, fitted to the measurements.

Table 2. Peak SNR Results for ViSta-ImFit

N_{Source}	S_{peak} (mJy/beam)	σ_{chan} (mJy/beam)	Peak SNR
1	3.3125	1.1191	2.9600
5	2.5168	0.4240	5.9358
10	2.4579	0.3009	8.1695
25	2.6655	0.1876	14.2058
50	2.4283	0.1404	17.3007

Note. — SNR results for ViSta-ImFit evaluated at the central channel. S_{peak} represents peak flux density and σ_{chan} denotes average channel RMS.

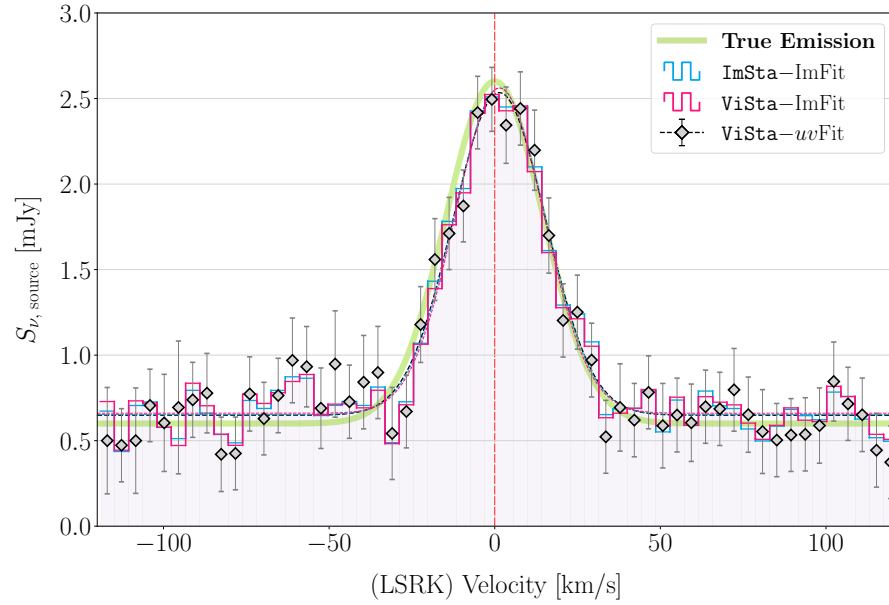


Figure 26: Comparison of recovered spectral line profiles across different stacking methods for unresolved sources. The green solid line represents the true input model. Blue and magenta lines indicate results from ImSta-ImFit and ViSta-ImFit respectively. Black diamonds display channel flux densities from ViSta-uvFit with 1σ uncertainties, and the black dashed line shows the corresponding best-fit Gaussian.

Figure 26 presents the recovered spectra across all three methods. The solid green profile denotes the true model. Black diamonds represent the per-channel flux densities recovered by ViSTA-uvFit, and the dashed black curve is the best-fit Gaussian to them, whose parameters and uncertainties are obtained by Monte Carlo sampling of the channel measurements within their errors. All three methods accurately recover the line profile, achieving approximately 90 % fidelity relative to the true model. No redshift rescaling was applied to the mock sources, whose input fluxes were set directly to the values above. The peak flux densities of Table 2 include the continuum, so that the expected value in the central channel is 2.6 mJy, slightly reduced by the rebinning of the line onto the coarser common channels. The small deviations of the continuum level from the input value are compatible with its uncertainty, since the faint continuum lies close to the noise floor and the baseline of the fit is poorly constrained.

At such low SNR, a positive bias could arise if the flux were measured from the modulus of the visibilities. The real and imaginary parts of each visibility carry independent Gaussian noise (Sec. 3.1.3), but the scalar amplitude $|V| = \sqrt{\text{Re}(V)^2 + \text{Im}(V)^2}$ is always positive, and its distribution is not Gaussian but follows a *Rice distribution* [175]. When the true amplitude is comparable to the noise, the noise contributes to the modulus regardless of its sign, and the mean measured amplitude exceeds the true one. For a vanishing signal, for instance, the mean amplitude is of the order of the noise rather than zero. This *Rician bias* decreases as the SNR increases, and the distribution becomes approximately Gaussian above an SNR of about five. None of the extraction methods used here relies on the scalar amplitudes. UVMultiFit fits the model to the complex visibilities, minimising the residuals of their real and imaginary parts, whose noise is Gaussian and symmetric, while the image-domain methods measure the flux on the image at the known source position. The measured fluxes are therefore unbiased even in the channels where the continuum is detected at only two or three times the noise.

4.3.3 Resolved source

For spatially resolved targets, stacking workflows remain conceptually similar, but signal extraction requires spatial flux integration rather than peak pixel measurements. I evaluate extended source recovery across three operational scenarios. First, I assume known intrinsic source sizes and measure flux within predefined apertures. Second, I assume no prior size knowledge and perform two dimensional Gaussian fitting using `casatasks.imfit()` or `uvmultifit()`. Third, I convolve the individual images to a common spatial resolution prior to stacking and flux extraction.

4.3.3.1 Native source size at native resolution

In the ideal case where source extents are known, circular apertures of the same size can be used for the whole sample, since rest-framing preserves the angular size of each source (Fig. 21). The circular apertures with a radius of 3σ based on the 5 arcsecond FWHM Gaussian profile are applied to encompass line emission while minimising background noise inclusion. For ViSTA-ImFit, flux density is integrated within this aperture across all channels of the stacked cube. For ViSTA-uvFit, spatial size parameters are fixed while channel flux densities vary. For ImSTA-ImFit, flux is integrated within individual images prior to channel averaging.

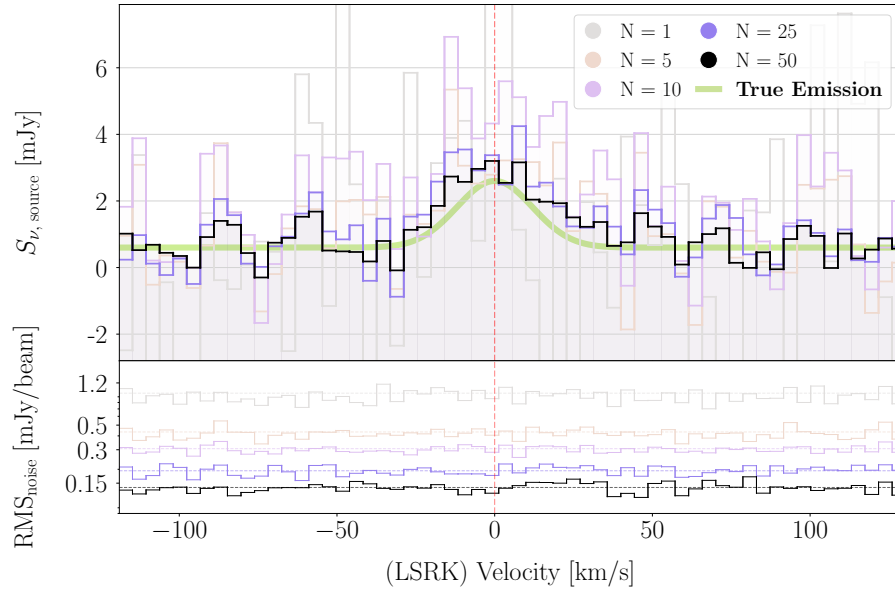


Figure 27: Top panel presents spectral line profiles recovered using ViSTA-ImFit for extended mock sources of known size across sample sizes $N = 1, 5, 10, 25, 50$ compared against the true input model. Bottom panel illustrates measured RMS noise scaling alongside the theoretical $1/\sqrt{N}$ reference curve.

Figure 27 shows noise scaling as a function of sample size for resolved sources. While RMS noise decreases as $1/\sqrt{N}$, recovering a well defined line profile requires a larger sample size than in the unresolved case due to noise accumulation across extended integration areas.

Figure 28 illustrates the resulting stacked spectra. Visibility domain fitting via ViSTA-uvFit provides the closest agreement with the true line profile. Conversely, ViSTA-ImFit slightly overestimates peak line intensity because concatenating heterogeneous array configurations alters the effective synthesised beam, concentrating flux into a central core. The traditional ImSTA-ImFit method exhibits increased channel

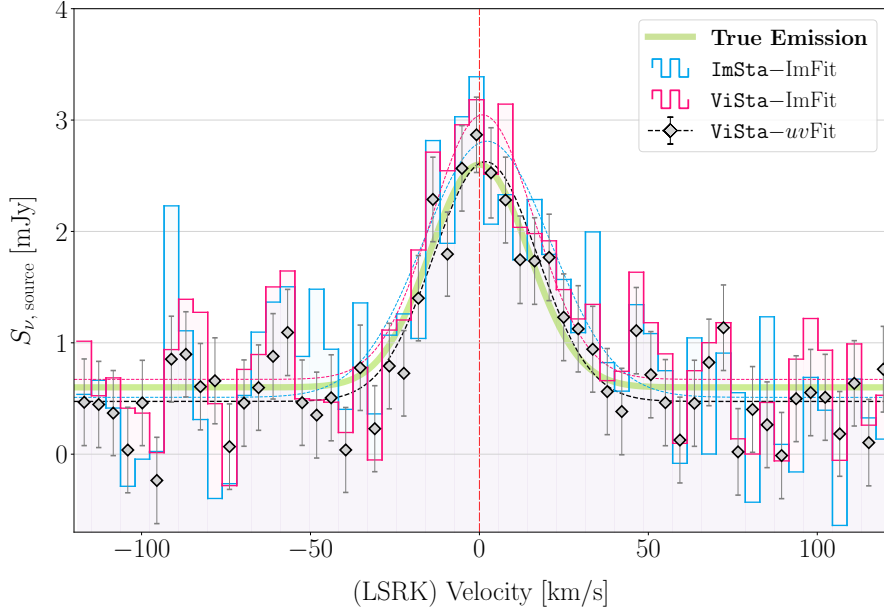


Figure 28: Comparison of recovered spectral line profiles for resolved sources using known intrinsic spatial extent. Green solid line indicates the true input model. Blue and magenta lines denote ImSta-ImFit and ViSta-ImFit results respectively. Black diamonds represent ViSta-uvFit measurements with 1σ uncertainties, alongside the best-fit Gaussian.

to channel scatter due to higher noise contamination during per image spatial integration. The continuum levels lie slightly below the input value, a difference compatible with their uncertainties.

4.3.3.2 Estimated source size at native resolution

I repeat the evaluation without assuming prior knowledge of source dimensions. In ViSta-uvFit, spatial parameters including major axis, minor axis, and position angle are fitted on channel averaged visibilities and fixed during subsequent channel by channel flux estimation. In ViSta-ImFit, two dimensional Gaussian fitting is performed on the moment zero map using `casatasks.imfit()`, returning deconvolved source dimensions used to define a 3σ flux extraction aperture across the spectral cube. The moment zero map is used only to raise the SNR enough for a reliable fit, and this approach assumes that the source size does not change across the line. In ImSta-ImFit, moment zero maps and Gaussian fits are generated for every individual observation, with integrated fluxes extracted per source before channel averaging across the sample.

Figure 29 demonstrates the performance of these size estimation approaches. Both ViSta-uvFit and ViSta-ImFit accurately recover continuum and line profiles by determining source dimensions on high significance stacked data. In contrast, ImSta-ImFit struggles to con-

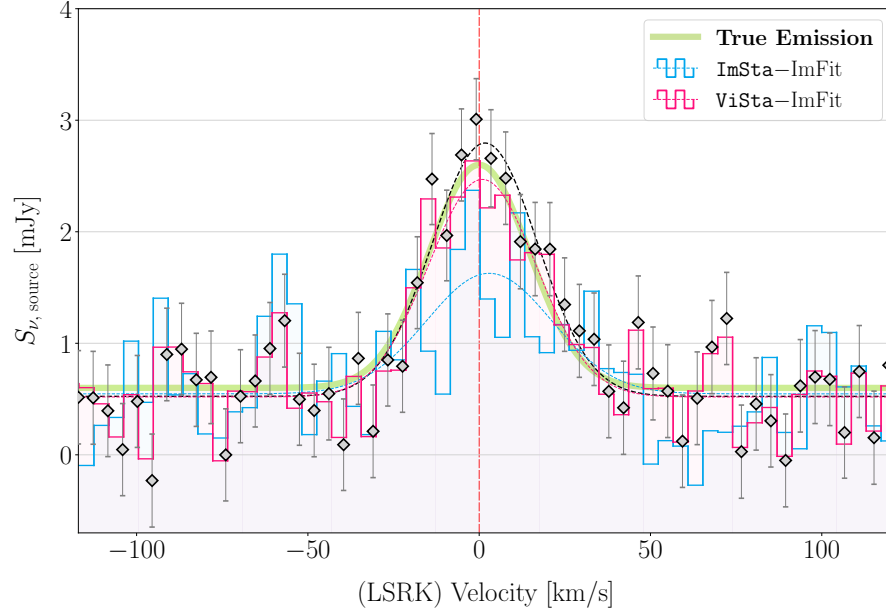


Figure 29: Comparison of recovered spectral line profiles for resolved sources following source size estimation. The green line represents the true profile, magenta and blue lines indicate ViSta-ImFit and ImSta-ImFit profiles, and black diamonds show ViSta-uvFit channel measurements with the corresponding best-fit Gaussian.

strain individual source sizes in low SNR regimes, leading to inaccurate aperture definitions, underestimated line peaks, and pronounced spectral fluctuations. This reflects the limitation of estimating the sizes source by source at low SNR, rather than an intrinsic limitation of image plane stacking. A common size could be estimated on the stacked image, but the individual images have different synthesised beams and can be combined only after convolving them to the coarsest resolution of the sample, which is the approach of the next section.

4.3.3.3 Estimated source size at lowest resolution

Finally, I evaluated the full sample after degrading all data to the lowest common angular resolution. In this configuration, both the integrated flux density and the noise level, which becomes uniform across all images in the dataset, were estimated using the ViSta-ImFit and ImSta-ImFit approaches.

Figure 30 presents the results of both image plane fitting methods applied to the two stacking strategies at the lowest common resolution. In this scenario, stacking directly in the uv domain yields a more uniform sampling of the Fourier plane and naturally down-weights the noisiest baselines, resulting in a higher fidelity reconstruction of the source core. Conversely, convolving each individual

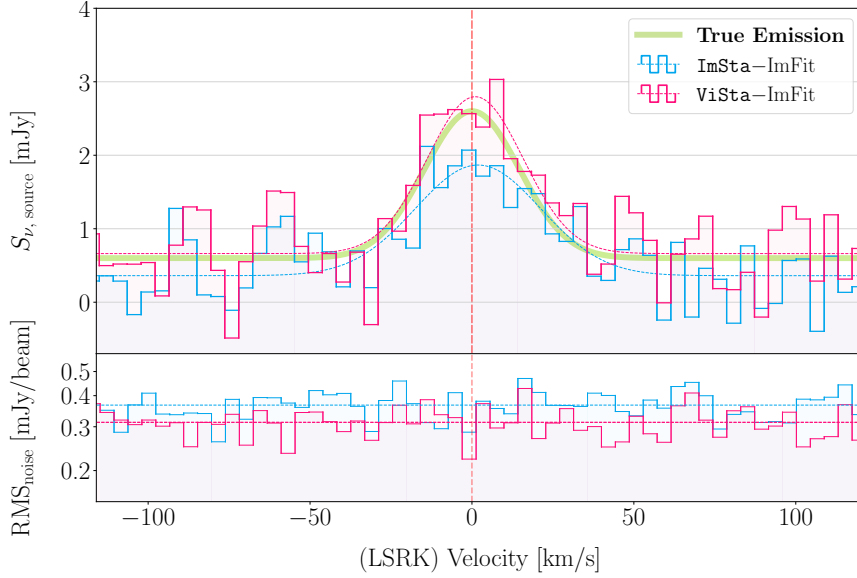


Figure 30: Comparison of recovered spectral line profiles for resolved sources using different stacking and fitting methods after convolving all sample images to the lowest common resolution. The green solid line represents the true input model. Blue and magenta lines denote ImSta-ImFit and ViSta-ImFit results respectively alongside their corresponding best-fit Gaussians.

image to the lowest resolution prior to stacking introduces additional noise when single observations possess low SNRs. This leads to increased scatter across the measured spectrum, a reduced peak amplitude, and a slightly elevated noise floor. Consequently, following resolution matching across low SNR observations, the ViSta framework consistently achieves a superior recovery of the intrinsic source profile.

Figure 31 displays the convolved stacked image for the central channel of the cube reconstructed with both approaches. The deconvolved source size recovered by both ViSta and ImSta aligns closely with the true input value, as shown by the fitted and effective source sizes overlaid on the two panels of Figure 31. However, image plane stacking exhibits a more pronounced noise pattern with an average central channel RMS of approximately 0.285 mJy per beam, whereas uv domain stacking produces a significantly smoother background with an RMS of 0.220 mJy per beam. Comparing the integrated flux densities obtained with the two methods, which reach 2.566 mJy for ViSta and 2.070 mJy for ImSta, against an input value of 2.6 mJy in the central channel, given by the 0.6 mJy continuum and the 2 mJy line peak, the ViSta approach provides a SNR gain of approximately 61% in the central channel relative to image plane stacking.

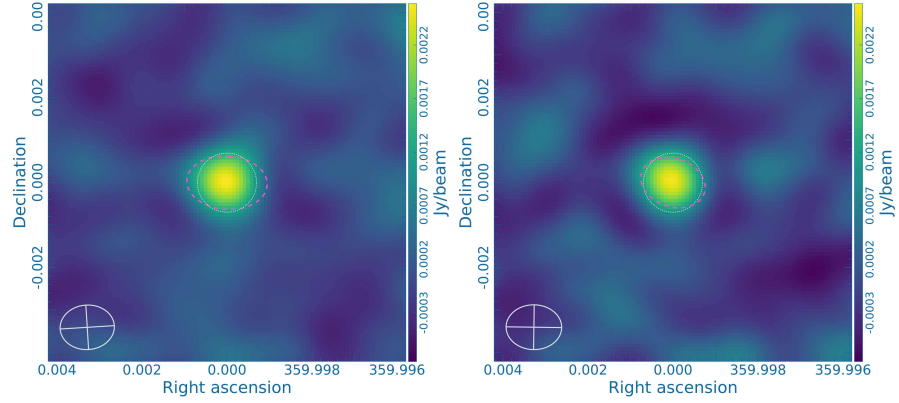


Figure 31: Stacked central channel images of the cube generated using two different resolution matching workflows. The left panel shows the image stacked in the uv plane using ViSTA and subsequently convolved to the coarsest synthesised beam. The right panel displays the image stacked in the image plane after convolving all input observations to the worst resolution. The purple dashed circle denotes the source size fitted via `casatasks.imfit()`, while the white dotted circle indicates the effective stacked source dimension. Both methods recover intrinsic source geometry, but visibility plane stacking achieves a smoother noise background.

4.3.3.4 Continuum extraction at lowest resolution

I evaluated the performance of ViSTA for continuum emission extraction to quantify the SNR gain offered by visibility domain processing. Continuum images were generated from the ViSTA stacked MS selecting exclusively line free channels, which were known prior to analysis in this synthetic dataset. In observational applications, continuum isolation requires subtracting line emission following the procedure detailed in Section 4.2.3.1.

The continuum comparison was executed by homogenising all datasets to the lowest angular resolution, ensuring a realistic and consistent noise comparison across both frameworks. While spectral cubes for line analysis were produced using `casatasks.tclean()` with `specmode = 'cube'`, continuum maps were synthesised using `specmode = 'mfs'` across line free channels. For both methods, the peak flux density was measured at the source position, while the RMS noise was measured far from the source, in an emission-free region of the map.

The reconstructed continuum maps are presented in Figure 32. Because the degraded restoring beam is comparable in size to the source rather than significantly larger, the source is marginally resolved and its peak flux does not represent the total integrated flux density. The

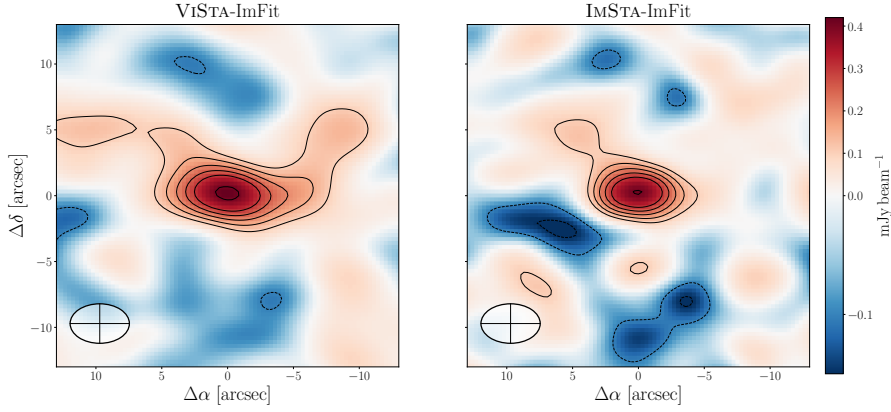


Figure 32: Stacked continuum emission maps obtained using two resolution matching procedures. The left panel shows the continuum image stacked in the uv plane using ViSTA and convolved to the coarsest beam. The right panel presents the image stacked in the image plane following individual image convolutions. Both maps are shown on the same colour scale, centred on zero so that negative values appear in blue, and as offsets from the phase centre. Contours are drawn at $\pm 2, 3, 4, 5, 6, 8 \times 0.05 \text{ mJy beam}^{-1}$, with negative levels dashed, and the ellipse in the lower left corner shows the synthesised beam. Stacking in the visibility plane yields a smoother background, with fewer and shallower negative features, and recovers a higher proportion of source emission.

peak is nevertheless used directly, since both methods share the same restoring beam and the peak-to-noise ratio is a detection SNR independent of the source morphology. This ensures that the SNR differences reflect the intrinsic performance of the methods rather than uncertainties in the fit of the spatial size, while the peak values are not expected to reproduce the input continuum of 0.6 mJy.

Under ImSta-ImFit, the measured peak flux reaches 0.404 mJy per beam with an RMS noise of 0.058 mJy per beam, yielding a SNR of 6.966. Under ViSta-ImFit, the peak flux reaches 0.419 mJy per beam with an RMS noise of 0.051 mJy per beam, resulting in a SNR of 8.216. This represents an improvement of approximately 18% in peak SNR, consistent with the gain factor of 1.2 reported in recent ALMA archival combination studies comparing visibility plane and image plane stacking [190].

4.4 APPLICATION TO A REAL CASE

After validating the framework on simulated datasets, the methodology is applied to a real sample of galaxies observed with ALMA, targeting the same CO(3–2) transition adopted for the simulations so that both tests probe the same low SNR regime. Over the redshift range of the sample, $z \sim 1.4$ –2.7, the line falls between 93 and 144 GHz, within ALMA Bands 3 and 4, and the archival observations covering it consist primarily of short spectroscopic scans rather than dedicated deep integrations.

The sample is built in two steps, starting from the ensemble assembled in Chapter 2. Among those sources, I first select the ones for which the ALMA archive contains observations covering the redshifted frequency of the CO(3–2) transition. I then inspect the angular resolution of each of these datasets, retaining only those in which the target appears unresolved or, at most, partially resolved, and discarding the observations where the emission is spatially resolved into the arcs, rings or multiple images of a lensing configuration, or into any other structure that the extraction strategies of Sections 3.5.1 and 3.5.2 would fail to reproduce.

The resulting sample comprises 21 galaxies observed in 25 datasets, four targets having been observed more than once, and preserves a broad spread in angular resolution ranging from 0.25 to 2.8 arcseconds, capturing precisely the structural heterogeneity that ViSta is built to handle. All of them belong to the HerBS catalogue and were followed up by the BEARS programme [162, 174, 191], from which their spectroscopic redshifts are drawn. Their properties are listed in Table 3, where the peak SNRs show that the line is at best marginally detected in the individual channels and undetected in most of them, which is precisely the regime in which stacking delivers the greatest benefit.

Table 3. Properties of the sources and of their ALMA observations

Source	z_{spec}	RA	Dec	Ang. res. (")	Peak SNR
HerBS-11	2.631	1:24:07.518	-28:14:34.84	0.317	1.500
HerBS-24	2.198	0:47:36.096	-27:29:51.943	2.445	1.106
HerBS-37	2.619	23:26:23.137	-34:26:43.993	0.361	—
HerBS-45	2.434	0:51:32.943	-30:18:49.566	2.366	—
HerBS-63	2.432	0:51:31.700	-30:20:20.383	2.367	0.626
HerBS-77	2.228	0:56:29.253	-31:12:07.431	2.459	0.741
HerBS-80	2.231	23:00:02.546	-31:50:08.808	2.505	—
HerBS-86	2.564	23:53:24.571	-33:11:11.766	2.661	1.008
HerBS-93	2.400	23:47:50.447	-35:29:29.932	0.250	—
HerBS-107	2.553	1:45:20.060	-31:38:32.954	2.731	—
HerBS-135	2.401	22:56:11.799	-32:56:51.802	0.348	—
HerBS-138	1.407	1:17:30.439	-32:07:22.445	1.663	1.262
HerBS-145	2.730	1:23:34.643	-31:46:23.483	0.319	—
—	—	—	—	2.744	—
HerBS-159	2.236	23:51:21.740	-33:29:00.400	2.669	1.756
HerBS-178A	2.658	1:18:50.254	-28:36:43.941	0.318	—
HerBS-178B	2.655	1:18:50.127	-28:36:40.863	0.318	—
HerBS-182	2.227	23:05:38.804	-31:22:05.351	2.471	—
HerBS-184	2.507	23:49:55.665	-33:08:34.370	0.253	—
—	—	—	—	2.544	1.298
HerBS-200	2.151	1:43:13.303	-33:26:33.125	2.725	—
HerBS-208	2.478	22:57:44.560	-32:42:33.566	1.527	—
—	—	—	—	0.350	—
HerBS-209	2.272	22:49:20.518	-33:29:41.137	2.508	—
—	—	—	—	2.802	—

Note. — Columns: spectroscopic redshift, RA, Dec, angular resolution of the observation, and per-channel peak SNR of the CO(3–2) line where measured. The moment zero maps yield a detection for all the targets. Rows with an em dash report additional observations of the source listed above, while an em dash in the last column indicates that the line peak is not detected in the individual channels.

Table 4. Best-fit Gaussian parameters of the stacked CO(3–2) line

Method	μ (km s ^{−1})	σ (km s ^{−1})	A (mJy)	C (mJy)
ViSTA-uvFit	−16.8	264.9	4.58	0.46
ViSTA-ImFit	−52.6	317.8	4.99	0.26
ImSTA-ImFit	3.3	262.5	5.07	0.35

Note. — Central velocity μ , velocity dispersion σ , line amplitude A and continuum level C recovered by each of the three stacking and extraction workflows.

Because the source sizes could be estimated from the moment zero maps, the image domain extraction followed the methodology described in Section 4.3.3.2. For ViSTA-uvFit, the Gaussian model was fitted to the stacked visibilities channel by channel, which does not require the sources to be individually detected. Although each source was selected to be at most partially resolved in its own observation, the stack combines sources of different size observed at different resolutions, so the spatial parameters of the model were kept unconstrained across the channels rather than fixed to the moment zero values, which also allows the spatial extent to vary across the line profile. As in the simulated case, the continuum was not subtracted from the individual sources and all visibilities were stacked together, fitting the full line-plus-continuum profile simultaneously (Sec. 4.2.3.1). This post-stacking choice is appropriate for this sample because, apart from the few targets with a measured peak SNR in Table 3, all sources have low individual SNRs and none exhibits a continuum bright enough to dominate and bias the stacked line profile. Consistently, natural weighting was adopted to maximise the sensitivity towards the faint stacked CO(3–2) line. Following the flux normalisation framework of Section 3.4.2, the visibility amplitudes of each source were rescaled to a common reference redshift $z_{\text{ref}} = 2.5$, close to the median of the sample and consistent with the high redshift DSFG stacking literature [130, 191], through the factor of Equation 29 computed for a flat Λ CDM cosmology.

The reconstructed spectra are presented in Figure 33, with the best-fit Gaussian parameters summarised in Table 4. All three methods, ViSTA-uvFit, ViSTA-ImFit, and ImSTA-ImFit, perform consistently: the line amplitudes agree within about 10 per cent, and the line centroids agree within three channels, a close alignment given the line widths and the possible spatial asymmetries. The continuum levels display larger relative variations, which is expected since the continuum emission sits at a lower SNR than the line. The largest deviation is found for ViSTA-ImFit, whose profile is broader and shifted by about −50

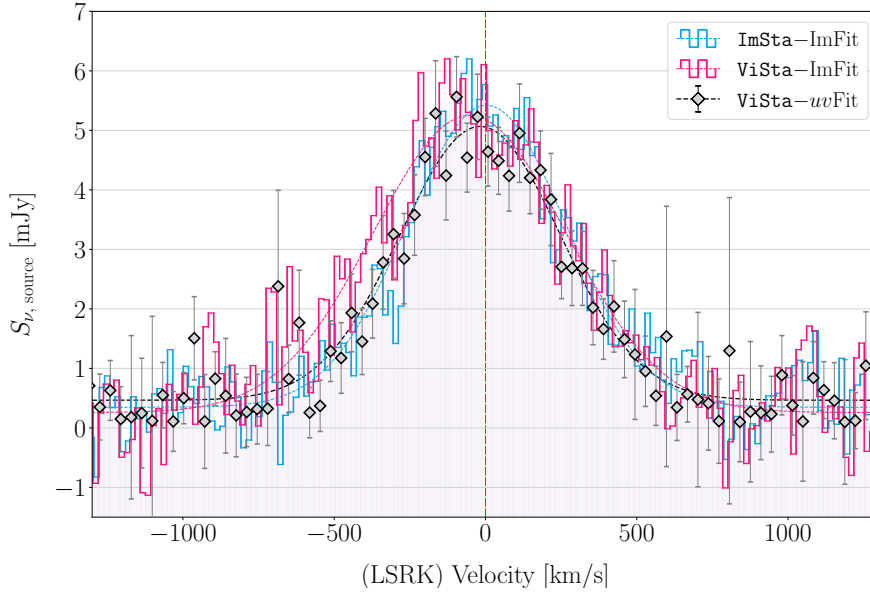


Figure 33: Stacked CO(3–2) spectra recovered with the three methods for the observational sample. Black diamonds show the per-channel flux densities from ViSTA-uvFit with 1σ uncertainties, together with the corresponding best-fit Gaussian. Blue and magenta curves show the ImSTA-ImFit and ViSTA-ImFit profiles respectively, with their best-fit Gaussians as dashed lines.

km s^{-1} . This offset is comparable to the width of the channels, and the width and centroid of a stacked line are less robust than its integrated emission, since the stacked profile is the average of lines of different widths and shapes and its Gaussian fit is sensitive to the few channels in the wings. This mutual agreement is not surprising, since the moment zero maps constrain the source morphology reliably for every target, and it is consistent with previous stacking experiments on local and Galactic sources [184, 192].

The recovered line parameters are in agreement with the literature values [130, 191]. Because the flux corrections were restricted to the luminosity distance rescaling, without incorporating the lensing magnification or an absolute flux normalisation, these results serve as a methodological validation of ViSTA rather than as a definitive astrophysical measurement of the line flux. As an additional test, the continuum was extracted as a post-processing step of the method, directly from the concatenated MS across the line free channels. The reconstructed map is shown in Figure 34 and yields a mean continuum flux density of 0.832 ± 0.064 mJy.

In the sample presented so far the moment zero maps, integrated over the entire spectral window, provide a detection for every target even where the per-channel peak of the line does not, which allows a robust characterisation of the line parameters and explains the mu-

tual consistency of the three methods. A more stringent test can be performed by artificially degrading the sensitivity of the individual observations, so that the sources are no longer detectable even in the integrated maps. Under these conditions ViSTA is expected to recover a larger fraction of the intrinsic signal and to achieve a higher SNR than conventional image stacking, as demonstrated on the simulated extended sources of Section 4.3.3.3. To explore this, one quarter of the available data was selected for each observation and the analysis was repeated, reducing both the uv coverage and the significance of the signal.

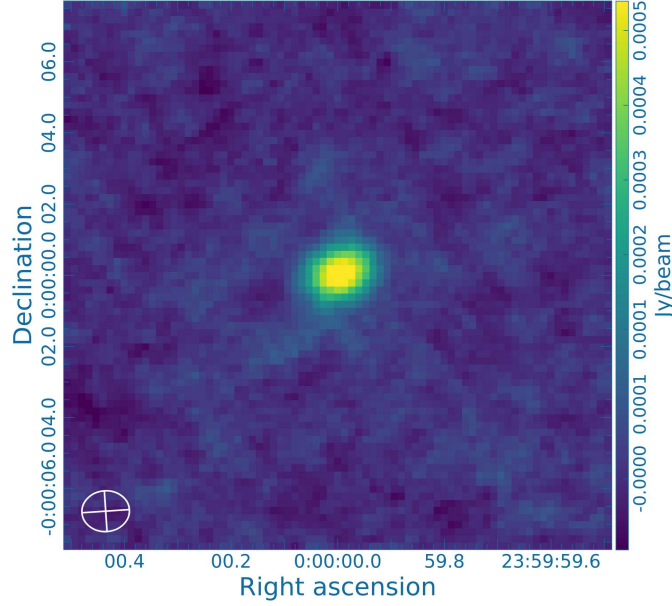


Figure 34: Stacked continuum image synthesised directly from the concatenated ViSTA visibilities across the line free channels, yielding a mean flux density of 0.832 ± 0.064 mJy.

At native resolution, the extraction apertures were defined from the moment zero maps of the full sensitivity observations, as done in the previous case. This test is restricted to the two image-domain workflows, since its aim is to compare stacking before and after imaging with the same extraction method, which also allows a direct visual comparison of the stacked cubes. Moreover, the SNR gain is quantified through the noise measured in the image plane, which is available for both methods only after imaging. Figure 35 shows that, as the sensitivity degrades and the line emission becomes increasingly embedded in the noise, image plane stacking produces a source that appears significantly more scattered and diffuse than the one recovered in the uv plane, so that under conditions approaching the detection limit ViSTA is more effective at capturing the line emission.

A consistent estimate of the actual SNR improvement, however, requires the noise to be evaluated coherently for both methods. Follow-

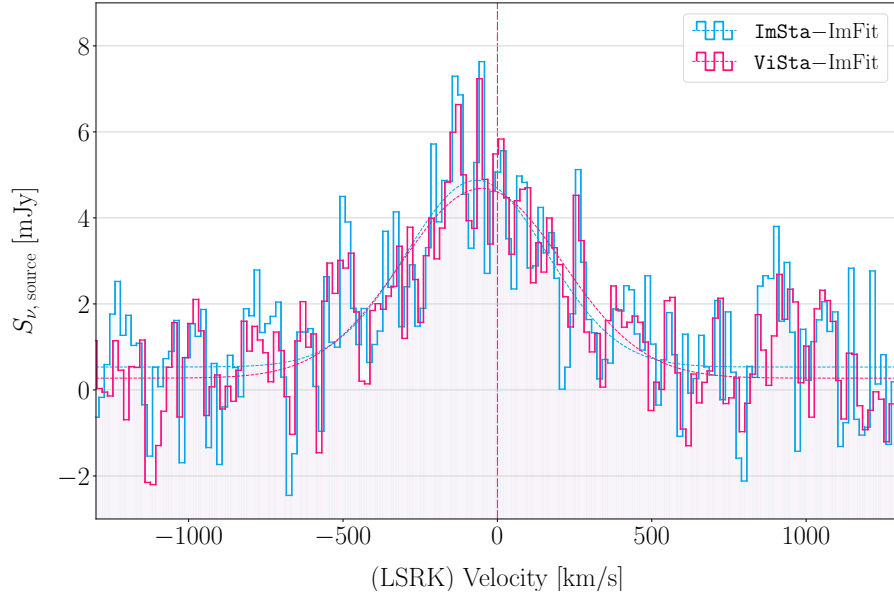


Figure 35: Stacked CO(3–2) spectra obtained with IMSTA-ImFit and ViSTA-ImFit at native resolution after degrading the sensitivity of the individual observations, by retaining one quarter of the data of each. Blue and magenta lines show the IMSTA-ImFit and ViSTA-ImFit profiles respectively, with their best-fit Gaussians as dashed lines. The image plane spectrum displays larger scatter and noise fluctuations, while the visibility domain stacking recovers a smoother profile.

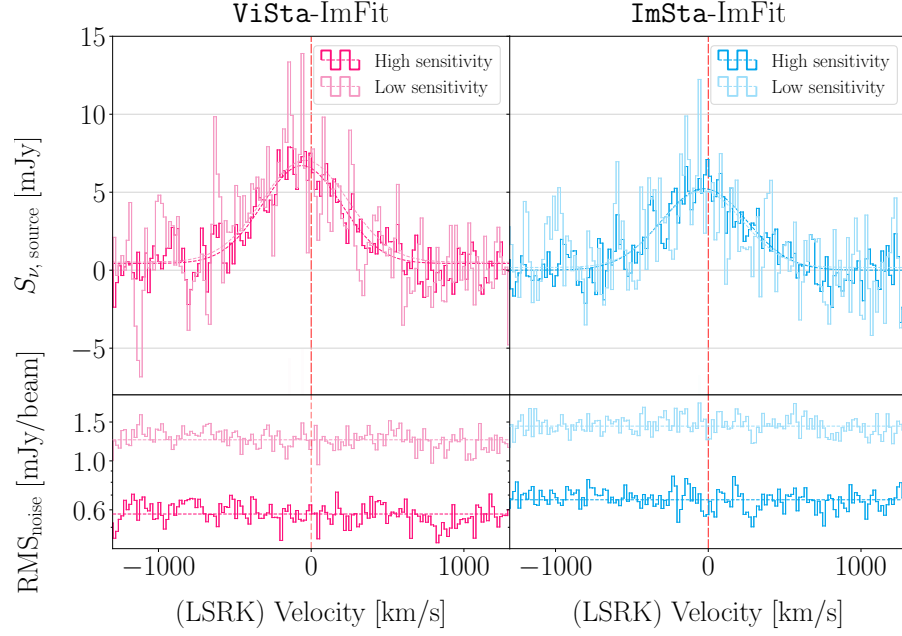


Figure 36: Comparison between the two stacking methods for datasets of different sensitivity. *Top:* flux densities extracted at the original and at the degraded sensitivity, the latter in lighter colours, with the best-fit Gaussians as dashed lines; the left panel refers to ViSta-ImFit and the right one to ImSta-ImFit. *Bottom:* corresponding channel RMS noise, evaluated across the line free channels after matching all datasets to a $5.063'' \times 2.820''$ beam. At both sensitivity levels the visibility domain stacking yields a higher peak flux and a lower noise scatter.

ing the procedure of Section 4.3.3.3, all datasets were therefore convolved to the worst angular resolution available, corresponding to a synthesised beam of $5.063'' \times 2.820''$. Figure 36 displays the resulting spectra and channel noise profiles for both sensitivity regimes. At the original sensitivity, the visibility domain stacking achieves a more effective noise suppression, raising the SNR from 7.8 with ImSTA-ImFit to 10.9 with ViSTA-ImFit, an improvement of about 39 per cent. Under the degraded sensitivity the SNR rises from 3.4 to 5.2, corresponding to a gain of about 53 per cent. The advantage of the visibility domain therefore grows as the data approach the noise dominated regime, consistent with the simulated case, and it stems from the ability of the method to preserve the phase coherence of the astronomical signal while avoiding the incoherent addition of noise that affects each image before image plane stacking.

4.5 WRAP-UP

This chapter presented the first implementation of ViSTA, built within the CASA environment and tailored to ALMA data, showing how the framework of Chapter 3 translates into an operational workflow. Starting from calibrated datasets, the pipeline brings every source onto a common spatial and spectral reference, combines them into a single stack, and extracts the resulting signal either in the image domain or directly in the uv plane. The implementation was validated on two simulated samples of 50 ALMA sources, representing unresolved and extended morphologies, and subsequently applied to a sample of 21 galaxies observed in the CO(3–2) transition. Across all tests, the stacked signal was reliably recovered by the three compared workflows, namely ViSTA-uvFit in the Fourier domain, ViSTA-ImFit in the image domain, and the traditional image plane stacking ImSTA-ImFit.

Together, the synthetic and observational benchmarks demonstrate that the advantage of visibility domain processing becomes increasingly pronounced as the data approach the noise floor. While both domains achieve comparable results for bright compact targets, ViSTA significantly outperforms image plane techniques when measuring faint or extended emission, recovering a larger fraction of the intrinsic flux with a lower noise level. The results on the real CO data reinforce this behaviour, showing that the performance gap widens as the data quality deteriorates. The method therefore proves particularly advantageous for analysing the shallow and heterogeneous archival datasets that motivate this work.

Although this implementation confirms the scientific viability of visibility stacking, its dependence on the CASA framework creates computational bottlenecks that scale poorly with expanding sample sizes. Each step of the workflow writes a new, complete copy of the MS to disk, including all its subtables and auxiliary columns, and it is

this repeated copying of the whole dataset, rather than the operations themselves, that dominates the execution time. These limitations directly motivate the optimised architecture introduced in the following chapter.

ViSTA 2.0: THE OPTIMISED FRAMEWORK

The initial implementation and validation of the ViSTA method presented in the previous chapter proved that visibility domain stacking is able to recover faint signals across diverse source shapes, redshifts, and array setups, and that its advantage over image domain methods grows as the observation sensitivity decreases. However, that first version scales poorly with sample size, making the analysis of large archives impractical, and it was designed and built primarily for ALMA data, since it relies entirely on the native software of the observatory. To overcome these bottlenecks, I developed a new optimised computational architecture. The mathematical operations remain identical to those described in Chapters 3 and 4, ensuring that the scientific behaviour of the method is fully preserved. Only the underlying processing engine changes, eliminating intermediate disk operations, reducing the memory footprint, and accelerating execution speed by orders of magnitude.

This chapter describes how the method has been reengineered to remove the performance limitations. Section 5.1 first analyses in detail the bottlenecks of the CASA-based version. Section 5.2 then presents the new self-contained pipeline of ViSTA 2.0, built around a lazy, streaming access to the data that replaces the sequential, disk-bound execution model of the original tool. Section 5.3 describes the new visibility-domain procedure for the extraction of the stacked flux density, and Section 5.4 validates the performance and the scientific correctness of the new implementation against the original version.

5.1 LIMITATIONS OF THE CASA-BASED IMPLEMENTATION

The original implementation of the ViSTA pipeline was built entirely upon the CASA package. While this modular structure successfully proved the scientific viability of visibility stacking, execution within CASA introduces fundamental computational and architectural bottlenecks.

A primary limitation stems from how CASA handles data persistence across consecutive processing steps. In that version the data are never loaded into Python arrays as a whole. Each

stage is either a CASA task, such as `casatasks.phaseshift()`, `casatasks.mstransform()` or `casatasks.concat()`, or a custom routine that edits the tables of an MS in place through the CASA table tools. Each pipeline stage is executed as an independent task call, requiring a full copy of the MS to be written to disk before the subsequent operation can begin. Although intermediate subproducts can be deleted programmatically, this sequential writing pattern generates extreme I/O overhead and multiplies local disk storage requirements by a factor proportional to the number of processing steps. Furthermore, CASA retains complete table structures during every duplication, copying calibration tables, pointing records, and auxiliary metadata columns that carry no functional value for visibility stacking. It is this repeated copying of the whole dataset, rather than the operations themselves, that dominates the execution time.

A further constraint comes from the architecture of CASA itself. The performance-critical routines are implemented in a compiled C++ core, `casacore`, but they are accessed through a Python interface designed for the interactive reduction of individual datasets. The parallelisation exposed at this level relies on MPI and is therefore tailored to distributed memory systems across cluster nodes, which is limited in scope within CASA and not guaranteed for custom visibility manipulation routines [188]. Deployment presents additional obstacles, as standard distributions depend on graphical libraries that are absent on headless cluster nodes, while modular pip wheels introduce fragile dependency chains that break easily during system environment updates.

These inefficiencies also undermine the core computational advantage of the visibility domain. While image domain stacking requires N individual deconvolutions for a sample of N observations, visibility stacking needs only a single deconvolution on the final combined dataset. However, this theoretical gain is completely negated if the pre-processing, intermediate disk operations, and memory overhead exceed the cost of those N image reconstructions.

These structural limitations motivate the second implementation of ViSTA, presented in the following section.

5.2 A SELF-CONTAINED PIPELINE

At the heart of the redesign is the choice to access the MS through `dask-ms` [193] rather than through the native CASA table system. `dask-ms` reads the tables of an MS and presents them as ordinary array structures in Python without loading their content directly into memory. Each column becomes an `xarray` [194] dataset backed by Dask [195] arrays, creating a multi-chunk structure where each segment can be read and processed independently. Operations

applied to such arrays are recorded into a computation graph rather than executed immediately, delaying input and output tasks until a `task.compute()` call is triggered. At that stage, the graph is evaluated, enabling the scheduler to read chunks in parallel, distribute workloads across available cores, and fuse successive operations into a single pass over the dataset. This lazy evaluation model restructures the entire procedure into a continuous stream where reading, processing, and writing occur in overlapping stages, replacing the sequential task execution that forced the original pipeline to serialise every step on disk.

Before any visibility is touched, the pipeline performs a sequential pass over the input datasets to build the common output grid and to assemble the static structures of the output MS. The list of datasets to be stacked is supplied through a plain text file in which every line specifies the path to the file, the source redshift, the sky coordinates of the target and, optionally, the identifiers of the field and of the spectral window to be selected within that file. Because each meta-data read is small and dominated by filesystem latency rather than by computation, the reads for all input datasets are dispatched concurrently through a `ThreadPoolExecutor`, a pool of worker threads scheduled by Python that execute independent tasks in parallel and collect their results once all of them complete. For every observation the `SPECTRAL_WINDOW`, `FIELD` and `DATA_DESCRIPTION` subtables are read to recover the per-channel frequency array, the phase-centre direction and the mapping between data description and spectral window for the selected field.

From these arrays the code derives, for each input dataset, the rest-framed frequency interval $[\nu_{lo}(1+z), \nu_{hi}(1+z)]$ and the rest-framed channel width $\Delta\nu_{rf} = \Delta\nu_{obs}(1+z)$. The common output channel width is then fixed to the widest of these values, $\Delta\nu_{new} = \max_i \Delta\nu_{rf,i}$, so that no dataset is ever oversampled beyond its native resolution, and the output frequency grid of each dataset is placed within its rest-framed interval, centred on the requested reference frequency or shifted to the nearest edge when the ideal window does not fit.

The static output subtables are then constructed entirely in memory. The `ANTENNA` table of the output is built by concatenating the antenna positions of all input datasets, each rescaled by $1/(1+z)$ as required by the rest-framing, and by assigning to every input dataset a contiguous block of antenna identifiers. The remaining fundamental subtables, namely `SPECTRAL_WINDOW`, `DATA_DESCRIPTION`, `FIELD`, `FEED`, `POLARIZATION` and `SOURCE`, are likewise assembled at this stage, with one spectral window per input dataset defined on its own output grid. The output is therefore a single combined MS, with one main table holding the visibilities of all sources and a single set of subtables, in which each input dataset is identified by its own spectral window and data description. The spectral windows are deliberately kept separate

rather than merged into a single window, since keeping them distinct is convenient for the subsequent analysis and merging them would bring no benefit either in memory footprint or at the imaging stage.

5.2.1 Building the stacked dataset with *dask-ms*

Once the preparatory pass is complete, the effective pipeline is organised as a small pool of specialised threads that communicate through bounded queues and follow a read, compute and write pattern, so that the three families of operations proceed concurrently and the latency of one is hidden behind the work of the others. The overall organisation of these threads is sketched in Fig. 37.

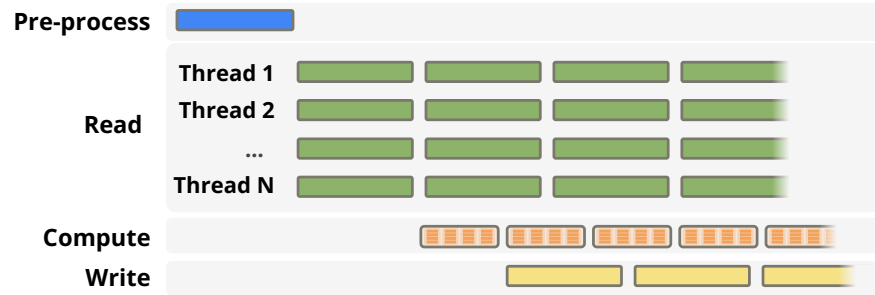


Figure 37: Threading scheme of the streaming pipeline. A configurable pool of reader threads (Thread 1 to Thread N_{readers}) opens subsets of the input datasets and materialises the visibilities, flags and UVW coordinates through a single `dask.compute()` call per spectral window. A single compute thread applies the rest-framing, centering and rebinning in the C++/OpenMP kernel, and a single writer thread flushes each processed chunk to the output MS through `xds_to_table()`. Read, compute and write stages communicate through bounded queues and proceed concurrently, so that the input and output latency is hidden behind the computation.

A configurable number of reader threads opens subsets of the input datasets and executes a single `dask.compute()` call per spectral window. This operation materialises the visibilities, the FLAG column, the UVW coordinates and a minimal set of scalar columns directly into standard NumPy arrays in a single graph evaluation. By evaluating these components together rather than column by column, the pipeline eliminates the overhead of repeatedly constructing lazy eval-

uation graphs and allows the scheduler to overlap the chunked reads, keeping several read graphs in flight and maintaining a continuous flow to the downstream processing stages.

A single dedicated execution thread manages the subsequent data manipulation. This thread dequeues each loaded block and passes it to the optimised core routines, which perform the rest framing, phase centering, and spectral rebinning in a single in-memory pass. A single management thread is sufficient because the parallelism is provided by OpenMP within each kernel call, which already saturates the `OMP_NUM_THREADS` cores available with a static schedule. Issuing concurrent kernel calls from several compute threads would instead oversubscribe the physical cores, degrading the performance through cache contention and scheduling overhead.

Upon completing chunk processing, a writer thread flushes the data to the single combined output MS via `xds_to_table()`, performing the inverse `dask-ms` operation to write xarray datasets back into the table columns chunk by chunk. A single writer thread is employed because CASA format subtables do not support concurrent writes across multiple streams without risking race conditions or table corruption, and because a single stream already saturates the available write bandwidth. As illustrated in Fig. 37, this architecture keeps reading, computation, and writing tasks continuously active, effectively concealing input and output latencies behind active processing.

5.2.2 *The compute core*

The compute core, invoked by the compute thread described above, is where the rest-framing, centering and rebinning of every visibility are actually carried out. It replaces in a single pass the steps that the first implementation performed with separate CASA tasks, namely the centering of Section 4.2.1.1, the rest-framing of Section 4.2.1.2, and the rebinning and regridding of Sections 4.2.1.3 and 4.2.1.4. It is implemented as a compiled kernel that processes the baseline rows in parallel, and it can optionally be offloaded to a GPU when one is available.

5.2.2.1 *The C++/OpenMP kernel*

The kernel is implemented as a pybind11 [196] extension module compiled with GCC and parallelised with OpenMP [197], a standard that distributes the iterations of a loop across the cores of a single machine, all of which share the same physical memory. It is the natural choice for the operations at hand, since each baseline row is processed independently of the others, which makes the workload embarrassingly parallel. The rows are distributed across the available threads to balance the load while keeping the synchronisation overhead negligible,

and for every row the kernel performs the rest-framing and centering of the UVW coordinates, the associated phase shift of the visibilities and the spectral rebinning onto the common grid.

For each row, the baseline vector is first scaled by $1/(1+z)$, which maps the observed coordinates onto the rest-frame spatial frequency, and is then rotated from the original phase centre (α_0, δ_0) to the common stacking position (α_1, δ_1) through a rotation matrix R built from the two pointing directions,

$$\begin{pmatrix} u_{\text{rf}} \\ v_{\text{rf}} \\ w_{\text{rf}} \end{pmatrix} = R \cdot \frac{1}{1+z} \begin{pmatrix} u_{\text{obs}} \\ v_{\text{obs}} \\ w_{\text{obs}} \end{pmatrix}. \quad (30)$$

Each visibility is then multiplied by the complex fringe corresponding to the same shift of the phase centre,

$$V_{\text{shifted}}(\nu) = V_{\text{obs}}(\nu) \cdot \exp \left[-2\pi i \frac{\nu}{c} (u_{\text{rf}} \Delta l + v_{\text{rf}} \Delta m + w_{\text{rf}} \Delta n) \right], \quad (31)$$

where Δl , Δm and Δn are the direction cosine differences between the two phase centres. This is a geometric operation, not a calibration step. The phase change depends on the baseline and is proportional to frequency, so it differs from channel to channel, and it is applied to each channel independently, as in `casatasks.phaseshift()`, on whose implementation the kernel is modelled. Across a spectral window of total bandwidth B centred at ν , the phase change varies by about $2\pi (B/\nu) (\theta_0/\theta_{\text{syn}})$ for a source offset by θ_0 from the phase centre, with θ_{syn} the synthesised beam. Applying a single phase per window would therefore smear the signal across the band as soon as the source lies several beams away from the phase centre. To avoid a trigonometric evaluation for each channel, the phase is updated incrementally through a recurrence relation: the per-channel phase increment $\Delta\phi \cdot \Delta\nu$ is precomputed once, and the phase of every subsequent channel is obtained by multiplying by the complex phasor step $= e^{i\Delta\phi \cdot \Delta\nu}$.

The phase-shifted visibilities are finally projected onto the common rest-frame output grid through a channel overlap weighting scheme identical to that of `casatasks.mstransform()` in frequency mode,

$$V_{\text{out}}(j) = \frac{\sum_k V_{\text{shifted}}(k) w_{kj}}{\sum_k w_{kj}}, \quad (32)$$

with weights

$$w_{kj} = \frac{\max(0, \min(\nu_k^+, \nu_j^+) - \max(\nu_k^-, \nu_j^-))}{\Delta\nu_k}, \quad (33)$$

where $\nu_k^\pm$ and $\nu_j^\pm$ denote the lower and upper edges of the input channel k and of the output channel j , and $\Delta\nu_k$ is the input channel

width, computed per channel as a centred finite difference rather than as a global median. Flagged channels are excluded from the interpolation, which falls back on the single valid neighbour when only one of the two is available, while output channels receiving no contribution at all are flagged in turn.

Although the visibilities are stored in the MS in single-precision `complex64`, the fringe phase of Eq. (31) involves products of baseline lengths and frequencies whose magnitude makes a double-precision representation essential, since evaluating it in single precision would introduce rounding errors large enough to partially decorrelate the stacked signal. The kernel therefore casts the visibilities to `float64` before any phase multiplication, and the conversion back to the native format happens only when the output is written.

5.2.2.2 GPU acceleration

The new version of ViSTA can optionally exploit a GPU, a processor built to execute the same instruction on thousands of data elements at once, which suits a workload where every row undergoes an identical sequence of operations. Support is enabled at build time through the dedicated `-DWITH_CUDA` compilation flag, and in its absence the extension module carries no CUDA [198] dependency whatsoever. When a device is available, the complete per-row pipeline, comprising the rest-frame correction of the UVW coordinates, the phase shift and the regridding, is offloaded to CUDA.

The organisation of the work differs from the CPU case. Whereas on the CPU each dataset is processed in turn and the OpenMP threads divide the rows of a single dataset among the cores, the GPU version packs a batch of N_{batch} input datasets into a single set of contiguous device buffers and launches one kernel that covers all of them at once. A single CUDA thread is assigned to each pair formed by a dataset and one of its rows, for a total of $N_{\text{batch}} \times N_{\text{row}}$ threads running concurrently, and the geometry and frequency-grid parameters of all the datasets in the batch are uploaded together and processed in parallel on the device. Each kernel launch is bracketed by the host-to-device and device-to-host transfers of its batch, and while the device processes the current batch the host continues to read and pre-process the next one, so that the data transfers and the input and output overlap with the GPU computation.

5.2.3 Post-processing

The post-processing operations, from the weighting of the visibilities to the recovery of the stacked flux, no longer act on the large and heterogeneous collection of input datasets but on the single combined MS produced by the pipeline. Since each input dataset keeps its own spectral window in this MS, every source remains addressable indi-

vidually, and each operation can be applied either to the whole stack or source by source. The data access relies on the table tools of the `casacore` framework, used only to read the columns of the combined MS one spectral window at a time and to update in place the visibility weights of each source. For operations as simple as these, accessing the data through `casacore` or through a dedicated framework such as `dask-ms` is equivalent and brings no real advantage. All the numerical operations proper, such as the weight renormalisation, the continuum subtraction and the flux extraction, are then carried out directly on the arrays returned by these reads.

The weighting and the flux normalisation follow the schemes described in Sections 3.4.1 and 3.4.2 and implemented in Chapter 4, but they are applied differently here. Both the weight renormalisation and the flux normalisation factors are never applied to the visibilities stored on disk. Instead, they are computed once and applied only at the extraction stage, to the compressed statistics of Section 5.3, so that the combined MS always keeps its native amplitudes and the original files are never corrupted, and the same stacked dataset can be reused to build different subsamples and to run separate tests without ever being rewritten. This also removes the need for the two parallel datasets of Section 4.2.2, since the noise recorded in the combined MS is always the native one and no unnormalised copy has to be carried through the analysis.

The continuum subtraction can be performed directly on the stacked dataset at the moment of the flux extraction, through the optional constant term of the fit described in Section 5.3, which models the continuum of the stack as a whole. On the other hand, if the scientific case requires the subtraction before the stacking, the operation is still performed on the combined MS, source by source. Indeed, as described above, the spectral windows of the input datasets are deliberately kept separate in the output, so that each source remains addressable through its own data descriptor and can be treated independently even after the combination.

The subtraction follows the same logic as the standard `casatasks.uvcontsub()` task, fitting a low-order polynomial to the line-free channels separately for the real and the imaginary part of the visibilities and then subtracting the fitted model from every channel. The fit is performed individually, spectral window by spectral window, and hence source by source, reading the visibilities in chunks of rows so that even a large stack never has to be held in memory at once. This is therefore a per-source subtraction and not a fit of the stacked continuum. The line-free channels are selected automatically. The channels within a velocity window centred on the target line are excluded from the fit, and this window is enlarged to also exclude any other known line that falls within the band. Flagged and empty samples are excluded as well, so that none of

them biases the estimated continuum. The order of the polynomial is a constant by default, which avoids introducing a spectral slope that the narrow line-free bandwidth cannot constrain, but a first-order term can be enabled, either fixed or selected automatically for each spectral window by comparing the two orders through their Bayesian information criterion, so that a linear continuum is preferred only when the data genuinely support it.

The fitted continuum is always written to the `MODEL_DATA` column of the MS, while the line visibilities with the continuum removed are written to the `CORRECTED_DATA` column and the original `DATA` column is left untouched, as in the standard use of the MS columns. Line quantities can therefore be measured on the subtracted column and continuum quantities on the line-free channels of the original one, and the whole step can be rerun on the same data.

5.3 FLUX DENSITY EXTRACTION

After the post-processing operations are complete, the stacked signal can be recovered. The final dataset is a standard interferometric product, so in principle the stacked flux density could be extracted with the same two approaches already used in Sections 4.2.3.4 and 4.2.3.3, that is either by imaging the final MS with `casatasks.tclean()` and measuring the flux on the restored map, or by fitting a parametric source model directly to the visibilities with a tool such as `UVMultiFit` [199]. However, both routes become problematic on a stack of highly heterogeneous and wide datasets. Imaging with `casatasks.tclean()` inherits the problem of the deconvolution step, and when the sample spans a wide range of angular resolutions and only partially overlapping uv coverages, the reconstruction is governed by an ill-defined hybrid beam, so that forcing a common resolution correlates the noise and dilutes the compact flux (Sec. 3.5.1). Fitting the visibilities with `UVMultiFit`, on the other hand, processes everything in memory, so that large concatenated stacks require a preliminary time averaging. This averaging reduces the number of rows only by the ratio between the averaging interval and the integration time, and it must be kept short to avoid time-average smearing (Sec. 4.2.3.3), so that a stack of many observations still leaves millions of rows to be fitted, and the fit has to be repeated in full for every test or resampling.

For these reasons, the stacked flux density is instead recovered through a new visibility-domain procedure I implemented in this pipeline, bypassing the image plane and never loading the full visibility set into memory. Compact and extended configurations sample the same visibility function $V(b)$ at different baseline lengths and jointly constrain the total flux and the source size without any loss of resolution. The extraction relies on a compact set of sufficient statis-

tics computed once per source, and all subsequent fits and error estimations run on this set.

5.3.1 *Compression into sufficient statistics*

The extraction begins by compressing the visibilities of each source into a compact set of sufficient statistics. As in the standard radial binning of visibilities in the uv plane, for every spectral channel the visibilities are binned radially into logarithmically spaced annuli of projected baseline length in the rest frame, and for each annulus five accumulators are stored, namely the weighted sum of the real part of the visibility, the weighted sum of the imaginary part, the sum of the weights, the weighted sum of the baseline length and the number of visibilities that fell in the bin. It is at this stage that the weighting and the flux normalisation factors (Sec. 5.2.3) enter, so that the compressed statistics already carry the correct relative weighting and the common flux scale while the MS on disk keeps its native values. Once these quantities have been accumulated, every subsequent operation, from the spectra to the radial profiles, the model fits and the error estimation, runs on this compact representation rather than on the full visibility set, which is what makes the repeated resampling described below inexpensive. This compression is a more aggressive form of averaging than time averaging, which reduces the number of rows only by the ratio $t_{\text{avg}}/t_{\text{int}}$ between the averaging interval and the integration time. Here the whole stack is reduced to $N_{\text{sources}} \times N_{\text{channels}} \times N_{\text{bins}} \times 5$ numbers, which makes it inexpensive to repeat the full chain for every bootstrap realisation. If the model depends only on the baseline length, as for a circular source at the phase centre, and varies little across each annulus, the weighted sums within each annulus are sufficient statistics for the fit, and no information relevant to the model is lost. The residual error, due to the variation of the model within an annulus, is reduced by the logarithmic spacing of the bins and by evaluating the model at the weighted mean baseline length of each annulus. Sufficiency is lost instead for elliptical or lensed morphologies and for sources that are not correctly centred, for which the radial profile is only an approximation. Each entry in the input list, corresponding to one MS, field and spectral window, is compressed as an independent descriptor, and the several sets of accumulators produced by different observations of the same source are combined together into a single measurement before the population fit. As a by-product, the imaginary part of the stacked signal provides a null test. For a source correctly centred on the phase centre and symmetric, the imaginary part vanishes, so that a significant imaginary component would signal a residual offset, due for instance to inaccurate coordinates, or an asymmetric morphology.

The phase centring of every source is therefore verified individually before the stacking.

5.3.2 Line profile and uv-plane flux fitting

The compressed statistics of the individual sources are combined into the stack through a weighted average, retaining only the radial bins to which at least a minimum number of sources contribute, so that no point in the profile is dominated by a single, potentially unrepresentative source. The stacked profile is then fitted directly in the visibility plane in three steps, illustrated in Fig. 38. First, the line shape is determined by extracting a spatially integrated and one-dimensional spectrum, obtained by averaging the real part of the visibilities across all annuli, and fitting it by weighted nonlinear least squares over a window wide enough to capture the full profile without truncating its wings. The real part is used intentionally throughout, both for the spectrum and for the radial profile, rather than the amplitude of the visibilities. For a source correctly centred, the vector-averaged amplitude coincides with the real part, and its noise bias vanishes at high SNR. However, the SNR relevant here is that of each annulus and channel, which is often low, especially at long baselines and in the line wings. There the amplitude would be positively biased by the Rician noise (Sec. 4.3.2) and would include the noise of the imaginary part, whereas the real part remains unbiased and keeps the noise symmetric, so that it can also become negative. The price is that the real part requires an accurate centring, since any residual phase moves signal into the imaginary part. A single profile is used for an isolated line and two profiles are fitted together for a blended pair, and only the shape parameters (the centroid and the width of each component) are retained from this step. Second, with the line shape held fixed, the line amplitude alone is then fitted independently in each annulus by weighted linear least squares, so that the flux in each annulus follows analytically from the fitted amplitude as

$$F(b) = A(b) \sigma \sqrt{2\pi}, \quad (34)$$

where σ is the fixed line width. Fixing the shape isolates the line component of a blend analytically, avoiding the truncation and the contamination between components that a fixed velocity window would introduce, and an optional constant term can be included to absorb any residual flat continuum. Third, the resulting profile of flux against baseline length is fitted with a circular Gaussian source model,

$$V(b) = F_{\text{tot}} \exp \left[-\frac{(\pi \theta_{\text{FWHM}} b)^2}{4 \ln 2} \right], \quad (35)$$

by weighted nonlinear least squares, which yields the total flux $F_{\text{tot}} = V(b = 0)$ and the effective source size θ_{FWHM} . When the stacked pro-

file shows no significant decrease with baseline length, namely when the stacked source is unresolved at the available resolution, the Gaussian is replaced by a fixed point-source model, which avoids introducing a size parameter that the data cannot constrain and that a free Gaussian fit would tend to overfit. The fit also allows the size to be constrained rather than left completely free, either by fixing θ_{FWHM} or by adding a soft prior that penalises solutions far from a reference size, for instance the size measured on the continuum of the same band. As a model-independent cross-check, the weighted average of the flux over all annuli is also computed, which is the estimate that a point-source model would give and which does not rely on the extrapolation to zero baseline. The continuum flux is obtained in the same way, from the weighted average of the line-free channels, together with the effective frequency of each band.

5.3.3 *Uncertainties and null tests*

All the uncertainties on these quantities are obtained by bootstrap resampling of the source sample. In each bootstrap realisation, N sources are drawn with replacement from the N sources of the sample, so that some sources appear more than once and others not at all. The accumulators of the drawn sources are recombined, and the whole chain is repeated, from the combination of the statistics through the line shape, the amplitudes in each annulus and the uv -plane fit. The uncertainty on each parameter is the standard deviation of its distribution across the N_{boot} realisations, while the 16th and 84th percentiles are also recorded to check for asymmetries. Since the realisations differ in which sources enter the stack, this uncertainty includes both the thermal noise and the intrinsic source-to-source variance, and it therefore represents the error on the population mean, which is the relevant quantity for a stack.

Sources that share the same execution block also share their calibration errors (Sec. 3.1.4), and can optionally be resampled together as a single unit, a scheme known as block bootstrap. For samples with few sources or blocks, roughly ten or less, the number of distinct resampling configurations is intrinsically limited, and the bootstrap distribution underestimates the true population variance regardless of the number of realisations. In these cases the bootstrap standard deviation should be regarded as a lower limit on the uncertainty rather than as a definitive estimate.

As a final null test, the imaginary part of the visibilities is processed through the same chain as the real part, and it is verified to be consistent with zero within the bootstrap uncertainty of the real signal, confirming both that the sources are correctly centred and that the stack is free from significant contamination.

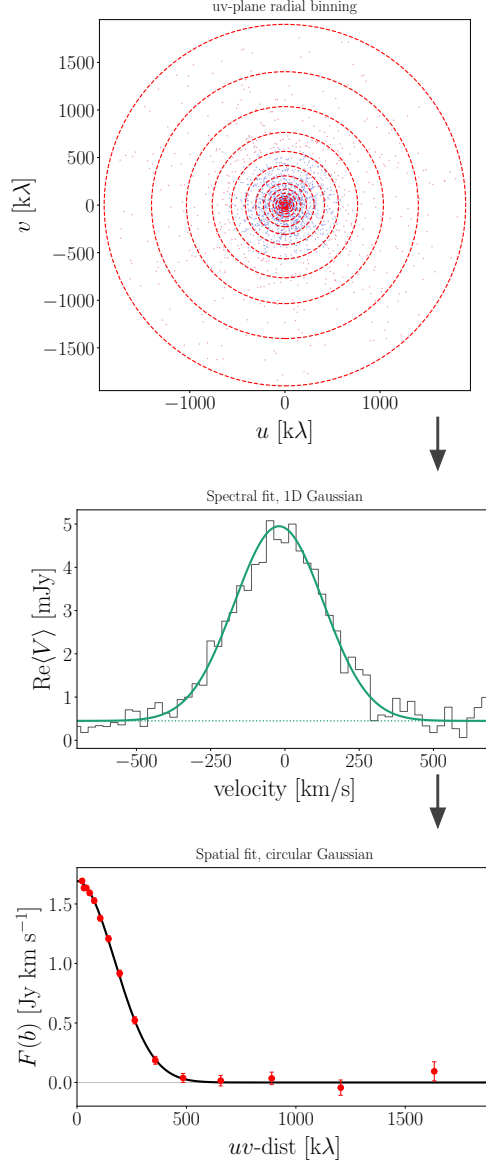


Figure 38: Illustration of the visibility-domain flux extraction on a representative stacked source, following the order of the three steps described in the text. *Top:* the visibilities are binned radially into logarithmically spaced annuli in the uv plane. *Middle:* the spatially integrated 1D spectrum (grey histogram) with the fitted line profile (green curve). Only the shape parameters, the centroid and the width, are retained from this fit and then held fixed. *Bottom:* with the shape fixed, the line amplitude alone is fitted annulus by annulus, and the resulting profile of the velocity-integrated flux of the real part as a function of uv distance b (red points, with bootstrap uncertainties) is fitted with a circular Gaussian source model (black curve), whose zero-baseline value gives the total flux $F_{\text{tot}} = V(b=0)$ and whose width gives the effective source size θ_{FWHM} .

5.4 PERFORMANCE VALIDATION

The new pipeline was tested against the original CASA-based implementation on the same sample and the same hardware, a single node of the Leonardo pre-exascale system at CINECA equipped with an NVIDIA A100 GPU and a multi-core Intel Xeon CPU. On the standard configuration adopted as a reference, the CASA-based pipeline required a total wall-clock time of 570.1 s, whereas the new CPU-only implementation completed the same task in $\sim 42\text{--}43$ s, a speed-up of about $13.6\times$. The full parameter scan underlying this figure, together with the description of the computing infrastructure and of the mock sample used for the tests, is reported in Appendix B.

Beyond the raw execution time, the two implementations were verified to produce the same physical result. The output spectral resolution agrees to within 0.01% between the two pipelines, the mean RMS noise of the stacked product agrees to within about one per cent, and the same holds for the noise level measured after imaging the two stacked datasets, with channel-to-channel differences consistent with the expected statistical fluctuations of a local RMS estimate rather than with a systematic bias. This confirms that the reengineering described in this chapter has changed only the computational engine, and not the scientific behaviour of the method, as intended.

The scaling of the new pipeline with the main configuration parameters, namely the number of input datasets, the number of output channels, the number of reader threads, and the degree of parallelism of the compute kernel, shows a consistent picture. The CPU configuration is comfortably I/O-bound, with the cumulative time spent reading the input datasets dominating the wall-clock time and the compute kernel contributing only a secondary fraction except at low thread counts, where the compute thread becomes slow enough to starve the reader and writer threads. The GPU configuration, on the other hand, is writer-bound: the compute kernel is so fast that the single writer thread saturates well before the available GPU throughput is exploited, so that additional reader threads or larger GPU batches bring no further improvement once the writer becomes the limiting stage. The CPU-only configuration is already highly efficient for the data volumes considered here, while the GPU path is expected to become increasingly advantageous for larger samples, where the compute savings accumulate, or on workstations with fewer CPU cores available.

5.5 WRAP-UP

This chapter presented the high-performance reimplementations of ViSTA, which removes the computational bottlenecks of the CASA-based version of Chapter 4 while leaving the method of

Chapter 3 unchanged. The sequential, disk-bound execution of the original tool is replaced by a streaming model built on `dask-ms`, where reader, compute and writer threads work concurrently and the per-row rest-framing, centering and rebinning are carried out by a compiled C++/OpenMP kernel, optionally offloaded to a GPU. The post-processing and the extraction then act on the single combined MS, whose native amplitudes are never overwritten, and the stacked flux is recovered directly in the uv domain from a compact set of sufficient statistics.

Tested on the same sample and hardware, the new pipeline is more than an order of magnitude faster than the CASA implementation. The stacked noise level remains consistent to within less than one percent, confirming that only the processing engine has changed. Full benchmark results are reported in Appendix B. The natural next step for this work is to abandon the MS format altogether, as it represents the main bottleneck for data writing, thereby eliminating all remaining dependencies on CASA.

With both implementations of the method now in place, I am finally in a position to address the scientific case that motivated this work. In the next chapter I therefore apply the high-performance pipeline described here to the ensemble assembled in Chapter 2, and use it to stack the archival visibilities and reconstruct the average CO SLED of a sample of DSFGs.

THE AVERAGE CO SLED OF DSFGs AT COSMIC NOON

In the preceding chapters I have built the foundational components of this work step by step, establishing the sample of DSFGs and developing the visibility domain stacking method ViSTA, together with its software implementations. In this chapter I bring these elements together in the first scientific application of ViSTA, namely the reconstruction and the interpretation of the average CO SLED of this population, whose diagnostic power was introduced in Section 1.2.2.

Section 6.1 describes the final sample used in this analysis, obtained by cross-matching the parent *Herschel* ensemble with archival ALMA data containing CO measurements, and it details the classification of targets into lensed and unlensed systems while presenting their key observational properties. Section 6.2 explains how the stacking is carried out with ViSTA, the specific choices adopted for this dataset and how the stacked signal is extracted from the combined visibilities. Section 6.3 presents the resulting average CO SLED and compares the lensed and unlensed subsamples directly. Section 6.4 models the stacked ladder with radiative transfer calculations to derive the average physical properties of the molecular gas. Finally, in Section 6.5 I compare the results of this analysis with other studies in the literature.

6.1 THE ALMA SAMPLE

I start from the parent sample of 202 *Herschel*-selected sources built in Chapter 2, selected with a flux threshold of $S_{500} \geq 80$ mJy and requiring a secure spectroscopic redshift. To specialise this selection for the present study, I search within the ALMA Science Archive public observations covering these targets. I query the archive using *astroquery* [200] through a dedicated program described in Appendix A.2, which automatically identifies the lines present in ALMA observations and builds the corresponding catalogues. For each candidate, I search for all observations where at least one spectral window covers a CO transition in the rest frame. This search yields sufficient statistics for transitions from CO(3 – 2) up to CO(9 – 8), whereas outside these limits fewer than five targets are available in each of the lensed and unlensed subsamples. Such a small number could still allow a stacked detection of the line, but not a meaningful compar-

ison between the two subsamples, which is one of the goals of this analysis. In addition to a secure spectroscopic redshift, an accurate measurement of the source position is a fundamental prerequisite for performing a stacking analysis. I therefore retain only those sources for which at least one archival ALMA observation of their continuum emission allows me to precisely determine the spatial coordinates.

This selection yields 104 sources with at least one CO observation, corresponding to 99 distinct *Herschel* detections. Table 5 lists all the selected targets alongside their observed CO transitions. I assign to each source an identifier from S_{001} to S_{104} , which is used throughout the analysis. The lensing classification is described in detail in Section 6.1.1, while Section 6.1.2 presents the distribution of the sample across its physical and observational properties.

Table 5: ALMA sample of *Herschel*-selected DSFGs.

ID	Survey ID	Target	z	CO transitions (J_{up})	L/U
S_{001}	HATLASJ000007.4-334059	HerBS-42B	3.314	4, 5, 6, 9	U
S_{002}	–	HerBS-42A	3.307	4, 5, 6	U
S_{003}	HATLASJ000124.9-354212	HerBS-41A	4.098	4, 5, 6, 7, 9	L
S_{004}	HATLASJ000330.6-321136	HerBS-155A	3.077	4, 5, 8	L
S_{005}	HATLASJ000455.4-330812	HerBS-170	4.182	5, 6, 7	U
S_{006}	HATLASJ000745.8-342014	HerBS-163A	3.140	4, 5	U
S_{007}	HATLASJ000806.8-351205	HerBS-117A	4.526	5, 7	L
S_{008}	HERMESJ001615.8+032433	HELMS51	2.765	7, 9	L
S_{009}	HERMESJ001626.0+042613	HELMS6	2.509	7, 9	L
S_{010}	HATLASJ001802.2-313505	HerBS-106	2.369	3, 6, 7	L
S_{011}	HATLASJ001838.7-354133	HerBS-104B	1.536	3	U
S_{012}	HATLASJ002054.6-312752	HerBS-81A	3.160	4, 7	U
S_{013}	–	HerBS-81B	2.588	3, 5, 6, 7	U
S_{014}	HELMSJ002220.9-015524	HELMS34	5.161	5, 6	L
S_{015}	HATLASJ002624.8-341737	HerBS-22A	3.050	4, 5	L
S_{016}	HATLASJ003207.7-303724	HerBS-56C	2.561	3, 5	U
S_{017}	HATLASJ003717.0-323307	HerBS-122A	2.883	5, 9	L
S_{018}	HELMSJ003814.1-002252	HELMS24	4.984	5, 6	L
S_{019}	HELMSJ003929.6+002426	HELMS11	2.848	3	U
S_{020}	HERMESJ004410.2+011821	HELMS28	4.162	8, 9	L

continued on next page

Table 5 – continued

ID	Survey ID	Target	z	CO transitions (J_{up})	L/U
S ₀₂₁	HELMSJ004622.3+073509	HELMS46	2.577	9	U
S ₀₂₂	HERMESJ004714.1+032453	HELMS5	2.292	5, 6, 7, 8, 9	L
S ₀₂₃	HERMESJ004723.3+015749	HELMS49	1.441	5, 6, 7	L
S ₀₂₄	HATLASJ004736.0-272951	HerBS-24	2.198	3, 4, 7, 8	L
S ₀₂₅	HATLASJ004853.3-303110	HerBS-57	3.265	4, 6	L
S ₀₂₆	HATLASJ005132.0-302011	HerBS-63A	2.432	3, 7, 8, 9	U
S ₀₂₇	HATLASJ005132.8-301848	HerBS-45A	2.434	3, 7, 8, 9	U
S ₀₂₈	HATLASJ005506.5-300027	HerBS-207	1.569	6, 9	L
S ₀₂₉	HATLASJ005629.6-311206	HerBS-77A	2.228	3, 4, 7	U
S ₀₃₀	HATLASJ005659.4-295039	HerBS-90A	3.992	4, 7	L
S ₀₃₁	HATLASJ005724.2-273122	HerBS-60	3.261	4, 6	L
S ₀₃₂	HATLASJ011014.5-314814	HerBS-160	3.955	4, 6	L
S ₀₃₃	HATLASJ011424.0-333614	HerBS-27	4.509	5, 7	L
S ₀₃₄	HERSJ011722.3+005624	HERS10	2.469	9	L
S ₀₃₅	HATLASJ011730.3-320719	HerBS-138	1.407	3	U
S ₀₃₆	HATLASJ012222.3-274456	HerBS-120A	3.125	4, 5, 9	U
S ₀₃₇	–	HerBS-120B	3.124	4, 5, 9	U
S ₀₃₈	HATLASJ012335.1-314619	HerBS-145A	2.730	3, 5	U
S ₀₃₉	HATLASJ012407.3-281434	HerBS-11	2.631	3, 5, 8	L
S ₀₄₀	HERSJ012521.0+011724	HERS13	2.476	9	L
S ₀₄₁	HATLASJ012853.0-332719	HerBS-73	3.026	5	U
S ₀₄₂	HERSJ013212.2+001754	HERS18	1.693	7	L
S ₀₄₃	HATLASJ013239.9-330906	HerBS-40	1.971	4, 8	U
S ₀₄₄	HATLASJ013840.4-281855	HerBS-14	3.782	4, 6	L
S ₀₄₅	HATLASJ014313.2-332633	HerBS-200	2.151	3, 4, 7, 8	U
S ₀₄₆	HATLASJ014520.0-313835	HerBS-107	2.553	3	L
S ₀₄₇	HERSJ020941.2+001558	Hers1	2.553	3, 4, 5, 7, 8, 9	L
S ₀₄₈	HERMESJ021831-053131	xmm06	3.395	6, 8, 9	L
S ₀₄₉	HATLASJ083051.0+013225	G09v1.97	3.634	6, 9	L
S ₀₅₀	HATLASJ085358.9+015537	G09v1.40	2.089	6, 9	L
S ₀₅₁	HATLASJ090302.9-014127	SDP.17	2.308	7, 9	L

continued on next page

Table 5 – continued

ID	Survey ID	Target	z	CO transitions (J_{up})	L/U
S ₀₅₂	HATLASJ090311.6+003907	SDP.81	3.042	3, 9	L
S ₀₅₃	HATLASJ090453.2+022017	HerBS-183	1.891	7	L
S ₀₅₄	HATLASJ090740.0-004200	SDP.9	1.577	6	L
S ₀₅₅	HATLASJ091043.0-000322	SDP.11	1.786	7	L
S ₀₅₆	HATLASJ113526.2-014606	G12v2.43	3.128	8, 9	L
S ₀₅₇	HATLASJ113803.6-011737	HerBS-96	3.150	4	U
S ₀₅₈	HATLASJ120127.6-014043	–	3.729	4	L
S ₀₅₉	HATLASJ125632.4+233627	NCv1.143	3.565	9	L
S ₀₆₀	HATLASJ125652.5+275903	HerBS-31	2.790	3, 5	U
S ₀₆₁	HATLASJ132630.1+334407	NAv1.195	2.951	3, 7	L
S ₀₆₂	HATLASJ132859.2+292326	NAv1.177	2.778	3	L
S ₀₆₃	HATLASJ133008.5+245900	NBv1.78	3.111	9	L
S ₀₆₄	HATLASJ141351.9-000026	G15v2.235	2.478	3, 7, 9	L
S ₀₆₅	HATLASJ141810.0-003747	–	2.241	3	U
S ₀₆₆	HATLASJ142413.9+022303	G15v2.779	4.243	7, 9	L
S ₀₆₇	HERMESJ142824+352620	J1428	1.325	7, 8	L
S ₀₆₈	HATLASJ145135.2-011418	–	2.588	3, 9	U
S ₀₆₉	HATLASJ223615.2-343301	HerBS-121A	3.741	4, 6, 8	L
S ₀₇₀	HATLASJ223753.8-305828	–	2.719	3, 5, 8, 9	L
S ₀₇₁	HATLASJ223942.4-333304	HerBS-111	2.371	3, 6, 7	U
S ₀₇₂	HATLASJ224759.7-310135	HerBS-141	2.085	3, 4, 7	L
S ₀₇₃	HATLASJ224920.6-332940	HerBS-209A	2.272	3, 7	U
S ₀₇₄	HATLASJ225045.5-304719	HerBS-168A	2.583	3, 5, 6, 9	L
S ₀₇₅	HATLASJ225250.7-313657	HerBS-47	2.433	3, 4, 7	L
S ₀₇₆	HATLASJ225324.2-323504	HerBS-103	2.942	3, 5, 7, 8	U
S ₀₇₇	HATLASJ225339.1-325550	HerBS-131B	2.197	3, 4, 7, 8	U
S ₀₇₈	HATLASJ225600.7-313232	HerBS-189	3.300	4, 5, 6	L
S ₀₇₉	HATLASJ225611.7-325653	HerBS-135A	2.401	3, 6, 7	U
S ₀₈₀	HATLASJ225744.6-324231	HerBS-208A	2.478	3, 4, 7	U
S ₀₈₁	–	HerBS-208B	2.483	3, 4, 7	U
S ₀₈₂	HATLASJ230002.6-315005	HerBS-80A	2.231	3, 4, 7	U

continued on next page

Table 5 – continued

ID	Survey ID	Target	z	CO transitions (J_{up})	L/U
S ₀₈₃	–	HerBS-80B	1.968	4, 5, 6	U
S ₀₈₄	HATLASJ230538.5-312204	HerBS-182	2.227	3, 4, 7	U
S ₀₈₅	HATLASJ230815.5-343801	HerBS-28	3.925	4, 6	L
S ₀₈₆	HATLASJ231205.2-295027	HerBS-132	2.473	3, 7	U
S ₀₈₇	HATLASJ232210.9-333749	HerBS-146B	2.003	3, 4, 5, 6	U
S ₀₈₈	HATLASJ232419.8-323926	HerBS-18	2.182	3, 4, 6, 7	L
S ₀₈₉	HERMESJ232439.4-043934	HELMS45	2.473	5, 6, 7, 8, 9	L
S ₀₉₀	HATLASJ232623.0-342642	HerBS-37	2.619	3, 5, 6	L
S ₀₉₁	HELMSJ232831.8-004035	HELMS55	2.283	9	L
S ₀₉₂	HATLASJ232900.6-321744	HerBS-39	3.229	4, 6, 7, 8	L
S ₀₉₃	HATLASJ233024.1-325032	HerBS-102A	3.287	4, 6, 8	U
S ₀₉₄	HATLASJ233037.2-331217	HerBS-123	2.170	3, 4, 6, 7	L
S ₀₉₅	HERMESJ233255.5-031134	HELMS44	2.690	6, 7, 8, 9	L
S ₀₉₆	HERMESJ233255.7-053424	HELMS61	2.402	5, 6, 7, 8, 9	L
S ₀₉₇	HERMESJ233620.7-060826	HELMS59	3.435	8, 9	L
S ₀₉₈	HERMESJ234051.3-041937	HELMS62	3.503	7, 8, 9	L
S ₀₉₉	HATLASJ234418.1-303936	HerBS-21A	3.323	4, 5, 6, 9	L
S ₁₀₀	HATLASJ234750.5-352931	HerBS-93	2.400	3, 4, 6, 7	U
S ₁₀₁	HATLASJ234955.7-330833	HerBS-184	2.507	3, 7	U
S ₁₀₂	HATLASJ235324.7-331111	HerBS-86	2.564	3, 5, 6, 7	U
S ₁₀₃	HATLASJ235623.1-354119	HerBS-36	3.095	4, 5, 8	L
S ₁₀₄	HATLASJ235827.6-323244	HerBS-25	2.912	3, 5, 7	L

Columns: the source identifier (ID) adopted throughout the analysis; the *Herschel* survey identifier (Survey ID); the target name; the spectroscopic redshift (z); the observed CO rotational transitions; and the lensing classification (L/U, see Sect. 6.1.1). The CO transitions are listed by their upper level J_{up} , so that each entry corresponds to the transition $\text{CO}(J_{\text{up}} - J_{\text{up}} - 1)$ and, for example, $J_{\text{up}} = 3$ denotes $\text{CO}(3 - 2)$. A transition covered by more than one archival observation is listed only once. Sources sharing a single *Herschel* detection resolved by ALMA into multiple components have the Survey ID reported only for the first component.

6.1.1 Lensing classification

Given the sample size, I adopt a binary classification scheme, distinguishing between *confirmed lensed systems* (L) and *unconfirmed systems* (U). Many targets benefit from repeated observations and dedicated

lens-modelling analyses that provide independent evidence of their gravitationally lensed nature. For all other sources lacking dedicated lens models, the classification relies on high-angular-resolution interferometric images drawn from the ALMA archive, including the ANGELS pilot survey for HerBS sources [159] alongside other published high-resolution studies and lens models [167, 201–206].

A source is assigned to class L when its (sub-)millimetre morphology exhibits unambiguous signatures of strong gravitational lensing, such as arcs, complete or partial Einstein rings, or counter-images opposite a candidate deflector (Sec. 1.1.3). Where available, I check this assignment against the independent classification of Borsato et al. [156], based on multi-band *HST* imaging with foreground deflector modelling and subtraction, finding agreement for all shared sources.

Conversely, a source is assigned to class U when visual inspection reveals a single compact object with no obvious signs of lensing. This class is intentionally heterogeneous: it encompasses genuinely unlensed DSFGs, sources resolved into physically associated multiple components, and systems that may be weakly lensed or possess image separations below the resolution limit of the available data. The apparent flux density alone cannot resolve this ambiguity, since the intrinsic luminosities of the population span a range comparable to typical lensing magnifications. Consequently, this population should be understood as “not confirmed to be lensed” rather than a definitive statement that lensing is absent. Splitting the sample any further would leave subsamples too small to reach a sufficient SNR in the stacking analysis and to maintain statistical robustness.

Of the 104 targets, 64 are classified as L and 40 as U. The two classes differ systematically in apparent brightness, since the confirmed lensed systems have a median flux density of $S_{500} = 140$ mJy against 93 mJy for the unconfirmed ones, in agreement with the theoretical number counts of Negrello et al. [153] discussed in Section 2.1.1. For simplicity, throughout the rest of this analysis I refer to the unconfirmed category simply as the *unlensed* sample.

6.1.2 Sample distributions

The intrinsic heterogeneity of the archival dataset—and the distribution of its statistical weight across the sample—is best appreciated through the observational properties summarised in Figs. 39 to 42. Figures 39 to 41 describe the sample as a whole, while Figs. 42 and 43 separate the lensed and unlensed subsamples and the two arrays.

Figure 39 illustrates the redshift distribution of the sample. Spanning $z \sim 1.3$ to 5.2, the sources cluster around a median redshift of $\bar{z} = 2.6$, placing the bulk of the population squarely at cosmic noon.

Figure 40 presents the distribution of angular resolutions categorised by array configuration. The two ALMA arrays cover very

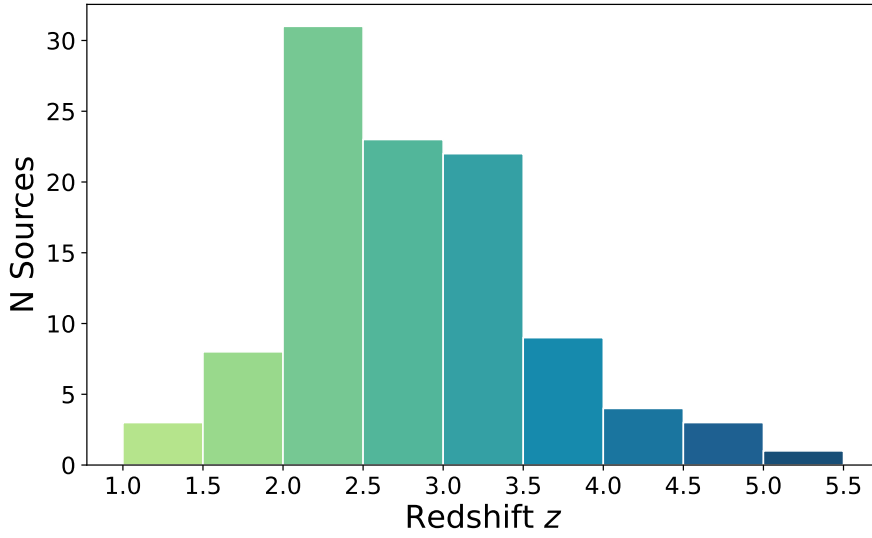


Figure 39: Redshift distribution of the 104 components of the sample. The population clusters around the median $\bar{z} = 2.6$, placing the bulk of the sources at cosmic noon.

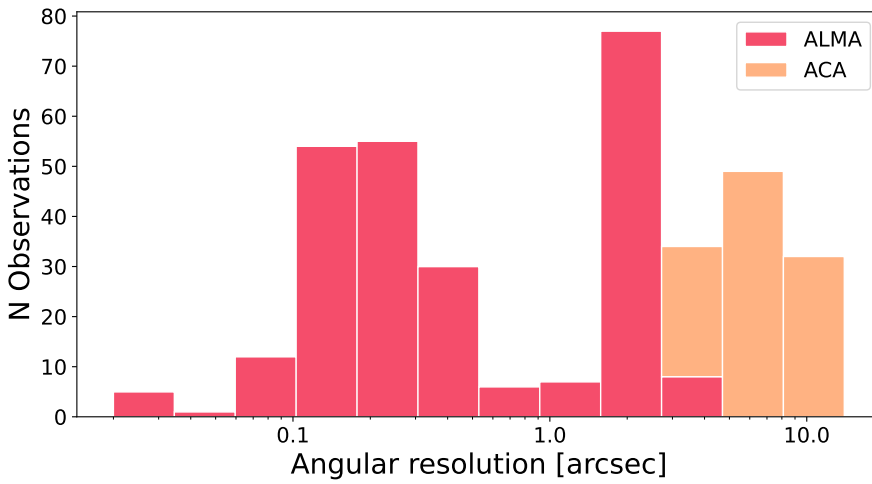


Figure 40: Distribution of the angular resolution of the archival observations, separated into the 12 m ALMA array and the 7 m ACA. The two arrays populate distinct resolution regimes, spanning from ~ 0.02 to ~ 13 arcsec.

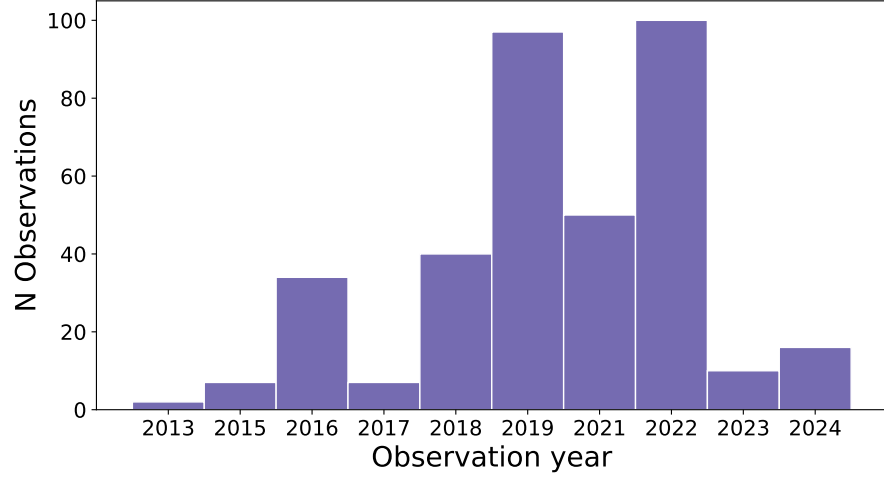


Figure 41: Number of observations as a function of observing year, illustrating how the sample accumulates across ALMA Cycles 2 to 11 (2013–2024).

different spatial resolutions, as the 12m array reaches high subarc-second resolution, whereas the ACA provides coarser resolution spanning several arcseconds. This is precisely the challenge ViSTA is designed to solve. Indeed, attempting to perform the stacking in the image plane would require degrading the high resolution data down to the coarsest beam size in the sample, often by more than an order of magnitude, erasing fine spatial detail and rendering the joint combination of 12m ALMA and ACA observations, or even 12m data taken across vastly different configurations, completely impractical.

Figure 41 displays the distribution of observations as a function of observing year. Spanning over a decade of ALMA operations (2013–2024), the dataset reflects both the steady growth of the archive and the frequent re-observation of these bright, compact targets across independent science programmes.

Figure 42 breaks down the number of unique CO observations per transition by lensing status, array configuration, and their cross-combination. Observational coverage is non-uniform and decreases toward higher J_{up} transitions, which fall in the high-frequency ALMA bands and require the rarest weather conditions at the site [207]. As expected, the unlensed subsample is systematically less represented than the lensed one. Finally, Fig. 43 shows the angular resolution of each observation as a function of the transition. These last two figures show that at the lowest transitions both subsamples are covered by the ACA and by compact 12m configurations. In contrast, at $J_{\text{up}} = 8$ and 9 the unlensed sources lack ACA observations completely and rely almost entirely on 12m data at resolutions of $\sim 0.1''$. This re-

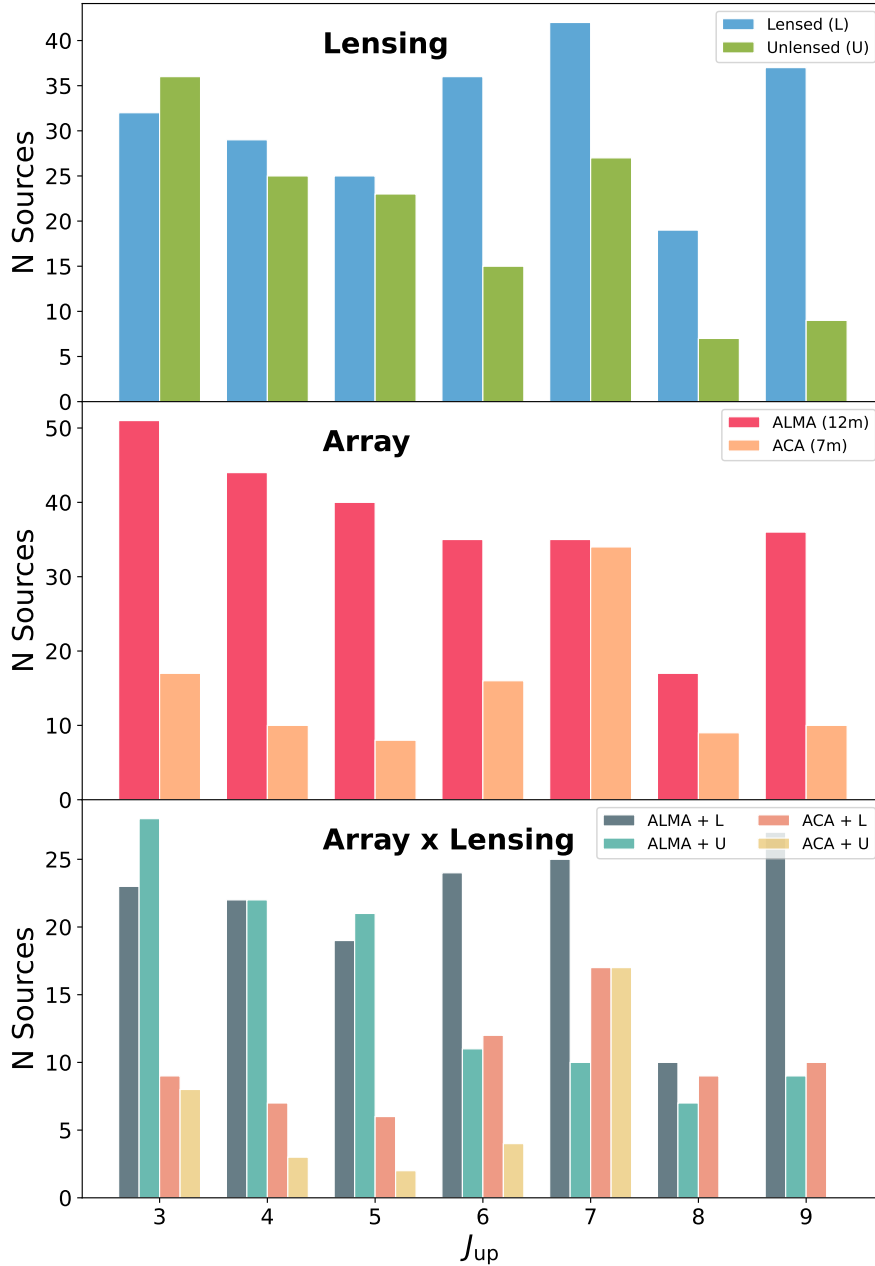


Figure 42: Number of unique CO observations per transition, split by lensed/unlensed status (top), by ALMA/ACA array configuration (middle), and by their cross-combination (bottom). The ACA + unlensed combination is entirely absent for CO(8 – 7) and CO(9 – 8), reflecting the lack of ACA follow-up of unlensed sources at these high- J transitions.

duces the source count and the SNR for these points, making them more vulnerable to spatial flux filtering (Section 6.3.1).

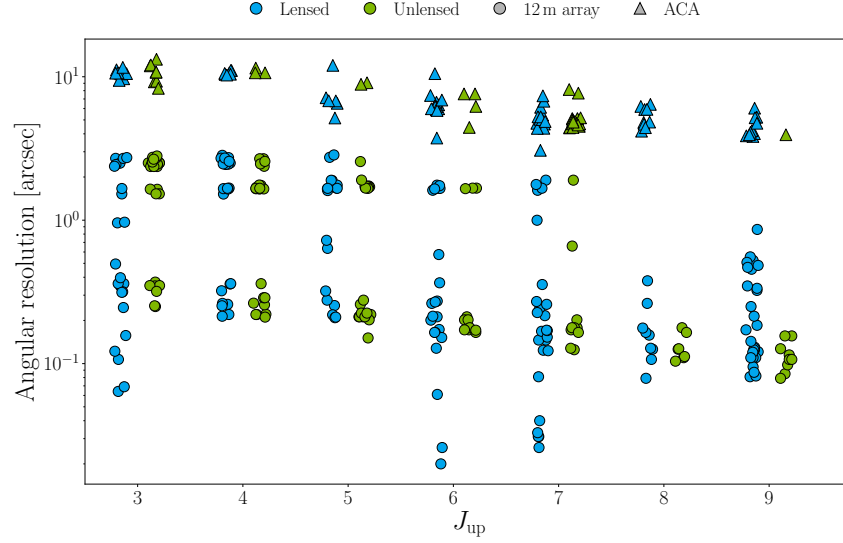


Figure 43: Angular resolution of the archival observations contributing to each CO transition, separated into lensed (blue) and unlensed (green) sources, with circles for the 12 m array and triangles for the ACA. At $J_{\text{up}} = 8$ and 9 the unlensed sources are covered almost exclusively by high-resolution 12 m observations, whereas the lensed ones retain ACA coverage.

6.2 STACKING WITH VISTA

The average CO SLED is reconstructed by stacking the calibrated archival visibilities directly in the uv plane with the high-performance implementation of ViSTA described in Chapter 5. The full formalism of the method has already been established in the preceding chapters, and I refer to Chapter 3 for the theoretical foundations of visibility stacking; this section therefore concentrates only on the choices that are specific to the CO SLED analysis, namely the weighting scheme, the normalisation procedure and the extraction technique used to retrieve the stacked signal.

6.2.1 Data retrieval and spectral window selection

The stacking is performed independently for each CO transition. For every rotational line I build one combined MS gathering all the observations that cover that transition in the rest frame, so that the ladder from CO(3 – 2) to CO(9 – 8) is reconstructed as a sequence

of independent stacks. The calibrated data, the inputs of the ViSta pipeline, are obtained through the calibrated MS service of the European ALMA Regional Centre network at ESO, which delivers calibrated MSs for specific Member Observation Unit Sets upon request. For the oldest observations I instead rely on the Additional Representative Images for Legacy (ARI-L) project, which restored the calibration delivered with the original data and reimaged the observations from Cycles 2 through 4 with modern pipeline standards [186], bringing them to a degree of homogeneity comparable to the recent cycles. This effort processed about 88.5% of the eligible early cycle data, exceeding its initial 70% target, and it delivered more than 3,000 reprocessed observation sets and over 150,000 images. I incorporate these reprocessed products whenever early cycle data enter the sample.

Within each combined MS, ViSta shifts the phase centre of every observation to the nominal position of the corresponding source, transforms the frequency axis to the source rest frame, and regrids all observations to the widest velocity resolution present in that transition (Section 3.3). Each individual observation is retained as a separate spectral window of the combined MS, with no averaging across sources at this stage, which preserves the ability to apply per-source corrections and to resample the dataset for the statistical tests described in the following sections (Section 5.2.3).

I select a velocity range of $\pm 2500 \text{ km s}^{-1}$ around the CO rest frequency (or the maximum range allowed by the raw data) to ensure complete line coverage and leave sufficient line-free channels to constrain the underlying continuum. In most cases, achieving this required combining two consecutive spectral windows from the original MS. I also discard the observations where the line falls so close to the window edge that line-free channels are available on one side only. In the stack, such profiles would be truncated, and they would bias both the continuum subtraction and the resulting stacked line flux.

Within these same spectral windows, two further transitions frequently fall very close to the target CO line and are often detected in the individual observations, so that they must be accounted for in the analysis. The $[\text{CI}](^3P_2-^3P_1)$ fine-structure line, at a rest frequency of 809.342 GHz, lies close to CO(7 – 6) at 806.652 GHz, so that the two are almost always observed together within the same spectral window. Similarly, the $\text{OH}^+(1_1-0_1)$ line at 1033.119 GHz falls near CO(9 – 8) at 1036.912 GHz. In high-redshift starbursts this line is commonly observed in absorption against the dust continuum, often with blueshifted or redshifted components that trace outflows or inflows of diffuse gas [208–211]. Because these lines sit within the fitted velocity range of the corresponding CO transition, they are relevant both for the continuum subtraction, where their emission must be excluded from the line-free channels (Section 6.2.4), and for the flux extraction,

where they are deblended from the CO line and measured jointly with it (Section 6.2.5).

6.2.2 *Weighting scheme: Democratic*

A key decision in the stacking procedure is determining the relative weight assigned to each source. As introduced in Section 3.4.1, ViSTA supports both inverse-variance (natural) weighting, which maximises the formal SNR, and democratic weighting, which equalises individual contributions across the sample regardless of integration time. Natural weighting is optimal only when uniform observations sample comparable intrinsic emission with similar sensitivity. However, this archival dataset is inherently heterogeneous, with integration times and velocity resolutions spanning more than an order of magnitude. In this regime, natural weighting would allow a small subset of deeply observed and/or most strongly-lensed targets to dominate the combined signal, and the result would reflect those few objects rather than be a representative population average. I therefore adopt democratic weighting across different targets, rescaling visibility weights so that every source contributes equally to the stack. Because the CO transitions analysed here are intrinsically bright, the dataset can afford to trade a portion of formal SNR for a genuinely representative population average.

To further validate this choice, I test its impact directly on the CO(3 – 2) stack of the full sample, including both lensed and unlensed sources, using non-normalised data to observe the effect in its purest form. In the visibility plane (Fig. 44), the natural-weighting fit is dominated by the single most sensitive source in the sample, causing the stacked profile to trace that individual object rather than the ensemble average and failing to yield a well-behaved profile even at the shortest baselines. Conversely, democratic weighting distributes contributions evenly, recovering a smooth profile to which all sources contribute comparably. This behaviour is even more pronounced in the image plane (Fig. 45), where the naturally weighted stack displays a highly irregular pattern heavily biased toward deeply observed, strongly lensed targets, making any reliable reconstruction impossible, whereas the democratically weighted stack remains well behaved.

6.2.3 *Flux normalisation: rescaling by infrared luminosities*

Two forms of heterogeneity affect the visibility amplitudes and must be corrected before averaging. Both corrections follow the framework presented in Section 3.4.2, to which I refer for their full derivation.

The first correction accounts for the wide redshift range of the sample ($z \simeq 1.5\text{--}5$), over which identical intrinsic line emission appears

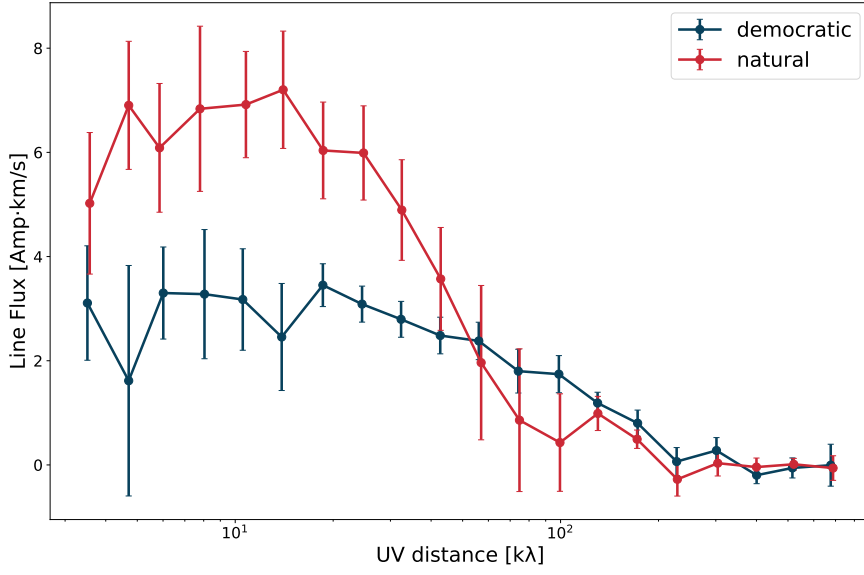


Figure 44: Fit in the uv -plane comparing the natural and democratic weighting methods on the CO(3 – 2) stack, with the most sensitive source of the sample also shown for reference. Under natural weighting, the fit is entirely dominated by this single high-sensitivity source, whereas democratic weighting recovers a well-behaved profile representative of the whole sample.

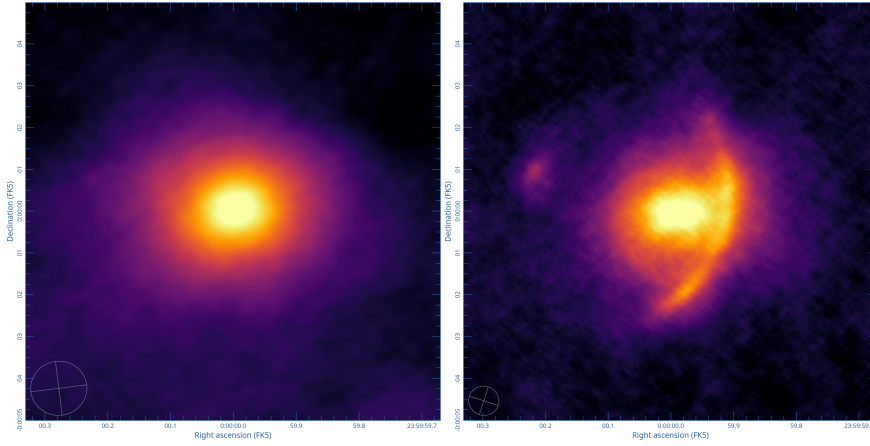


Figure 45: Image-plane comparison of the two weighting schemes for the CO(3 – 2) stack: democratic weighting (left) and natural weighting (right). The naturally weighted image is highly irregular and biased toward the deeply observed, strongly lensed objects, while the democratic image is well behaved.

with amplitudes differing by nearly an order of magnitude purely due to variations in luminosity distance. Each source is shifted using the factor α_i from Eq. 29. Although the actual median redshift of the sample is 2.6, I adopted $z_{\text{med}} = 2.5$ for all transitions and subsamples to maintain consistency with other stacking studies used for comparison, and this choice ensures that the stacked amplitude corresponds to a well-defined line luminosity.

The second correction addresses the fact that the sample spans more than an order of magnitude in luminosity (L_{IR}). Because CO line luminosity scales roughly linearly with L_{IR} , unrescaled bright objects would otherwise dominate the stack and distort the excitation ladder, reflecting the luminosity distribution of contributing sources rather than their intrinsic gas excitation. I therefore normalise each source to a common reference luminosity ($L_{\text{IR,ref}}$), so that all rescaled data points describe an equivalent source at fixed luminosity, allowing line ratios to be interpreted directly as excitation conditions. The visibility amplitudes are multiplied by $L_{\text{IR,ref}}/L_{\text{IR},i}$, yielding a total scaling factor for source i of:

$$f_i = \alpha_i \frac{L_{\text{IR,ref}}}{L_{\text{IR},i}}. \quad (36)$$

Ideally, this normalisation would use the CO(1 – 0) luminosity, which directly traces the total molecular gas mass rather than a proxy linked to star formation and dust heating. However, the archive lacks sufficient CO(1 – 0) coverage across the sample, making L_{IR} the most viable practical alternative [following 191]. As outlined in Section 3.4.2, this luminosity normalisation remains optional and both normalised and non-normalised stacks are presented in Section 6.3, as their comparison provides valuable insight into differential magnification between lensed and unlensed subsamples.

The infrared luminosities L_{IR} used in the normalisation are apparent values (uncorrected for lensing magnification μ) and are derived homogeneously across the sample using a single, internally consistent approach. Each source is fitted with a single-temperature, optically thin MBB of the form:

$$S_\nu \propto \nu^\beta B_\nu(T_{\text{dust}}), \quad (37)$$

where $B_\nu(T_{\text{dust}})$ is the Planck function and β is the dust emissivity index. To maintain a uniform constraint on the Rayleigh–Jeans tail across the sample, I adopt a fixed observed wavelength window of $100 \mu\text{m} \leq \lambda_{\text{obs}} \leq 2100 \mu\text{m}$. This window consistently includes the *Herschel*/SPIRE 250, 350, and 500 μm bands sampling the SED peak, alongside any available longer-wavelength photometry (SCUBA-2, SMA, MAMBO, ALMA). The free parameters T_{dust} and the SED normalisation are fitted via maximum likelihood, accounting for the effect of the CMB on both dust heating and observed contrast

following Cunha et al. [88]. Photometric uncertainties are combined in quadrature with a 15% error floor, non-detections enter the likelihood as censored data, and detections deviating by more than 5σ are iteratively clipped. Uncertainties on T_{dust} and L_{IR} are propagated through Monte Carlo resampling of the flux densities, and L_{IR} is obtained by integrating the best-fit MBB over the rest-frame FIR range. The complete table containing the individual line fluxes for all sources is provided in Appendix C.1.

The emissivity index is fixed to $\beta = 2$ for all sources rather than left free, following the analyses of Bendo et al. [174] and Hagimoto et al. [191]. Free- β fits to these samples show that roughly half of the sources are consistent with $\beta = 2$, with a mean of 2.0 ± 0.3 . In particular, observed-frame 151/101 GHz flux density ratios—which trace the Rayleigh–Jeans slope independently of single-temperature fitting and are largely insensitive to T_{dust} —show no trend with redshift and remain consistent with $\beta \simeq 2$ across the entire sample. This indicates that the apparent redshift dependence of β in free- β fits is an artifact of the temperature- β degeneracy, which worsens at high redshift as the observed 250–350 μm bands sample progressively warmer dust. Fixing $\beta = 2$ thus yields a more robust and internally consistent baseline for sample-wide comparisons.

Systems with multiple, blended components unresolved by *Herschel* are treated following Hagimoto et al. [191]. Because single-dish photometry for these sources represents the combined emission of the system and cannot be uniquely deblended, I derive L_{IR} for each individual component directly from its deblended ~ 151 GHz (rest-frame $\sim 1985 \mu\text{m}$) ALMA flux density. This calculation assumes the same MBB shape with a fixed dust temperature of $T_{\text{dust}} = 35$ K—consistent with the mean of 32.3 ± 5.5 K [174] for single-component sources—and $\beta = 2$. Representative SED fits are shown in Fig. 46, and the resulting temperatures and apparent luminosities for the full sample are compiled in Table 6. From this analysis, I choose to adopt $L_{\text{IR,ref}} = 3 \times 10^{13} L_{\odot}$, a value close to the median apparent infrared luminosity of the sample and identical to the reference luminosity adopted by Hagimoto et al. [191], so that the stacked fluxes reported here can be compared directly with that work.

Table 6: MBB temperature (T) and apparent infrared luminosity ($\log(\mu L_{\text{IR}})$) for the sample.

ID	T [K]	$\log(\mu L_{\text{IR}}/L_{\odot})$	ID	T [K]	$\log(\mu L_{\text{IR}}/L_{\odot})$
S ₀₀₁	–	12.91	S ₀₅₃	25.4	12.87
S ₀₀₂	–	13.45	S ₀₅₄	40.5	13.75
S ₀₀₃	–	13.67	S ₀₅₅	38.0	13.77

continued on next page

Table 6 – continued

ID	T [K]	$\log(\mu L_{\text{IR}}/L_{\odot})$	ID	T [K]	$\log(\mu L_{\text{IR}}/L_{\odot})$
S ₀₀₄	–	13.54	S ₀₅₆	41.8	13.99
S ₀₀₅	35.8	13.56	S ₀₅₇	35.4	13.39
S ₀₀₆	–	12.84	S ₀₅₈	34.0	13.55
S ₀₀₇	–	13.60	S ₀₅₉	37.8	14.02
S ₀₀₈	40.8	13.68	S ₀₆₀	31.9	13.47
S ₀₀₉	35.8	13.41	S ₀₆₁	31.3	13.73
S ₀₁₀	–	13.46	S ₀₆₂	39.3	13.90
S ₀₁₁	–	13.12	S ₀₆₃	40.7	13.97
S ₀₁₂	–	13.08	S ₀₆₄	38.7	13.58
S ₀₁₃	–	13.11	S ₀₆₅	29.5	13.03
S ₀₁₄	42.5	13.89	S ₀₆₆	36.5	13.93
S ₀₁₅	–	13.70	S ₀₆₇	33.4	13.27
S ₀₁₆	–	12.78	S ₀₆₈	34.2	13.21
S ₀₁₇	–	12.80	S ₀₆₉	–	13.43
S ₀₁₈	39.5	13.93	S ₀₇₀	34.1	13.44
S ₀₁₉	35.6	13.52	S ₀₇₁	31.8	13.20
S ₀₂₀	34.1	13.91	S ₀₇₂	33.5	13.17
S ₀₂₁	27.7	13.21	S ₀₇₃	–	13.13
S ₀₂₂	45.8	13.79	S ₀₇₄	–	13.70
S ₀₂₃	28.1	13.27	S ₀₇₅	32.8	13.32
S ₀₂₄	28.8	13.33	S ₀₇₆	36.6	13.49
S ₀₂₅	33.6	13.57	S ₀₇₇	–	13.22
S ₀₂₆	–	13.32	S ₀₇₈	37.4	13.56
S ₀₂₇	–	13.13	S ₀₇₉	–	13.21
S ₀₂₈	25.1	12.76	S ₀₈₀	–	13.22
S ₀₂₉	–	13.35	S ₀₈₁	–	13.04
S ₀₃₀	–	13.62	S ₀₈₂	–	12.77
S ₀₃₁	31.8	13.41	S ₀₈₃	–	12.82
S ₀₃₂	32.8	13.50	S ₀₈₄	29.8	13.07
S ₀₃₃	34.2	13.85	S ₀₈₅	32.8	13.71
S ₀₃₄	32.6	13.26	S ₀₈₆	33.4	13.18
S ₀₃₅	23.2	12.67	S ₀₈₇	–	12.95
S ₀₃₆	–	13.29	S ₀₈₈	37.5	13.52
S ₀₃₇	–	13.31	S ₀₈₉	31.4	13.52
S ₀₃₈	–	13.11	S ₀₉₀	34.3	13.48
S ₀₃₉	32.2	13.67	S ₀₉₁	28.4	13.13
S ₀₄₀	32.8	13.40	S ₀₉₂	34.0	13.56
S ₀₄₁	33.7	13.47	S ₀₉₃	–	13.32
S ₀₄₂	34.4	13.20	S ₀₉₄	29.5	13.09

continued on next page

Table 6 – continued

ID	T [K]	$\log(\mu L_{\text{IR}}/L_{\odot})$	ID	T [K]	$\log(\mu L_{\text{IR}}/L_{\odot})$
S ₀₄₃	29.6	13.08	S ₀₉₅	39.9	13.85
S ₀₄₄	33.0	13.80	S ₀₉₆	30.8	13.40
S ₀₄₅	32.7	13.13	S ₀₉₇	36.7	13.86
S ₀₄₆	33.9	13.24	S ₀₉₈	41.7	13.85
S ₀₄₇	32.7	14.17	S ₀₉₉	34.2	13.67
S ₀₄₈	31.7	13.54	S ₁₀₀	29.2	13.07
S ₀₄₉	40.5	14.10	S ₁₀₁	30.1	13.19
S ₀₅₀	37.5	13.80	S ₁₀₂	29.8	13.15
S ₀₅₁	38.5	13.84	S ₁₀₃	29.9	13.53
S ₀₅₂	30.9	13.63	S ₁₀₄	29.2	13.47

Note: the symbol “–” indicates a blended system, for which T was fixed at 35 K following Bendo et al. [174].

6.2.4 Continuum subtraction and treatment of multiple sources

In the (sub-)millimetre regime, the CO lines sit on top of the dust continuum. As discussed in Section 3.4.3 and in line with the processing pipeline (Sec. 5.2.3), I subtract the continuum from each source individually prior to stacking, rather than from the combined spectrum afterwards. Postponing continuum subtraction until after stacking would allow a few disproportionately continuum bright sources to bias the stacked line profile. This effect is particularly pronounced for strongly lensed galaxies at high J_{up} transitions, where dust continuum levels can rival or exceed the line flux. In practice, the continuum is subtracted per source with a zeroth-order fit to the line-free channels, excluding a velocity window around the CO line that is enlarged where a neighbouring feature falls within the fitted range, as for the [C I](2 – 1) line near CO(7 – 6) and the OH⁺ feature near CO(9 – 8). A constant fit is preferred over a linear or higher-order one to avoid introducing free parameters that the narrow line-free bandwidth cannot constrain, and that would risk a degeneracy between a genuine spectral slope and channel-to-channel noise, and over such a narrow bandwidth any real continuum slope is in any case expected to be negligible. The constant is fitted separately to the real and imaginary parts of the visibilities, as in Section 5.2.3, so that it is not affected by the Rician bias of the amplitudes (Sec. 4.3.2). Its noise is reduced by the number of line-free channels with respect to that of a single channel, and it enters each source as a small random offset that averages out in the stack rather than as a systematic bias.

One might argue that normalising sources by their infrared luminosity and redshift (Section 6.2.3) compresses the spread in continuum brightness across the sample, thereby mitigating the bias from

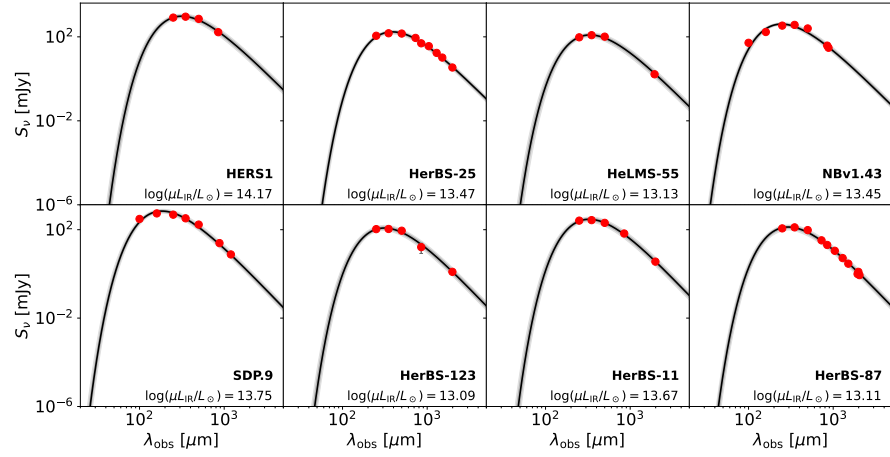


Figure 46: Examples of modified-blackbody SED fits for individual components and blended systems. The panels show the observed flux density S_ν as a function of observed wavelength λ_{obs} , with the derived apparent infrared luminosity $\log(\mu L_{\text{IR}}/L_\odot)$ computed from the deblended ALMA flux densities assuming $T_{\text{dust}} = 35$ K and $\beta = 2$.

bright continuum targets and theoretically allowing subtraction after stacking. However, performing continuum subtraction on a per source basis prior to stacking remains essential, as it provides a robust framework for handling fields with multiple sources. Indeed, several targets feature companion galaxies within the primary beam whose emission would otherwise contaminate the target stacked signal. The treatment of each companion depends on its redshift relative to the target, a problem that the per source continuum fit naturally helps resolve.

When the per-source continuum fit is carried out, any companion that does not lie at the same redshift as the target, or more generally has no line emission falling within the fitted velocity window, contributes continuum emission only within that window and is therefore removed together with the continuum by the same fit. The line signal of the target is thus intrinsically unaffected by such external sources, with no further step required. Companions confirmed at the same redshift as the target, by contrast, carry genuine line flux within the same velocity window, so they cannot be removed by the continuum fit alone and their line contribution is identified and subtracted separately, at their fixed known sky positions, in the extraction stage described in Section 6.2.5.

6.2.5 Signal extraction

I recover the stacked signal directly in the visibility domain without forming images, following the extraction procedure described in Sec-

tion 5.3. To recall the basic steps of the procedure, for each source and spectral channel, visibilities are radially binned into logarithmically spaced baseline annuli, compressing the data into five accumulators and providing a compact set of sufficient statistics for all subsequent analysis. The compressed statistics of all sources are then combined into the stack, and the stacked profile is fitted in three steps. First, I determine the spectral shape of the line, namely its centroid and width, from the spatially integrated spectrum. Holding this spectral shape fixed, I then fit the line amplitude per annulus to obtain the flux as a function of baseline length. Finally, a circular Gaussian fit to the visibility profile yields the total flux F_{tot} and the effective source size.

Figure 47 illustrates the stacked CO(6 – 5) line profile alongside the corresponding circular Gaussian fit to the real part of the visibilities as a function of uv distance. For the two blended transitions, CO(7 – 6) + [C I](2 – 1) and CO(9 – 8) + OH⁺(1 – 0), two components are fitted jointly to isolate the CO flux analytically from its neighbouring feature. For ACA subsamples, where sources remain unresolved, the Gaussian is replaced by a fixed point source model. The real part of the visibilities carries the stacked signal, provided that the sources are correctly centred, while the imaginary part serves as a null test to confirm correct source centering and the absence of residual contamination. Uncertainties on all recovered quantities are derived via bootstrap resampling of the source sample ($N_{\text{boot}} = 500$), grouping the sources that share the same execution block, so as to capture both the measurement noise and the intrinsic population variance. Continuum fluxes are measured analogously from the weighted average of line free channels, recording the effective frequency of each band for dust SED analysis.

Beyond the target line, signal extraction accounts for companion galaxies within the primary beam, as introduced in Section 6.2.4. Emission from these companions is modelled at their known sky positions and removed from the visibilities of each field, in which the target lies at the phase centre, via a joint weighted least squares fit prior to radial binning, performed separately on line and continuum visibility columns. For each target, the target and all catalogued companions within the field are modelled simultaneously as point sources at fixed (l, m) offsets from the phase centre. The target is always retained as a free component, ensuring that only companion amplitudes are subtracted from the data.

For the line column, I subtract only companions with known spectroscopic redshifts by mapping their expected line frequencies onto the target rest frame channel grid. A per-channel complex amplitude is fitted jointly for the target and all such companions within that velocity window. When multiple companions with known redshifts fall in the same window, their amplitudes are fitted jointly to appor-

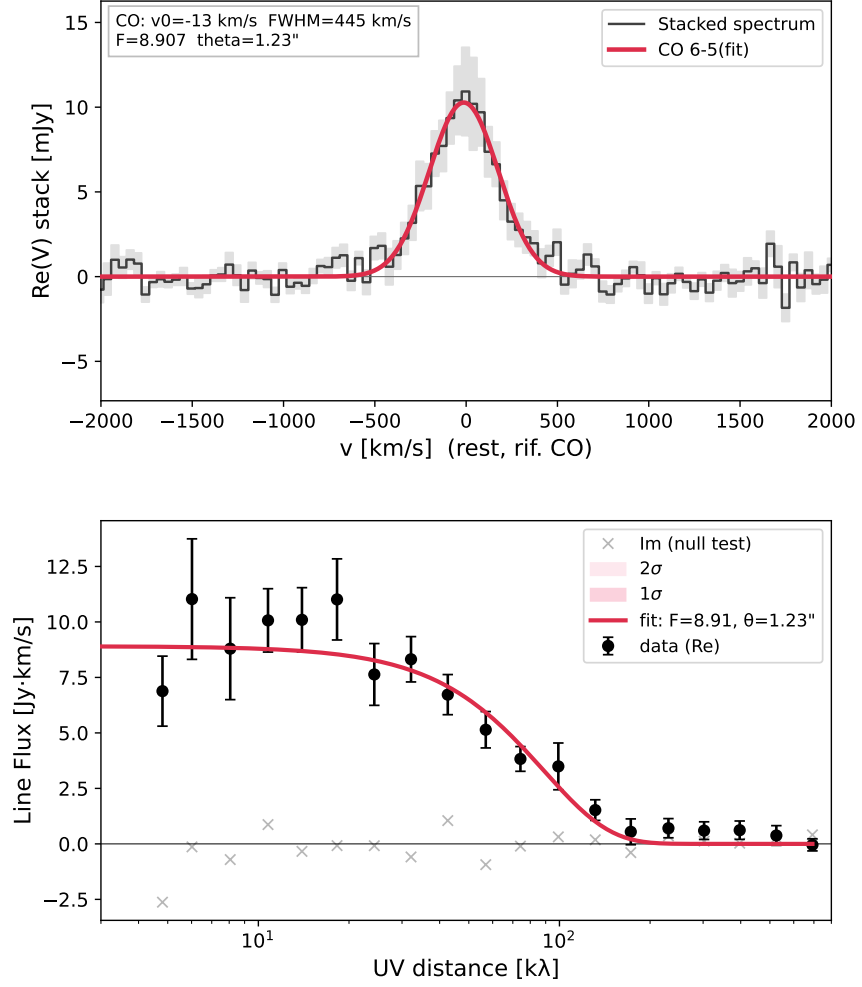


Figure 47: Example uv fit for CO(6 – 5), presenting the stacked line spectrum with the best fit Gaussian profile in the top panel and the real part of the visibilities versus uv distance with the best fit circular Gaussian model and 1σ and 2σ confidence bands in the bottom panel. The real part of the visibilities contains the stacked signal, whereas the imaginary part shown as grey crosses is consistent with zero as a null test. Error bars represent bootstrap standard deviations across the source sample ($N_{\text{boot}} = 500$).

tion flux accurately rather than double-subtracting, removing each fitted amplitude strictly within its respective line window. Companions lacking known redshifts are omitted from the line fit entirely, as unconstrained line frequencies could lead to inadvertent subtraction of target line flux.

For the continuum column, the target and all companions, regardless of redshift availability, are fitted jointly using a single low-order polynomial per source across line-free channels outside the line windows of both the target and any known redshift companions. Companion amplitudes are then subtracted before radial binning. Fitted continuum amplitudes are evaluated against a minimum SNR threshold to avoid injecting noise from companions consistent with zero continuum flux. In all cases, joint fit reliability is constrained by angular separation relative to observational resolution, monitored through the correlation of each companion with the target beam response and the associated variance inflation factor. Companions exceeding a variance inflation threshold, most frequently in lower resolution ACA data, are considered excessively degenerate with the target and are excluded from the fit rather than subtracted with uncertain amplitudes.

6.3 AVERAGE CO SLED

Figures 48 and 49 show the average CO SLEDs derived from the stacking analysis described in the previous section. The thick black line represents the average across the full sample, while lighter lines illustrate the trends for individual subsamples split by lensing status (lensed, L, probable unlensed, U) and array configuration (ALMA 12 m, ACA 7 m). Figure 48 presents the L_{IR} normalised CO SLED, whereas Fig. 49 shows the corresponding unnormalised SLED. The bottom panel of each figure reports the number of sources contributing to each transition, defined as targets with at least one observation in the respective category. In both cases, line fluxes are corrected for luminosity distance and referred to a common redshift $z = 2.5$, following the normalisation procedure outlined in Section 6.2.3.

Figure 50 displays the continuum flux densities as a function of rest frame frequency extrapolated in the vicinity of each line for both normalisation schemes, adopting the same colour coding and sample splits as the CO SLED figures. The same colour scheme is used in Figures 51 and 52, which present the retrieved line and continuum source sizes for both cases. All line and continuum flux densities, along with the apparent source sizes (θ_{as} , defined as the Gaussian FWHM in arcseconds) used in this analysis, are reported in Tables 7 and 8. As detailed in Section 6.2.5, the ACA only subsample has $\theta_{\text{as}} = 0''$ by construction, as these sources were fitted with a point source model.

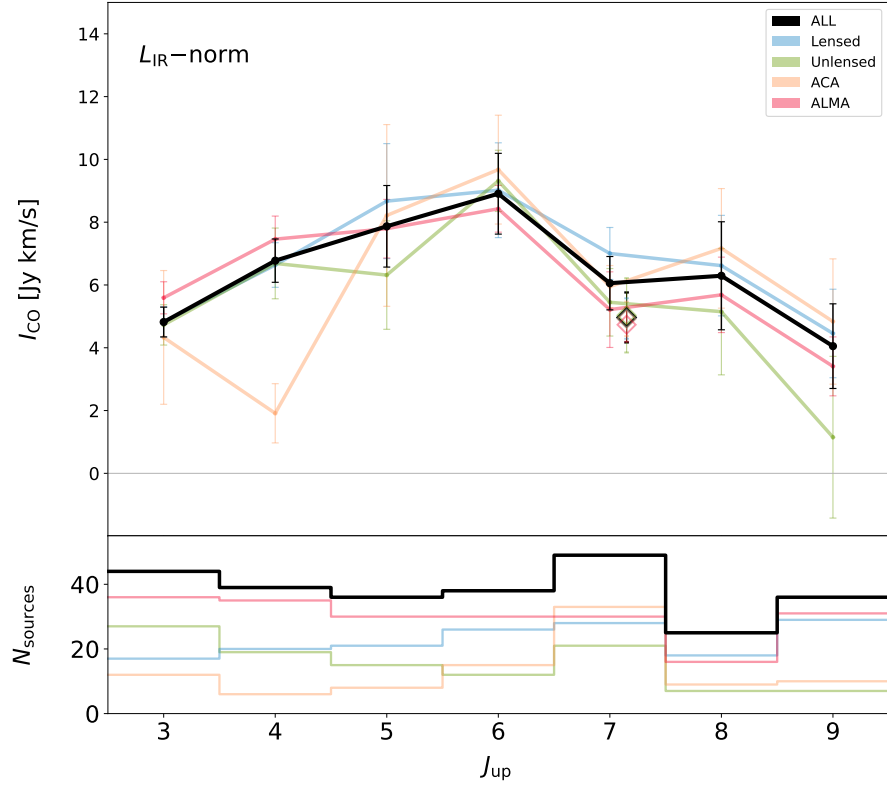


Figure 48: L_{IR} -normalised CO SLED for the different subsamples, with the source counts per transition in the bottom panel. *Top panel:* integrated CO line intensity, I_{CO} , as a function of the upper rotational quantum number J_{up} , from stacking normalised by the infrared luminosity L_{IR} of the contributing sources. The thick black curve (circles) is the mean SLED of the full sample (All); thinner curves show the breakdown by lensing status, lensed (blue) and unlensed (green), and by array, ACA (orange) and ALMA (red). Error bars are the 1σ uncertainty on the stacked line intensity. Open diamonds at $J_{\text{up}} = 7.15$ (a fiducial position for display only) show the L_{IR} -normalised stacked flux of the [C I](1–0) transition, following the same colour scheme. *Bottom panel:* number of sources N_{sources} contributing to each J_{up} bin, colour-coded as above, illustrating the non-uniform spectral coverage of the sample and hence the statistical robustness of each SLED point.

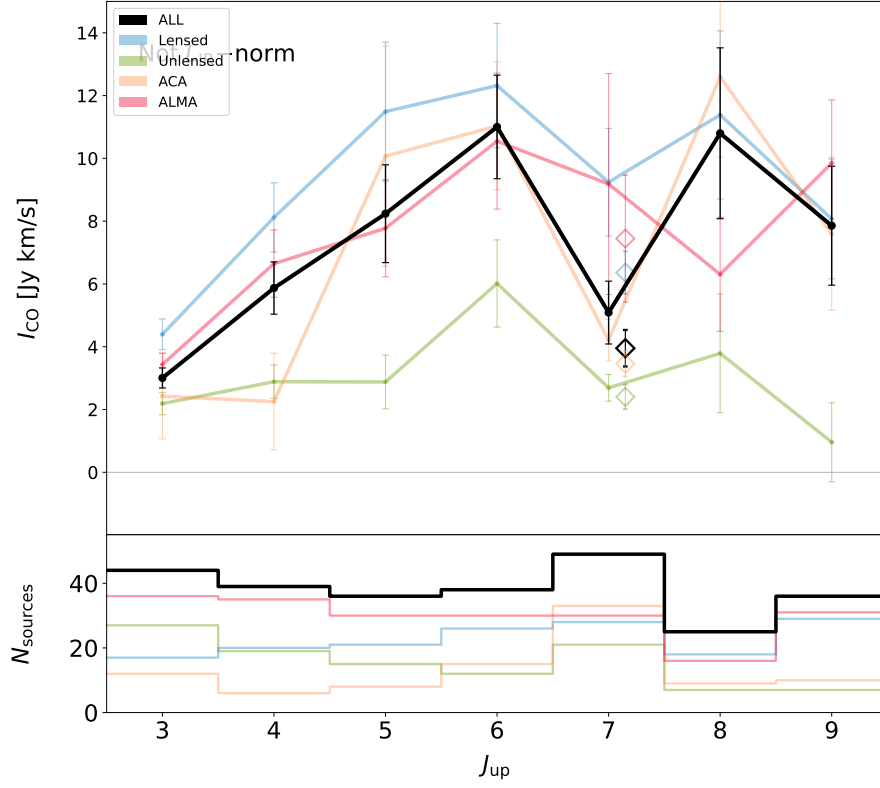


Figure 49: Not-normalised CO SLED for the different subsamples, with the source counts per transition in the bottom panel. *Top panel:* integrated CO line intensity, I_{CO} , as a function of the upper rotational quantum number J_{up} , without infrared luminosity normalisation. The thick black curve (circles) is the mean SLED of the full sample (All); thinner curves show the breakdown by lensing status, lensed (blue) and unlensed (green), and by array, ACA (orange) and ALMA (red). Error bars are the 1σ uncertainty on the stacked line intensity. Open diamonds at $J_{\text{up}} = 7.15$ (a fiducial position for display only) show the stacked flux of the [CI](1 – 0) transition, following the same colour scheme. *Bottom panel:* number of sources N_{sources} contributing to each J_{up} bin, colour-coded as above, illustrating the non-uniform spectral coverage of the sample and hence the statistical robustness of each SLED point.

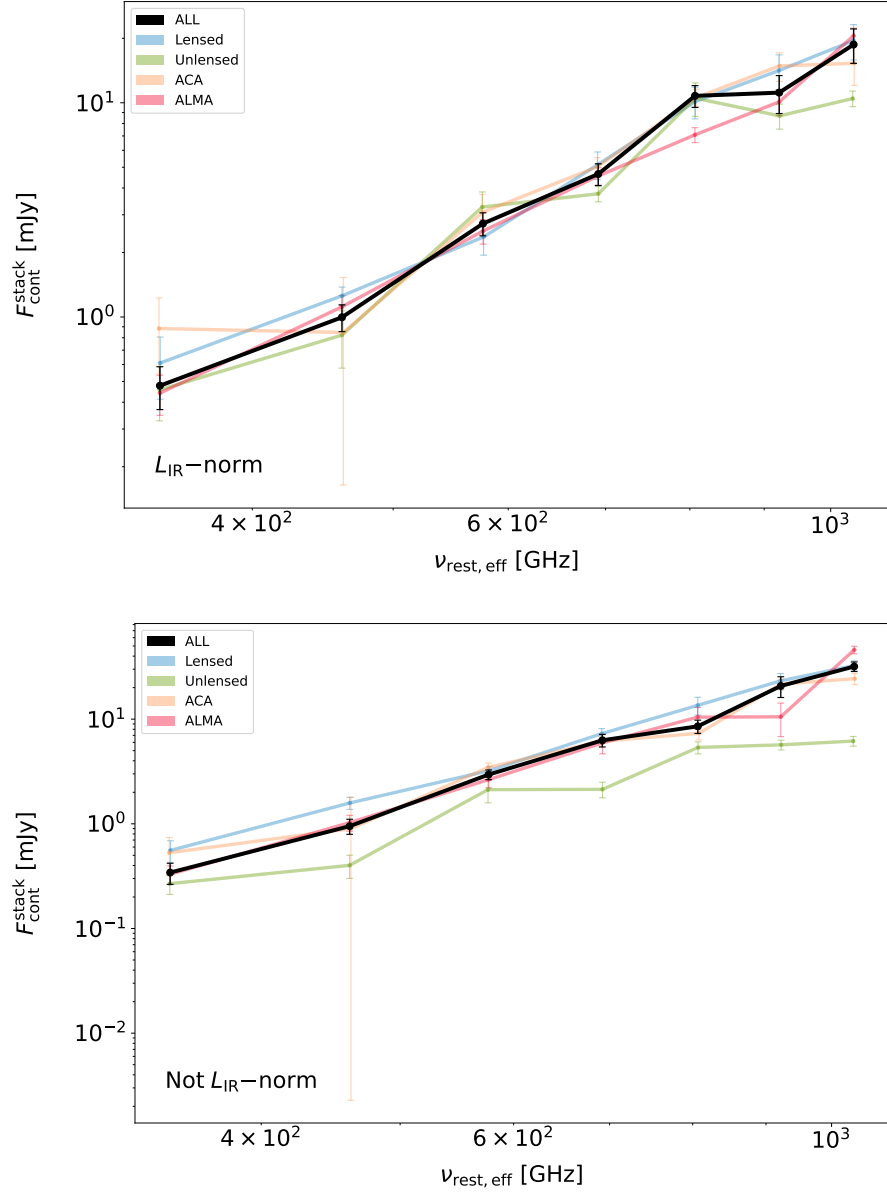


Figure 50: Stacked continuum SED, obtained from stacking the observations associated with each spectral window used for the CO line extraction, shown as a function of the rest-frame effective frequency ($\nu_{\text{rest, eff}}$), in the L_{IR} -normalised (top) and not-normalised (bottom) cases. Each curve represents a different subsample: the full sample (black, $\alpha = 1$), lensed sources (blue), unlensed sources (green), and the array-based subsamples ACA (orange) and ALMA (red). Vertical error bars represent the 1σ uncertainty on the stacked continuum flux density in each spectral window.

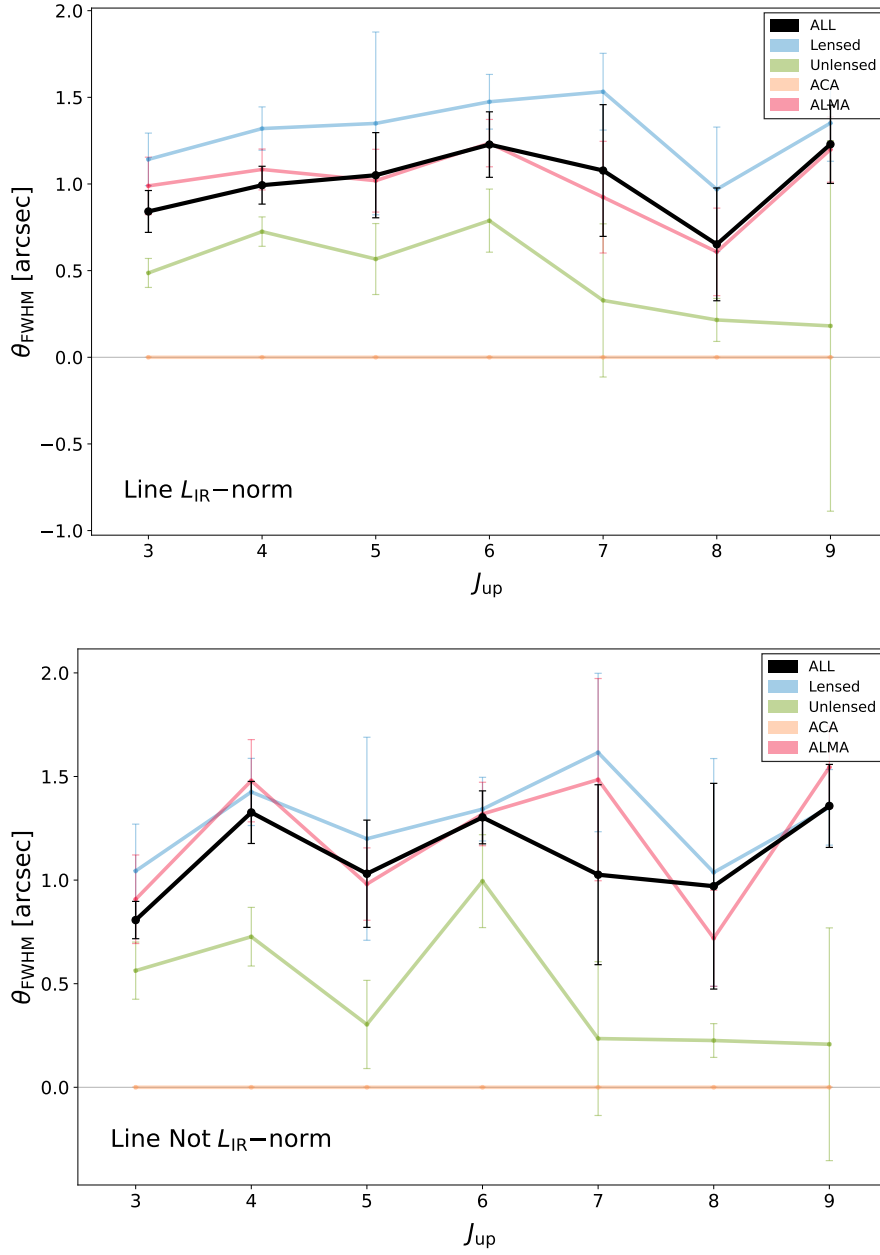


Figure 51: Apparent line source sizes (θ_{as} , Gaussian FWHM in arcsec) recovered from the uv -plane fit as a function of J_{up} , in the L_{IR} -normalised (top) and not-normalised (bottom) cases. Curves represent the different subsamples: the full sample (black), lensed (blue), unlensed (green), ACA (orange) and ALMA (red). The ACA-only subsample has $\theta_{\text{as}} = 0$ by construction, as these sources were fitted with a constant point-source model. Error bars are the 1σ bootstrap uncertainties.

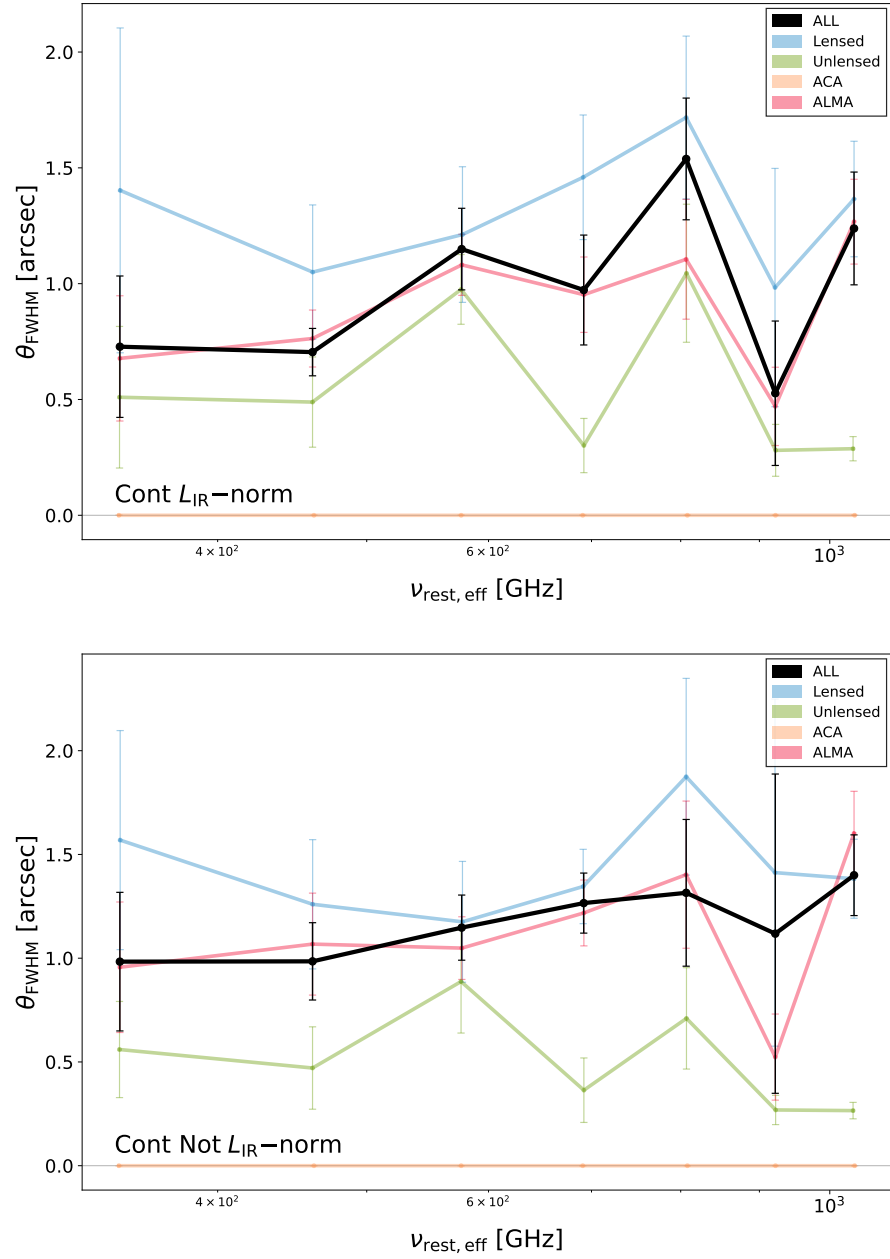


Figure 52: Apparent continuum source sizes (θ_{as} , Gaussian FWHM in arcsec) measured in each CO band as a function of J_{up} , in the L_{IR} -normalised (top) and not-normalised (bottom) cases. Curves represent the different subsamples: the full sample (black), lensed (blue), unlensed (green), ACA (orange) and ALMA (red). The ACA-only subsample has $\theta_{\text{as}} = 0$ by construction, as these sources were fitted with a constant point-source model. Error bars are the 1σ bootstrap uncertainties.

Both normalisations peak around $J = 5$ to 6 , indicating a mid J excited SLED typical of intensely star-forming, moderately dense molecular gas, in contrast to the low J peaked ($J \sim 2$ to 3) SLEDs characteristic of extended, quiescent disks. However, the two stacks diverge in their high- J behaviour. The unnormalised stack exhibits a local dip at $J = 7$ ($F = 5.26 \text{ Jy km s}^{-1}$) followed by a secondary peak at $J = 8$ ($F = 10.8 \text{ Jy km s}^{-1}$), producing a mild double humped profile. This feature largely disappears once the stack is normalised to a common L_{IR} , after which the SLED declines smoothly beyond $J = 6$ ($F = 6.28, 6.29$, and $4.05 \text{ Jy km s}^{-1}$ for $J = 7, 8$, and 9 , respectively). The $J = 8$ bump is most plausibly driven by a small number of bright, highly magnified sources rather than a population-wide physical trait. This interpretation is reinforced by the drop in contributing sources at $J = 8$ ($n_{\text{src}} = 25$, compared to 36 and 49 in the two adjacent transitions, and 36 to 49 across the rest of the ladder). The OH^+ line deblended from $\text{CO}(9 - 8)$ has a negative stacked flux in all the subsamples (Table 7), which indicates that it is seen on average in absorption against the dust continuum rather than in emission. The joint fit keeps this component separate from the $\text{CO}(9 - 8)$ line, so that the absorption does not bias the CO flux. For the full sample, the apparent source size (θ_{as}) remains remarkably stable at $\approx 1''$ across all J transitions, displaying no systematic trends or erratic scatter that would suggest incomplete flux recovery (Fig. 51). I emphasise, however, that these sizes should not be interpreted as intrinsic physical scales, rather they represent an effective operational parameter designed to capture the total integrated stacked flux. For the smallest subsamples, in particular the ACA ones, the number of contributing sources is of the order of ten or fewer. In this regime the number of distinct bootstrap configurations is intrinsically limited, so the corresponding uncertainties should be regarded as lower limits rather than as definitive estimates.

Table 7: CO/[C I]/ OH^+ line SLED normalised to L_{IR} ($L_{\text{IR}} = 3 \times 10^{13} L_{\odot}$, $z = 2.5$).

Sample	J_{up}	F [Jy km/s]	σ_F	θ_{as}	σ_{θ}
All	3	4.8198	0.4740	0.841	0.121
	4	6.7752	0.6900	0.993	0.109
	5	7.8690	1.3000	1.051	0.246
	6	8.9068	1.2900	1.227	0.189
	7	6.2814	0.8100	1.149	0.343
	8	6.2930	1.7200	0.652	0.325
	9	4.0512	1.3500	1.229	0.226
	[C I](2-1)	5.4554	0.7578	1.526	–

continued on next page

Table 7 – continued

Sample	J_{up}	F [Jy km/s]	σ_F	θ_{as}	σ_θ
	OH ⁺ (1–0)	-1.9834	1.4160	1.046	–
L	3	4.8228	0.4820	1.142	0.152
	4	6.6355	0.7110	1.320	0.125
	5	8.6710	1.8300	1.349	0.527
	6	9.0188	1.5100	1.474	0.158
	7	7.1181	0.8580	1.553	0.223
	8	6.6166	1.6000	0.969	0.359
	9	4.4555	1.4100	1.351	0.220
	[C I](2–1)	5.1880	0.6459	1.695	–
	OH ⁺ (1–0)	-2.1866	1.6155	1.147	–
U	3	4.7281	0.6420	0.487	0.084
	4	6.6865	1.1300	0.725	0.085
	5	6.3151	1.7300	0.567	0.205
	6	9.3160	0.9750	0.788	0.182
	7	5.5975	1.1300	0.378	0.425
	8	5.1486	2.0100	0.215	0.124
	9	1.1503	2.5800	0.181	1.069
	[C I](2–1)	6.0558	1.0489	1.277	–
	OH ⁺ (1–0)	-5.0929	3.3438	3.176	–
ACA	3	4.3294	2.1300	–	–
	4	1.9123	0.9430	–	–
	5	8.2145	2.8900	–	–
	6	9.6742	1.7300	–	–
	7	5.9496	0.6690	–	–
	8	7.1661	1.9000	–	–
	9	4.8358	2.0000	–	–
	[C I](2–1)	4.9206	0.5487	–	–
	OH ⁺ (1–0)	-0.9156	0.9337	–	–
ALMA	3	5.5916	0.5110	0.989	0.166
	4	7.4598	0.7350	1.084	0.118
	5	7.7927	0.9330	1.019	0.181
	6	8.4267	0.7440	1.236	0.137
	7	5.2141	1.2000	0.924	0.322
	8	5.6851	1.2000	0.608	0.253
	9	3.4075	0.9390	1.199	0.189
	[C I](2–1)	4.7335	0.5420	1.308	–
	OH ⁺ (1–0)	-2.6762	1.0411	0.920	–

Fluxes F are reported in Jy km/s. The [C I](2 – 1) and OH⁺(1 – 0) measurements are derived from simultaneous deblending (rest frequencies 809.342 and 1033.119 GHz, respectively), with σ_F estimated as $(p_{84} - p_{16})/2$ from the bootstrap distributions. Negative OH⁺ fluxes indicate absorption against the dust continuum. The θ uncertainties are not available for [C I] and OH⁺, as no bootstrap analysis was performed for the size parameter of the second line in the joint fit.

Table 8: Continuum flux densities in Jy km/s of the stack normalised to L_{IR} ($L_{\text{IR}} = 3 \times 10^{13} L_{\odot}$, $z = 2.5$), reported in mJy at the frequency ν of each CO band, together with the apparent source sizes (θ_{as} , Gaussian FWHM in arcsec) and their uncertainties.

Sample	J_{up}	ν [GHz]	F	σ_F	θ_{as}	σ_{θ}
All	3	345.537	0.480	0.110	0.728	0.305
	4	461.051	1.000	0.140	0.704	0.102
	5	576.381	2.730	0.340	1.149	0.176
	6	691.914	4.650	0.540	0.973	0.237
	7	806.578	10.770	1.260	1.538	0.263
	8	921.652	11.150	2.240	0.527	0.312
	9	1036.985	18.700	3.420	1.238	0.243
L	3	345.692	0.610	0.200	1.403	0.701
	4	461.128	1.260	0.120	1.050	0.290
	5	577.050	2.360	0.420	1.212	0.293
	6	691.336	5.140	0.760	1.459	0.269
	7	806.578	10.120	1.700	1.717	0.351
	8	921.289	14.140	2.600	0.983	0.514
	9	1036.985	19.510	3.660	1.365	0.249
U	3	345.352	0.460	0.130	0.510	0.306
	4	461.051	0.820	0.250	0.489	0.195
	5	575.832	3.260	0.570	0.975	0.151
	6	692.058	3.760	0.310	0.301	0.117
	7	806.930	10.510	1.880	1.045	0.298
	8	922.136	8.690	1.160	0.280	0.112
	9	1035.469	10.460	0.880	0.287	0.052
ACA	3	345.044	0.880	0.340	–	–
	4	462.013	0.850	0.680	–	–
	5	575.928	3.080	0.660	–	–
	6	690.892	5.020	0.520	–	–
	7	808.037	10.570	1.000	–	–
	8	921.975	14.850	2.270	–	–
	9	1037.886	15.290	3.230	–	–
ALMA	3	345.537	0.440	0.090	0.677	0.270
	4	461.051	1.120	0.140	0.763	0.123
	5	576.506	2.520	0.330	1.081	0.132
	6	692.058	4.560	0.420	0.952	0.163
	7	806.508	7.080	0.570	1.105	0.259
	8	921.652	10.090	1.370	0.470	0.169
	9	1036.985	20.550	1.740	1.268	0.183

6.3.1 *Lensed vs unlensed sample*

Having examined the overall behaviour of the full sample, I now focus on the differences between the lensed and unlensed populations. In the unnormalised stack (Fig. 49), the lensed curve lies systematically above the unlensed one across all transitions, with $F_L/F_U \sim 2$ for low J transitions ($J_{\text{up}} = 3, 4$) and stabilising around 4 at higher transitions. In the L_{IR} normalised stack (Fig. 48), the two curves become mutually consistent within uncertainties across the entire ladder, with the largest separation occurring at $J = 9$.

Regarding spatial extension, the apparent source sizes (θ_{as}) for both populations remain relatively stable within observational uncertainties across the entire J range (Fig. 51). Lensed sources display mean values of $\theta_{\text{as}} \approx 1.0$ to $1.5''$, whereas unlensed sources cluster around $\theta_{\text{as}} \approx 0.5''$, making lensed sources systematically larger across all transitions. However, the unlensed sizes at high J are essentially unconstrained ($\sigma_{\theta} \sim \theta_{\text{as}}$).

For the $J = 8 - 7$ and $J = 9 - 8$ transitions of the unlensed sample, the observational constraints are indeed limited by the smaller number of sources, by the lower SNR of the available observations and by the lack of short-baseline data. This is quantified in Figs. 42 and 43. At these two transitions, the unlensed subsample is observed almost exclusively at high resolution with the 12 m array, with synthesised beams of 0.10 – $0.18''$ at $J = 8 - 7$ and no ACA data, and a median beam of $0.11''$ at $J = 9 - 8$ with a single ACA dataset. These beams are several times smaller than the apparent size of the stacked unlensed source, so that the emission is resolved across multiple synthesised beams. Since the maximum recoverable scale of these configurations is comparable to the source size itself, a fraction of the extended emission may be filtered out. The lensed subsample, in contrast, retains ACA coverage at both transitions, which leaves it far less susceptible to this effect.

Taken together, these results point toward the absence of significant differential lensing. Indeed, if strong differential magnification were driving the emission, the more compact high- J emitting gas would experience enhanced localised magnification relative to the extended low J reservoir. This would produce a steady, monotonically increasing trend in the F_L/F_U ratio with excitation in both unnormalised and L_{IR} normalised SLEDs, particularly toward high- J . In the unnormalised SLEDs, such a monotonic rise is absent. Although F_L/F_U is lower at the lowest J transitions, it flattens across intermediate transitions rather than steadily climbing. Moreover, the lower ratio at low J cannot be securely interpreted as an effect of differential lensing, as low J lines are observed at lower frequencies where line flux densities are intrinsically lower, yielding lower SNR and making the separation between line and continuum emission less robust. In the L_{IR}

normalised case, any variation in the F_L/F_U ratio remains small and fully consistent within measurement uncertainties, and any minor discrepancy between the two subsamples is indistinguishable from the intrinsic scatter in the CO SLED of the overall population and the substantial measurement uncertainties characterising high- z unlensed sources. In addition, as seen in the L_{IR} normalised dust continuum profile (Fig. 50), a similar flattening is present in the continuum emission of the unlensed population at high frequencies. This reinforces the conclusion that the observed behaviour is driven by observational artifacts rather than differential lensing. Deeper observations targeting both low ($J_{\text{up}} < 3$) and high- J ($J_{\text{up}} > 9$) transitions will therefore be essential to probe subtle differential magnification effects in the warmest, densest gas component at high- J . With the currently available observations of high- J CO lines, the mutual compatibility of the two SLEDs at fixed L_{IR} validates the use of strongly lensed DSFGs as representative proxies for the ISM conditions of the broader, unlensed high- z DSFGs population.

6.4 AVERAGE PHYSICAL PROPERTIES OF THE MOLECULAR GAS

As discussed in Section 1.2.2, interpreting the physical conditions of the molecular gas from observed CO SLEDs requires non-LTE LVG radiative transfer modelling. When only a handful of mid- J transitions are available, studies typically rely on a single-component LVG statistical equilibrium approach [e.g., 130]. With more comprehensive rotational ladder sampling, analyses generally decompose the emission into cool, extended and warm, dense gas phases [e.g., 74, 101, 117, 121, 127], or alternatively describe the ISM using a continuous, turbulence-motivated density distribution [105]. However, within this sample, fully unravelling a multi-component gas structure with high precision is intrinsically challenging due to the limited number of detected J -transitions and the observational uncertainties at the highest excitation levels. Consequently, constraining a complex parameter space runs the risk of introducing severe parameter degeneracies. Because the derived stacked CO SLED is itself the product of a heterogeneous, multi-source stacking procedure with sparse low- J coverage but a mid-to-high- J ladder extending up to CO(9 – 8), I adopt an approach closer to that of Yang et al. [121]. I first fit a single-component LVG model (Section 6.4.2) and then test whether the data require a second, cooler phase (Section 6.4.3). I model the stacked L_{IR} -normalised CO SLEDs using the non-LTE radiative transfer code RADEX [126], interfaced via its Python wrapper SPECTRALRADEX. I adopt the LVG method assuming an expanding spherical geometry

with an escape probability given by $\beta = (1 - e^{-\tau})/\tau$. Collisional excitation rates with H_2 are retrieved from the LAMDA database [212].

6.4.1 Dust continuum of the stack and gas mass cross-check

The seven continuum measurements of the stack (one per CO band, spanning $\lambda_{\text{rest}} \simeq 290\text{--}870\ \mu\text{m}$ at z_{med}) are fitted with an optically thin MBB with the same CMB corrections as in Section 6.2.3, a dust opacity $\kappa_{250\mu\text{m}} = 0.45\ \text{cm}^2\ \text{g}^{-1}$ scaling as ν^β , and free $\log M_{\text{dust}}$, T_{d} and β . Since all points lie on the Rayleigh–Jeans tail, T_{d} and M_{dust} are degenerate; I therefore adopt a Gaussian prior on $T_{\text{d}} = 38 \pm 8\ \text{K}$. I performed the MBB fits for the complete sample (L+U), lensed (L), and unlensed (U) subsamples. The MBB fits are excellent for the L+U ($\chi^2_{\text{red}} = 1.04$) and L subsamples ($\chi^2_{\text{red}} = 0.44$), with $\beta \simeq 1.8\text{--}1.9$, band-to-band shape probabilities $p_{\text{shape}} = 0.60$ and 0.93 respectively, and $T_{\text{dust}} \sim 38\ \text{K}$. The U continuum is in mild tension ($p_{\text{shape}} = 0.18$): with β free, the fit becomes unstable and rails to $\beta \simeq 1.6$ and $L_{\text{IR}}^{\text{MBB}}/L_{\text{IR}}^{\text{ref}} = 0.5$ because the U data suffer from a SNR that is too low and a resolution that is too high to yield a stable fit. Indeed, by removing the outermost poorly sampled points from the short-baseline observations, the fit becomes compatible with the other two subsamples. I also note that $L_{\text{IR}}^{\text{MBB}}/L_{\text{IR}}^{\text{ref}}$ is not uniform across subsamples, reaching higher values up to 1.4, so absolute dust-mass comparisons between subsamples should be made with care. This variation must be interpreted keeping in mind that I have fitted with multiple free parameters without explicitly constraining the dust peak. Furthermore, the MBB fit used to derive the L_{IR} imposed a value of $\beta = 2$ (see Section 6.2.3), which might not be ideally suited for all subsamples individually, even though it represented the best choice for continuity across the sample.

6.4.2 Single-component LVG modelling

I first describe the stacked CO ladder (CO(3–2) to CO(9–8)) with a one-zone, LVG calculation performed with RADEX through SPECTRALRADEX, adopting the expanding-sphere escape probability appropriate for the supersonic velocity fields ($\Delta v = 400\ \text{km s}^{-1}$) of these systems. The background is fixed to the CMB at the median redshift, $T_{\text{bg}} = 2.725(1 + z_{\text{med}}) = 9.5\ \text{K}$. The four free parameters are the gas density $\log(n_{\text{H}_2}/\text{cm}^{-3})$, the kinetic temperature T_{kin} , the CO column per unit line width $\log(N_{\text{CO}}/\Delta v)$, and the apparent solid angle $\log \Omega_{\text{app}}$, which absorbs the absolute flux normalisation so that the excitation parameters are independent of magnification and of the reference L_{IR} to which the stack is scaled. I impose a log-normal prior on T_{kin} centred at 55 K (0.25 dex wide), motivated by the typical ratio $T_{\text{kin}}/T_{\text{dust}} \sim 1.5$ found by Harrington et al. [105] combined with the $T_{\text{dust}} = 38\ \text{K}$ measured from the FIR continuum (Section 6.4.1); a

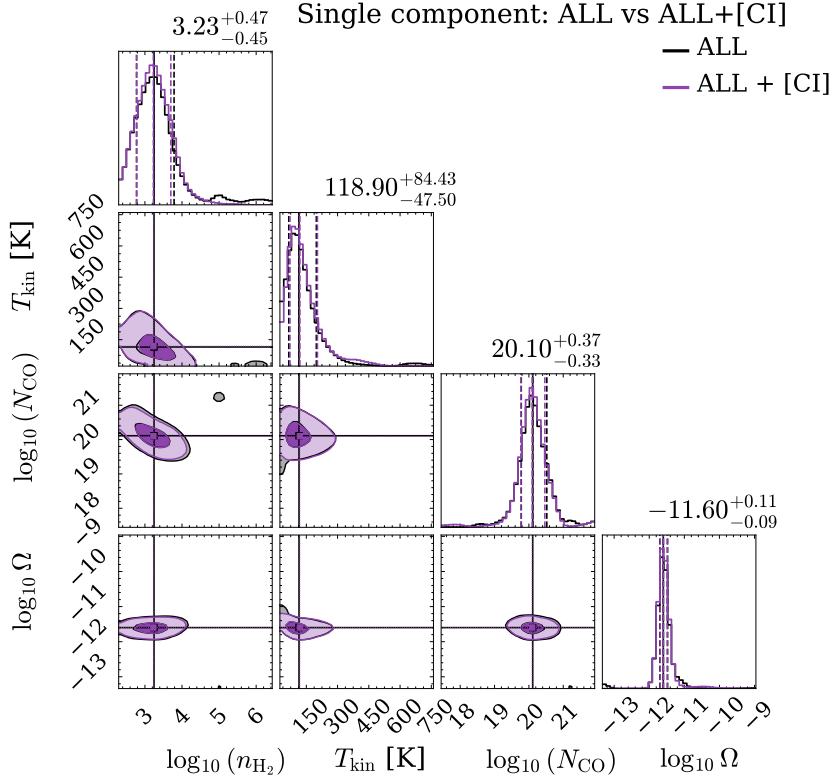


Figure 53: Marginalised one- and two-dimensional posteriors of the single-component fit for the complete sample. The single warm phase reproduces the mid- and high- J ladder but leaves the anchored CO(1–0) point under-predicted at $+1.5\sigma$. The n_{H_2} – T_{kin} anti-correlation visible here is the density–temperature degeneracy that leaves T_{kin} and n_{H_2} individually poorly constrained while keeping their product, the thermal pressure $P_{\text{th}}/k_{\text{B}}$, well determined. Contours are the 1 and 2σ credible regions; black shows the fit without the [CI](2 – 1) datum, purple the fit including it.

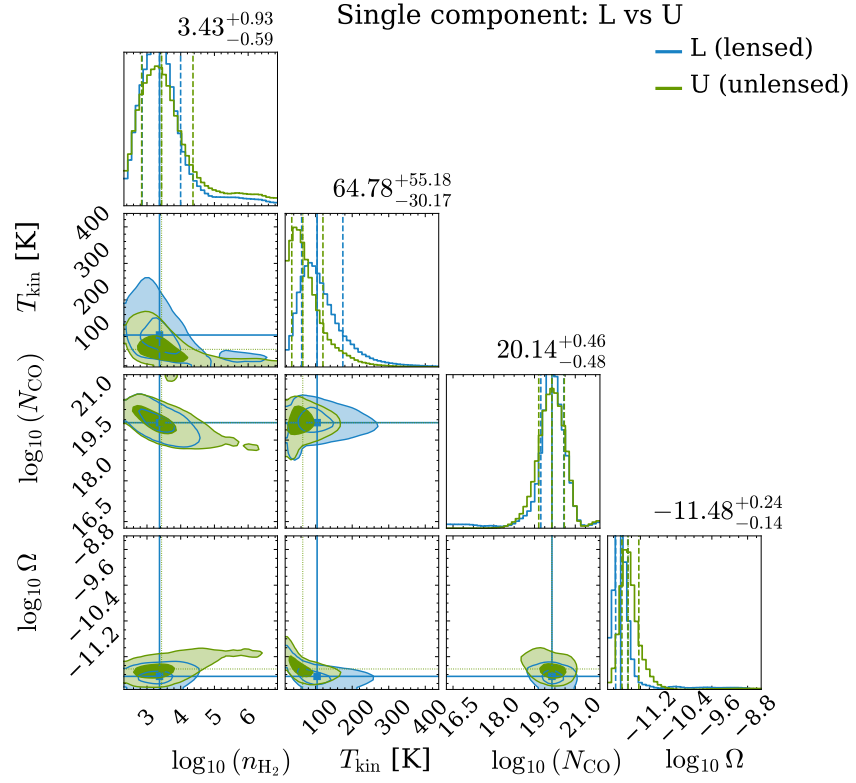


Figure 54: Lensed vs unlensed posteriors, single-component fit. Overlaid 1–2 σ contours for the L (red) and U (blue) subsamples. The mutual overlap in every projection quantifies the absence of a significant lensing-induced difference in the excitation parameters.

flat prior on $\log \Omega_{\text{app}}$; and a broad flat prior on $\log(N_{\text{CO}}/\Delta v)$ supplemented by the physical requirement that the emitting column length not exceed the source diameter ($\lesssim 7$ kpc). The lowest transition is anchored by treating the CO(1–0) line as a fixed synthetic datum, obtained from the stacked CO(3–2) flux as

$$F_{\text{CO}(1-0)} = F_{\text{CO}(3-2)} \frac{\nu_{\text{CO}(1-0)}^2}{\nu_{\text{CO}(3-2)}^2} \frac{1}{r_{31}}, \quad (38)$$

where $r_{31} \equiv L'_{\text{CO}(3-2)}/L'_{\text{CO}(1-0)}$ is the CO(3–2)-to-CO(1–0) luminosity ratio and the ν^2 factor converts it into the corresponding flux ratio. This anchoring forces the fit to reproduce the low- J turnover rather than merely constraining it through a soft prior. I adopt $r_{31} = 0.66 \pm 0.15$, consistent with the typical CO(3–2)/CO(1–0) luminosity ratio for this population [33, 105, 121, 191]. Because the stack combines heterogeneous observations, I add in quadrature a population-variance term $\sigma_{\text{sys},J} = \ln(10) \sigma_{\text{pop}} F_J / \sqrt{n_{\text{src},J}}$ with $\sigma_{\text{pop}} = 0.25$ dex and a 5% calibration floor, automatically down-weighting transitions built on fewer objects. Posteriors are obtained with EMCEE [213]; I quote the median and 16th–84th percentile of each marginalised distribution, and evaluate the reduced χ^2 on the data alone (keeping the prior penalty separate).

The resulting parameters are collected in the upper block of Table 9. Across the three splits the CO ladder is reproduced by warm, moderately dense gas: $\log(n_{\text{H}_2}/\text{cm}^{-3}) \simeq 3.3$, $T_{\text{kin}} \simeq 60\text{--}110$ K and $\log(N_{\text{CO}}/\Delta v) \simeq 17.5$. The best-determined combination is the thermal pressure $P_{\text{th}}/k_{\text{B}} = n_{\text{H}_2} T_{\text{kin}}$, which is stable across splits at $\log(P_{\text{th}}/k_{\text{B}} [\text{K cm}^{-3}]) \simeq 5.3$, i.e. $P_{\text{th}}/k_{\text{B}} \approx (1.8\text{--}2.5) \times 10^5 \text{ K cm}^{-3}$. As a consistency check on the LVG solution I derive the intrinsic virial parameter $K_{\text{vir}}^{\text{LVG}} = (dv/dr)_{\text{LVG}}/(dv/dr)_{\text{vir}}$, with $(dv/dr)_{\text{vir}} = 0.65 (n_{\text{H}_2}/10^3 \text{ cm}^{-3})^{0.5} \text{ km s}^{-1} \text{ pc}^{-1}$ [214], finding $K_{\text{vir}}^{\text{LVG}} \simeq 2$, consistent with gravitationally bound, marginally supra-virial gas. The single-phase solution reproduces the mid- and high- J ladder well but leaves a systematic positive residual on the anchored CO(1–0) point of $+1.3\text{--}1.5\sigma$ across all splits: a single warm phase that fits $J \geq 3$ slightly under-produces the lowest transition. This mild tension is the physical motivation for the two-temperature model of Section 6.4.3. In addition, I also test the model with a $J \geq 3$ -only ladder; without a low- J anchor, however, the SLED alone drives the single phase towards the warm, thermalised branch, predicting an unphysically large $r_{31}^{\text{model}} \simeq 1.0\text{--}1.1$ (i.e. almost no CO(1–0) excess) and biasing T_{kin} upwards and N_{CO} downwards. Fixing CO(1–0) as a datum breaks in part this degeneracy, so that n_{H_2} , T_{kin} and N_{CO} are constrained by the full shape of the SLED rather than by its high- J portion alone. The residual r_{31} tension that survives ($+1.3\text{--}1.5\sigma$) is then a genuine, quantified signal rather than a hidden freedom absorbed by an extreme line ratio. The same limitation is

confirmed by the $[\text{C I}](2-1)$ line, which I include in the fit for the complete sample: a single warm phase under-predicts it dramatically ($F_{\text{model}} \simeq 0.02$ vs $F_{\text{obs}} \simeq 5 \text{ Jy km s}^{-1}$, a $+2.6\sigma$ residual), since it lacks the cool, extended gas that produces most of the $[\text{C I}]$ emission.

The marginalised posterior distributions of the single-component fit are shown in Fig. 53 for the complete sample and compared between the lensed and unlensed subsamples in Fig. 54. The corner plot makes the density–temperature degeneracy explicit: n_{H_2} and T_{kin} are individually poorly constrained and broaden along their anti-correlation, while their product, the thermal pressure, remains well determined. The overlaid L and U contours mutually overlap in every projection, quantifying the absence of a significant lensing-induced difference in the recovered excitation parameters.

6.4.3 Double component LVG modelling

I next extend the single phase description of Section 6.4.2 to a two component model, in which the observed emission is the sum of a cool and a warm phase, each characterised by its own $(n_{\text{H}_2}, T_{\text{kin}}, N_{\text{CO}}, \Omega_{\text{app}})$ and sharing a common line width, and this doubles the number of free parameters to eight. The fitting parameters are otherwise unchanged from the single component case, as the same RADEX expanding sphere escape probability, CMB background, CO(1–0) anchoring strategy, population variance weighting and EMCEE sampling are applied to both phases simultaneously. What changes is the treatment of the temperature prior, which can no longer be captured by a single log normal centred on $1.5 T_{\text{dust}}$. I instead adopt a log normal prior centred on the dust temperature ($\sim 38 \text{ K}$) for the cool phase and on 120 K for the warm phase, and I require the warm component to be no cooler than the cool one, and I also impose $\Omega_{\text{app}}^{\text{cool}} \geq \Omega_{\text{app}}^{\text{warm}}$, which reflects the physical expectation that the cool and more extended phase encloses the compact warm one. The CO(1–0) datum and all other settings are identical to the single component fit, and thus the two models differ only in the number of phases. The complete results are tabulated in the lower block of Table 9. The decomposition is robust and physically sensible on every split, showing a *cool* phase at $T_{\text{cool}} \simeq 35\text{--}38 \text{ K}$, tied to the dust temperature, and a *warm* phase at $T_{\text{warm}} \simeq 125\text{--}150 \text{ K}$ that carries the bulk of the mid and high J ladder. The warm phase is systematically hotter than the single-component value ($T_{\text{kin}} \simeq 66\text{--}112 \text{ K}$), which is consistent with the single-phase solution representing a luminosity-weighted average biased towards the warm gas that dominates the CO ladder, while the cool phase is required specifically to reproduce the low J emission. The gas densities are $\log(n_{\text{H}_2}^{\text{cool}}/\text{cm}^{-3}) \simeq 3.6\text{--}4.0$ and $\log(n_{\text{H}_2}^{\text{warm}}/\text{cm}^{-3}) \simeq 3.7\text{--}4.1$, and the warm phase is the denser and more compact one in the complete and lensed samples. The

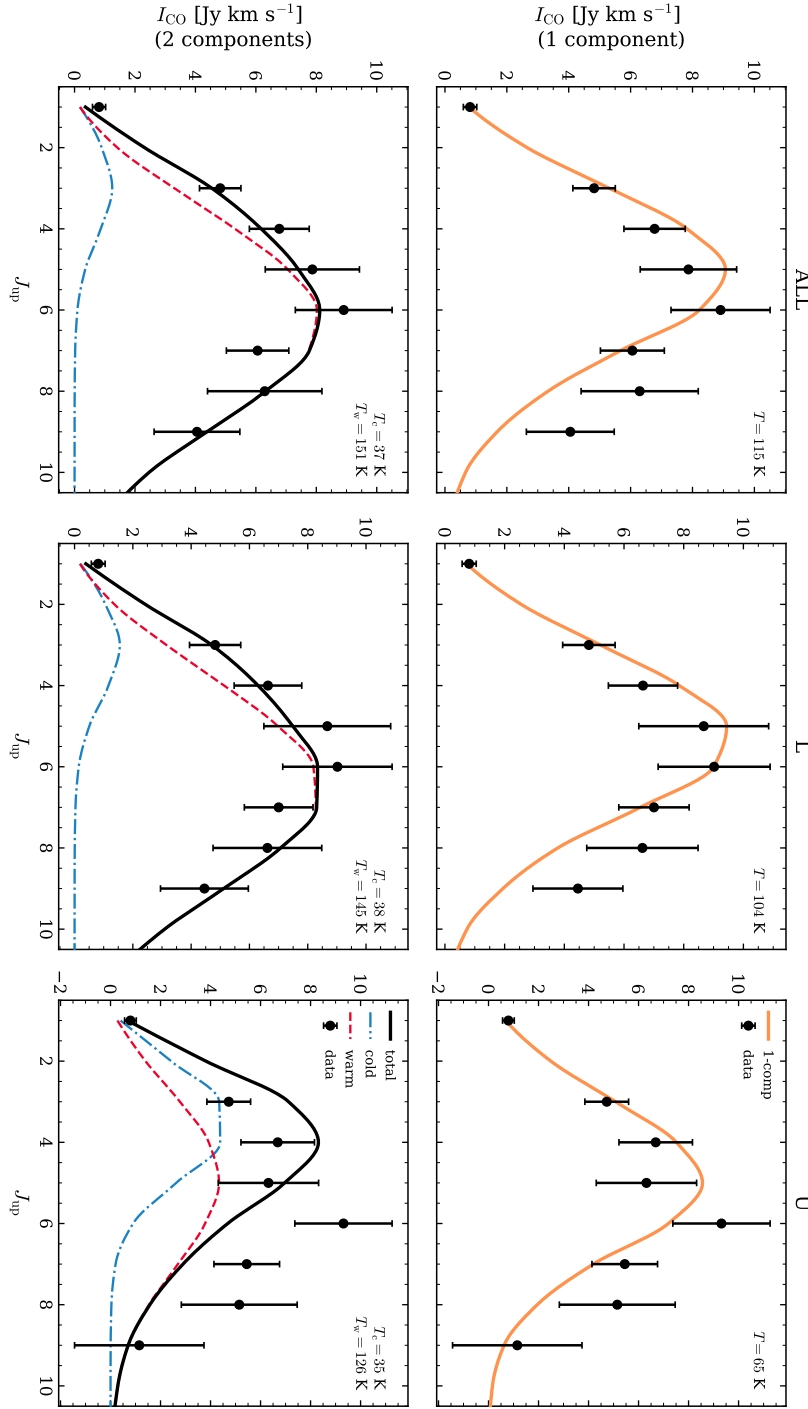


Figure 55: CO SLED fits across splits. Top row: single-component fit for the complete, lensed and unlensed sample; bottom row: two-temperature fit with cool (dash-dotted) and warm (dashed) phases. Comparing columns shows the lensed/unlensed consistency; comparing rows shows the improvement at low J from the cool phase.

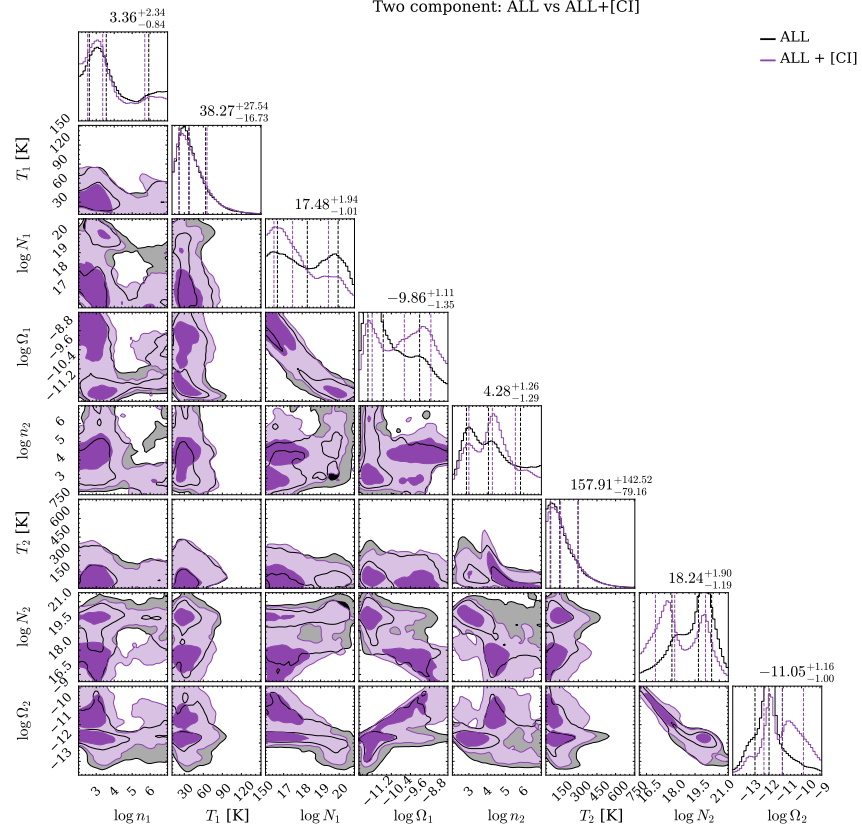


Figure 56: Marginalised one- and two-dimensional posteriors of the two-temperature decomposition for the complete sample, showing the cool- and warm-phase parameters jointly. The cool phase ($T_{\text{cool}} \simeq 37$ K, tied to the dust temperature) fills in the low- J flux and relaxes the CO(1–0) residual to $+0.2\sigma$, while the warm phase carries the mid/high- J ladder. Within each phase the same n_{H_2} – T_{kin} anti-correlation is present, so the individual densities and temperatures remain broad (and prior-dependent) while the thermal pressure stays the better-determined combination. Contours are the 1 and 2σ credible regions; black shows the fit without the [C I](2 – 1) datum, purple the fit including it.

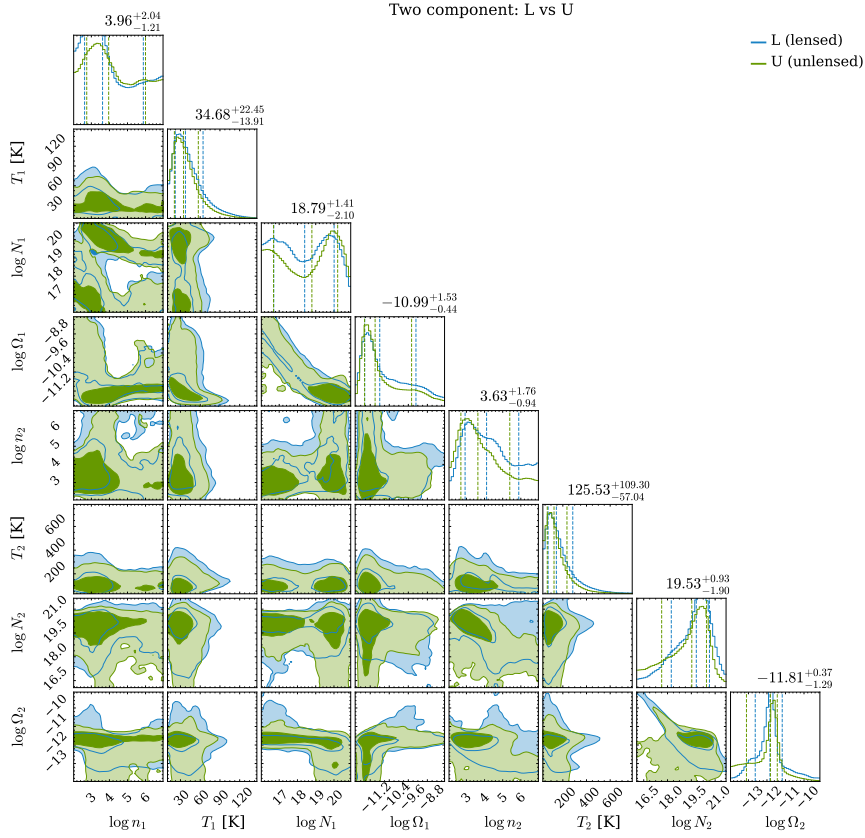


Figure 57: Lensed vs unlensed posteriors, two-component (cool+warm) fit. Overlaid 1–2 σ contours for the L (red) and U (blue) subsamples. The mutual overlap in every projection quantifies the absence of a significant lensing-induced difference in the excitation parameters.

decisive result is the behaviour of the anchored CO(1–0) point. While the single-phase model leaves it underfit at $+1.3\text{--}1.5\sigma$, the two-temperature model reproduces it to $+0.2\sigma$ on both the lensed and unlensed subsamples, because the cool component fills in exactly the low J flux that a single warm phase cannot. The model predicted line ratio relaxes from $r_{31}^{\text{model}} \simeq 1.0\text{--}1.1$ in the single phase case towards the observed value once the cool phase is present. This is the clearest evidence in the data that a distinct low-excitation reservoir accompanies the warm and star-forming gas. In this two-temperature model, anchoring CO(1–0) turns the low J turnover into a measurement of the cool reservoir, and this breaks the degeneracy and relaxes the prior tension.

The two phases are clearly separated in thermal pressure. For the complete and lensed samples, the cool component sits at $\log(P_{\text{th}}/k_{\text{B}} [\text{K cm}^{-3}]) \simeq 5.2$ and the warm one at $\simeq 6.3$, which shows a difference of about an order of magnitude that reflects the warmer, denser and more compact nature of the high excitation gas. In the unlensed sample, the separation is milder, with the cool and warm phases at $\simeq 5.5$ and $\simeq 5.8$ respectively, which is driven by a denser cool component and a cooler, less dense warm one, and given the larger uncertainties of the U subsample, this is fully consistent with the behaviour of the other splits. In all cases, the warm phase carries the higher pressure, as expected for the compact gas that dominates the mid and high J ladder.

Comparing the lensed and unlensed subsamples, I find no significant difference in either the single component or the two component fits. The cool phase sits at $\sim T_{\text{dust}}$ on every subsample, and the warm phase of U is marginally cooler than that of L (127 vs 148 K) but fully consistent within the uncertainties, which are over 50 K. As shown in the bottom right panel of Fig. 55, the two temperature decomposition holds also for the U subsample, because despite the larger statistical uncertainties associated with the smaller number of unlensed sources and the correspondingly noisier SLED, a distinct cool component (dash dotted) is still clearly required to reproduce the low J turnover, and this closely mirrors the behaviour seen for the lensed and the complete sample, while the warm phase (dashed) again dominates the mid and high J ladder. This indicates that the need for a second cool excitation component is not an artifact of the higher SNR of the lensed sample, but a genuine feature of the unlensed population as well. I conclude that differential lensing does not detectably bias the recovered excitation, and that the two subsamples trace the same molecular gas conditions. For the complete sample, I also include the [C I](2 – 1) line, which is associated with the cool phase in the fit. Whereas the single-phase model fails to reproduce it ($+2.6\sigma$), the two-component model matches it to within -0.5σ , and the implied excitation temperature ($T_{\text{ex}} \sim 22$ K) becomes physically sensible.

The [CI] line therefore corroborates, independently of the anchored CO(1–0) point, the need for a distinct cool and low-excitation reservoir. The marginalised posteriors of the two temperature fit are shown in Fig. 56 for the complete sample, and they are compared between the lensed and unlensed subsamples in Fig. 57. Within each phase, the same density and temperature anti-correlation seen in the single component case is present, so the individual densities and temperatures remain broad and prior dependent, while the thermal pressure stays the better determined combination, and the L and U contours again overlap in every projection.

Finally, it is important to notice that the individual excitation parameters of both the single component and two component fits, in particular $\log(n_{\text{H}_2}/\text{cm}^{-3})$ and T_{kin} , are still appreciably prior dependent. Their marginalised posteriors are broad, and in the two-component case they are only weakly peaked away from the temperature priors and from the ordering constraint $\Omega_{\text{app}}^{\text{cool}} \geq \Omega_{\text{app}}^{\text{warm}}$. This is not unexpected, given that the stacked SLED comprises only seven line points (plus the anchored CO(1–0) datum) built from a stacking analysis rather than from a large number of independent and high SNR detections per source, so the priors are doing genuine work in regularising an otherwise poorly constrained parameter space. Two main results are robust against this prior dependence. First, the thermal pressure $P_{\text{th}}/k_{\text{B}} = n_{\text{H}_2}T_{\text{kin}}$ is tightly constrained by the well-determined density and temperature anti-correlation, and for each phase it remains stable across the complete, L, and U subsamples. Second, the data themselves require a cool phase to reproduce the anchored CO(1–0) point, and this requirement is driven by the measurements rather than by the priors. I regard the individual n_{H_2} and T_{kin} values, especially those of the warm phase, as indicative rather than as tightly constrained measurements.

6.5 DISCUSSION

6.5.1 Stacked CO SLED

Figure 58 compares the average L_{IR} -normalised CO SLED of Section 6.3 with the statistical SLEDs available in the literature, all of them normalised to the CO(3 – 2) transition so as to allow a direct comparison of the shapes. A direct comparison of the absolute levels is difficult, since all the SLEDs presented are constructed with different normalisations, which alone can shift the line ratios by 25–65% [36]. Among the DSFG samples, Bothwell et al. [32] normalised by the FIR luminosity and Spilker et al. [130] by the 1.4 mm flux density, while Yang et al. [121], Boogaard et al. [118], Birkin et al. [33] and this work adopt a common L_{IR} normalisation. Some of these SLEDs are averages of individually measured sources [32, 33, 105, 121, 131], while

Table 9: LVG excitation parameters of the stacked CO ladder.

Parameter	Complete	L	U
Single component			
$\log(n_{\text{H}_2}/\text{cm}^{-3})$	$3.26^{+0.56}_{-0.48}$	$3.38^{+0.65}_{-0.52}$	$3.43^{+0.90}_{-0.59}$
$T_{\text{kin}} [\text{K}]$	112^{+80}_{-49}	103^{+69}_{-43}	66^{+54}_{-31}
$\log(N_{\text{CO}}/\Delta v)$	$17.51^{+0.41}_{-0.34}$	$17.53^{+0.44}_{-0.43}$	$17.53^{+0.45}_{-0.49}$
$\log(P_{\text{th}}/k_{\text{B}})$	5.31	5.40	5.25
$K_{\text{vir}}^{\text{LVG}}$	2.0	2.2	2.3
χ^2_{v}	2.1	1.8	1.3
resid ₁₀	+1.5 σ	+1.5 σ	+1.3 σ
Double component			
Cool phase			
$\log(n_{\text{H}_2}/\text{cm}^{-3})$	$3.60^{+2.38}_{-0.97}$	$3.62^{+2.24}_{-0.99}$	$3.99^{+1.99}_{-1.22}$
$T_{\text{kin}} [\text{K}]$	37^{+26}_{-15}	38^{+27}_{-16}	35^{+22}_{-14}
$\log(N_{\text{CO}}/\Delta v)$	$18.3^{+1.6}_{-1.6}$	$18.4^{+1.6}_{-1.7}$	$18.8^{+1.4}_{-2.1}$
Warm phase			
$\log(n_{\text{H}_2}/\text{cm}^{-3})$	$4.12^{+1.74}_{-1.24}$	$4.12^{+1.79}_{-1.19}$	$3.65^{+1.75}_{-0.95}$
$T_{\text{kin}} [\text{K}]$	146^{+143}_{-75}	148^{+139}_{-72}	127^{+105}_{-57}
$\log(N_{\text{CO}}/\Delta v)$	$19.7^{+0.8}_{-1.7}$	$19.8^{+0.9}_{-1.6}$	$19.5^{+1.0}_{-1.9}$
χ^2_{v}	1.5	1.4	2.0
resid ₁₀	+0.2 σ	+0.2 σ	+0.2 σ

Note: Top: single-component fit; Bottom: two-temperature decomposition (cool/warm phase). Values are posterior medians with 16–84th percentile uncertainties. $P_{\text{th}}/k_{\text{B}}$ is the thermal pressure $n_{\text{H}_2} T_{\text{kin}}$; $K_{\text{vir}}^{\text{LVG}}$ the intrinsic virial parameter; “resid₁₀” the residual on the anchored CO(1–0) point; χ^2_{v} and resid₁₀ in the two-component table refer to the joint two-phase fit.

others are obtained by stacking [36, 130, 191]. These samples also span different populations, moving from the largely lensed sources of Yang et al. [121] to the mostly unlensed ones of Boogaard et al. [131] and Birkin et al. [33].

The SLED presented here rises to a peak at $J_{\text{up}} \sim 6$ and turns over sharply at $J_{\text{up}} \geq 7$. Up to $J_{\text{up}} \sim 4\text{--}5$ it follows the bulk of the DSFG composites closely, and at the peak it is among the more excited samples, comparable to the SPT stacks [36, 130] and to the HerBS one from Hagimoto et al. [191]. Its distinctive feature is the pronounced high- J turnover: whereas the SPT and HerBS composites remain elevated out to $J_{\text{up}} \sim 7$, this ladder declines beyond its peak and converges at $J_{\text{up}} = 7\text{--}8$ with the least excited samples, namely the turbulence-modelled one of Harrington et al. [105] and the blind, gas-selected one of Boogaard et al. [131]. This turnover most plausibly reflects the excitation of the dominant molecular phase rather than the presence of a more extreme, highly excited gas component. The lensed ladder (dashed line in Fig. 58) stays elevated at high J and closely resembles the largely lensed, *Herschel*/H-ATLAS-selected sample of Yang et al. [121] and the SPT stacks of Spilker et al. [130] and Reuter et al. [36], whereas the unlensed one shows the sharper turnover. The difference between the two subsamples may nonetheless be a matter of statistics and of spatial filtering, since the unlensed sources are systematically under-represented with respect to the lensed ones, especially at high J , where they are observed only at high angular resolution (Fig. 43), so that claiming a genuine departure from these composites would require better statistics; within the current uncertainties, the two ladders remain compatible.

6.5.2 Gas derived properties

These LVG results can be compared with the body of CO excitation studies of high- z DSFGs. Spilker et al. [130], modelling the stacked ^{12}CO and ^{13}CO ladder of the SPT sample with RADEX, found that a single representative phase with $T_{\text{kin}} \sim T_{\text{dust}}$ requires moderately high densities, $\log(n_{\text{H}_2}/\text{cm}^{-3}) \sim 4\text{--}5$, comparable to the warm-phase densities recovered here. The single-component fit of Section 6.4.2 ($\log(n_{\text{H}_2}/\text{cm}^{-3}) \simeq 3.3$, $T_{\text{kin}} \simeq 60\text{--}110$ K, $\log(N_{\text{CO}}/\Delta v) \simeq 17.5$) is consistent with this picture and reproduces the mid- and high- J ladder with warm, moderately dense gas. I note that Spilker et al. [130] additionally included the ^{13}CO isotopologue and the dense-gas tracers (HCN, HNC, HCO^+ , CN), modelled as a separate dense-gas component; the ^{13}CO lines in particular constrain the line optical depth and the $[^{12}\text{CO}/^{13}\text{CO}]$ abundance, breaking part of the $N_{\text{CO}}\text{--}\tau$ degeneracy that a ^{12}CO -only fit cannot address. The quantity that is best determined here, and stable across the complete, lensed and unlensed splits, is the thermal pressure $\log(P_{\text{th}}/k_{\text{B}} [\text{K cm}^{-3}]) \simeq 5.3$. It

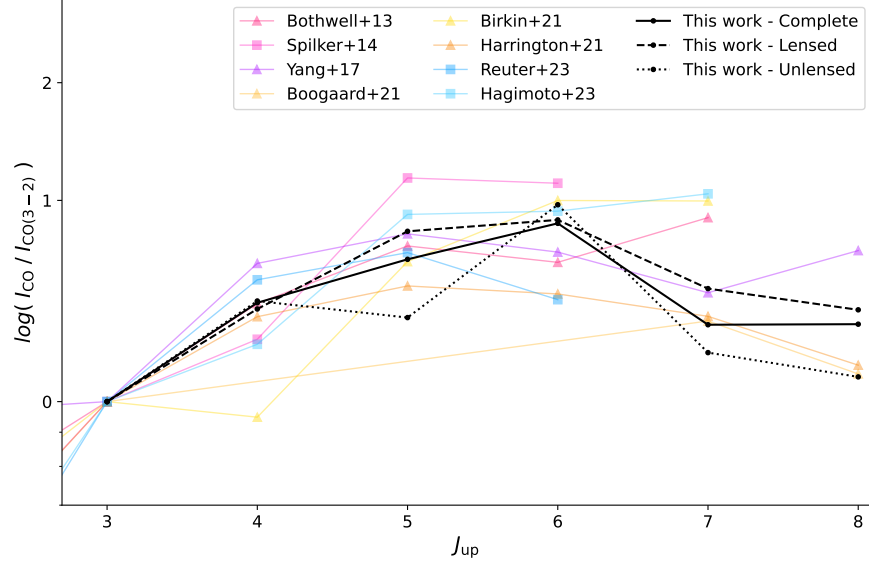


Figure 58: Resulting average CO SLED compared with statistical SLEDs from the literature, all normalised to the CO(3 – 2) transition ($\log[I_{\text{CO}}/I_{\text{CO}(3-2)}]$) to allow a direct comparison of shapes independent of the different absolute normalisations. *Triangles* denote ensemble averages of individually-measured sources [32, 33, 105, 121, 131], while *squares* denote SLEDs obtained from spectral or visibility stacking [36, 130, 191]. The results of this analysis are shown in black: the complete sample (solid), the L subsample (dashed), and the U subsample (dotted). This ladder peaks at $J_{\text{up}} \sim 6$; the sharp high- J turnover is driven by the unlensed subsample, which lacks short-baseline ACA data at $J_{\text{up}} \geq 8$, whereas the lensed ladder declines more gently and closely resembles the composite of Yang et al. [121].

falls squarely within the range $P_{\text{th}}/k_{\text{B}} \sim 10^5\text{--}10^6 \text{ K cm}^{-3}$ derived by Yang et al. [121] for individually modelled lensed SMGs, which indicates that the average excitation of this population is representative of the class. As Spilker et al. [130] also cautioned, describing a heterogeneous, multi-source stack with a single phase is a deliberate simplification, and the systematic under-prediction of the anchored CO(1–0) point ($+1.3\text{--}1.5\sigma$) is precisely the signature of this limitation, which motivates the two-component description.

The two-temperature decomposition recovers a cool phase at $T_{\text{cool}} \simeq 32\text{--}38 \text{ K}$, similar to the dust temperature, and a warm phase at $T_{\text{warm}} \simeq 125\text{--}150 \text{ K}$ carrying the bulk of the mid- and high- J emission. This agrees well with the two-component analysis of Yang et al. [121], which retrieved a low-excitation phase at $T_{\text{k}} \sim 20\text{--}30 \text{ K}$, less correlated with the star formation, and a high-excitation phase at $T_{\text{k}} \sim 60\text{--}400 \text{ K}$ tied to the ongoing activity; that modelling, which also incorporates the [CI](2–1) line observed together with CO(7–6), yields densities $\log(n_{\text{H}_2}/\text{cm}^{-3}) \sim 2.7\text{--}4.6$ that bracket the values of both phases recovered here. A different modelling philosophy was adopted by Harrington et al. [105], who described the unresolved ISM of the *Planck*-selected lensed starbursts not with one or two discrete phases but with a continuous, turbulence-motivated log-normal density distribution. Despite the very different parameterisation, the sample-averaged density $\langle \log(n_{\text{H}_2}/\text{cm}^{-3}) \rangle \sim 4.3$ and kinetic temperature $\langle T_{\text{kin}} \rangle \sim 119 \text{ K}$ of that work bracket the warm-phase values found here. A further difference is that Harrington et al. [105] fitted the two [CI] fine-structure lines and the dust continuum *simultaneously* with the CO ladder. It is precisely this joint constraint on the atomic gas and on the dust that allows that analysis to *derive* $\langle T_{\text{kin}}/T_{\text{dust}} \rangle \sim 2.5$ from the data, whereas in the ^{12}CO -only fit presented here the same ratio is effectively *set* by the log-normal prior on T_{kin} , centred on $1.5 T_{\text{dust}}$. These data cannot therefore constrain $T_{\text{kin}}/T_{\text{dust}}$ independently, and the apparent difference reflects a prior choice rather than a physical discrepancy.

6.6 WRAP-UP

This chapter has presented the first scientific application of ViSTA, namely the average CO SLED of 104 DSFGs at $z \sim 1.5\text{--}5$, reconstructed from $J_{\text{up}} = 3$ to $J_{\text{up}} = 9$ by stacking directly in the uv plane a heterogeneous set of archival ALMA observations that no image-plane combination could have handled.

The average ladder peaks at $J_{\text{up}} \sim 5\text{--}6$, which is characteristic of moderately dense, intensely star-forming gas, and places this population among the more excited DSFG composites in the literature. Once normalised to a common infrared luminosity, the lensed and unlensed subsamples are fully compatible across the whole ladder

and yield consistent excitation parameters, with no compelling evidence of differential lensing, which supports the use of strongly lensed DSFGs as representative tracers of the underlying population. The radiative transfer modelling shows that a single warm phase reproduces the mid- and high- J transitions but systematically underpredicts $\text{CO}(1-0)$, whereas a two-temperature decomposition, with a cool phase tied to the dust temperature and a warm one carrying the bulk of the high- J emission, resolves the tension and is independently corroborated by the $[\text{C I}](2-1)$ line. Being based on ^{12}CO alone and derived from a heterogeneous stack, the individual densities and temperatures should be regarded as indicative of the average conditions rather than as a definitive characterisation.

The reconstruction of the average SLED, the comparison between lensed and unlensed populations, and the attempt to determine the average gas properties presented in this chapter mark the first complete scientific application of ViSTA. The final chapter summarises these results, retraces the path from sample selection to software development, and outlines future prospects, such as targeting new emission lines or incorporating data from other interferometers.

CONCLUSIONS AND PROSPECTS

The work presented in this thesis arises from the desire to uncover the physical conditions of DSFGs at cosmic noon using (sub-)millimetre line emission from their ISM, taking full advantage of rich archival datasets. Tracing this gas requires a statistical stacking approach, both to recover faint, individually undetectable lines and to characterise the average properties of the broader population (Ch. 1). To this end, I built a sample from wide-area *Herschel* surveys combined with archival ALMA data, yielding a final set of 202 sources spanning different redshifts and comprising both strongly lensed and unlensed galaxies (Ch. 2). While this ensemble provides excellent statistical power, the underlying observations proved highly heterogeneous in angular resolution, spectral setup, sensitivity, and redshift, such that combining them directly in the image plane would introduce severe artifacts and beam discrepancies. To solve this challenge, I developed ViSTA, a stacking framework that operates natively in the visibility domain, combining disparate observations directly *before* rather than after image reconstruction (Ch. 3). By working on the visibilities themselves, ViSTA preserves the Gaussian statistics of the thermal noise and remains consistent across angular scales, avoiding the artifacts that such a heterogeneous archive would otherwise introduce once reconstructed into images. On the methodological side, the validation of ViSTA on simulated and real observations established that faint line signals buried in noise are recovered reliably (Ch. 4). The advantage of visibility domain stacking over image plane combination scales inversely with the sensitivity of individual observations, meaning that shallow, faint, and non-uniform data benefit the most from this approach. The high-performance reimplementation (Ch. 5) replaced the sequential disk-bound execution model with a streaming architecture backed by a compiled C++ and OpenMP kernel, with possible Graphics Processing Unit (GPU) acceleration. This version improved throughput by more than an order of magnitude and enabled archive-scale analyses. Furthermore, the resulting framework enables the analysis of archival data originating also from other standard radio or (sub-)millimetre interferometers.

To test its scientific capabilities, I used ViSTA to reconstruct and model the average CO SLED of 104 *Herschel*-selected galaxies at $z \sim 1.5\text{--}5$ using archival ALMA observations from CO(3 – 2) up to

CO(9 – 8), along with the detection of the [C I](2 – 1) line (Ch. 6). The resulting ladder peaks around $J_{\text{up}} \sim 5\text{--}6$, which indicates moderately dense and intensely star-forming gas. The analysis aimed to retrieve the representative average properties of the entire galaxy population. Splitting the sample into confirmed lensed and unconfirmed systems revealed consistent CO SLEDs and excitation parameters across both groups. Moreover, the absence of strong evidence for differential lensing supports the validity of strongly lensed galaxies as representative proxies for the broader population. Non-LTE radiative transfer modelling indicated that a single warm phase underpredicts the lowest CO transitions, whereas a two-temperature decomposition consisting of a cool phase at $T_{\text{cool}} \simeq 32\text{--}38\text{ K}$ and a warm phase at $T_{\text{warm}} \simeq 125\text{--}150\text{ K}$ resolves this tension, yielding results consistent with previous literature studies. Overall, this modelling approach and the resulting constraints provide a useful benchmark for studying the ISM of high-redshift galaxies and demonstrate a complete end-to-end application of visibility-domain stacking on real data, paving the way toward new scientific investigations as discussed in the following section.

7.1 FUTURE PROSPECTS

7.1.1 *Expanding the CO ladder*

The ensemble of 202 sources serves as the source pool for every stacking experiment, with subsamples defined by the relevant transition. The CO analysis presented here focuses on the 104 sources that have both archival CO coverage and accurate positions. The remaining targets feature coverage of other spectral lines or of the dust continuum only, and remain available for future applications. The logical continuation of this work is to expand the observational coverage of the CO SLED. The analysis presented in this thesis restricted the target transitions to the range between CO(3–2) and CO(9–8), as these lines represented the observational boundary where archival ALMA coverage provided a sufficiently large sample size to maintain statistical power across both lensed and unlensed galaxies. Extending this excitation ladder requires acquiring a statistically robust dataset across both lower and higher rotational transitions. On the low-excitation side, fundamental lines such as CO(1–0) and CO(2–1) can be directly probed at cosmic noon using the newly operational ALMA Band 1 and Band 2 receivers, together with the ongoing Wideband Sensitivity Upgrade [215]. Obtaining these low- J observations will replace the synthetic r_{31} prior used in the radiative transfer modelling with a direct empirical measurement, thereby resolving the residual tension on the lowest transition and tightening the constraint on the cool gas mass. On the high-excitation side, transitions above CO(9–

8) can be systematically targeted using the higher-frequency ALMA bands, providing a complete view of the warm and dense gas phase. In addition, securing uniform line coverage across all targets will be essential to eliminate sampling biases. This will ensure that reconstructed excitation profiles reflect genuine physical diversity, rather than observational incompleteness.

In parallel with these ALMA upgrades, the new implementation of ViSTA enables the exploitation of additional astronomical archives. By natively ingesting data in the MS standard format, the software can stack visibilities from any facility adopting this standard across different spectral setups and epochs onto a common rest-frame uv -plane. As illustrated in Figure 59, the radio and (sub-)millimetre observational landscape comprises a vast network of facilities spanning several orders of magnitude in frequency, ranging from low-frequency arrays such as LOFAR, MeerKAT, eMERLIN, and SKA, through centimetre facilities such as the VLA and ATCA, up to high-frequency interferometers such as NOEMA and the SMA. Combining these multi-facility datasets will effectively broaden the dynamic range of the average CO SLED and establish a complete multi-phase model of the ISM in high-redshift galaxies. Ultimately, future extensions and computational optimisations of the codebase could allow the framework to detach even further from specific file formats like the MS, paving the way for a fully format-agnostic architecture capable of processing raw visibility streams from next-generation observatories.

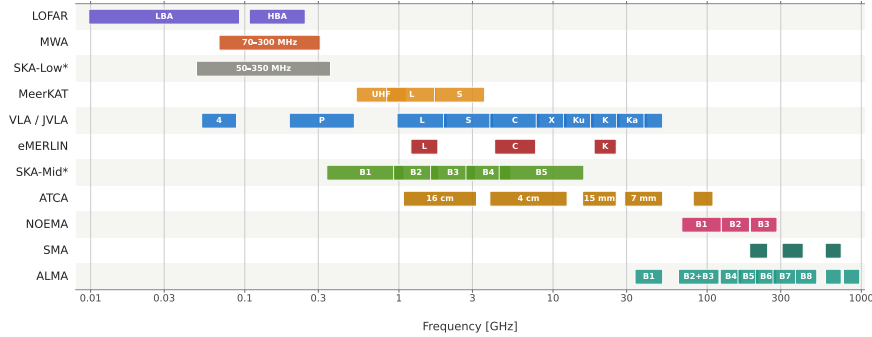


Figure 59: Frequency coverage of major radio, millimetre, and sub-millimetre interferometers. The native visibility-domain framework of ViSTA is not restricted to ALMA, enabling the direct combination and stacking of archival visibilities across disparate facilities ranging from low-frequency centimetre arrays to high-frequency submillimetre telescopes.

7.1.2 *Beyond CO: Additional tracers*

A second natural application of the stacking framework is extending the analysis beyond the ^{12}CO rotational ladder to the complementary diagnostics introduced in Section 1.2.3. These tracers probe the multi-phase ISM and break the physical degeneracies inherent to CO-only modelling. Because many of these species are intrinsically faint and often fall below single source detection thresholds, visibility stacking provides the ideal tool to recover their emission. As illustrated in Figure 60, the ALMA archival dataset assembled in this work offers a wealth of serendipitous and secondary coverage for these transitions, including fine-structure atomic carbon, ionised carbon, dense gas molecules, water transitions, and faint isotopologues.

Each tracer targets distinct physical regimes. The [CII] line provides the largest number of available observations in this archive, having been heavily targeted in high-redshift studies as the principal ISM coolant and star-formation indicator [216, 217]. Optically thin [CI] fine-structure lines trace the bulk cold gas mass independently of CO excitation [109, 115], while fainter ^{13}CO and C^{18}O isotopologues constrain line optical depths and abundance ratios, directly lifting column density degeneracies and probing the stellar initial mass function [130, 218]. High critical density species such as HCN, HCO^+ , and HNC isolate dense star-forming cores, where line ratios distinguish between UV-driven PDRs and XDRs powered by AGN [219, 220]. In addition, water transitions trace warm FIR radiation fields [136, 137], and molecular ions like H_2O^+ provide rare constraints on cosmic-ray and X-ray ionisation rates [209, 221].

In addition to these primary diagnostics, stacking other, sometimes fainter, spectral transitions would extend the analysis toward a comprehensive chemical inventory of the ISM at cosmic noon [222]. Probes such as CH^+ and SiO pinpoint turbulent dissipation and fast shocks [223, 224], while radicals like CN and CCH map radiation-damaged environments [225, 226]. Similarly, outflow tracers like OH track galactic feedback mechanisms [227], and millimetre recombination lines offer dust-unbiased measures of ionised gas [228]. Combining these diagnostics through visibility stacking transforms incidental archival data into a unified, population-averaged view of high-redshift interstellar physics.

7.1.3 *Redshift evolution and subsamples*

As soon as a statistically significant sample is assembled, additional subsample divisions beyond the lensed and unlensed comparison become feasible. The most natural axis for this division is redshift. Binning sources in redshift and stacking each interval makes it possible to track how the average CO SLED evolves over cosmic time.

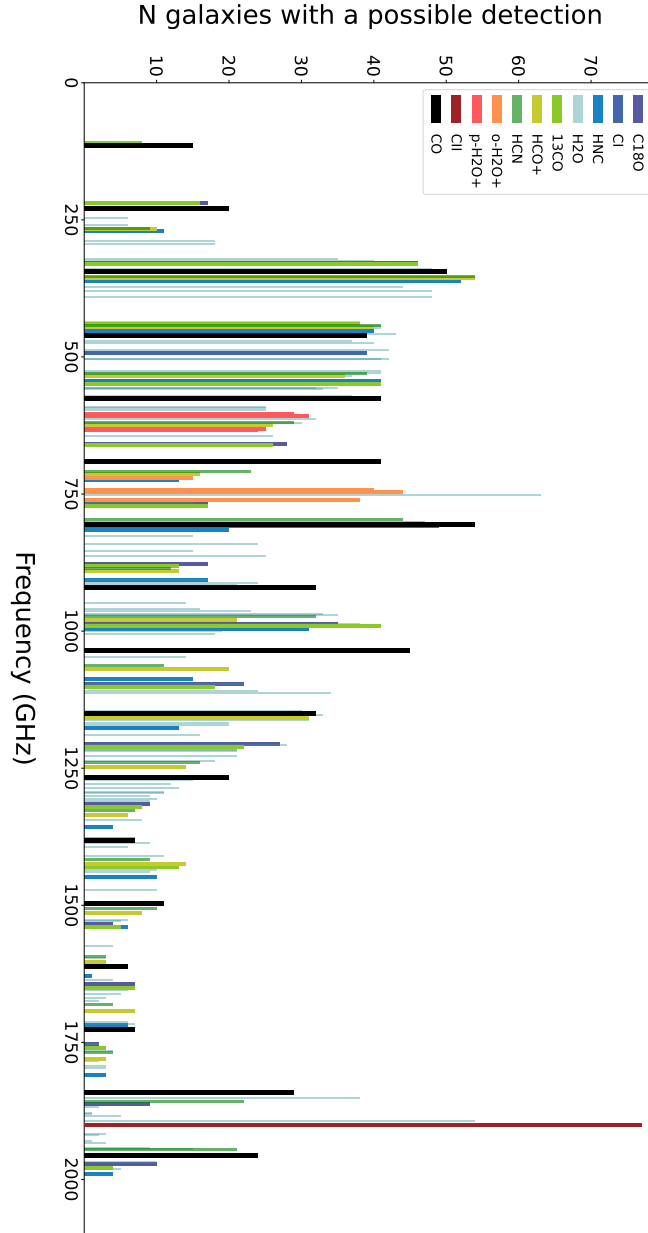


Figure 60: Number of galaxies in the sample for which a possible detection of a given emission line is present in the archival data, as a function of the line rest-frame frequency. In addition to the CO rotational ladder, the sample provides serendipitous or secondary coverage of the atomic carbon line [C I], the fine structure line [C II], the water line H_2O and its ions $\text{o-H}_2\text{O}^+$ and $\text{p-H}_2\text{O}^+$, the dense-gas tracers HCN, HCO^+ and HNC, and the isotopologues ^{13}CO and C^{18}O . These lines, individually too faint for a source-by-source study, are natural targets for the same visibility-domain stacking approach applied here to CO.

This links the stacked ladders to the excitation evolution observed in main-sequence galaxies [117]. I did not perform this differential analysis on the CO sample because the available statistics remain insufficient. Even in the most populated range between $z \sim 2$ and 4, isolating the true impact of redshift evolution remains difficult, as the reconstructed CO SLED suffers from non-uniform transition coverage across targets.

This framework applies equally well to dust continuum emission. Indeed, by binning the sample in redshift, I can use ViSTA to study the evolution of the population's average SED across cosmic time. This evolution, especially the shift in the SED peak alongside the rise in dust temperature and mean radiation field with redshift, serves as a well-established diagnostic of galaxy evolution. Similar stacking studies binned by redshift and stellar mass have mapped these trends extensively [87, 133, 229]. Stacked SEDs also provide empirical templates and priors for fitting individual and poorly sampled sources, effectively bridging population-averaged results with single-object studies.

Beyond redshift, this stacking strategy can easily extend to other physical properties. Splitting the sample by infrared luminosity, dust temperature, apparent compactness, or AGN presence and activity helps isolate individual drivers of excitation that a global average necessarily blends together. Indeed, composite stacking studies show that dust temperature correlates directly with ladder excitation [36, 130]. In this framework, the lensed-versus-unlensed comparison is just the first step toward a strategy that will become increasingly powerful as sample sizes grow.

7.1.4 *Beyond DSFGs: A general framework for visibility stacking*

While the primary focus of this work has been dedicated to the dense interstellar medium of DSFGs at cosmic noon, the architectural foundation of ViSTA is fundamentally general. Wherever observations allow for spectral rest-framing and spatial configuration matching, operating natively in the visibility domain is entirely feasible and should function effectively. Consequently, the methodology is not restricted to high-redshift dusty starbursts and can in principle be adapted to a wide array of astronomical populations.

A natural extension is the population of ultraviolet-selected Lyman Break Galaxies (LBGs) at $z \sim 4 - 8$, where individual dust-continuum and $[\text{C II}]/[\text{O III}]$ line detections are rare and continuum and line stacking is already standard practice, as in the ALPINE and REBELS large programmes [230–233]. In those cases the stacked datasets come from a single programme, with a homogeneous spectral setup and array configuration, so that the data can be combined directly. The contribution of ViSTA in this context is not the stacking

itself but the ability to combine, in the visibility plane, observations that are heterogeneous in redshift, spectral setup and angular resolution, which is what makes an archival, multi-programme sample stackable at all. A second, methodologically close application is HI 21 cm emission stacking, routinely performed directly on interferometric visibilities from the (u)GMRT and MeerKAT to recover the average cold-gas content of star-forming galaxies at $z \sim 0.2 - 1.4$, where per-source detections are below the noise [234–236]. More broadly, this framework extends to any population of sources that are individually faint yet statistically numerous, leveraging visibilities from facilities such as ALMA, NOEMA, or JVLA. Potential applications range from molecular-gas and dust stacking in AGN and quasar host galaxies to the study of quiescent and dwarf galaxies in the local Universe, and even to Galactic sources, where the rest-framing is not needed but ViSTA can still match the sampled angular scales across heterogeneous observations. However, a systematic exploration across these diverse astrophysical domains lies beyond the scope of the present work, leaving them as compelling directions for future methodological applications.

7.1.5 ViSTA in the SKAO era

The clear advantages of visibility stacking over traditional image-plane combination for noisy data demonstrated by ViSTA underscore its strategic relevance for the future of interferometric archives. This paradigm becomes particularly critical when considering next-generation observatories, most notably the Square Kilometre Array Observatory. As the largest radio telescope array ever constructed, spanning SKA Low in Australia for low-frequency operations and SKA Mid in South Africa for mid-frequency bands, the SKAO is optimised to revolutionise our understanding of the high-redshift Universe. Among its primary scientific objectives are mapping neutral hydrogen through the 21 cm line to trace the Epoch of Reionisation, tracking the structural assembly of galaxies across early cosmic epochs, and probing the cold gas reservoirs that fuelled star formation in the most distant systems. Achieving these transformative goals requires processing an unprecedented data stream, with raw visibility throughput reaching hundreds of terabits per second and generating petabytes of calibrated data daily. Because the long-term retention of raw uv data across decades of operations is computationally and financially unfeasible, current SKAO operational models plan to archive only reconstructed image-plane products, discarding raw visibility streams after standard processing pipeline execution.

Within this changing context, ViSTA provides a double contribution. On one hand, it maximises the scientific return of public archives

where raw visibilities are preserved, recovering faint and subthreshold signals that remain invisible to image-based reanalyses. On the other hand, it offers an essential diagnostic and benchmarking framework. Indeed, by enabling direct and controlled comparisons between visibility-domain and image-domain stacking on identical target samples, ViSTA yields the quantitative empirical evidence needed to measure the precise information loss incurred when transitioning exclusively to the image plane. It thus equips the astronomical community with the tools to critically assess whether discarding raw visibilities is a sustainable long-term strategy, or if preserving uv data remains indispensable for subthreshold science cases in the next generation of radio astronomy. A further possibility, intermediate between preserving the raw visibilities and keeping only the images, is to archive the visibilities pre-averaged onto a common, controlled grid, or the compressed statistics of Section 5.3, which are far smaller than the original data and remain sufficient for the models considered here. Using a tool like ViSTA keeps such a reduction reproducible across heterogeneous datasets, and comparing the two domains quantifies the information lost in the process.

APPENDIX



SUPPORTING CODE FOR THE ANALYSIS

This appendix describes two ancillary codes developed specifically for this thesis. These tools operate independently of ViSTA but they remain essential for the preceding analysis. The first code is the simulator used to generate the synthetic ALMA datasets of Chapter 4. The second code is the routine building the archival line coverage catalogue. This algorithm establishes which transitions fall inside the observed bandpass for every galaxy of the sample and for every archival observation. It therefore defines the exact datasets eligible for a stack.

A.1 THE MOCK OBSERVATION SIMULATOR

The mock observations are generated with the `simobserve` task of CASA v.6 [187, 188]. This task takes as input a sky model alongside an antenna configuration file and an observing date and a total integration time. It then returns a simulated MS with the same structure as an actual ALMA observation. The task can also add thermal noise on the basis of realistic atmospheric models of the site following the radiometer equation already introduced as Equation 16. However, its corruption algorithm computes a single average T_{sys} for the whole execution, it assigns unit weights and uncertainties to all the visibilities and introduces no antenna-dependent phase delay. The resulting noise distribution is uniform and does not reflect the real conditions where every antenna looks through a different atmosphere. This means the stacking cannot be tested against realistic fluctuations. For this reason, the MS is generated in its noiseless configuration with `thermalnoise = ''` and the corruption is applied afterwards through custom Python routines built on the same atmospheric libraries.

The sky model is a cube of $N_x \times N_y$ pixels and N_c channels. It features a pixel scale that oversamples the synthesised beam and a channel width matching the spectral setup to be reproduced. Its spectral axis is centred on the observed frequency of the target transition $\nu_0 = \nu_{\text{rest}} / (1 + z)$. The rest frequency is kept identical for every mock galaxy while the redshift is drawn within the interval sampled by the mock population. All the lines therefore coincide in the rest frame by

construction. In each channel, the source emits a flat continuum plus a Gaussian line

$$S(c) = S_{\text{cont}} + S_{\text{line}} \exp \left[-\frac{(c - c_0)^2}{2\sigma_c^2} \right], \quad (39)$$

where c_0 is the central channel and σ_c is the line width expressed in channels. This flux is either deposited into a single pixel to reproduce a source much smaller than the synthesised beam or spread over a circular Gaussian of assigned angular size normalised to unit sum. In both cases, the flux integrated over the source equals $S(c)$ exactly in every channel. The recovered profile can then be compared with the input model without any residual calibration factor. The source is then displaced from the centre of the map by an integer number of pixels. The same simulator is also used with bright sources during the development of the pipeline, where each function can be validated against the corresponding CASA task and the recovered fluxes compared with the input model without the interference of noise. This reproduces the off-centre positioning typical of archival data and exercises the centering and the primary beam correction of Section 4.2.1.1. Its true coordinates are stored for later comparison with the recovered ones. The cube is written as a CASA image through the `ia` toolkit. The spectral reference value and the rest frequency are both set to ν_0 alongside a brightness unit of Jy/beam and a small circular restoring beam attached to the header. It is finally exported to FITS and passed to `simobserve`.

The corruption starts from the atmosphere, which is described with the CASA atmosphere toolkit. This is an implementation of the Atmospheric Transmission at Microwaves model [237] and serves as the reference radiative transfer model adopted by ALMA. A single spectral window centred on ν_0 is initialised over the atmospheric profile of the site. The precipitable water vapour is fixed to the value tabulated for the weather octile of the simulated band. The position of every antenna is read from the MS together with its pad elevation offset so each element is placed at its own altitude. For every integration, the direction of the source is converted to the horizontal frame of each antenna using `Astropy` [238, 239]. The airmass $A_i = \sec z_i$ follows from its own zenith angle and evolves along the elevation track of the source. The zenith opacity returned by the toolkit alongside the ambient and receiver temperatures are perturbed by random factors of up to ten per cent. These factors are drawn independently for each antenna and each mock observation so no two simulated executions share identical conditions. The system temperature of the antenna i becomes

$$T_{\text{sys},i} = T_{\text{CMB}} + \eta_s T_{\text{atm}} \left(e^{\tau_\nu A_i} - 1 \right) + (1 - \eta_s) T_{\text{amb}} e^{\tau_\nu A_i} + T_{\text{RX}} e^{\tau_\nu A_i}, \quad (40)$$

where η_s is the spillover efficiency and T_{atm} is the atmospheric temperature, while T_{amb} and T_{RX} are the ambient ground temperature and the receiver noise temperature, respectively. The code evaluates the sky contribution through the photon occupation number rather than through the brightness temperature since the Rayleigh-Jeans approximation is not accurate at submillimetre frequencies. Equation 40 is the limit to which that calculation reduces. It makes explicit that the sky and the spillover alongside the receiver noise are all amplified by the atmospheric attenuation. Antennas observing at lower elevation are assigned a systematically higher T_{sys} because the airmass appears in the exponent. This is precisely the antenna dependence missing from the native corruption.

The same airmass drives the atmospheric phase. The excess path length introduced by the water vapour along the line of sight of an antenna i is taken proportional to the water vapour column and to the airmass. The associated phase $\phi_i = 2\pi\Delta L_i/\lambda$ is wrapped in the $[0, 2\pi)$ interval. An interferometer measures phase differences so the visibility of the baseline (i, j) is corrupted by the differential term

$$V'_{ij} = V_{ij} e^{i(\phi_i - \phi_j)}, \quad (41)$$

which introduces a baseline-dependent and time-dependent phase. This term reproduces only the systematic component of the atmospheric delay, driven by the different elevation at which each antenna observes, and not the turbulent component associated with a moving phase screen. Since all the antennas observe the source at almost the same elevation, the resulting differential phases reach at most a few degrees on the longest baselines and the associated decorrelation is negligible. It should be stressed that a term of this kind is in any case removed by the standard phase calibration, so these mocks do not test the effect of residual phase errors on the stack (Sec. 3.1.4). Simulating those residuals realistically would require a turbulent phase screen with a prescribed phase RMS over a given timescale, applied to a mock calibrator as well as to the target and then calibrated out in the same way as real data, which is beyond the scope of these tests.

The thermal noise is then computed baseline by baseline from Equation 16 by replacing the single T_{sys} with the geometric mean of the two antenna values

$$\sigma_{ij} = \frac{4\sqrt{2}k}{\eta_a \eta_c \pi d^2 \sqrt{\Delta\nu} \Delta t} \sqrt{T_{\text{sys},i} T_{\text{sys},j}}, \quad (42)$$

and the real and imaginary parts of each visibility are subsequently perturbed independently with zero-mean Gaussian deviates of dispersion σ_{ij} . The SIGMA and WEIGHT entries are updated to σ_{ij} and σ_{ij}^{-2} . The statistical weights carried by the mock data now reflect the actual sensitivity of each individual baseline rather than the unit weights returned by the standard simulator. This allows the weighting schemes

of Section 4.2.3.2 to be exercised meaningfully. The corrupted visibilities are written to the `CORRECTED_DATA` column and the noiseless `DATA` column is left untouched. This ensures the true model remains available within the same dataset for direct comparison. The main table is processed in blocks of rows and the atmospheric quantities are recomputed only when the `TIME` column changes. This means they are updated once per integration rather than once per baseline. This caching makes the corruption of a full execution tractable.

Each mock source is corrupted and stored in its own MS, exactly as an individual archival observation would be. The corrupted dataset is finally imaged with `casatasks.tclean()` in `specmode = 'cube'` using the corrected column and natural weighting. The process employs a common restoring beam across channels and primary beam correction with `niter = 0` so the recovered signal is not altered by deconvolution. The phase centre is set to the true position of the source. The cell size is chosen as one fifth of the nominal angular resolution λ/B_{max} of the configuration in agreement with the oversampling criterion adopted in Section 4.2.3.4.

Nothing in the procedure is specific to a particular instrument or transition. The array enters only through the configuration file which is a plain text table listing the coordinates alongside the dish diameter and the pad name of each element. CASA is distributed with such files for the ALMA main array and the ACA as well as for the VLA and ATCA alongside the SMA and several other facilities. Switching interferometers therefore only requires substituting that file and updating the diameter entering Equation 42. The atmospheric and receiver quantities must instead be chosen consistently with the band in which ν_0 falls. This is done by adopting the values tabulated by the ALMA Observing Tool and by the ALMA Sensitivity Calculator for the appropriate weather octile. The rest frequency and the redshift interval alongside the continuum and line fluxes describe the modelled population. The source size and its offset from the map centre do the same and all these parameters can be reassigned without modifying the logic of the code.

A.2 BUILDING THE ARCHIVAL LINE COVERAGE CATALOGUE

Designing a stacking experiment requires mapping which spectral lines are covered by archival observations for a given target sample. To address this, the proposed routine cross-matches a catalog of galaxies with known spectroscopic redshifts against the ALMA archive to identify all covered transitions. It systematically evaluates every archival observation for each source and extracts the relevant data identifiers. The routine tracks spectral coverage rather than detection,

flagging a line whenever its redshifted frequency falls within an observed spectral window, regardless of its signal strength. This logic directly aligns with the operational paradigm of ViSta.

The catalogue is built programmatically by querying the ALMA Science Archive [240] alongside Splatalogue and VizieR [241] through astroquery [200]. Splatalogue [242] acts as the sole rest frequency reference. It is the compilation maintained by NRAO that merges the major spectroscopic databases. It flags one recommended rest frequency per transition so no line is counted twice. It is also the database that feeds the ALMA Observing Tool and CASA. This ensures the frequencies used here to predict the coverage match the ones used by the observatory to define the searched spectral setups.

To focus on extragalactic science, the reference line list retains only key transitions, including the primary component of rotational transitions where hyperfine splitting is unresolved at high-redshift line widths, the ground-state ladders of vibrationally or torsionally excited species, and the α -series recombination lines. The archive is then queried target by target using a cone search restricted to science observations, with the search radius set to a small fraction of the primary beam to ensure matches lie well within the field of view rather than at its noisy edges. Each returned observation generates a distinct table entry—allowing multiply observed galaxies to contribute several rows—from which the spectral window containing the representative line frequency is identified along with its frequency limits and channel width. A transition k is then flagged as covered when

$$\nu_{\min}^{\text{obs}} < \frac{\nu_k}{1 + z_g} < \nu_{\max}^{\text{obs}}, \quad (43)$$

with strict inequalities. In any case, a line falling at the exact boundary of a spectral window would be unusable, as its wings would be truncated, leaving no adjacent channels for continuum estimation. While the rest frame channel width is broadened by the factor $(1 + z_g)$, the velocity resolution remains invariant and must therefore be homogenized across the sample during the rebinning stage (Section 4.2.1.3).

Finally, each row is complemented with ancillary far-infrared and submillimetre photometry retrieved from survey catalogs following a fixed priority order. Unmatched entries are deliberately left undefined rather than assigned upper limits, preserving the distinction between an absent measurement and a true non-detection.

The resulting table compiles target properties, alongside observational metadata and covered transitions for each species. Sorting rows by rest frame frequency groups together all observations sampling the same spectral regime, regardless of individual source redshifts. Furthermore, aggregating the distinct galaxies that satisfy Equation 43 ensures that multiply observed targets are counted only once, yielding the actual number of unique sources available per transition.

A.3 IMAGE-PLANE STACKING

The image-plane stacking used as a benchmark in Chapter 4, denoted `ImSta`, follows the conventional approach in which each observation is imaged individually and the combination is performed on the reconstructed cubes [e.g., 130, 180]. To make the comparison with `ViSta` as fair as possible, the two methods share the same input data and the same pre-processing, and differ only in the domain in which the sources are combined. The procedure is implemented with the `spectral-cube` [243] and `radio-beam` [244] packages, and consists of the following steps.

IMAGING. Each dataset is first centred on its target and rest-framed as described in Sections 4.2.1.1 and 4.2.1.2, so that the source lies at the centre of the field and the spectral axis is expressed in the rest frame. Each dataset is then imaged separately with `casatasks.tclean()`, using its native uv coverage, natural weighting, and a pixel size set by its own synthesised beam, and deconvolved down to a threshold of a few times its noise level. The resulting cubes therefore differ in pixel size, synthesised beam, channel width and spectral coverage. In the image plane the primary beam correction is applied correctly in this step rather than as the single scalar factor used for the visibilities (Sec. 4.2.1.1). At the end of the stacking procedure, the stacked, primary beam corrected cube is used to measure the flux of the source, while the uncorrected cube can be used to estimate the noise level correctly.

SPECTRAL AXIS. Each cube is read with `SpectralCube.read()`, and its spectral axis is converted into radio velocity relative to the rest frequency of the target transition with `with_spectral_unit()`, so that the line lies at zero velocity for all sources. Since the synthesised beam of a cube generally varies slightly from channel to channel, a single beam per cube is defined with `beams.common_beam()`, the smallest one that encloses the beams of all its channels. The cubes are then cut to the common velocity range of the sample with `spectral_slab()`, keeping one additional channel at the edges so that the subsequent interpolation is not affected by edge effects. As in the uv plane (Sec. 3.3.3), all cubes are brought to the coarsest channel width of the sample, $\Delta v_{\max}$. Each finer cube, with channel width Δv , is first smoothed with `spectral_smooth()`, using a Gaussian kernel of width $\sqrt{\Delta v_{\max}^2 - \Delta v^2}$, which degrades its spectral resolution to that of the coarsest cube, and is then resampled onto the velocity axis of the coarsest cube with `spectral_interpolate()`, so that all cubes share the same velocity grid.

SPATIAL AXES. At native resolution, the cubes keep their own synthesised beams, and the flux is extracted from each cube separately, within an aperture or through a two-dimensional Gaussian fit (Sec. 3.5.1), before the individual spectra are averaged channel by channel. At the lowest common resolution, the smallest beam that encloses the synthesised beams of the whole sample is computed with `radio-beam`, every cube is convolved to it with `convolve_to()`, and is then reprojected onto the pixel grid of a reference cube with `reproject()`. After this step the cubes share the same beam, pixel grid and spectral axis, and can be combined pixel by pixel.

COMBINATION. The cubes, or the extracted spectra, are combined through a weighted mean,

$$\bar{I}(x, y, v) = \frac{\sum_i w_i f_i I_i(x, y, v)}{\sum_i w_i}, \quad (44)$$

where the sums run over all the non-flagged data. The weights w_i reproduce the schemes of Section 3.4.1. For natural weighting, w_i is the inverse of the noise variance of the whole cube, while for democratic weighting $w_i = 1$ for every source, split among its observations when a source has been observed more than once. The factors f_i include all the flux normalisations applied to the source, such as the redshift rescaling of Equation 29 and the normalisation to a common infrared luminosity, exactly as in the uv plane. The stacked flux is finally extracted from $\bar{I}$ with the same methods adopted for `ViSTA-ImFit`, so that any difference between the two results reflects the domain of the combination rather than the extraction.

BENCHMARKING OF THE HPC PIPELINE

This appendix reports the full benchmarking campaign used to validate the performance of the reengineered ViSTA pipeline described in Chapter 5, and it complements the summary given in Section 5.4. The benchmarks were run on a sample of 100 mock measurement sets of 200 channels and about 10000 rows each, spanning a range of redshifts, telescopes and array configurations, generated with the simulator described in Appendix A.1.

B.1 BENCHMARK SETUP

The benchmarks were carried out on Leonardo, the pre-exascale Tier-0 supercomputer of the EuroHPC Joint Undertaking, hosted by CINECA at the Bologna Data Center. The GPU-accelerated booster partition (`boost_usr_prod`) was used, whose nodes are each equipped with four NVIDIA A100 GPUs of 64 GB HBM2e each and a 32-core Intel Xeon Platinum 8358 CPU, connected via NVLink and to the rest of the system through a high-bandwidth NVIDIA Mellanox HDR100 InfiniBand network. All runs were performed on a single node, with the input and output MS files stored on the Lustre-based parallel filesystem.

The original CASA-based pipeline was run on the same system and on the same input sample to provide a reference baseline. On the standard configuration described below, it required a total wall-clock time of 570.1 s, or 9.5 minutes, broken down into 160.8 s for rest-framing, 150.7 s for centering, 227.6 s for rebinning and 31.0 s for the final stacking step, and it produced three intermediate output files for every input file in addition to the final combined one.

For the new version of the code, four scaling tests were performed, each varying a single pipeline configuration parameter at a time while keeping the others fixed at a standard reference configuration, on the mock sample introduced above. For every run, four quantities are reported, namely the cumulative time spent by the reader thread or threads inside `dask.compute()`, the time spent by the compute thread inside the C++/OpenMP or CUDA kernel, the time spent by the writer thread inside `xds_to_table()` and the total wall-clock time. The standard reference configuration consists of $N_{\text{MS}} = 50$ input datasets, $N_{\text{chan}} = 1000$ channels in the common rest-frame output

spectral grid, $N_{\text{readers}} = 2$ reader threads, `OMP_NUM_THREADS` equal to 32 and, for the GPU runs, a batch size of $N_{\text{batch}} = 20$. Starting from this configuration, each of the four tests below varies one parameter at a time, and for the CPU-only kernel `OMP_NUM_THREADS` was varied while for the GPU kernel the batch size N_{batch} was varied instead.

B.2 RESULTS

Figure 61 summarises the wall-clock breakdown of the four scaling tests, whose numerical values are reported in Tables 10 to 14; the main trends are discussed below.

Standard configuration: $N_{\text{MS}} = 50$, $N_{\text{chan}} = 1000$, readers = 2, `OMP_NUM_THREADS` = 32 (with GPU: $N_{\text{batch}} = 20$)

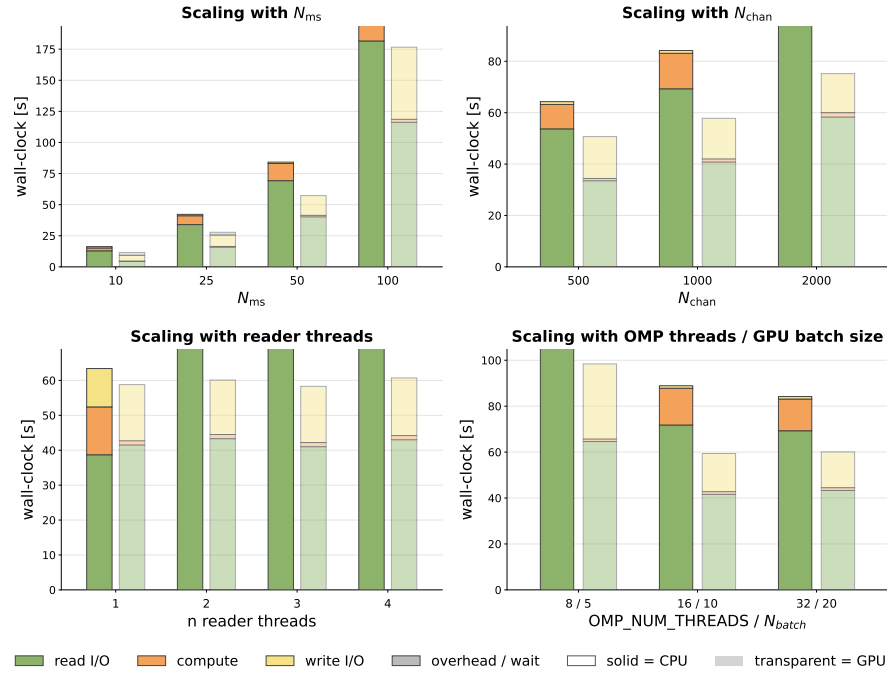


Figure 61: Wall-clock time breakdown for reading, computing and writing, together with the overhead and wait time, for the CPU and GPU implementations, as a function of N_{MS} , N_{chan} , number of reader threads and `OMP_NUM_THREADS` for the CPU case or GPU batch size N_{batch} for the GPU case. The standard configuration is $N_{\text{MS}} = 50$, $N_{\text{chan}} = 1000$, $N_{\text{readers}} = 2$ and `OMP_NUM_THREADS` equal to 32, with $N_{\text{batch}} = 20$ for the GPU runs.

- **Scaling with N_{MS} (Table 10).** On both CPU and GPU, `read_s` and `wall_s` scale close to linearly with N_{MS} , since both the total data volume and the number of per-dataset kernel calls grow linearly with the sample size. The GPU `write_s` instead grows faster than linearly between $N_{\text{MS}} = 50$ and 100, because

Table 10: Pipeline timing breakdown as a function of N_{MS} .

N_{MS}	CPU				GPU			
	read	compute	write	wall	read	compute	write	wall
10	12.9	2.3	1.1	9.4	4.4	0.2	4.7	11.6
25	34.0	7.1	1.1	22.0	15.8	0.6	9.2	27.8
50	69.3	13.8	1.1	43.4	40.2	1.2	15.9	50.4
100	181.6	30.9	1.3	112.8	116.2	2.4	58.0	151.0

Table 11: Pipeline timing breakdown as a function of N_{chan} .

N_{chan}	CPU				GPU			
	read	compute	write	wall	read	compute	write	wall
500	53.7	9.5	1.1	34.2	33.5	0.9	16.3	41.9
1000	69.3	13.8	1.1	43.4	40.8	1.2	15.8	50.6
2000	94.2	25.3	1.2	62.3	58.3	1.7	15.2	73.1

Table 12: Pipeline timing breakdown as a function of N_{readers} .

N_{readers}	CPU				GPU			
	read	compute	write	wall	read	compute	write	wall
1	38.7	13.7	11.0	46.6	41.5	1.2	16.1	52.2
2	69.3	13.8	1.1	43.4	43.3	1.2	15.6	53.8
3	82.6	17.2	1.3	43.3	41.0	1.2	16.1	51.1
4	97.1	18.4	1.4	44.5	43.0	1.2	16.5	53.2

Table 13: Pipeline timing breakdown as a function of OMP_NUM_THREADS (CPU).

OMP	read	compute	write	wall
8	121.7	20.1	6.9	81.8
16	71.8	16.0	1.1	43.3
32	69.3	13.8	1.1	43.4

Table 14: Pipeline timing breakdown as a function of N_{batch} (GPU).

N_{batch}	read	compute	write	wall
5	64.6	1.1	32.7	76.3
10	41.6	1.2	16.6	51.9
20	43.3	1.2	15.6	53.8

the GPU processes batches of 20 datasets almost instantly, so the writer queue backs up and the run becomes bound by the writer.

- **Scaling with N_{chan} (Table 11).** The CPU `compute_s` scales close to linearly with N_{chan} , consistent with the rebinning operation being proportional to N_{chan} per row, and `read_s` increases correspondingly since the output buffers scale with the output grid size. On GPU, `compute_s` is nearly unchanged because the rebinning loop is fully parallel across output channels on the device, while `write_s` stays roughly constant, depending mainly on the number of rows rather than on N_{chan} .
- **Scaling with N_{reader} (Table 12).** On CPU, a single reader leaves the writer thread repeatedly stalling while waiting for data, draining the pipeline serially at the end of the run and inflating the writer-side contribution to the wall time, and moving to $N_{\text{readers}} = 2$ removes this stall almost entirely and gives the best overall wall time. Beyond two readers, `read_s` keeps increasing while `wall_s` stays essentially flat, since extra readers only add filesystem contention once the compute and writer stages are no longer starved. On GPU, the pipeline is bound by the writer regardless of N_{readers} , so feeding the compute stage faster has no measurable effect on the wall-clock time.
- **Scaling with `OMP_NUM_THREADS` / N_{batch} (Tables 13 and 14).** These two parameters control different aspects of the pipeline and are compared only for visual convenience. `OMP_NUM_THREADS` sets the fine-grained parallelism of the C++ kernel across baseline rows, while N_{batch} controls how many datasets are grouped into a single host-to-device kernel launch on the GPU, and therefore the granularity of the overlap between reading, computing and writing. On CPU, `compute_s` decreases with the thread count but with strongly diminishing returns beyond 16 threads, consistent with the per-row kernel being bound by memory bandwidth rather than by compute at high thread counts, and the larger improvement between 8 and 16 threads is also reflected in `read_s` and `write_s`, since with too few threads the compute stage is slow enough that the reader and writer threads queue up waiting on it. On GPU, `compute_s` is essentially insensitive to N_{batch} , since the kernel launch overhead is amortised regardless of batch size at the sample size tested here, while `write_s` is highest at the smallest batch size and drops substantially for larger batches, since very small batches return results too frequently to overlap efficiently with the surrounding read and write cycle. Since the GPU run is bound by the writer rather than by the compute stage, increasing `OMP_NUM_THREADS` in the GPU configuration

would be expected to reduce `read_s` only marginally, with no measurable effect on `wall_s`, and for this reason it was not varied independently in the GPU runs.

These tests confirm that the new pipeline achieves substantial speed-ups over the original CASA-based implementation, with the CPU configuration comfortably bound by input and output operations and the GPU configuration bound by the writer, while compute plays only a secondary role except at very low thread counts.

PHOTOMETRIC FLUXES AND ALMA ARCHIVAL DATA

This appendix collects, for reference, the full set of 100–2100 μm observed flux densities used to build the spectral energy distributions of the sample (Table 15), and the ALMA archival observations covering a CO transition for each source, used for the visibility–stacking analysis presented in this thesis (Table 16). Both tables are set in landscape orientation to accommodate the large number of columns.

c.1 100–2100 μm OBSERVED FLUXES

Table 15: 100–2100 μm observed flux densities for the sources in the sample.

ID	$S_{100-160}^{11}$	S_{250}	$S_{350}^{2,6}$	S_{500}	$S_{700-900}^{1,4,5,7,8,9,10,14,17}$	$S_{1000-1300}^{5,13,14,16}$	$S_{1300-1600}^3$	$S_{1900-2100}^{3,12,15}$
S ₀₀₁		130.3±5.8	160±6.1	116.2±6.8	10.1±1.3 / 50.4±6.1	4.8±0.7 / 2.6±0.3	1.6±0.2	0.54±0.02
S ₀₀₂		130.3±5.8	160±6.1	116.2±6.8	42±2 / 50.4±6.1	14.9±1.1 / 6.9±0.6	4.2±0.3	1.84±0.03 / 2.47±0.2
S ₀₀₃		63.3±6.2	91.1±6.1	121.7±7.4	50±3 / 31.8±5.9	23.8±1.4 / 13.8±0.7	8.8±0.3	3.79±0.03
S ₀₀₄		59.9±5.8	94.2±5.8	85.6±7.2	50±2 / 33.3±6.5	18.3±1.7 / 9.6±0.8	5.6±0.4	2.16±0.04
S ₀₀₅		61.9±5.4	78.8±6	83.8±7	39±4 / 45.8±6.4	24±3 / 12.6±1.4	7.4±0.4	3.5±0.03
S ₀₀₆		92.7±5.9	92.6±5.9	84.5±7.6	5.6±6.9			0.43±0.02
S ₀₀₇		81±5.6	112.7±5.9	91.6±6.9	39.4±6.9			3.46±0.02
S ₀₀₈		195±6	221±6	149±7				1.19±0.17
S ₀₀₉		117±7	151±6	127±7				
S ₀₁₀		126.7±5.8	125.6±5.9	93.1±7.4	41±3 / 29.7±6.2	14.8±1.5 / 8.5±1.1	4.5±0.3	1.89±0.26 / 1.45±0.03
S ₀₁₁		134±5.6	128.5±6.1	93.4±6.9	23±2 / 20.1±6.7	6.3±1 / 2.5±0.5	1.4±0.2	0.52±0.02
S ₀₁₂		82.8±5.6	114.8±5.9	97.5±7.2	19±3 / 19.4±6.6	8.6±1.3 / 3.2±0.6	2.3±0.3	0.76±0.03
S ₀₁₃		82.8±5.6	114.8±5.9	97.5±7.2	17.2±1 / 19.4±6.6	6±0.6 / 4.4±0.4	2.11±0.19	0.68±0.02
S ₀₁₄		66±6	102±6	121±7	52.4±4.4			
S ₀₁₅		137.7±5.6	185.9±6.1	148.8±7.2	80±3 / 79.2±8.8	29±3 / 13.6±1.2	9.2±0.5	3.1±0.02
S ₀₁₆		80.3±5.4	106.1±5.5	105.8±6.7	24.6±6.5			0.32±0.03
S ₀₁₇		73.7±5.7	95.8±6	90.3±7.6	17.2±5.5			0.37±0.03
S ₀₁₈		82±6	120±6	126±8		40.88±5.88	23.91±2.65	
S ₀₁₉		140±7	157±7	154±8		10±2.07		
S ₀₂₀		113±7	177±6	209±8	82±3.9	72.32±6.26	35.11±2.62	
S ₀₂₁		82±9	113±9	105±10				2.18±0.2

Continued on next page

Table 15 – *continued*

ID	$S_{100-160}^{11}$	S_{250}	$S_{350}^{2,6}$	S_{500}	$S_{700-900}^{1,4,5,7,8,9,10,14,17}$	$S_{1000-1300}^{5,13,14,16}$	$S_{1300-1600}^3$	$S_{1900-2100}^{3,12,15}$
S_{022}		312 ± 6	244 ± 7	168 ± 8				0.61 ± 0.2
S_{023}		398 ± 6	320 ± 6	164 ± 8				
S_{024}		170.9 ± 5.7	197.1 ± 6.3	145.6 ± 7.4	64.8 ± 7.8		2.32 ± 0.21 / 2.7 ± 0.03 / 1.89 ± 0.13	
S_{025}		118.1 ± 4.9	147.3 ± 5.2	105.4 ± 6.4	60.7 ± 8.5			3.09 ± 0.03
S_{026}		119.3 ± 5.4	121 ± 6	102 ± 7	33.9 ± 6.9			1.07 ± 0.03
S_{027}		164.6 ± 5.8	160.2 ± 6.1	113.1 ± 7.6	18.4 ± 7.3			0.68 ± 0.03
S_{028}		96.9 ± 5.9	121.7 ± 6.1	80.2 ± 7.5	33.1 ± 6.3			0.88 ± 0.02
S_{029}		93.2 ± 5.8	135.2 ± 5.9	98.3 ± 7.7	14.3 ± 7.5			1.08 ± 0.03
S_{030}		59.5 ± 5.9	96.9 ± 6.2	95.6 ± 7.4	34.5 ± 6.1			3.23 ± 0.03
S_{031}		73.3 ± 5.8	101.2 ± 6.1	103.6 ± 7.5	39.8 ± 5.7			2.56 ± 0.03
S_{032}		48.6 ± 5.6	84.2 ± 6	84.8 ± 7.1	36.6 ± 6.4			4.2 ± 0.04
S_{033}		72.2 ± 5.3	129.8 ± 5.6	138.6 ± 7	90.5 ± 6.3			8.76 ± 0.04
S_{034}		105 ± 6	125 ± 6	117 ± 7				1.07 ± 0.24
S_{035}		120.4 ± 5.8	111.2 ± 6.4	87.4 ± 7.8	18.6 ± 6			
S_{036}		61.8 ± 5.9	101.3 ± 6.4	90.7 ± 7.6	17.5 ± 6.5			1.21 ± 0.03
S_{037}		61.8 ± 5.9	101.3 ± 6.4	90.7 ± 7.6	17.5 ± 6.5			1.28 ± 0.03
S_{038}		54.7 ± 6	67.4 ± 6.2	86.8 ± 7.7	13.4 ± 7			0.72 ± 0.03
S_{039}		257.5 ± 6.4	271.1 ± 6.3	204 ± 7.2	67.3 ± 6.3			3.59 ± 0.03
S_{040}		165 ± 8	153 ± 7	114 ± 10				1.94 ± 0.3
S_{041}		117.1 ± 6	129 ± 6.2	99.6 ± 7.4	49.4 ± 6.8			2.08 ± 0.03
S_{042}		176 ± 7	175 ± 6	104 ± 8				0.87 ± 0.29 / 0.44 ± 0.1
S_{043}		112 ± 5.9	148.8 ± 6.5	117.7 ± 7.3	23.8 ± 6.9			0.96 ± 0.03

Continued on next page

Table 15 – *continued*

ID	$S_{100-160}^{11}$	S_{250}	$S_{350}^{2,6}$	S_{500}	$S_{700-900}^{1,4,5,7,8,9,10,14,17}$	$S_{1000-1300}^{5,13,14,16}$	$S_{1300-1600}^3$	$S_{1900-2100}^{3,12,15}$
S ₀₄₄		116.3±6.1	177±6.3	179.3±7.5	77.9±6.4			7.43±0.06
S ₀₄₅		107.1±6.1	109.7±6	80.5±7.5	11.3±6.6			0.79±0.02
S ₀₄₆		97.3±6.1	99.1±6.4	93.1±7.8	15.3±5.1			1.01±0.04
S ₀₄₇		826±7	912±7	718±8	167±4 / 160±3			
S ₀₄₈		92±7	122±8	113±7	66±5.4			
S ₀₄₉		248.5±7.5	305.3±8.1	269.1±8.7	98.7±5.4			
S ₀₅₀		389±7	381±8	241±9	61.4±2.9			
S ₀₅₁	78±55 / 182±56	328±50	308±47	220±34	54.7±3.1	15.3±1.3		
S ₀₅₂		133.2±7.4	186.1±8.2	165.2±8.8	87.7±6.3			
S ₀₅₃		87±7.2	98.2±8	82.3±8.8	33.2±6.2		1.36±0.15 / 1.75±0.15 / 0.88±0.17	
S ₀₅₄	187±57 / 416±94	485±73	323±49	175±28	24.8±3.3	7.6±0.8		
S ₀₅₅	198±55 / 397±90	442±67	363±55	238±37	30.6±2.4	12.2±1.2		
S ₀₅₆		278.8±7.4	282.9±8.2	204±8.6	73.2±5.3			
S ₀₅₇		85.1±7.3	98.4±8.2	94.8±8.8	11.3±5.1			2.09±0.12
S ₀₅₈		67.4±6.5	112.1±7.4	103.9±7.7	38.1±4.8			3.19±0.24
S ₀₅₉		209.3±5.6	288.5±6	264±7	97.6±5			
S ₀₆₀		133.9±5.8	164.1±6	131.8±7.4	41.5±4.9			
S ₀₆₁		190.5±5.6	281.3±5.9	278.6±7.5	73.8±5			
S ₀₆₂		268.4±4.4	296.3±4.8	248.9±5.9	77.4±6.5			
S ₀₆₃		271.2±5.4	278.2±5.9	203.5±6.9	67.7±6.4			
S ₀₆₄		188.6±7.4	217±8.1	176.4±8.7	44±4.9			0.93±0.14
S ₀₆₅		77.7±6.5	97.3±7.4	87.1±7.9	16.4±3.8			1.05±0.18

Continued on next page

Table 15 – *continued*

ID	$S_{100-160}^{11}$	S_{250}	$S_{350}^{2,6}$	S_{500}	$S_{700-900}^{1,4,5,7,8,9,10,14,17}$	$S_{1000-1300}^{5,13,14,16}$	$S_{1300-1600}^3$	$S_{1900-2100}^{3,12,15}$
S ₀₆₆		112.2±7.3	182.2±8.2	193.3±8.5	89.7±5.4			
S ₀₆₇		323±6	244±7	140±33	18.4±2.5			
S ₀₆₈		81.9±7.2	95.9±8.2	89.8±8.8	22.3±5.7			0.71±0.08
S ₀₆₉		85.4±6	99.1±6.3	90.6±7.2	45.3±7.5			1.97±0.06
S ₀₇₀		139.1±5.3	144.8±5.4	100.5±6.6	46.7±7.7			1.62±0.04
S ₀₇₁		105.9±6.5	115.6±6.2	92.7±7.4	13.3±6.9		1.46±0.07 / 1.32±0.04 / 0.79±0.18	
S ₀₇₂		122.1±6.1	124.4±6.5	87.3±7.5	16.3±7			0.73±0.05
S ₀₇₃		85.6±6	102.6±6.3	80.1±7.5				1.59±0.19 / 0.65±0.04
S ₀₇₄		65.5±6.1	88.1±6.1	84±7.5	58±7.4			2.66±0.03
S ₀₇₅		127.4±4.6	138.7±5.1	111.4±6.3	27.4±7.1			1.28±0.03
S ₀₇₆		126.1±5.3	131.2±5.7	93.5±7	33.7±7.6			1.42±0.03
S ₀₇₇		85.5±5.2	99.7±5.5	88±6.9	19.1±6.6			0.79±0.04
S ₀₇₈		119.5±5.9	132.1±6.2	81.7±7.6	39.5±6.6			1.89±0.05
S ₀₇₉		85.4±5.5	96.7±6.2	87.8±7.5				0.82±0.03
S ₀₈₀		69.4±5.1	91.9±5.5	80.1±6.6				0.93±0.14 / 0.86±0.03
S ₀₈₁		69.4±5.1	91.9±5.5	80.1±6.6				0.57±0.03 / 3.49±0.21
S ₀₈₂		122.7±5.7	122.1±6.3	97.7±7.6	20.6±6.1			0.28±0.02
S ₀₈₃		122.7±5.7	122.1±6.3	97.7±7.6	20.6±6.1			0.29±0.02
S ₀₈₄		89±5.7	109.1±6.2	82.3±7.9				1.09±0.03
S ₀₈₅		79.4±5.8	135.4±6	140±7.4	79.4±7.8			5.56±0.03
S ₀₈₆		86.7±5.8	102.6±6	90.6±7.8	16.6±7.4			0.87±0.02
S ₀₈₇		122.4±5.2	134.6±5.4	86.6±6.8	24.1±6.9			0.4±0.02

Continued on next page

Table 15 – *continued*

ID	$S_{100-160}^{11}$	S_{250}	$S_{350}^{2,6}$	S_{500}	$S_{700-900}^{1,4,5,7,8,9,10,14,17}$	$S_{1000-1300}^{5,13,14,16}$	$S_{1300-1600}^3$	$S_{1900-2100}^{3,12,15}$
S ₀₈₈		212.9±4.7	244.2±5	169.4±6.2	52.9±6.1			0.88±0.22 / 2.7±0.03 / 0.75±0.05
S ₀₈₉		214±7	218±7	172±9	36±4			
S ₀₉₀		153.7±4.8	178.4±5.2	123.5±6.6	39.8±8.2			1.6±0.03
S ₀₉₁		95±7	120±6	102±7				1.65±0.23
S ₀₉₂		118.3±5.1	141.2±5.5	119.7±6.8	36.5±7			2.98±0.03
S ₀₉₃		74.5±5.7	100.2±6	93.7±7.5	31.7±7.2			1.38±0.04
S ₀₉₄		106.2±5.9	107.9±6	90±7.5	16.3±7.9			1.24±0.05
S ₀₉₅		271±6	336±6	263±8	75±8			
S ₀₉₆		148±6	187±6	147±9	48±5			
S ₀₉₇		193±7	252±6	202±8				
S ₀₉₈		151±6	209±6	205±8	41±4			
S ₀₉₉		125.8±5.5	185.5±5.8	155.1±7.4	51.3±6.3			
S ₁₀₀		77.3±5.4	87.3±5.7	94.8±7	32±2 / 22.8±5.7	10.9±1.3 / 6.1±0.8	3.4±0.3	1.37±0.02
S ₁₀₁		91.9±5.9	107.6±6	82.3±7.1	45±2	15±1.1 / 6.5±0.7	4.4±0.4	1.47±0.02
S ₁₀₂		77.4±5.6	90.7±5.8	96±7.4	35±2 / 33.8±6.1	13±1.1 / 6.1±0.8	3.9±0.4	1.53±0.02
S ₁₀₃		121.5±6.1	161±6.7	125.5±7.7	76±2 / 64±8.8	34.9±1.3 / 19±0.9	11.8±0.3	4.81±0.03
S ₁₀₄		112.5±5	148±5.4	143.4±6.5	89±3 / 49.2±5.7	36±2 / 17.2±1.8	10.3±0.5	3.46±0.03

Columns: the internal source identifier and the observed flux densities across the different wavelength ranges. Superscripts indicate specific bibliographic references or notes associated with each measurement: (1) Liu et al. [245]; (2) Nayyeri et al. [158]; (3) Ismail et al. [246]; (4) Bakx et al. [161]; (5) Ivison et al. [247]; (6) Negrello et al. [153]; (7) Wardlow et al. [157]; (8) Bussmann et al. [201]; (9) Duivenvoorden et al. [248]; (10) Maresca et al. [206]; (11) Zhang et al. [218]; (12) Yang, C. et al. [249]; (13) Gralla et al. [250]; (14) Bendo et al. [251]; (15) Bendo et al. [174]; (16) Su et al. [252]; (17) Bussmann et al. [167].

C.2 ALMA ARCHIVAL OBSERVATIONS WITH A POSSIBLE CO EMISSION LINE DETECTION

For each source, Table [16](#) lists every archival ALMA observation covering a CO rotational transition, together with the array configuration, achieved angular resolution, and project code, used to build the stacked visibilities discussed in the main text.

Table 16: ALMA archival observations with a possible CO emission line detection by source ID.

ID	Transition	Array	Res. ["]	Project ID	ID	Transition	Array	Res. ["]	Project ID
S ₀₀₁	4-3	ACA	10.646	2016.2.00133.S	S ₀₅₆	9-8	ALMA	0.455	2019.1.00663.S
-	5-4	ALMA	0.212	2021.1.01628.S	S ₀₅₇	4-3	ACA	11.495	2021.2.00052.S
-	6-5	ALMA	1.673	2019.1.01477.S	S ₀₅₈	4-3	ACA	11.046	2018.1.00804.S
-	6-5	ACA	7.589	2023.1.00322.S	S ₀₅₉	9-8	ALMA	0.324	2018.1.00966.S
-	9-8	ALMA	0.127	2021.1.01628.S	S ₀₆₀	3-2	ACA	13.234	2019.2.00155.S
S ₀₀₂	4-3	ACA	10.646	2016.2.00133.S	-	5-4	ACA	9.078	2022.1.00224.S
-	5-4	ALMA	0.212	2021.1.01628.S	S ₀₆₁	3-2	ALMA	0.495	2017.1.01214.S
-	6-5	ALMA	1.673	2019.1.01477.S	-	7-6	ALMA	0.216	2024.1.01396.S
-	6-5	ACA	7.589	2023.1.00322.S	S ₀₆₂	3-2	ALMA	0.958	2022.1.00172.S
S ₀₀₃	4-3	ALMA	2.832	2019.1.01477.S	S ₀₆₃	9-8	ALMA	0.524	2018.1.00966.S
-	5-4	ALMA	0.254	2021.1.01628.S	S ₀₆₄	3-2	ALMA	0.968	2022.1.00172.S
-	6-5	ALMA	0.212	2021.1.01628.S	-	7-6	ALMA	0.146	2018.1.00861.S
-	7-6	ALMA	1.901	2019.1.01477.S	-	9-8	ALMA	0.121	2024.1.01570.S
-	9-8	ALMA	0.172	2021.1.01628.S	S ₀₆₅	3-2	ACA	11.907	2018.1.00804.S
S ₀₀₄	4-3	ALMA	0.259	2021.1.01628.S	-	3-2	ACA	9.210	2016.2.00133.S
-	5-4	ALMA	1.892	2019.1.01477.S	S ₀₆₆	7-6	ALMA	0.146	2016.1.00450.S
-	8-7	ALMA	0.126	2021.1.01628.S	-	7-6	ALMA	0.168	2018.1.00861.S
S ₀₀₅	5-4	ALMA	0.259	2021.1.01628.S	-	9-8	ALMA	0.555	2018.1.00966.S
-	5-4	ALMA	2.562	2019.1.01477.S	S ₀₆₇	7-6	ALMA	0.244	2021.1.00353.S
-	6-5	ALMA	0.212	2021.1.01628.S	-	8-7	ALMA	0.378	2021.1.00353.S
-	7-6	ALMA	1.901	2019.1.01477.S	S ₀₆₈	3-2	ACA	9.300	2016.2.00133.S

Continued on next page

Table 16 – *continued*

ID	Transition	Array	Res. ["]	Project ID	ID	Transition	Array	Res. ["]	Project ID
-	7-6	ACA	7.685	2021.2.00052.S	-	9-8	ALMA	0.115	2024.1.01570.S
-	7-6	ACA	8.144	2022.1.00224.S	S ₀₆₉	4-3	ALMA	1.521	2019.1.01477.S
S ₀₀₆	4-3	ALMA	2.570	2019.1.01477.S	-	6-5	ALMA	0.219	2022.1.00432.S
-	5-4	ALMA	1.904	2019.1.01477.S	-	6-5	ALMA	1.678	2019.1.01477.S
-	5-4	ACA	8.856	2022.1.00224.S	-	8-7	ALMA	0.203	2022.1.00432.S
S ₀₀₇	5-4	ALMA	2.850	2019.1.01477.S	S ₀₇₀	3-2	ACA	10.589	2016.2.00133.S
-	7-6	ALMA	1.774	2019.1.01477.S	-	3-2	ACA	11.619	2018.1.00804.S
-	7-6	ACA	7.359	2023.1.00322.S	-	5-4	ALMA	1.667	2019.1.01477.S
S ₀₀₈	7-6	ALMA	0.033	2016.1.01188.S	-	8-7	ALMA	0.177	2022.1.00432.S
-	7-6	ACA	4.871	2016.2.00105.S	-	9-8	ALMA	0.110	2024.1.01570.S
-	9-8	ACA	3.941	2018.1.00922.S	S ₀₇₁	3-2	ALMA	1.642	2019.1.01477.S
S ₀₀₉	7-6	ALMA	0.031	2016.1.01188.S	-	6-5	ALMA	0.202	2022.1.00432.S
-	7-6	ACA	5.392	2016.2.00105.S	-	7-6	ACA	5.031	2021.2.00059.S
-	9-8	ALMA	0.130	2022.1.00145.S	S ₀₇₂	3-2	ALMA	1.524	2019.1.01477.S
S ₀₁₀	3-2	ALMA	2.698	2019.1.01477.S	-	4-3	ALMA	1.674	2019.1.01477.S
-	6-5	ALMA	0.173	2021.1.01628.S	-	7-6	ACA	4.587	2022.1.00064.S
-	7-6	ALMA	0.124	2021.1.01628.S	S ₀₇₃	3-2	ALMA	2.508	2019.1.01477.S
-	7-6	ACA	4.414	2022.1.00064.S	-	7-6	ALMA	0.178	2022.1.00432.S
S ₀₁₁	3-2	ACA	8.298	2022.1.00224.S	-	7-6	ACA	4.508	2022.1.00064.S
S ₀₁₂	4-3	ALMA	0.264	2021.1.01628.S	S ₀₇₄	3-2	ALMA	2.496	2019.1.01477.S
-	4-3	ALMA	2.375	2019.1.01477.S	-	5-4	ALMA	1.673	2019.1.01477.S
-	7-6	ALMA	0.172	2021.1.01628.S	-	6-5	ALMA	0.201	2022.1.00432.S

Continued on next page

Table 16 – *continued*

ID	Transition	Array	Res. ["]	Project ID	ID	Transition	Array	Res. ["]	Project ID
S ₀₁₃	3-2	ALMA	2.470	2019.1.01477.S	-	9-8	ALMA	0.110	2024.1.01570.S
-	5-4	ALMA	1.732	2019.1.01477.S	S ₀₇₅	3-2	ALMA	0.346	2022.1.00432.S
-	6-5	ALMA	0.172	2021.1.01628.S	-	3-2	ACA	11.180	2016.2.00133.S
-	7-6	ALMA	0.125	2021.1.01628.S	-	3-2	ACA	9.676	2016.2.00133.S
S ₀₁₄	5-4	ALMA	0.636	2013.1.00449.S	-	4-3	ALMA	0.220	2022.1.00432.S
-	6-5	ALMA	0.576	2013.1.00449.S	-	7-6	ACA	4.734	2022.1.00064.S
S ₀₁₅	4-3	ALMA	0.263	2021.1.01628.S	S ₀₇₆	3-2	ALMA	0.348	2022.1.00432.S
-	5-4	ALMA	1.713	2019.1.01477.S	-	5-4	ALMA	0.220	2022.1.00432.S
S ₀₁₆	3-2	ALMA	0.370	2022.1.00432.S	-	5-4	ALMA	1.691	2019.1.01477.S
-	3-2	ALMA	2.463	2019.1.01477.S	-	7-6	ALMA	0.202	2022.1.00432.S
-	5-4	ALMA	0.277	2022.1.00432.S	-	8-7	ALMA	0.178	2022.1.00432.S
-	5-4	ALMA	1.734	2019.1.01477.S	S ₀₇₇	3-2	ALMA	2.513	2019.1.01477.S
S ₀₁₇	5-4	ALMA	0.277	2022.1.00432.S	-	4-3	ALMA	0.221	2022.1.00432.S
-	5-4	ALMA	1.754	2019.1.01477.S	-	4-3	ALMA	1.692	2019.1.01477.S
-	9-8	ALMA	0.124	2022.1.00432.S	-	7-6	ACA	4.396	2022.1.00064.S
S ₀₁₈	5-4	ACA	11.997	2018.1.00804.S	-	8-7	ALMA	0.110	2024.1.01570.S
-	6-5	ACA	10.507	2018.1.00804.S	S ₀₇₈	4-3	ALMA	2.504	2019.1.01477.S
S ₀₁₉	3-2	ACA	12.079	2018.1.00804.S	-	5-4	ALMA	0.219	2022.1.00432.S
S ₀₂₀	8-7	ACA	6.208	2021.2.00011.S	-	6-5	ALMA	1.671	2019.1.01477.S
-	9-8	ALMA	0.185	2021.1.01116.S	-	6-5	ACA	6.365	2021.2.00036.S
-	9-8	ACA	6.036	2018.1.00922.S	S ₀₇₉	3-2	ALMA	0.348	2022.1.00432.S
S ₀₂₁	9-8	ALMA	0.085	2022.1.00145.S	-	3-2	ALMA	1.646	2019.1.01477.S

Continued on next page

Table 16 – *continued*

ID	Transition	Array	Res. ["]	Project ID	ID	Transition	Array	Res. ["]	Project ID
S ₀₂₂	5-4	ACA	6.524	2018.1.00922.S	-	3-2	ACA	10.746	2021.2.00052.S
-	6-5	ACA	5.979	2018.1.00922.S	-	6-5	ALMA	0.202	2022.1.00432.S
-	7-6	ACA	4.641	2016.2.00105.S	-	7-6	ACA	4.684	2022.1.00064.S
-	8-7	ACA	4.172	2021.2.00011.S	S ₀₈₀	3-2	ALMA	0.350	2022.1.00432.S
-	9-8	ACA	3.833	2016.2.00105.S	-	3-2	ALMA	1.527	2019.1.01477.S
S ₀₂₃	5-4	ACA	5.157	2016.2.00105.S	-	4-3	ALMA	0.220	2022.1.00432.S
-	6-5	ACA	3.738	2016.2.00105.S	-	7-6	ALMA	0.178	2022.1.00432.S
-	7-6	ACA	3.068	2016.2.00105.S	-	7-6	ACA	5.111	2022.1.00064.S
S ₀₂₄	3-2	ALMA	2.445	2019.1.01477.S	S ₀₈₁	3-2	ALMA	0.350	2022.1.00432.S
-	4-3	ALMA	1.648	2019.1.01477.S	-	3-2	ALMA	1.527	2019.1.01477.S
-	7-6	ACA	4.676	2022.1.00064.S	-	4-3	ALMA	0.220	2022.1.00432.S
-	8-7	ALMA	0.079	2022.1.00145.S	-	7-6	ALMA	0.178	2022.1.00432.S
-	8-7	ALMA	0.107	2022.1.00432.S	-	7-6	ACA	5.111	2022.1.00064.S
S ₀₂₅	4-3	ALMA	2.458	2019.1.01477.S	S ₀₈₂	3-2	ALMA	2.505	2019.1.01477.S
-	6-5	ALMA	0.273	2022.1.00432.S	-	4-3	ALMA	1.748	2019.1.01477.S
-	6-5	ALMA	1.617	2019.1.01477.S	-	7-6	ACA	4.431	2022.1.00064.S
-	6-5	ACA	6.649	2023.1.00322.S	S ₀₈₃	4-3	ALMA	1.649	2019.1.01477.S
S ₀₂₆	3-2	ALMA	2.367	2019.1.01477.S	-	5-4	ALMA	0.202	2022.1.00432.S
-	7-6	ACA	4.692	2022.1.00064.S	-	6-5	ALMA	0.178	2022.1.00432.S
-	8-7	ALMA	0.126	2022.1.00432.S	-	6-5	ACA	4.431	2022.1.00064.S
-	9-8	ALMA	0.098	2022.1.00145.S	S ₀₈₄	3-2	ALMA	2.471	2019.1.01477.S
-	9-8	ALMA	0.107	2022.1.00432.S	-	4-3	ALMA	1.652	2019.1.01477.S

Continued on next page

Table 16 – *continued*

ID	Transition	Array	Res. ["]	Project ID	ID	Transition	Array	Res. ["]	Project ID
S ₀₂₇	3-2	ALMA	2.366	2019.1.01477.S	-	7-6	ACA	4.429	2022.1.00064.S
-	7-6	ACA	5.056	2022.1.00064.S	S ₀₈₅	4-3	ALMA	0.359	2022.1.00432.S
-	8-7	ALMA	0.127	2022.1.00432.S	-	4-3	ACA	10.384	2016.2.00133.S
-	9-8	ALMA	0.079	2022.1.00145.S	-	4-3	ACA	10.650	2018.1.00804.S
-	9-8	ALMA	0.107	2022.1.00432.S	-	4-3	ACA	11.060	2016.2.00133.S
S ₀₂₈	9-8	ALMA	0.095	2022.1.00432.S	-	6-5	ALMA	0.263	2022.1.00432.S
-	6-5	ALMA	0.128	2022.1.00432.S	-	6-5	ALMA	1.755	2019.1.01477.S
S ₀₂₉	3-2	ALMA	2.459	2019.1.01477.S	S ₀₈₆	3-2	ALMA	2.367	2019.1.01477.S
-	4-3	ALMA	1.656	2019.1.01477.S	-	7-6	ALMA	0.165	2022.1.00432.S
-	7-6	ACA	4.745	2022.1.00064.S	-	7-6	ACA	5.168	2021.2.00059.S
S ₀₃₀	4-3	ALMA	2.705	2019.1.01477.S	S ₀₉₉	4-3	ACA	10.410	2016.2.00133.S
-	7-6	ALMA	0.271	2022.1.00432.S	-	4-3	ACA	10.542	2018.1.00804.S
-	7-6	ALMA	1.616	2019.1.01477.S	-	5-4	ALMA	0.209	2021.1.01628.S
S ₀₃₁	4-3	ALMA	2.446	2019.1.01477.S	-	6-5	ALMA	1.665	2019.1.01477.S
-	6-5	ALMA	0.269	2022.1.00432.S	-	6-5	ACA	7.400	2023.1.00322.S
-	6-5	ACA	6.177	2023.1.00322.S	-	9-8	ALMA	0.128	2021.1.01628.S
S ₀₃₂	4-3	ALMA	0.322	2022.1.00432.S	S ₀₈₇	3-2	ACA	10.770	2022.1.00224.S
-	4-3	ALMA	2.729	2019.1.01477.S	-	4-3	ALMA	0.257	2022.1.00432.S
-	6-5	ALMA	0.214	2022.1.00432.S	-	4-3	ALMA	1.758	2019.1.01477.S
-	6-5	ALMA	1.740	2019.1.01477.S	-	5-4	ALMA	0.151	2022.1.00432.S
S ₀₃₃	5-4	ALMA	0.321	2022.1.00432.S	-	6-5	ALMA	0.165	2022.1.00432.S
-	5-4	ALMA	2.742	2019.1.01477.S	S ₀₈₈	3-2	ALMA	0.361	2022.1.00432.S

Continued on next page

Table 16 – *continued*

ID	Transition	Array	Res. ["]	Project ID	ID	Transition	Array	Res. ["]	Project ID
-	7-6	ALMA	1.670	2019.1.01477.S	-	3-2	ACA	10.456	2016.2.00133.S
-	7-6	ACA	6.119	2023.1.00322.S	-	3-2	ACA	10.588	2018.1.00804.S
S ₀₃₄	9-8	ALMA	0.081	2022.1.00145.S	-	4-3	ALMA	1.659	2019.1.01477.S
S ₀₃₅	3-2	ALMA	1.663	2019.1.01477.S	-	6-5	ALMA	0.165	2022.1.00432.S
S ₀₃₆	4-3	ALMA	0.288	2024.1.00278.S	-	7-6	ACA	4.371	2022.1.00064.S
-	4-3	ALMA	2.681	2019.1.01477.S	S ₀₈₉	5-4	ACA	7.134	2018.1.00922.S
-	5-4	ALMA	0.212	2022.1.00432.S	-	6-5	ACA	6.285	2018.1.00922.S
-	5-4	ALMA	0.225	2024.1.00278.S	-	7-6	ALMA	0.040	2017.1.00027.S
-	5-4	ALMA	1.706	2019.1.01477.S	-	7-6	ACA	5.252	2016.2.00105.S
-	9-8	ALMA	0.156	2022.1.00432.S	-	8-7	ACA	4.800	2021.2.00011.S
S ₀₃₇	4-3	ALMA	0.288	2024.1.00278.S	-	9-8	ACA	4.010	2018.1.00922.S
-	4-3	ALMA	2.681	2019.1.01477.S	S ₀₉₀	3-2	ALMA	0.361	2022.1.00432.S
-	5-4	ALMA	0.212	2022.1.00432.S	-	3-2	ACA	10.517	2016.2.00133.S
-	5-4	ALMA	0.225	2024.1.00278.S	-	3-2	ACA	10.643	2018.1.00804.S
-	5-4	ALMA	1.706	2019.1.01477.S	-	5-4	ALMA	1.653	2019.1.01477.S
-	9-8	ALMA	0.156	2022.1.00432.S	-	6-5	ALMA	0.152	2022.1.00432.S
S ₀₃₈	3-2	ALMA	0.319	2022.1.00432.S	S ₀₉₁	9-8	ALMA	0.125	2024.1.01570.S
-	3-2	ALMA	2.744	2019.1.01477.S	S ₀₉₂	4-3	ALMA	0.361	2022.1.00432.S
-	5-4	ALMA	1.724	2019.1.01477.S	-	4-3	ALMA	2.468	2019.1.01477.S
S ₀₃₉	3-2	ALMA	0.317	2022.1.00432.S	-	4-3	ACA	10.269	2018.1.00804.S
-	3-2	ALMA	2.696	2019.1.01477.S	-	6-5	ACA	6.168	2023.1.00322.S
-	5-4	ALMA	1.609	2019.1.01477.S	-	7-6	ALMA	0.153	2022.1.00432.S

Continued on next page

Table 16 – *continued*

ID	Transition	Array	Res. ["]	Project ID	ID	Transition	Array	Res. ["]	Project ID
-	8-7	ALMA	0.158	2022.1.00432.S	-	8-7	ALMA	0.165	2022.1.00432.S
S ₀₄₀	9-8	ALMA	0.082	2022.1.00145.S	S ₀₉₃	4-3	ALMA	0.361	2022.1.00432.S
S ₀₄₁	5-4	ALMA	1.666	2019.1.01477.S	-	4-3	ALMA	2.469	2019.1.01477.S
S ₀₄₂	7-6	ALMA	0.081	2022.1.00145.S	-	6-5	ALMA	1.659	2019.1.01477.S
S ₀₄₃	4-3	ALMA	1.667	2019.1.01477.S	-	6-5	ACA	6.208	2023.1.00322.S
-	8-7	ALMA	0.112	2024.1.01570.S	-	8-7	ALMA	0.165	2022.1.00432.S
S ₀₄₄	4-3	ALMA	2.711	2019.1.01477.S	S ₀₉₄	3-2	ALMA	0.361	2022.1.00432.S
-	6-5	ALMA	1.652	2019.1.01477.S	-	3-2	ALMA	2.377	2019.1.01477.S
S ₀₄₅	3-2	ALMA	2.725	2019.1.01477.S	-	4-3	ALMA	1.668	2019.1.01477.S
-	4-3	ALMA	1.669	2019.1.01477.S	-	6-5	ALMA	0.165	2022.1.00432.S
-	7-6	ACA	4.609	2022.1.00064.S	-	7-6	ACA	4.362	2022.1.00064.S
-	8-7	ALMA	0.104	2022.1.00432.S	S ₀₉₅	6-5	ACA	5.870	2018.1.00922.S
S ₀₄₆	3-2	ALMA	2.731	2019.1.01477.S	-	7-6	ALMA	0.026	2016.1.01188.S
S ₀₄₇	3-2	ALMA	0.064	2017.1.01230.S	-	7-6	ACA	5.050	2016.2.00105.S
-	3-2	ALMA	0.107	2019.1.01197.S	-	8-7	ACA	4.668	2021.2.00011.S
-	3-2	ALMA	0.157	2023.1.00432.S	-	9-8	ACA	4.215	2018.1.00922.S
-	3-2	ALMA	0.314	2017.1.01214.S	S ₀₉₆	5-4	ACA	6.783	2018.1.00922.S
-	3-2	ALMA	0.397	2019.1.01197.S	-	6-5	ACA	5.979	2018.1.00922.S
-	4-3	ALMA	0.214	2017.1.00814.S	-	7-6	ALMA	0.031	2017.1.00027.S
-	5-4	ALMA	0.724	2019.1.01365.S	-	7-6	ACA	5.025	2016.2.00105.S
-	5-4	ACA	6.796	2016.2.00105.S	-	8-7	ACA	4.828	2021.2.00011.S
-	7-6	ALMA	0.170	2024.1.01396.S	-	9-8	ACA	3.873	2018.1.00922.S

Continued on next page

Table 16 – *continued*

ID	Transition	Array	Res. ["]	Project ID	ID	Transition	Array	Res. ["]	Project ID
-	7-6	ALMA	0.999	2024.1.01396.S	S ₀₉₇	8-7	ACA	5.908	2021.2.00011.S
-	7-6	ACA	5.250	2016.2.00105.S	-	9-8	ALMA	0.120	2021.1.01116.S
-	8-7	ACA	4.413	2021.2.00011.S	-	9-8	ACA	5.220	2018.1.00922.S
-	9-8	ALMA	0.087	2022.1.00145.S	S ₀₉₈	7-6	ACA	6.747	2018.1.00922.S
-	9-8	ALMA	0.214	2021.1.01116.S	-	8-7	ACA	6.415	2021.2.00011.S
-	9-8	ACA	4.095	2018.1.00922.S	-	9-8	ALMA	0.143	2021.1.01116.S
S ₀₄₈	6-5	ACA	6.878	2016.2.00105.S	-	9-8	ACA	5.173	2018.1.00922.S
-	8-7	ACA	5.925	2021.2.00011.S	S ₁₀₀	3-2	ALMA	0.250	2021.1.01628.S
-	9-8	ALMA	0.508	2015.1.01042.S	-	3-2	ALMA	2.802	2019.1.01477.S
-	9-8	ACA	4.737	2018.1.00922.S	-	4-3	ALMA	0.211	2021.1.01628.S
S ₀₄₉	6-5	ALMA	0.061	2018.1.01710.S	-	6-5	ALMA	0.172	2021.1.01628.S
-	6-5	ALMA	0.366	2015.1.01320.S	-	7-6	ACA	4.941	2022.1.00064.S
-	9-8	ALMA	0.335	2019.1.00663.S	S ₁₀₁	3-2	ALMA	0.253	2021.1.01628.S
-	9-8	ALMA	0.484	2019.1.00533.S	-	3-2	ALMA	2.544	2019.1.01477.S
S ₀₅₀	6-5	ALMA	0.026	2017.1.00027.S	-	7-6	ACA	4.824	2022.1.00064.S
-	9-8	ALMA	0.469	2015.1.01042.S	S ₁₀₂	3-2	ALMA	2.661	2019.1.01477.S
S ₀₅₁	7-6	ALMA	0.259	2015.1.01455.S	-	5-4	ALMA	1.674	2019.1.01477.S
-	9-8	ALMA	0.348	2015.1.01042.S	-	6-5	ALMA	0.171	2021.1.01628.S
-	9-8	ALMA	0.862	2015.1.01042.S	-	7-6	ALMA	0.128	2021.1.01628.S
S ₀₅₂	3-2	ALMA	0.069	2016.1.01231.S	-	7-6	ALMA	0.660	2018.1.00526.S
-	3-2	ALMA	0.122	2016.1.00633.S	S ₁₀₃	4-3	ALMA	0.253	2021.1.01628.S
-	3-2	ALMA	0.246	2018.1.00747.S	-	4-3	ALMA	2.573	2019.1.01477.S

Continued on next page

Table 16 – *continued*

ID	Transition	Array	Res. ["]	Project ID	ID	Transition	Array	Res. ["]	Project ID
-	9-8	ALMA	0.250	2018.1.00966.S	-	5-4	ALMA	1.898	2019.1.01477.S
S ₀₅₃	7-6	ALMA	0.123	2024.1.01570.S	-	8-7	ALMA	0.128	2021.1.01628.S
S ₀₅₄	6-5	ALMA	0.020	2015.1.00415.S	S ₁₀₄	3-2	ACA	9.422	2016.2.00133.S
S ₀₅₅	7-6	ALMA	0.227	2019.1.00129.S	-	5-4	ALMA	0.211	2021.1.01628.S
-	7-6	ALMA	0.356	2016.1.00282.S	-	5-4	ALMA	1.667	2019.1.01477.S
S ₀₅₆	8-7	ALMA	0.263	2018.1.00861.S	-	7-6	ALMA	0.171	2021.1.01628.S

Column: the internal source identifier (see main sample table), CO transition covered by the observation, array configuration, angular resolution and ALMA project code.

BIBLIOGRAPHY

- [1] Piero Madau, Lucia Pozzetti, and Mark Dickinson. “The Star Formation History of Field Galaxies.” In: *ApJ* 498.1 (1998). doi: [10.1086/305523](https://doi.org/10.1086/305523).
- [2] Robert C. Kennicutt Jr. “Star Formation in Galaxies Along the Hubble Sequence.” In: *ARA&A* 36 (1998). doi: [10.1146/annurev.astro.36.1.189](https://doi.org/10.1146/annurev.astro.36.1.189).
- [3] Gilles Chabrier. “Galactic Stellar and Substellar Initial Mass Function.” In: *PASP* 115.809 (2003). doi: [10.1086/376392](https://doi.org/10.1086/376392).
- [4] Piero Madau et al. “High-redshift galaxies in the Hubble Deep Field: colour selection and star formation history to $z \sim 4$.” In: *MNRAS* 283.4 (1996). doi: [10.1093/mnras/283.4.1388](https://doi.org/10.1093/mnras/283.4.1388).
- [5] Andrew M. Hopkins and John F. Beacom. “On the Normalization of the Cosmic Star Formation History.” In: *ApJ* 651.1 (2006). doi: [10.1086/506610](https://doi.org/10.1086/506610).
- [6] Piero Madau and Mark Dickinson. “Cosmic Star-Formation History.” In: *ARA&A* 52 (2014). doi: [10.1146/annurev-astro-081811-125615](https://doi.org/10.1146/annurev-astro-081811-125615).
- [7] C. Gruppioni et al. “The Herschel PEP/HerMES luminosity function - I. Probing the evolution of PACS selected Galaxies to $z \simeq 4$.” In: *MNRAS* 432.1 (2013). doi: [10.1093/mnras/stt308](https://doi.org/10.1093/mnras/stt308).
- [8] Caitlin M. Casey, Desika Narayanan, and Asantha Cooray. “Dusty star-forming galaxies at high redshift.” In: *Phys. Rep.* 541.2 (2014). doi: [10.1016/j.physrep.2014.02.009](https://doi.org/10.1016/j.physrep.2014.02.009).
- [9] J. A. Zavala et al. “The Evolution of the IR Luminosity Function and Dust-obscured Star Formation over the Past 13 Billion Years.” In: *ApJ* 909.2, 165 (2021). doi: [10.3847/1538-4357/abdb27](https://doi.org/10.3847/1538-4357/abdb27).
- [10] L. Barrufet et al. “Unveiling the nature of infrared bright, optically dark galaxies with early JWST data.” In: *MNRAS* 522.1 (2023). doi: [10.1093/mnras/stad947](https://doi.org/10.1093/mnras/stad947).
- [11] Margherita Talia et al. “Illuminating the Dark Side of Cosmic Star Formation Two Billion Years after the Big Bang.” In: *ApJ* 909.1, 23 (2021). doi: [10.3847/1538-4357/abd6e3](https://doi.org/10.3847/1538-4357/abd6e3).
- [12] Meriem Behiri et al. “Illuminating the Dark Side of Cosmic Star Formation. II. A Second Date with RS-NIRdark Galaxies in COSMOS.” In: *ApJ* 957.2, 63 (2023). doi: [10.3847/1538-4357/acf616](https://doi.org/10.3847/1538-4357/acf616).

- [13] M. Michałowski, J. Hjorth, and D. Watson. “Cosmic evolution of submillimeter galaxies and their contribution to stellar mass assembly.” In: *A&A* 514, A67 (2010). doi: [10.1051/0004-6361/200913634](https://doi.org/10.1051/0004-6361/200913634).
- [14] S. Toft et al. “Submillimeter Galaxies as Progenitors of Compact Quiescent Galaxies.” In: *ApJ* 782.2, 68 (2014). doi: [10.1088/0004-637X/782/2/68](https://doi.org/10.1088/0004-637X/782/2/68).
- [15] Andrew W. Blain, Ian Smail, R. J. Ivison, J.-P. Kneib, and David T. Frayer. “Submillimeter galaxies.” In: *Phys. Rep.* 369.2 (2002). doi: [10.1016/S0370-1573\(02\)00134-5](https://doi.org/10.1016/S0370-1573(02)00134-5).
- [16] J. A. Hodge and E. da Cunha. “High-redshift star formation in the Atacama large millimetre/submillimetre array era.” In: *Royal Society Open Science* 7.12, 200556 (2020). doi: [10.1098/rsos.200556](https://doi.org/10.1098/rsos.200556).
- [17] T. Ronconi et al. “GalaPy: A highly optimised C++/Python spectral modelling tool for galaxies. I. Library presentation and photometric fitting.” In: *A&A* 685, A161 (2024). doi: [10.1051/0004-6361/202346978](https://doi.org/10.1051/0004-6361/202346978).
- [18] Marika Giulietti et al. “ALMA Resolves the First Strongly Lensed Optical/Near-IR-dark Galaxy.” In: *ApJ* 943.2, 151 (2023). doi: [10.3847/1538-4357/aca53f](https://doi.org/10.3847/1538-4357/aca53f).
- [19] D. B. Sanders and I. F. Mirabel. “Luminous Infrared Galaxies.” In: *ARA&A* 34 (1996). doi: [10.1146/annurev.astro.34.1.749](https://doi.org/10.1146/annurev.astro.34.1.749).
- [20] J.-L. Puget et al. “Tentative detection of a cosmic far-infrared background with COBE.” In: *A&A* 308 (1996).
- [21] D. J. Fixsen, E. Dwek, J. C. Mather, C. L. Bennett, and R. A. Shafer. “The Spectrum of the Extragalactic Far-Infrared Background from the COBE* FIRAS Observations.” In: *The Astrophysical Journal* 508.1 (1998). doi: [10.1086/306383](https://doi.org/10.1086/306383).
- [22] Ian Smail, R. J. Ivison, and A. W. Blain. “A Deep Submillimeter Survey of Lensing Clusters: A New Window on Galaxy Formation and Evolution.” In: *ApJ* 490.1 (1997). doi: [10.1086/311017](https://doi.org/10.1086/311017).
- [23] A. J. Barger et al. “Submillimetre-wavelength detection of dusty star-forming galaxies at high redshift.” In: *Nature* 394.6690 (1998). doi: [10.1038/28338](https://doi.org/10.1038/28338).
- [24] David H. Hughes et al. “High-redshift star formation in the Hubble Deep Field revealed by a submillimetre-wavelength survey.” In: *Nature* 394.6690 (1998). doi: [10.1038/28328](https://doi.org/10.1038/28328).
- [25] S. Eales et al. “The Herschel ATLAS.” In: *PASP* 122.891 (2010). doi: [10.1086/653086](https://doi.org/10.1086/653086).

- [26] E. Valiante et al. "The Herschel-ATLAS data release 1 - I. Maps, catalogues and number counts." In: *MNRAS* 462.3 (2016). doi: [10.1093/mnras/stw1806](https://doi.org/10.1093/mnras/stw1806).
- [27] S. J. Oliver et al. "The Herschel Multi-tiered Extragalactic Survey: HerMES." In: *MNRAS* 424.3 (2012). doi: [10.1111/j.1365-2966.2012.20912.x](https://doi.org/10.1111/j.1365-2966.2012.20912.x).
- [28] M. P. Viero et al. "The Herschel Stripe 82 Survey (HerS): Maps and Early Catalog." In: *ApJS* 210.2, 22 (2014). doi: [10.1088/0067-0049/210/2/22](https://doi.org/10.1088/0067-0049/210/2/22).
- [29] G. L. Pilbratt et al. "Herschel Space Observatory. An ESA facility for far-infrared and submillimetre astronomy." In: *A&A* 518, L1 (2010). doi: [10.1051/0004-6361/201014759](https://doi.org/10.1051/0004-6361/201014759).
- [30] J. D. Vieira et al. "Dusty starburst galaxies in the early Universe as revealed by gravitational lensing." In: *Nature* 495.7441 (2013). doi: [10.1038/nature12001](https://doi.org/10.1038/nature12001).
- [31] R. Cañameras et al. "Planck's dusty GEMS: The brightest gravitationally lensed galaxies discovered with the Planck all-sky survey." In: *A&A* 581, A105 (2015). doi: [10.1051/0004-6361/201425128](https://doi.org/10.1051/0004-6361/201425128).
- [32] M. S. Bothwell et al. "A survey of molecular gas in luminous sub-millimetre galaxies." In: *MNRAS* 429.4 (2013). doi: [10.1093/mnras/sts562](https://doi.org/10.1093/mnras/sts562).
- [33] Jack E. Birkin et al. "An ALMA/NOEMA survey of the molecular gas properties of high-redshift star-forming galaxies." In: *MNRAS* 501.3 (2021). doi: [10.1093/mnras/staa3862](https://doi.org/10.1093/mnras/staa3862).
- [34] T. B. Miller et al. "A massive core for a cluster of galaxies at a redshift of 4.3." In: *Nature* 556.7702 (2018). doi: [10.1038/s41586-018-0025-2](https://doi.org/10.1038/s41586-018-0025-2).
- [35] U. Dudzevičiūtė et al. "An ALMA survey of the SCUBA-2 CLS UDS field: physical properties of 707 sub-millimetre galaxies." In: *MNRAS* 494.3 (2020). doi: [10.1093/mnras/staa769](https://doi.org/10.1093/mnras/staa769).
- [36] C. Reuter et al. "The Rest-frame Submillimeter Spectrum of High-redshift, Dusty, Star-forming Galaxies from the SPT-SZ Survey." In: *ApJ* 948.1, 44 (2023). doi: [10.3847/1538-4357/acaf51](https://doi.org/10.3847/1538-4357/acaf51).
- [37] A. Weiß et al. "ALMA Redshifts of Millimeter-selected Galaxies from the SPT Survey: The Redshift Distribution of Dusty Star-forming Galaxies." In: *ApJ* 767.1, 88 (2013). doi: [10.1088/0004-637X/767/1/88](https://doi.org/10.1088/0004-637X/767/1/88).
- [38] J. A. Hodge et al. "Kiloparsec-scale Dust Disks in High-redshift Luminous Submillimeter Galaxies." In: *ApJ* 833.1, 103 (2016). doi: [10.3847/1538-4357/833/1/103](https://doi.org/10.3847/1538-4357/833/1/103).

- [39] B. Gullberg et al. "An ALMA survey of the SCUBA-2 Cosmology Legacy Survey UKIDSS/UDS field: high-resolution dust continuum morphologies and the link between sub-millimetre galaxies and spheroid formation." In: *MNRAS* 490.4 (2019). DOI: [10.1093/mnras/stz2835](https://doi.org/10.1093/mnras/stz2835).
- [40] F. Rizzo et al. "A dynamically cold disk galaxy in the early Universe." In: *Nature* 584.7820 (2020). DOI: [10.1038/s41586-020-2572-6](https://doi.org/10.1038/s41586-020-2572-6).
- [41] F. Rizzo et al. "Dynamical characterization of galaxies up to $z \sim 7$." In: *A&A* 667, A5 (2022). DOI: [10.1051/0004-6361/202243582](https://doi.org/10.1051/0004-6361/202243582).
- [42] Ryan C. Hickox et al. "The LABOCA survey of the Extended Chandra Deep Field-South: clustering of submillimetre galaxies." In: *MNRAS* 421.1 (2012). DOI: [10.1111/j.1365-2966.2011.20303.x](https://doi.org/10.1111/j.1365-2966.2011.20303.x).
- [43] S. C. Chapman, A. W. Blain, Ian Smail, and R. J. Ivison. "A Redshift Survey of the Submillimeter Galaxy Population." In: *ApJ* 622.2 (2005). DOI: [10.1086/428082](https://doi.org/10.1086/428082).
- [44] I. Delvecchio et al. "Tracing the cosmic growth of supermassive black holes to $z \sim 3$ with Herschel." In: *MNRAS* 439.3 (2014). DOI: [10.1093/mnras/stu130](https://doi.org/10.1093/mnras/stu130).
- [45] Philip F. Hopkins et al. "A Unified, Merger-driven Model of the Origin of Starbursts, Quasars, the Cosmic X-Ray Background, Supermassive Black Holes, and Galaxy Spheroids." In: *ApJS* 163.1 (2006). DOI: [10.1086/499298](https://doi.org/10.1086/499298).
- [46] Rachel S. Somerville and Romeel Davé. "Physical Models of Galaxy Formation in a Cosmological Framework." In: *ARA&A* 53 (2015). DOI: [10.1146/annurev-astro-082812-140951](https://doi.org/10.1146/annurev-astro-082812-140951).
- [47] C. M. Baugh et al. "Can the faint submillimetre galaxies be explained in the Λ cold dark matter model?" In: *MNRAS* 356.3 (2005). DOI: [10.1111/j.1365-2966.2004.08553.x](https://doi.org/10.1111/j.1365-2966.2004.08553.x).
- [48] Romeel Davé et al. "The nature of submillimetre galaxies in cosmological hydrodynamic simulations." In: *MNRAS* 404.3 (2010). DOI: [10.1111/j.1365-2966.2010.16395.x](https://doi.org/10.1111/j.1365-2966.2010.16395.x).
- [49] Christopher C. Hayward et al. "Submillimetre galaxies in a hierarchical universe: number counts, redshift distribution and implications for the IMF." In: *MNRAS* 428.3 (2013). DOI: [10.1093/mnras/sts222](https://doi.org/10.1093/mnras/sts222).
- [50] Desika Narayanan et al. "The formation of submillimetre-bright galaxies from gas infall over a billion years." In: *Nature* 525.7570 (2015). DOI: [10.1038/nature15383](https://doi.org/10.1038/nature15383).

- [51] Claudia del P. Lagos et al. "Physical properties and evolution of (sub-)millimetre-selected galaxies in the galaxy formation simulation SHARK." In: *MNRAS* 499.2 (2020). doi: [10.1093/mnras/staa2861](https://doi.org/10.1093/mnras/staa2861).
- [52] Lennox L. Cowie, Antoinette Songaila, Esther M. Hu, and J. G. Cohen. "New Insight on Galaxy Formation and Evolution From Keck Spectroscopy of the Hawaii Deep Fields." In: *AJ* 112 (1996). doi: [10.1086/118058](https://doi.org/10.1086/118058).
- [53] Daniel Thomas, Claudia Maraston, Ralf Bender, and Claudia Mendes de Oliveira. "The Epochs of Early-Type Galaxy Formation as a Function of Environment." In: *ApJ* 621.2 (2005). doi: [10.1086/426932](https://doi.org/10.1086/426932).
- [54] A. Sandage. "Star formation rates, galaxy morphology and the Hubble sequence." In: *A&A* 161 (1986).
- [55] Simon J. Lilly et al. "The Canada-United Kingdom Deep Submillimeter Survey. II. First Identifications, Redshifts, and Implications for Galaxy Evolution." In: *ApJ* 518.2 (1999). doi: [10.1086/307310](https://doi.org/10.1086/307310).
- [56] A. M. Swinbank et al. "The link between submillimetre galaxies and luminous ellipticals: near-infrared IFU spectroscopy of submillimetre galaxies." In: *MNRAS* 371.1 (2006). doi: [10.1111/j.1365-2966.2006.10673.x](https://doi.org/10.1111/j.1365-2966.2006.10673.x).
- [57] C. Gómez-Guijarro et al. "Starburst to Quiescent from HST/ALMA: Stars and Dust Unveil Minor Mergers in Submillimeter Galaxies at $z \sim 4.5$." In: *ApJ* 856.2, 121 (2018). doi: [10.3847/1538-4357/aab206](https://doi.org/10.3847/1538-4357/aab206).
- [58] A. Lapi et al. "The Dramatic Size and Kinematic Evolution of Massive Early-type Galaxies." In: *ApJ* 857.1 (2018). doi: [10.3847/1538-4357/aab6af](https://doi.org/10.3847/1538-4357/aab6af).
- [59] L. Pantoni, A. Lapi, M. Massardi, S. Goswami, and L. Danese. "New Analytic Solutions for Galaxy Evolution: Gas, Stars, Metals, and Dust in Local ETGs and Their High- z Star-forming Progenitors." In: *ApJ* 880.2, 129 (2019). doi: [10.3847/1538-4357/ab2adc](https://doi.org/10.3847/1538-4357/ab2adc).
- [60] K. G. Noeske et al. "Star Formation in AEGIS Field Galaxies since $z=1.1$: The Dominance of Gradually Declining Star Formation, and the Main Sequence of Star-Forming Galaxies." In: *ApJ* 660.1 (2007). doi: [10.1086/517926](https://doi.org/10.1086/517926).
- [61] D. Elbaz et al. "GOODS-Herschel: an infrared main sequence for star-forming galaxies." In: *A&A* 533, A119 (2011). doi: [10.1051/0004-6361/201117239](https://doi.org/10.1051/0004-6361/201117239).

- [62] J. S. Speagle, C. L. Steinhardt, P. L. Capak, and J. D. Silverman. "A Highly Consistent Framework for the Evolution of the Star-Forming "Main Sequence" from $z \sim 0-6$." In: *ApJS* 214.2, 15 (2014). doi: [10.1088/0067-0049/214/2/15](https://doi.org/10.1088/0067-0049/214/2/15).
- [63] Cedric G. Lacey et al. "A unified multiwavelength model of galaxy formation." In: *MNRAS* 462.4 (2016). doi: [10.1093/mnras/stw1888](https://doi.org/10.1093/mnras/stw1888).
- [64] A. Dekel et al. "Cold streams in early massive hot haloes as the main mode of galaxy formation." In: *Nature* 457.7228 (2009). doi: [10.1038/nature07648](https://doi.org/10.1038/nature07648).
- [65] E. Daddi et al. "Very High Gas Fractions and Extended Gas Reservoirs in $z = 1.5$ Disk Galaxies." In: *ApJ* 713.1 (2010). doi: [10.1088/0004-637X/713/1/686](https://doi.org/10.1088/0004-637X/713/1/686).
- [66] Andrea Cimatti, Filippo Fraternali, and Carlo Nipoti. "Introduction to Galaxy Formation and Evolution. From Primordial Gas to Present-Day Galaxies." In: *arXiv e-prints*, arXiv:1912.06216 (2019). doi: [10.48550/arXiv.1912.06216](https://doi.org/10.48550/arXiv.1912.06216).
- [67] G. Rodighiero et al. "The Lesser Role of Starbursts in Star Formation at $z = 2$." In: *ApJ* 739.2, L40 (2011). doi: [10.1088/2041-8205/739/2/L40](https://doi.org/10.1088/2041-8205/739/2/L40).
- [68] R. Genzel et al. "A study of the gas-star formation relation over cosmic time." In: *MNRAS* 407.4 (2010). doi: [10.1111/j.1365-2966.2010.16969.x](https://doi.org/10.1111/j.1365-2966.2010.16969.x).
- [69] Desika Narayanan et al. "The formation of high-redshift submillimetre galaxies." In: *MNRAS* 401.3 (2010). doi: [10.1111/j.1365-2966.2009.15790.x](https://doi.org/10.1111/j.1365-2966.2009.15790.x).
- [70] L. J. Tacconi et al. "High molecular gas fractions in normal massive star-forming galaxies in the young Universe." In: *Nature* 463.7282 (2010). doi: [10.1038/nature08773](https://doi.org/10.1038/nature08773).
- [71] R. J. Ivison et al. "Gas-rich mergers and feedback are ubiquitous amongst starbursting radio galaxies, as revealed by the VLA, IRAM PdBI and Herschel." In: *MNRAS* 425.2 (2012). doi: [10.1111/j.1365-2966.2012.21544.x](https://doi.org/10.1111/j.1365-2966.2012.21544.x).
- [72] M. Aravena et al. "A survey of the cold molecular gas in gravitationally lensed star-forming galaxies at $z > 2$." In: *MNRAS* 457.4 (2016). doi: [10.1093/mnras/stw275](https://doi.org/10.1093/mnras/stw275).
- [73] Alberto D. Bolatto, Mark Wolfire, and Adam K. Leroy. "The CO-to-H₂ Conversion Factor." In: *ARA&A* 51.1 (2013). doi: [10.1146/annurev-astro-082812-140944](https://doi.org/10.1146/annurev-astro-082812-140944).
- [74] C. L. Carilli and F. Walter. "Cool Gas in High-Redshift Galaxies." In: *ARA&A* 51.1 (2013). doi: [10.1146/annurev-astro-082812-140953](https://doi.org/10.1146/annurev-astro-082812-140953).

- [75] Stephen Serjeant. “Strong biases in infrared-selected gravitational lenses.” In: *MNRAS* 424.4 (2012). doi: [10.1111/j.1365-2966.2012.20761.x](https://doi.org/10.1111/j.1365-2966.2012.20761.x).
- [76] Yashar D. Hezaveh, Daniel P. Marrone, and Gilbert P. Holder. “Size Bias and Differential Lensing of Strongly Lensed, Dusty Galaxies Identified in Wide-Field Surveys.” In: *ApJ* 761.1, 20 (2012). doi: [10.1088/0004-637X/761/1/20](https://doi.org/10.1088/0004-637X/761/1/20).
- [77] B. T. Draine. “Interstellar Dust Grains.” In: *ARA&A* 41 (2003). doi: [10.1146/annurev.astro.41.011802.094840](https://doi.org/10.1146/annurev.astro.41.011802.094840).
- [78] Bruce T. Draine. *Physics of the Interstellar and Intergalactic Medium*. 2011.
- [79] Katia M. Ferrière. “The interstellar environment of our galaxy.” In: *Reviews of Modern Physics* 73.4 (2001). doi: [10.1103/RevModPhys.73.1031](https://doi.org/10.1103/RevModPhys.73.1031).
- [80] Houjun Mo, Frank C. van den Bosch, and Simon White. *Galaxy Formation and Evolution*. 2010. doi: [10.1017/CB09780511807244](https://doi.org/10.1017/CB09780511807244).
- [81] Maarten Schmidt. “The Rate of Star Formation.” In: *ApJ* 129 (1959). doi: [10.1086/146614](https://doi.org/10.1086/146614).
- [82] Robert C. Kennicutt Jr. “The Global Schmidt Law in Star-forming Galaxies.” In: *ApJ* 498.2 (1998). doi: [10.1086/305588](https://doi.org/10.1086/305588).
- [83] Roberto Decarli et al. “The ALMA Spectroscopic Survey in the Hubble Ultra Deep Field: Multiband Constraints on Line-luminosity Functions and the Cosmic Density of Molecular Gas.” In: *ApJ* 902.2, 110 (2020). doi: [10.3847/1538-4357/abaa3b](https://doi.org/10.3847/1538-4357/abaa3b).
- [84] Dušan Kereš, Neal Katz, David H. Weinberg, and Romeel Davé. “How do galaxies get their gas?” In: *MNRAS* 363.1 (2005). doi: [10.1111/j.1365-2966.2005.09451.x](https://doi.org/10.1111/j.1365-2966.2005.09451.x).
- [85] Amélie Saintonge et al. “Validation of the Equilibrium Model for Galaxy Evolution to $z \sim 3$ through Molecular Gas and Dust Observations of Lensed Star-forming Galaxies.” In: *ApJ* 778.1, 2 (2013). doi: [10.1088/0004-637X/778/1/2](https://doi.org/10.1088/0004-637X/778/1/2).
- [86] L. J. Tacconi et al. “Phibss: Molecular Gas Content and Scaling Relations in $z \sim 1-3$ Massive, Main-sequence Star-forming Galaxies.” In: *ApJ* 768.1, 74 (2013). doi: [10.1088/0004-637X/768/1/74](https://doi.org/10.1088/0004-637X/768/1/74).
- [87] Matthieu Béthermin et al. “Evolution of the dust emission of massive galaxies up to $z = 4$ and constraints on their dominant mode of star formation.” In: *A&A* 573, A113 (2015). doi: [10.1051/0004-6361/201425031](https://doi.org/10.1051/0004-6361/201425031).

- [88] Elisabete da Cunha et al. "ON THE EFFECT OF THE COSMIC MICROWAVE BACKGROUND IN HIGH-REDSHIFT (SUB-)MILLIMETER OBSERVATIONS." In: *ApJ* 766.1 (2013). DOI: [10.1088/0004-637X/766/1/13](https://doi.org/10.1088/0004-637X/766/1/13).
- [89] Zhi-Yu Zhang et al. "Gone with the heat: a fundamental constraint on the imaging of dust and molecular gas in the early Universe." In: *Royal Society Open Science* 3.6, 160025 (2016). DOI: [10.1098/rsos.160025](https://doi.org/10.1098/rsos.160025).
- [90] Derek Wilson et al. "Stacked Average Far-infrared Spectrum of Dusty Star-forming Galaxies from the Herschel/SPIRE Fourier Transform Spectrometer." In: *ApJ* 848.1, 30 (2017). DOI: [10.3847/1538-4357/aa8cc7](https://doi.org/10.3847/1538-4357/aa8cc7).
- [91] Matus Rybak et al. "Strong Far-ultraviolet Fields Drive the [C II]/Far-infrared Deficit in $z \sim 3$ Dusty, Star-forming Galaxies." In: *ApJ* 876.2, 112 (2019). DOI: [10.3847/1538-4357/ab0e0f](https://doi.org/10.3847/1538-4357/ab0e0f).
- [92] Mark G. Wolfire, Livia Vallini, and Mélanie Chevance. "Photodissociation and X-Ray-Dominated Regions." In: *ARA&A* 60 (2022). DOI: [10.1146/annurev-astro-052920-010254](https://doi.org/10.1146/annurev-astro-052920-010254).
- [93] Thomas G. Bisbas, Padelis P. Papadopoulos, and Serena Viti. "Effective Destruction of CO by Cosmic Rays: Implications for Tracing H₂ Gas in the Universe." In: *ApJ* 803.1, 37 (2015). DOI: [10.1088/0004-637X/803/1/37](https://doi.org/10.1088/0004-637X/803/1/37).
- [94] Thomas G. Bisbas et al. "Cosmic-ray Induced Destruction of CO in Star-forming Galaxies." In: *ApJ* 839.2, 90 (2017). DOI: [10.3847/1538-4357/aa696d](https://doi.org/10.3847/1538-4357/aa696d).
- [95] Marco Padovani et al. "Impact of Low-Energy Cosmic Rays on Star Formation." In: *Space Sci. Rev.* 216.2, 29 (2020). DOI: [10.1007/s11214-020-00654-1](https://doi.org/10.1007/s11214-020-00654-1).
- [96] D. Donevski et al. "In pursuit of giants. I. The evolution of the dust-to-stellar mass ratio in distant dusty galaxies." In: *A&A* 644, A144 (2020). DOI: [10.1051/0004-6361/202038405](https://doi.org/10.1051/0004-6361/202038405).
- [97] Carlos De Breuck et al. "A dense, solar metallicity ISM in the $z = 4.2$ dusty star-forming galaxy SPT 0418-47." In: *A&A* 631, A167 (2019). DOI: [10.1051/0004-6361/201936169](https://doi.org/10.1051/0004-6361/201936169).
- [98] Jack E. Birkin et al. "JWST's TEMPLATES for Star Formation: The First Resolved Gas-phase Metallicity Maps of Dust-obscured Star-forming Galaxies at $z \sim 4$." In: *ApJ* 958.1, 64 (2023). DOI: [10.3847/1538-4357/acf712](https://doi.org/10.3847/1538-4357/acf712).
- [99] Alain Omont. "Molecules in galaxies." In: *Reports on Progress in Physics* 70.7 (2007). DOI: [10.1088/0034-4885/70/7/R03](https://doi.org/10.1088/0034-4885/70/7/R03).
- [100] P. M. Solomon and P. A. Vanden Bout. "Molecular Gas at High Redshift." In: *ARA&A* 43.1 (2005). DOI: [10.1146/annurev-astro.43.051804.102221](https://doi.org/10.1146/annurev-astro.43.051804.102221).

- [101] A. Weiß et al. “Highly-excited CO emission in APM 08279+5255 at $z = 3.9$.” In: *A&A* 467.3 (2007). doi: [10.1051/0004-6361:20066117](https://doi.org/10.1051/0004-6361:20066117).
- [102] P. M. Solomon, A. R. Rivolo, J. Barrett, and A. Yahil. “Mass, Luminosity, and Line Width Relations of Galactic Molecular Clouds.” In: *ApJ* 319 (1987). doi: [10.1086/165493](https://doi.org/10.1086/165493).
- [103] A. W. Strong and J. R. Mattox. “Gradient model analysis of EGRET diffuse Galactic γ -ray emission.” In: *A&A* 308 (1996).
- [104] D. Downes and P. M. Solomon. “Rotating Nuclear Rings and Extreme Starbursts in Ultraluminous Galaxies.” In: *ApJ* 507.2 (1998). doi: [10.1086/306339](https://doi.org/10.1086/306339).
- [105] Kevin C. Harrington et al. “Turbulent Gas in Lensed Planck-selected Starbursts at $z \sim 1$ -3.5.” In: *ApJ* 908.1, 95 (2021). doi: [10.3847/1538-4357/abcc01](https://doi.org/10.3847/1538-4357/abcc01).
- [106] Mark G. Wolfire, David Hollenbach, and Christopher F. McKee. “The Dark Molecular Gas.” In: *ApJ* 716.2 (2010). doi: [10.1088/0004-637X/716/2/1191](https://doi.org/10.1088/0004-637X/716/2/1191).
- [107] Rowan J. Smith, Simon C. O. Glover, Paul C. Clark, Ralf S. Klessen, and Volker Springel. “CO-dark gas and molecular filaments in Milky Way-type galaxies.” In: *MNRAS* 441.2 (2014). doi: [10.1093/mnras/stu616](https://doi.org/10.1093/mnras/stu616).
- [108] A. Weiß, C. Henkel, D. Downes, and F. Walter. “Gas and dust in the Cloverleaf quasar at redshift 2.5.” In: *A&A* 409 (2003). doi: [10.1051/0004-6361:20031337](https://doi.org/10.1051/0004-6361:20031337).
- [109] P. P. Papadopoulos, W.-F. Thi, and S. Viti. “CI lines as tracers of molecular gas, and their prospects at high redshifts.” In: *MNRAS* 351.1 (2004). doi: [10.1111/j.1365-2966.2004.07762.x](https://doi.org/10.1111/j.1365-2966.2004.07762.x).
- [110] Kasper E. Heintz and Darach Watson. “Direct Measurement of the [C I] Luminosity to Molecular Gas Mass Conversion Factor in High-redshift Star-forming Galaxies.” In: *ApJ* 889.1, L7 (2020). doi: [10.3847/2041-8213/ab6733](https://doi.org/10.3847/2041-8213/ab6733).
- [111] Thomas G. Bisbas, Jonathan C. Tan, and Kei E. I. Tanaka. “Photodissociation region diagnostics across galactic environments.” In: *MNRAS* 502.2 (2021). doi: [10.1093/mnras/stab121](https://doi.org/10.1093/mnras/stab121).
- [112] J. Stutzki et al. “Atomic Carbon in M82: Physical Conditions Derived from Simultaneous Observations of the [C I] Fine-Structure Submillimeter-Wave Transitions.” In: *ApJ* 477.1 (1997). doi: [10.1086/310514](https://doi.org/10.1086/310514).
- [113] F. Walter, A. Weiß, D. Downes, R. Decarli, and C. Henkel. “A Survey of Atomic Carbon at High Redshift.” In: *ApJ* 730.1, 18 (2011). doi: [10.1088/0004-637X/730/1/18](https://doi.org/10.1088/0004-637X/730/1/18).

- [114] Padelì P. Papadopoulos and Thomas R. Greve. "C I Emission in Ultraluminous Infrared Galaxies as a Molecular Gas Mass Tracer." In: *ApJ* 615.1 (2004). DOI: [10.1086/426059](https://doi.org/10.1086/426059).
- [115] M. S. Bothwell et al. "ALMA observations of atomic carbon in $z \sim 4$ dusty star-forming galaxies." In: *MNRAS* 466.3 (2017). DOI: [10.1093/mnras/stw3270](https://doi.org/10.1093/mnras/stw3270).
- [116] Francesco Valentino et al. "A Survey of Atomic Carbon [C I] in High-redshift Main-sequence Galaxies." In: *ApJ* 869.1, 27 (2018). DOI: [10.3847/1538-4357/aaeb88](https://doi.org/10.3847/1538-4357/aaeb88).
- [117] E. Daddi et al. "CO excitation of normal star-forming galaxies out to $z = 1.5$ as regulated by the properties of their interstellar medium." In: *A&A* 577, A46 (2015). DOI: [10.1051/0004-6361/201425043](https://doi.org/10.1051/0004-6361/201425043).
- [118] Leindert A. Boogaard et al. "The ALMA Spectroscopic Survey in the Hubble Ultra Deep Field: CO Excitation and Atomic Carbon in Star-forming Galaxies at $z = 1-3$." In: *ApJ* 902.2 (2020). DOI: [10.3847/1538-4357/abb82f](https://doi.org/10.3847/1538-4357/abb82f).
- [119] Desika Narayanan and Mark R. Krumholz. "A theory for the excitation of CO in star-forming galaxies." In: *MNRAS* 442.2 (2014). DOI: [10.1093/mnras/stu834](https://doi.org/10.1093/mnras/stu834).
- [120] Mark R. Krumholz and Todd A. Thompson. "The Relationship between Molecular Gas Tracers and Kennicutt-Schmidt Laws." In: *ApJ* 669.1 (2007). DOI: [10.1086/521642](https://doi.org/10.1086/521642).
- [121] C. Yang et al. "Molecular gas in the Herschel-selected strongly lensed submillimeter galaxies at z 2-4 as probed by multi-J CO lines." In: *A&A* 608, A144 (2017). DOI: [10.1051/0004-6361/201731391](https://doi.org/10.1051/0004-6361/201731391).
- [122] R. Meijerink and M. Spaans. "Diagnostics of irradiated gas in galaxy nuclei. I. A far-ultraviolet and X-ray dominated region code." In: *A&A* 436.2 (2005). DOI: [10.1051/0004-6361:20042398](https://doi.org/10.1051/0004-6361:20042398).
- [123] V. V. Sobolev. "The Brightness of a Spherical Nebula." In: *Soviet Ast.* 4 (1960).
- [124] Peter Goldreich and John Kwan. "Molecular Clouds." In: *ApJ* 189 (1974). DOI: [10.1086/152821](https://doi.org/10.1086/152821).
- [125] N. Z. Scoville and P. M. Solomon. "Radiative Transfer, Excitation, and Cooling of Molecular Emission Lines (CO and CS)." In: *ApJ* 187 (1974). DOI: [10.1086/181398](https://doi.org/10.1086/181398).
- [126] F. F. S. van der Tak, J. H. Black, F. L. Schöier, et al. "A computer program for fast non-LTE analysis of interstellar line spectra. With diagnostic plots to interpret observed line intensity ratios." In: *A&A* 468.2 (2007). DOI: [10.1051/0004-6361:20066820](https://doi.org/10.1051/0004-6361:20066820).

- [127] Daizhong Liu et al. "High-J CO versus Far-infrared Relations in Normal and Starburst Galaxies." In: *ApJ* 810.2, L14 (2015). DOI: [10.1088/2041-8205/810/2/L14](https://doi.org/10.1088/2041-8205/810/2/L14).
- [128] P. P. van der Werf et al. "Black hole accretion and star formation as drivers of gas excitation and chemistry in Markarian 231." In: *A&A* 518, L42 (2010). DOI: [10.1051/0004-6361/201014682](https://doi.org/10.1051/0004-6361/201014682).
- [129] M. J. F. Rosenberg et al. "The Herschel Comprehensive (U)LIRG Emission Survey (HERCULES): CO Ladders, Fine Structure Lines, and Neutral Gas Cooling." In: *ApJ* 801.2, 72 (2015). DOI: [10.1088/0004-637X/801/2/72](https://doi.org/10.1088/0004-637X/801/2/72).
- [130] J. S. Spilker et al. "The Rest-frame Submillimeter Spectrum of High-redshift, Dusty, Star-forming Galaxies." In: *ApJ* 785.2, 149 (2014). DOI: [10.1088/0004-637X/785/2/149](https://doi.org/10.1088/0004-637X/785/2/149).
- [131] Leindert A. Boogaard et al. "Measuring the Average Molecular Gas Content of Star-forming Galaxies at $z = 3-4$." In: *ApJ* 916.1, 12 (2021). DOI: [10.3847/1538-4357/ac01d7](https://doi.org/10.3847/1538-4357/ac01d7).
- [132] Adam K. Leroy et al. "The CO-to-H₂ Conversion Factor from Infrared Dust Emission across the Local Group." In: *ApJ* 737.1, 12 (2011). DOI: [10.1088/0004-637X/737/1/12](https://doi.org/10.1088/0004-637X/737/1/12).
- [133] Georgios E. Magdis et al. "The Evolving Interstellar Medium of Star-forming Galaxies since $z = 2$ as Probed by Their Infrared Spectral Energy Distributions." In: *ApJ* 760.1, 6 (2012). DOI: [10.1088/0004-637X/760/1/6](https://doi.org/10.1088/0004-637X/760/1/6).
- [134] N. Scoville et al. "ISM Masses and the Star formation Law at $Z = 1$ to 6: ALMA Observations of Dust Continuum in 145 Galaxies in the COSMOS Survey Field." In: *ApJ* 820.2, 83 (2016). DOI: [10.3847/0004-637X/820/2/83](https://doi.org/10.3847/0004-637X/820/2/83).
- [135] R. H. Hildebrand. "The determination of cloud masses and dust characteristics from submillimetre thermal emission." In: *QJRAS* 24 (1983).
- [136] E. González-Alfonso, J. Fischer, S. Aalto, and N. Falstad. "Modeling the H₂O submillimeter emission in extragalactic sources." In: *A&A* 567, A91 (2014). DOI: [10.1051/0004-6361/201423980](https://doi.org/10.1051/0004-6361/201423980).
- [137] A. Omont et al. "H₂O emission in high- z ultra-luminous infrared galaxies." In: *A&A* 551, A115 (2013). DOI: [10.1051/0004-6361/201220811](https://doi.org/10.1051/0004-6361/201220811).
- [138] C. Yang et al. "Submillimeter H₂O and H₂O⁺ emission in lensed ultra- and hyper-luminous infrared galaxies at $z = 2-4$." In: *A&A* 595, A80 (2016). DOI: [10.1051/0004-6361/201628160](https://doi.org/10.1051/0004-6361/201628160).

- [139] G. C. Jones, R. Maiolino, P. Caselli, and S. Carniani. "Detection of a high-redshift molecular outflow in a primeval hyperstarburst galaxy." In: *A&A* 632, L7 (2019). DOI: [10.1051/0004-6361/201936989](https://doi.org/10.1051/0004-6361/201936989).
- [140] Francesca Perrotta et al. "The Way of Water: ALMA Resolves H₂O Emission Lines in a Strongly Lensed Dusty Star-forming Galaxy at $z \sim 3.1$." In: *ApJ* 952.1, 90 (2023). DOI: [10.3847/1538-4357/accd72](https://doi.org/10.3847/1538-4357/accd72).
- [141] L. Liu et al. "HIFI Spectroscopy of H₂O Submillimeter Lines in Nuclei of Actively Star-forming Galaxies." In: *ApJ* 846.1, 5 (2017). DOI: [10.3847/1538-4357/aa81b4](https://doi.org/10.3847/1538-4357/aa81b4).
- [142] Yu Gao and Philip M. Solomon. "HCN Survey of Normal Spiral, Infrared-luminous, and Ultraluminous Galaxies." In: *ApJS* 152.1 (2004). DOI: [10.1086/383003](https://doi.org/10.1086/383003).
- [143] Jens Kauffmann et al. "Molecular Line Emission as a Tool for Galaxy Observations (LEGO). I. HCN as a tracer of moderate gas densities in molecular clouds and galaxies." In: *A&A* 605, L5 (2017). DOI: [10.1051/0004-6361/201731123](https://doi.org/10.1051/0004-6361/201731123).
- [144] M. Rybak et al. "PRUSSIC. I. A JVLA survey of HCN, HCO⁺, and HNC (1-0) emission in $z \sim 3$ dusty galaxies: Low dense-gas fractions in high-redshift star-forming galaxies." In: *A&A* 667, A70 (2022). DOI: [10.1051/0004-6361/202243894](https://doi.org/10.1051/0004-6361/202243894).
- [145] G. C. Privon et al. "Excitation Mechanisms for HCN (1-0) and HCO⁺ (1-0) in Galaxies from the Great Observatories All-sky LIRG Survey." In: *ApJ* 814.1, 39 (2015). DOI: [10.1088/0004-637X/814/1/39](https://doi.org/10.1088/0004-637X/814/1/39).
- [146] Takuma Izumi et al. "Submillimeter-HCN Diagram for Energy Diagnostics in the Centers of Galaxies." In: *ApJ* 818.1, 42 (2016). DOI: [10.3847/0004-637X/818/1/42](https://doi.org/10.3847/0004-637X/818/1/42).
- [147] A. L. R. Danielson et al. "¹³CO and C¹⁸O emission from a dense gas disc at $z = 2.3$: abundance variations, cosmic rays and the initial conditions for star formation." In: *MNRAS* 436.3 (2013). DOI: [10.1093/mnras/stt1775](https://doi.org/10.1093/mnras/stt1775).
- [148] M. Béthermin et al. "Dense-gas tracers and carbon isotopes in five $2.5 < z < 4$ lensed dusty star-forming galaxies from the SPT SMG sample." In: *A&A* 620, A115 (2018). DOI: [10.1051/0004-6361/201833081](https://doi.org/10.1051/0004-6361/201833081).
- [149] Amélie Saintonge et al. "COLD GASS, an IRAM legacy survey of molecular gas in massive galaxies - I. Relations between H₂, H I, stellar content and structural properties." In: *MNRAS* 415.1 (2011). DOI: [10.1111/j.1365-2966.2011.18677.x](https://doi.org/10.1111/j.1365-2966.2011.18677.x).

- [150] Amélie Saintonge et al. “xCOLD GASS: The Complete IRAM 30 m Legacy Survey of Molecular Gas for Galaxy Evolution Studies.” In: *ApJS* 233.2, 22 (2017). DOI: [10.3847/1538-4365/aa97e0](https://doi.org/10.3847/1538-4365/aa97e0).
- [151] Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, and et al. “The FAIR Guiding Principles for scientific data management and stewardship.” In: *Scientific Data* 3 (2016). DOI: [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18).
- [152] Ulisses Barres de Almeida. “The Southern Wide Field Gamma ray Observatory.” In: *Astronomische Nachrichten* 342.1-2 (2021). DOI: [10.1002/asna.202113946](https://doi.org/10.1002/asna.202113946).
- [153] M. Negrello et al. “The Herschel-ATLAS: a sample of 500 μ m-selected lensed galaxies over 600 deg².” In: *MNRAS* 465.3 (2017). DOI: [10.1093/mnras/stw2911](https://doi.org/10.1093/mnras/stw2911).
- [154] Mattia Negrello et al. “The Detection of a Population of Submillimeter-Bright, Strongly Lensed Galaxies.” In: *Science* 330.6005 (2010). DOI: [10.1126/science.1193420](https://doi.org/10.1126/science.1193420).
- [155] Tom J. L. C. Bakx et al. “The Herschel Bright Sources (HerBS): sample definition and SCUBA-2 observations.” In: *MNRAS* 473.2 (2018). DOI: [10.1093/mnras/stx2267](https://doi.org/10.1093/mnras/stx2267).
- [156] E. Borsato, L. Marchetti, M. Negrello, E. M. Corsini, D. Wake, et al. “Characterization of Herschel-selected strong lens candidates through HST and sub-mm/mm observations.” In: *MNRAS* 528.4 (2024). DOI: [10.1093/mnras/stad3381](https://doi.org/10.1093/mnras/stad3381).
- [157] Julie L. Wardlow et al. “HerMES: CANDIDATE GRAVITATIONALLY LENSED GALAXIES AND LENSING STATISTICS AT SUBMILLIMETER WAVELENGTHS.” In: *ApJ* 762.1 (2012). DOI: [10.1088/0004-637X/762/1/59](https://doi.org/10.1088/0004-637X/762/1/59).
- [158] H. Nayyeri et al. “Candidate Gravitationally Lensed Dusty Star-forming Galaxies in the Herschel Wide Area Surveys.” In: *ApJ* 823.1, 17 (2016). DOI: [10.3847/0004-637X/823/1/17](https://doi.org/10.3847/0004-637X/823/1/17).
- [159] T. J. L. C. Bakx et al. “A novel high-z submm galaxy efficient line survey in ALMA Bands 3 through 8 - an ANGELS pilot.” In: *MNRAS* 535.2 (2024). DOI: [10.1093/mnras/stae2409](https://doi.org/10.1093/mnras/stae2409).
- [160] Jean-Baptiste Jolly, Kirsten K. Knudsen, and Flora Stanley. “LINESTACKER: a spectral line stacking tool for interferometric data.” In: *MNRAS* 499.3 (2020). DOI: [10.1093/mnras/staa2908](https://doi.org/10.1093/mnras/staa2908).
- [161] Tom J. L. C. Bakx et al. “Erratum: The Herschel Bright Sources (HerBS): sample definition and SCUBA-2 observations.” In: *MNRAS* 494.1 (2020). DOI: [10.1093/mnras/staa658](https://doi.org/10.1093/mnras/staa658).

- [162] S A Urquhart et al. "The bright extragalactic ALMA redshift survey (BEARS) I: redshifts of bright gravitationally lensed galaxies from the Herschel ATLAS." In: *Monthly Notices of the Royal Astronomical Society* 511.2 (2022). DOI: [10.1093/mnras/stac150](https://doi.org/10.1093/mnras/stac150).
- [163] R. Neri et al. "NOEMA redshift measurements of bright Herschel galaxies." In: *A&A* 635, A7 (2020). DOI: [10.1051/0004-6361/201936988](https://doi.org/10.1051/0004-6361/201936988).
- [164] P. Cox et al. "z-GAL: A NOEMA spectroscopic redshift survey of bright Herschel galaxies. I. Overview." In: *A&A* 678, A26 (2023). DOI: [10.1051/0004-6361/202346801](https://doi.org/10.1051/0004-6361/202346801).
- [165] D. T. Frayer et al. "Green Bank Telescope Zpectrometer CO(1-0) Observations of the Strongly Lensed Submillimeter Galaxies from the Herschel ATLAS." In: *ApJ* 726.2, L22 (2011). DOI: [10.1088/2041-8205/726/2/L22](https://doi.org/10.1088/2041-8205/726/2/L22).
- [166] Hai Fu et al. "A Comprehensive View of a Strongly Lensed Planck-Associated Submillimeter Galaxy." In: *ApJ* 753.2, 134 (2012). DOI: [10.1088/0004-637X/753/2/134](https://doi.org/10.1088/0004-637X/753/2/134).
- [167] R. S. Bussmann et al. "Gravitational Lens Models Based on Submillimeter Array Imaging of Herschel-selected Strongly Lensed Sub-millimeter Galaxies at $z > 1.5$." In: *ApJ* 779.1, 25 (2013). DOI: [10.1088/0004-637X/779/1/25](https://doi.org/10.1088/0004-637X/779/1/25).
- [168] J. A. Calanog et al. "Lens Models of Herschel-selected Galaxies from High-resolution Near-IR Observations." In: *ApJ* 797.2, 138 (2014). DOI: [10.1088/0004-637X/797/2/138](https://doi.org/10.1088/0004-637X/797/2/138).
- [169] Hugo Messias et al. "Herschel-ATLAS and ALMA. HATLAS J142935.3-002836, a lensed major merger at redshift 1.027." In: *A&A* 568, A92 (2014). DOI: [10.1051/0004-6361/201424410](https://doi.org/10.1051/0004-6361/201424410).
- [170] M. Negrello et al. "Herschel *-ATLAS: deep HST/WFC3 imaging of strongly lensed submillimetre galaxies." In: *MNRAS* 440.3 (2014). DOI: [10.1093/mnras/stu413](https://doi.org/10.1093/mnras/stu413).
- [171] S. Dye et al. "Revealing the complex nature of the strong gravitationally lensed system H-ATLAS J090311.6+003906 using ALMA." In: *MNRAS* 452.3 (2015). DOI: [10.1093/mnras/stv1442](https://doi.org/10.1093/mnras/stv1442).
- [172] A. Remijan et al. *ALMA Technical Handbook*. Tech. rep. ALMA Doc. 7.3, ver. 1.1. ALMA, 2019.
- [173] Felix Stoehr et al. "ALMA Publication Statistics." In: *PASP* 138.7, 075001 (2026). DOI: [10.1088/1538-3873/ae7762](https://doi.org/10.1088/1538-3873/ae7762).
- [174] G J Bendo et al. "The bright extragalactic ALMA redshift survey (BEARS) – II. Millimetre photometry of gravitational lens candidates." In: *Monthly Notices of the Royal Astronomical Society* 522.2 (2022). DOI: [10.1093/mnras/stac3771](https://doi.org/10.1093/mnras/stac3771).

- [175] A. Richard Thompson, James M. Moran, and Jr. Swenson George W. *Interferometry and Synthesis in Radio Astronomy*. 3rd. Astronomy and Astrophysics Library. Cham: Springer International Publishing, 2017. doi: [10.1007/978-3-319-44431-4](https://doi.org/10.1007/978-3-319-44431-4).
- [176] Luke T. Maud, Yoshiharu Asaki, William R. F. Dent, et al. "ALMA High-frequency Long-baseline Campaign in 2017: An Investigation of Phase-referencing Cycle Times and Effective Baseline Lengths Using Band-to-band and In-band Phase Calibration Techniques." In: *ApJS* 259 (2022).
- [177] Crystal L. Brogan, Todd R. Hunter, and Ed B. Fomalont. "Advanced Gain Calibration Techniques in Radio Interferometry." In: *arXiv e-prints*, arXiv:1805.05266 (2018).
- [178] Daniel Shenon Briggs. "High Fidelity Deconvolution of Moderately Resolved Sources." PhD thesis. New Mexico Institute of Mining and Technology, 1995.
- [179] J. A. Högbom. "Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines." In: *A&AS* 15 (1974).
- [180] R. Decarli et al. "An ALMA Survey of Sub-millimeter Galaxies in the Extended Chandra Deep Field South: Sub-millimeter Properties of Color-selected Galaxies." In: *ApJ* 780.2, 115 (2014). doi: [10.1088/0004-637X/780/2/115](https://doi.org/10.1088/0004-637X/780/2/115).
- [181] S. Berta, D. Lutz, R. Genzel, N. M. Förster-Schreiber, and L. J. Tacconi. "Measures of galaxy dust and gas mass with Herschel photometry and prospects for ALMA." In: *A&A* 587, A73 (2016). doi: [10.1051/0004-6361/201527746](https://doi.org/10.1051/0004-6361/201527746).
- [182] J. S. Dunlop et al. "A deep ALMA image of the Hubble Ultra Deep Field." In: *MNRAS* 466.1 (2017). doi: [10.1093/mnras/stw3088](https://doi.org/10.1093/mnras/stw3088).
- [183] L. Lindroos, K. K. Knudsen, W. Vlemmings, J. Conway, and I. Martí-Vidal. "Stacking of large interferometric data sets in the image- and uv-domain - a comparative study." In: *MNRAS* 446.4 (2015). doi: [10.1093/mnras/stu2344](https://doi.org/10.1093/mnras/stu2344).
- [184] P. J. Hancock, B. M. Gaensler, and T. Murphy. "Visibility Stacking in the Quest for Type Ia Supernova Radio Emission." In: *ApJ* 735.2, L35 (2011). doi: [10.1088/2041-8205/735/2/L35](https://doi.org/10.1088/2041-8205/735/2/L35).
- [185] Kirsten K. Knudsen, Lukas Lindroos, Wouter H.T. Vlemmings, John E. Conway, and Ivan Marti-Vidal. "Stacking of SKA data: comparing uv-plane and image-plane stacking." In: *PoS AASKA14* (2015). doi: [10.22323/1.215.0168](https://doi.org/10.22323/1.215.0168).
- [186] M. Massardi et al. "The Additional Representative Images for Legacy (ARI-L) Project for the ALMA Science Archive." In: *PASP* 133.1026, 085001 (2021). doi: [10.1088/1538-3873/ac159c](https://doi.org/10.1088/1538-3873/ac159c).

- [187] J. P. McMullin, B. Waters, D. Schiebel, W. Young, and K. Golap. "CASA Architecture and Applications." In: *Astronomical Data Analysis Software and Systems XVI*. Vol. 376. ASP Conf. Ser. 2007.
- [188] The CASA Team et al. "CASA, the Common Astronomy Software Applications for Radio Astronomy." In: *Publications of the Astronomical Society of the Pacific* 134.1041 (2022). doi: [10.1088/1538-3873/ac9642](https://doi.org/10.1088/1538-3873/ac9642).
- [189] A. J. Kemball and M. H. Wieringa. *Measurement Set definition version 2.0*. NRAO report, January 21, 2000, 52 pages. 2000.
- [190] Ryley Hill et al. "An optimal ALMA image of the Hubble Ultra Deep Field in the era of JWST: obscured star formation and the cosmic far-infrared background." In: *MNRAS* 528.3 (2024). doi: [10.1093/mnras/stae346](https://doi.org/10.1093/mnras/stae346).
- [191] M Hagimoto et al. "Bright extragalactic ALMA redshift survey (BEARS) III: detailed study of emission lines from 71 Herschel targets." In: *Monthly Notices of the Royal Astronomical Society* 521.4 (2023). doi: [10.1093/mnras/stad784](https://doi.org/10.1093/mnras/stad784).
- [192] Lukas Neumann et al. "Spectral stacking of radio-interferometric data." In: *A&A* 675, A104 (2023). doi: [10.1051/0004-6361/202346129](https://doi.org/10.1051/0004-6361/202346129).
- [193] S. Perkins et al. *dask-ms: Dask-powered Measurement Sets*. 2022.
- [194] S. Hoyer and J. Hamman. "xarray: N-D labeled arrays and datasets in Python." In: vol. 5. 2017. doi: [10.5334/jors.148](https://doi.org/10.5334/jors.148).
- [195] M. Rocklin. "Dask: Parallel computation with blocked algorithms and task scheduling." In: *Proceedings of the 14th Python in Science Conference*. 2015. doi: [10.25080/Majora-7b98e3ed-013](https://doi.org/10.25080/Majora-7b98e3ed-013).
- [196] J. Wenzel, J. Rhinelanders, and D. Moldovan. *pybind11 – Seamless operability between C++11 and Python*. 2017.
- [197] OpenMP Architecture Review Board. *OpenMP Application Programming Interface, Version 5.2*. 2021.
- [198] NVIDIA Corporation. *CUDA Toolkit*. 2024.
- [199] Marti-Vidal, I., Vlemmings, W. H. T., Muller, S., and Casey, S. "UVMULTIFIT: A versatile tool for fitting astronomical radio interferometric data." In: *A&A* 563 (2014). doi: [10.1051/0004-6361/201322633](https://doi.org/10.1051/0004-6361/201322633).
- [200] Adam Ginsburg et al. "astroquery: An Astronomical Web-querying Package in Python." In: *AJ* 157.3 (2019). doi: [10.3847/1538-3881/aafc33](https://doi.org/10.3847/1538-3881/aafc33).
- [201] R. S. Bussmann et al. "HerMES: ALMA IMAGING OF HERSCHEL-SELECTED DUSTY STAR-FORMING GALAXIES*." In: *ApJ* 812.1 (2015). doi: [10.1088/0004-637X/812/1/43](https://doi.org/10.1088/0004-637X/812/1/43).

- [202] A. Amvrosiadis et al. “ALMA observations of lensed Herschel sources: testing the dark matter halo paradigm.” In: *MNRAS* 475.4 (2018). doi: [10.1093/mnras/sty138](https://doi.org/10.1093/mnras/sty138).
- [203] A. Enia et al. “The Herschel-ATLAS: magnifications and physical sizes of 500- μ m-selected strongly lensed galaxies.” In: *MNRAS* 475.3 (2018). doi: [10.1093/mnras/sty021](https://doi.org/10.1093/mnras/sty021).
- [204] S. Dye et al. “Modelling high-resolution ALMA observations of strongly lensed highly star-forming galaxies detected by Herschel.” In: *MNRAS* 476.4 (2018). doi: [10.1093/mnras/sty513](https://doi.org/10.1093/mnras/sty513).
- [205] S. Dye et al. “A high-resolution investigation of the multiphase ISM in a galaxy during the first two billion years.” In: *MNRAS* 510.3 (2022). doi: [10.1093/mnras/stab3569](https://doi.org/10.1093/mnras/stab3569).
- [206] Jacob Maresca et al. “Modelling high-resolution ALMA observations of strongly lensed dusty star-forming galaxies detected by Herschel.” In: *MNRAS* 512.2 (2022). doi: [10.1093/mnras/stac585](https://doi.org/10.1093/mnras/stac585).
- [207] Luke T. Maud, Andrés F. Pérez-Sánchez, Yoshiharu Asaki, et al. “Updates to ALMA Site Properties: using the ESO-Allegro Phase RMS database – ALMA Memo 624.” In: *arXiv e-prints*, arXiv:2304.08318 (2023).
- [208] Dominik A. Riechers et al. “A dust-obscured massive maximum-starburst galaxy at a redshift of 6.34.” In: *Nature* 496 (2013). doi: [10.1038/nature12050](https://doi.org/10.1038/nature12050).
- [209] Nick Indriolo et al. “Constraints on the Cosmic-Ray Ionization Rate in the $z \sim 2.3$ Lensed Galaxies SMM J2135-0102 and SDP 17b from Observations of OH⁺ and H₂O⁺.” In: *ApJ* 865.2, 127 (2018). doi: [10.3847/1538-4357/aad7b3](https://doi.org/10.3847/1538-4357/aad7b3).
- [210] Dominik A. Riechers et al. “The GADOT Galaxy Survey: Dense Gas and Feedback in Herschel-selected Starburst Galaxies at Redshifts 2 to 6.” In: *ApJ* 913 (2021).
- [211] Kirsty M. Butler et al. “Resolved Neutral Outflow from a Lensed Dusty Star-forming Galaxy at $z = 2.09$.” In: *ApJ* 919 (2021).
- [212] F. L. Schöier, F. F. S. van der Tak, et al. “An atomic and molecular database for analysis of submillimetre line observations.” In: *A&A* 432.1 (2005). doi: [10.1051/0004-6361:20041729](https://doi.org/10.1051/0004-6361:20041729).
- [213] Daniel Foreman-Mackey, David W. Hogg, et al. “emcee: The MCMC Hammer.” In: *PASP* 125.925 (2013). doi: [10.1086/670067](https://doi.org/10.1086/670067).

- [214] Padelì P. Papadopoulos and E. R. Seaquist. "The State of the Molecular Gas in a Luminous Starburst/Seyfert 2 Galaxy: NGC 1068 Revisited." In: *ApJ* 516.1 (1999). doi: [10.1086/307090](https://doi.org/10.1086/307090).
- [215] John Carpenter, Crystal Brogan, Daisuke Iono, and Tony Mroczkowski. "The ALMA2030 Wideband Sensitivity Upgrade." In: *arXiv e-prints*, arXiv:2211.00195 (2022).
- [216] G. J. Stacey et al. "A 158 μm [C II] Line Survey of Galaxies at $z \sim 1-2$: An Indicator of Star Formation in the Early Universe." In: *ApJ* 724.2 (2010). doi: [10.1088/0004-637X/724/2/957](https://doi.org/10.1088/0004-637X/724/2/957).
- [217] Ilse De Looze et al. "The applicability of far-infrared fine-structure lines as star formation rate tracers over wide ranges of metallicities and galaxy types." In: *A&A* 568, A62 (2014). doi: [10.1051/0004-6361/201322489](https://doi.org/10.1051/0004-6361/201322489).
- [218] Zhi-Yu Zhang et al. "Far-infrared Herschel SPIRE spectroscopy of lensed starbursts reveals physical conditions of ionized gas." In: *MNRAS* 481.1 (2018). doi: [10.1093/mnras/sty2082](https://doi.org/10.1093/mnras/sty2082).
- [219] Yu Gao and Philip M. Solomon. "The Star Formation Rate and Dense Molecular Gas in Galaxies." In: *ApJ* 606.1 (2004). doi: [10.1086/382999](https://doi.org/10.1086/382999).
- [220] I. Oteo et al. "High Dense Gas Fraction in Intensely Star-forming Dusty Galaxies." In: *ApJ* 850.2, 170 (2017). doi: [10.3847/1538-4357/aa8ee3](https://doi.org/10.3847/1538-4357/aa8ee3).
- [221] E. González-Alfonso et al. "Excited OH^+ , H_2O^+ , and H_3O^+ in NGC 4418 and Arp 220." In: *A&A* 550, A25 (2013). doi: [10.1051/0004-6361/201220466](https://doi.org/10.1051/0004-6361/201220466).
- [222] Chentao Yang et al. "SUNRISE: The rich molecular inventory of high-redshift dusty galaxies revealed by broadband spectral line surveys." In: *A&A* 680, A95 (2023). doi: [10.1051/0004-6361/202347610](https://doi.org/10.1051/0004-6361/202347610).
- [223] J. Martín-Pintado, R. Bachiller, and A. Fuente. "SiO emission as a tracer of shocked gas in molecular outflows." In: *A&A* 254 (1992).
- [224] E. Falgarone et al. "Large turbulent reservoirs of cold molecular gas around high-redshift starburst galaxies." In: *Nature* 548.7668 (2017). doi: [10.1038/nature23298](https://doi.org/10.1038/nature23298).
- [225] A. Fuente, J. Martín-Pintado, J. Cernicharo, and R. Bachiller. "A chemical study of the photodissociation region NGC 7023." In: *A&A* 276 (1993).
- [226] David S. Meier and Jean L. Turner. "Spatially Resolved Chemistry in Nearby Galaxies. I. The Center of IC 342." In: *ApJ* 618.1 (2005). doi: [10.1086/426499](https://doi.org/10.1086/426499).

- [227] E. Sturm et al. “Massive Molecular Outflows and Negative Feedback in ULIRGs Observed by Herschel-PACS.” In: *ApJ* 733.1, L16 (2011). doi: [10.1088/2041-8205/733/1/L16](https://doi.org/10.1088/2041-8205/733/1/L16).
- [228] N. Scoville and L. Murchikova. “Submillimeter Recombination Lines in Dust-obscured Starbursts and Active Galactic Nuclei.” In: *ApJ* 779.1, 75 (2013). doi: [10.1088/0004-637X/779/1/75](https://doi.org/10.1088/0004-637X/779/1/75).
- [229] C. Schreiber et al. “Near infrared spectroscopy and star-formation histories of $3 \leq z \leq 4$ quiescent galaxies.” In: *A&A* 618, A85 (2018). doi: [10.1051/0004-6361/201833070](https://doi.org/10.1051/0004-6361/201833070).
- [230] O. Le Fèvre et al. “The ALPINE-ALMA [CII] survey. Survey strategy, observations, and sample properties of 118 star-forming galaxies at $4 < z < 6$.” In: *A&A* 643, A1 (2020). doi: [10.1051/0004-6361/201936965](https://doi.org/10.1051/0004-6361/201936965).
- [231] M. Béthermin et al. “The ALPINE-ALMA [CII] survey: Data processing, catalogs, and statistical source properties.” In: *A&A* 643, A2 (2020). doi: [10.1051/0004-6361/202037649](https://doi.org/10.1051/0004-6361/202037649).
- [232] Y. Fudamoto et al. “The ALPINE-ALMA [CII] survey. Dust attenuation properties and obscured star formation at $z \sim 4.4$ -5.8.” In: *A&A* 643, A4 (2020). doi: [10.1051/0004-6361/202038163](https://doi.org/10.1051/0004-6361/202038163).
- [233] R. J. Bouwens et al. “Reionization Era Bright Emission Line Survey: Selection and Characterization of Luminous Interstellar Medium Reservoirs in the $z > 6.5$ Universe.” In: *ApJ* 931.2, 160 (2022). doi: [10.3847/1538-4357/ac5a4a](https://doi.org/10.3847/1538-4357/ac5a4a).
- [234] Apurba Bera, Nissim Kanekar, Jayaram N. Chengalur, and Jasjeet S. Bagla. “Atomic Hydrogen in Star-forming Galaxies at Intermediate Redshifts.” In: *ApJ* 882.1, L7 (2019). doi: [10.3847/2041-8213/ab3656](https://doi.org/10.3847/2041-8213/ab3656).
- [235] Aditya Chowdhury, Nissim Kanekar, Jayaram N. Chengalur, Shiv Sethi, and K. S. Dwarakanath. “H I 21-centimetre emission from an ensemble of galaxies at an average redshift of one.” In: *Nature* 586.7829 (2020). doi: [10.1038/s41586-020-2794-7](https://doi.org/10.1038/s41586-020-2794-7).
- [236] Francesco Sinigaglia et al. “MIGHTEE-HI: Evolution of HI Scaling Relations of Star-forming Galaxies at $z < 0.5$.” In: *ApJ* 935.1, L13 (2022). doi: [10.3847/2041-8213/ac85ae](https://doi.org/10.3847/2041-8213/ac85ae).
- [237] J. R. Pardo, J. Cernicharo, and E. Serabyn. “Atmospheric transmission at microwaves (ATM): an improved model for millimeter/submillimeter applications.” In: *IEEE Transactions on Antennas and Propagation* 49.12 (2001). doi: [10.1109/8.982447](https://doi.org/10.1109/8.982447).
- [238] Astropy Collaboration. “Astropy: A community Python package for astronomy.” In: *A&A* 558 (2013).

- [239] Astropy Collaboration. "The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package." In: *AJ* 156 (2018).
- [240] F. Stoehr, A. Manning, C. Moins, et al. "The ALMA Science Archive." In: *The Messenger* 167 (2017). DOI: [10.18727/0722-6691/5000](https://doi.org/10.18727/0722-6691/5000).
- [241] F. Ochsenbein, P. Bauer, and J. Marcout. "The VizieR database of astronomical catalogues." In: *A&AS* 143 (2000). DOI: [10.1051/aas:2000169](https://doi.org/10.1051/aas:2000169).
- [242] A. J. Remijan, A. Markwick-Kemper, and ALMA Working Group on Spectral Line Frequencies. "Splatalogue: Database for Astronomical Spectroscopy." In: *American Astronomical Society Meeting Abstracts*. Vol. 211. 2007, 132.11.
- [243] Adam Ginsburg, Eric Koch, Thomas Robitaille, Chris Beaumont, et al. *radio-astro-tools/spectral-cube: Release v0.4.5*. Version v0.4.5. 2019. DOI: [10.5281/zenodo.3558614](https://doi.org/10.5281/zenodo.3558614).
- [244] Eric Koch, Adam Ginsburg, Adam Leroy, Erik Rosolowsky, and Thomas Robitaille. *radio-beam: Tools for Beam IO and Manipulation*. 2018.
- [245] Bin Liu, N. Chartab, H. Nayyeri, and ... "Massive Molecular Gas Reservoir in a Luminous Submillimeter Galaxy during Cosmic Noon." In: *ApJ* 929.1 (2022). DOI: [10.3847/1538-4357/ac5745](https://doi.org/10.3847/1538-4357/ac5745).
- [246] D. Ismail et al. "z-GAL: A NOEMA spectroscopic redshift survey of bright Herschel galaxies. II. Dust properties." In: *A&A* 678, A27 (2023). DOI: [10.1051/0004-6361/202346804](https://doi.org/10.1051/0004-6361/202346804).
- [247] R. J. Ivison et al. "THE SPACE DENSITY OF LUMINOUS DUSTY STAR-FORMING GALAXIES AT $z > 4$: SCUBA-2 AND LABOCA IMAGING OF ULTRARED GALAXIES FROM HERSCHEL-ATLAS." In: *ApJ* 832.1 (2016). DOI: [10.3847/0004-637X/832/1/78](https://doi.org/10.3847/0004-637X/832/1/78).
- [248] S. Duivenvoorden et al. "Red, redder, reddest: SCUBA-2 imaging of colour-selected Herschel sources." In: *MNRAS* 477.1 (2018). DOI: [10.1093/mnras/sty691](https://doi.org/10.1093/mnras/sty691).
- [249] Yang, C. et al. "CO, H₂O, H₂O⁺ line and dust emission in a $z = 3.63$ strongly lensed starburst merger at sub-kiloparsec scales." In: *A&A* 624 (2019). DOI: [10.1051/0004-6361/201833876](https://doi.org/10.1051/0004-6361/201833876).
- [250] Megan B. Gralla et al. "Atacama Cosmology Telescope: Dusty Star-forming Galaxies and Active Galactic Nuclei in the Equatorial Survey." In: *ApJ* 893.2 (2020). DOI: [10.3847/1538-4357/ab7915](https://doi.org/10.3847/1538-4357/ab7915).

- [251] G J Bendo et al. "The dust emissivity index β in infrared-bright galaxies at $1.5 < z < 4.2$." In: *MNRAS* 540.2 (2025). doi: [10.1093/mnras/staf816](https://doi.org/10.1093/mnras/staf816).
- [252] T. Su et al. "On the redshift distribution and physical properties of ACT-selected DSFGs." In: *MNRAS* 464.1 (2017). doi: [10.1093/mnras/stw2334](https://doi.org/10.1093/mnras/stw2334).

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The majority of the plots reported in this thesis have been generated using the Python package `matplotlib`.

This work is licensed under a Creative Commons Attribution 4.0 International License.

Final Version as of September 23, 2026 (`classicthesis`).