



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN ASTROFISICA

CICLO XXXVIII

Settore Concorsuale: 02/C1 - ASTRONOMIA, ASTROFISICA, FISICA DELLA TERRA E DEI PIANETI
Settore Scientifico Disciplinare: FIS/05 - ASTRONOMIA E ASTROFISICA

**Detection and multi-wavelength
characterisation of galaxy clusters, groups
and protocluster cores up to $z = 3.7$**

TESI DI DOTTORATO

Presentata da: **Greta Toni**

Coordinatore Dottorato
Andrea Miglio

Supervisore
Lauro Moscardini

Co-supervisore
Matteo Maturi

Esame Finale Anno 2026

*Humans, like galaxies, tend to cluster
together in groups and families.*

*Humans, like galaxies, evolve and
transform throughout their lifetimes,
shaped by the interactions and
environments they encounter.*

Change is never a lonely journey.

To my family

Abstract

Galaxies in the Universe are distributed across an interconnected structure known as the cosmic web, formed through gravitational collapse, accretion, and merging. Within this framework, galaxy clusters and groups emerge as key environments where galaxies evolve under the combined influence of internal and environmental processes. This Thesis focuses on the detection, multi-wavelength characterisation, and scientific analysis of these systems, with particular emphasis on deep, small-area surveys capable of probing the high-redshift Universe. In this work, we employed the Adaptive Matched Identifier of Clustered Objects (AMICO) algorithm, here enhanced for application to deep, multi-band photometric data. Through careful data preparation, noise modelling, and masking strategies, we created two robust catalogues of galaxy clusters and groups up to high redshift. The AMICO-COSMOS catalogue comprises 1269 detections up to $z = 2$, complemented by X-ray characterisation that enabled the calibration of scaling relations between optical and X-ray mass proxies. The highest-redshift application of AMICO to date produced the COSMOS-Web catalogue—the largest deep sample of galaxy groups and protocluster cores produced to date, extending to $z = 3.7$. Among these, over 300 new high-redshift systems were identified, including several extended protocluster regions such as the AmicOne candidate, the largest among them. Leveraging these catalogues, this Thesis explores the evolution of the quiescent galaxy fraction and the red sequence over ~ 12 Gyr, revealing evidence of mature galaxy populations already at early cosmic times. The analysis of Brightest Group Galaxies (BGGs) uncovered systems undergoing transformation and rapid molecular gas depletion, indicative of quenching processes in action. Together, these results lay the groundwork for future studies of structure formation and galaxy evolution, particularly in view of forthcoming data from facilities such as Euclid, JWST, LSST, and Roman.

Contents

<i>Abstract</i>	i
<i>List of Figures</i>	vi
<i>List of Tables</i>	ix
Introduction	1
1 Galaxy clusters and groups	5
1.1 Clusters or groups of galaxies?	6
1.2 Galaxy population in the optical and NIR	8
1.2.1 Detection methods in optical and NIR surveys	12
1.2.2 Peculiar galaxies and seeded detection	17
1.3 The hot intracluster/group medium	18
1.3.1 X-ray emission of the ICM/IGM	18
1.3.2 Sunyaev-Zeldovich effect in the microwaves	21
1.4 Gravitational lensing	23
1.5 High-redshift groups and protoclusters	24
2 The AMICO algorithm	29
2.1 Mathematical formalism	30
2.2 Input data and model requirements	32
2.3 Detection workflow	36
2.4 Previous applications of AMICO	41
2.5 Selection for the Euclid mission	42
2.6 Novelty of this work and new features	44

I	Detection	51
3	The AMICO-COSMOS catalogue	53
3.1	COSMOS field data	53
3.2	Applying AMICO to the COSMOS data	57
3.3	The cluster and group catalogues	60
3.3.1	Results of the search	60
3.3.2	Matching the r , Y , and H catalogues	61
3.3.3	Examples of detections	62
3.3.4	Spectroscopic counterparts	67
3.4	Comparison with X-ray group catalogues	67
3.4.1	Quality flags	68
3.4.2	The r , Y , and H analyses when compared with X-rays	68
3.4.3	Mass-observable scaling relations	69
3.5	X-ray counterparts for new detections	71
3.5.1	Redshift dependence of the scaling relations	74
3.5.2	New detections without X-ray counterparts	74
3.6	Conclusions	79
4	The COSMOS-Web catalogue	81
4.1	COSMOS-Web data	82
4.1.1	Photometric redshifts	82
4.1.2	Data cleaning and visibility mask	83
4.2	Applying AMICO to the COSMOS-Web data	85
4.3	Catalogue of galaxy groups	86
4.3.1	Purity and completeness of the sample with SinFoniA	91
4.3.2	Spectroscopic counterparts	98
4.3.3	Detection flags	99
4.4	Candidates at $z \geq 2$	99
4.4.1	COSMOS protocluster compilation	100
4.4.2	Our candidates in the COSMOS-Web field	100
4.5	Conclusions	102
II	Characterisation	107
5	Quiescent population build-up	109
5.1	Catalogue of galaxies and galaxy groups	109
5.2	Galaxy classification with machine learning	110

5.2.1	Machine learning algorithms	111
5.2.2	Training and testing set	111
5.2.3	Feature engineering	113
5.2.4	Classification performance and results	116
5.3	Quiescent galaxy fractions	118
5.3.1	Pure membership and secondary association subtraction	119
5.3.2	Cylinder background subtraction	120
5.3.3	X-ray luminosity of groups and environmental effects	122
5.4	Red sequence	124
5.4.1	Observed red sequence with rest-frame matching	126
5.4.2	Results and comparison with rest-frame red sequence	129
5.5	Conclusions	133
6	Brightest Group Galaxies	137
6.1	Motivations and goals	140
6.2	Star-forming brightest group galaxies	141
6.2.1	AMICO-COSMOS BGG catalogue	141
6.2.2	Selection of targets	142
6.2.3	Star formation rates	143
6.3	IRAM observations and data reduction	151
6.3.1	Molecular gas reservoirs	153
6.4	Discussion	153
6.4.1	Molecular gas consumption and environment	157
6.4.2	Comparison with the literature	158
6.5	Conclusions	159
III	Epilogue	161
7	Conclusions	163
8	Future Prospects	171
IV	Appendix	177
A	AMICO run based on spectroscopic redshifts	179
B	BGG selection: a new hybrid method	183

C Clusters as first seen by Euclid	193
D Details on the ML training & cylinder method	201
Bibliography	211

List of Figures

1.1	The emergence of a massive cluster in a simulated dark matter density field . . .	6
1.2	Examples of how clusters are identified by four detection algorithms	16
1.3	Galaxy cluster in COSMOS-Web as seen by JWST, XMM, and Chandra	19
1.4	Schematic of the formation and evolution of galaxy clusters	25
2.1	Example of amplitude map cleaning	39
2.2	Purity–completeness curves for Euclid CFC cluster finders.	44
2.3	Survey footprint of the COSMOS field	46
2.4	Noise model for different datasets and cleaning methods	48
2.5	Example of constructing a composite visibility mask.	49
3.1	Distribution of the relative magnitude scatter in COSMOS2020 and 2015.	55
3.2	Distribution of selected galaxies as a function of redshift	56
3.3	Effect of noise regularisation on the filter constants	58
3.4	Number of AMICO-COSMOS detections and amplitude as a function of redshift	60
3.5	Number of matched/unmatched AMICO-COSMOS detections and amplitude as a function of redshift	62
3.6	Matched and unmatched detections for r vs H	63
3.7	Four examples of detections at different z as identified in the r -band run	64
3.8	(Super-)group/cluster with substructures at $z = 0.22$	65
3.9	Pair of detections aligned along the line of sight	66
3.10	Relation between X-ray mass and richness/amplitude for the X-ray matched detections	70
3.11	Distribution of the redshifts of new detections analysed in X-rays	72
3.12	Relation between X-ray mass and richness/amplitude for the new detections . . .	73
3.13	Relation between X-ray mass and richness/amplitude for the full sample	75
3.14	X-ray-luminosity–richness relation with and without point-source removal	77
3.15	Candidate cluster detected at $z = 0.68$ without significant X-ray emission	78

4.1	Magnitude and redshift distribution of the cleaned COSMOS-Web galaxy catalogue	84
4.2	Evolution of the galaxy magnitude and m_* in redshift	86
4.3	Redshift and cumulative signal-to-noise distribution for the COSMOS-Web group catalogue	87
4.4	Intrinsic richness, λ_* , <i>vs</i> redshift	87
4.5	Four examples of detections in the group catalogue at different redshifts	88
4.6	Cumulative distribution function of S/N_{nocl}	92
4.7	Distributions of spatial and redshift scatter for the mock matching	93
4.8	Relative scatter of the three observables (λ , λ_* and A) in the mock matching	93
4.9	Purity of the group sample evaluated against the mock catalogue	94
4.10	Completeness of the group sample evaluated against the mock catalogue	96
4.11	Relation between minimum S/N_{nocl} and purity of the sample	97
4.12	Two examples of high- z detections not known in the literature	101
4.13	AmicOne protocluster candidate, consisting of fourteen cores	105
5.1	Feature importance for an XGB training	112
5.2	Calibration curves of the bias between predicted probability and fraction of positives	114
5.3	Results of our probabilistic classification of COSMOS-Web galaxies	115
5.4	Purity-weighted quiescent fraction <i>vs.</i> redshift	120
5.5	Purity-weighted quiescent fraction with the cylinder method	123
5.6	Quiescent fraction difference between X-ray bright and faint groups	125
5.7	Filament signal probability density function for X-ray bright and faint groups	126
5.8	The sampling in the COSMOS-Web filters of the observed 4000 Å	128
5.9	Fraction of RS galaxies over the total and of groups with RS <i>vs</i> z	130
5.10	Observed-frame CMD of an overdensity with quiescent galaxies at $z = 3.4$	131
5.11	The main red-sequence parameters <i>vs</i> redshift	132
6.1	JWST NIRCcam colour-composite images of five BGGs in COSMOS	138
6.2	Colour composites of our three targets	145
6.3	Location of our three targets in the NUVrJ diagram	146
6.4	SEDs and respective LePhare best fitting modelling for our three targets	149
6.5	Baseline-subtracted CO(1→0) spectrum of BGG 1000+0207	152
6.6	Molecular gas properties: comparison with the literature	156
7.1	Concluding image of this Thesis	169
B.1	Selection of BGGs in the COSMOS-Web galaxy group catalogue	185
B.2	Comparison of BGG selection methods	186
B.3	Classified BGGs in the λ_* - z diagram	189

B.4	Size–mass relation of BGGs	190
B.5	Redshift evolution of the size–mass relation slope and intrinsic scatter	192
C.1	VIS and PZWav and AMICO-map images of a Q1 detection	195
C.2	Redshift and separation scatter of the AMICO–PZWav comparison	196
C.3	VIS/NISP combined colour postage stamps of five clusters	197
C.4	Zoomed-in image of a strong lens detected in Q1	198
C.5	Correlation of the Euclid richness against other mass proxies	199
D.1	Different classification thresholds in the sSFR– z plane	204
D.2	Classification variation test in bins of z and stellar mass	206
D.3	Distribution of RMS radii and trend with richness	208
D.4	Purity-weighted quiescent fractions using fixed radii or RMS radius	209

List of Tables

1.1	Some of the techniques used in the literature for cluster and group detection. . .	28
3.2	Mass-proxy scaling relation parameters and pivot values	69
3.3	Mass-proxy scaling relation parameters and pivot values including z -dependence based on the full sample	76
4.1	Compilation of protoclusters known in the literature for the COSMOS field . . .	104
4.2	Descriptions of columns for the group and member catalogue	106
5.1	Performance of the classification methods	116
5.2	Colour and magnitude choices by redshift interval for RS selection	127
6.1	Properties of our selected targets: three star-forming BGGs in COSMOS	144
6.2	SFRs and relative sSFR of our target BGGs.	147
6.3	Results of our CO observations	154
D.1	Galaxy counts with different classification thresholds	205

Introduction

Galaxies are distributed across a vast, interconnected structure of nodes, filaments, walls and voids, emerging as a consequence of gravitational collapse, accretion and merging (e.g. [Bond et al., 1996](#); [Cautun et al., 2014](#); [Libeskind et al., 2018](#)). Through this hierarchical assembly, the cosmic web grows increasingly complex, with filaments and walls funnelling matter into nodes, where the most massive clusters of galaxies form. As the culmination of cosmic structure formation, galaxy clusters have long been recognised as highly sensitive and powerful cosmological probes (see e.g. the reviews by [Allen et al., 2011](#); [Clerc & Finoguenov, 2023](#)). However, the bulk of galaxies in the Universe do not reside in these rare, dense peaks of the large-scale matter distribution, but in more numerous, smaller associations of up to a few tens of galaxies, known as galaxy groups. As the most common galaxy environment, these systems play a key role in hierarchical assembly and in driving the transformation of galaxies, which profoundly shapes their evolution (e.g. [Tully, 1987](#); [Eke et al., 2004](#); [Lietzen et al., 2012](#); [Paul et al., 2017](#); [Lovisari et al., 2021](#)).

Over the past decades, a large effort has been dedicated to the analysis of wide-area surveys, enabling the creation of extensive samples of galaxy clusters, specifically optimised for cosmological analyses (e.g. [Koester et al., 2007](#); [Rykoff et al., 2014](#); [Bleem et al., 2015](#); [Planck Collaboration et al., 2016](#); [Hilton et al., 2021](#); [Bulbul et al., 2024](#); [Maturi et al., 2025](#)). Yet, wide-field surveys typically do not reach the depth necessary to unveil the fainter and more complex phases of early structure formation. This leaves room for complementary analyses capable of probing both smaller systems, like galaxy groups, and earlier stages of structure formation. This is where deep-field surveys come into play: by investing in depth instead of large spatial coverage, deep fields offer a unique opportunity to study these processes in detail, capturing the emergence and evolution of structures across cosmic time, and the evolution of galaxies within them. Despite their key role in structure formation and galaxy evolution, galaxy groups have remained observationally challenging systems to detect and characterise. Their relatively low masses and shallow potential wells make them less prominent than massive clusters, particularly in X-ray and Sunyaev–Zeldovich surveys. However, recent advances in deep, multi-wavelength imaging and spectroscopy have greatly improved our ability to identify and study these low-mass, low-richness systems. The current generation of modern deep surveys now enables the construction of well-characterised

samples of galaxy groups and cluster progenitors, extending the detection to high redshifts where the structures we observe today were first taking shape.

Moving in this direction, this Thesis intends to bridge the gap between large statistical samples and in-depth environmental galaxy evolution and structure formation studies. This is achieved by applying the Adaptive Matched Identifier of Clustered Objects (AMICO; Bellagamba et al., 2018; Maturi et al., 2019)—a robust detection algorithm widely tested on wide-field survey data and simulations—to the COSMOS field (Scoville et al., 2007), a 2-deg² deep, multi-wavelength survey renowned for its exceptional data richness. This innovative application of the algorithm to a deep field resulted in two well-characterised catalogues of clusters, groups, and protocluster cores, enabling a detailed investigation of structure assembly and galaxy evolution over nearly 12 Gyr of cosmic history. This work not only advances methodological approaches for detecting low-mass and high-redshift systems through the enhancement and optimisation of the detection algorithm, but also provides valuable public resources, offering a foundation for future studies—spanning from the environmental processes driving galaxy transformation to the assembly of structures and the high-redshift Universe.

This Thesis focuses on the detection of galaxy clusters, groups, and protocluster cores, and on the creation of catalogues that serve as public resources for the scientific community. It also quantifies and investigates how these resources can be used to address open questions on the evolution of the overall galaxy population and on the properties of peculiar galaxies within these systems, thereby contributing to their characterisation. Following this two-stage approach, the body of the Thesis is organised into two main parts. **Part I**, titled *Detection*, will be focused on the primary goal of building robust and well-characterised catalogues of galaxy clusters, groups, and protocluster cores in the COSMOS and COSMOS-Web field with AMICO. **Part II**, titled *Characterisation*, will be dedicated to the first analyses of these catalogues, aiming to provide new insights into the evolution and transformation of galaxies in dense environments.

Before delving into this, **Chapter 1** will serve as a guide through the essential concepts and background underlying this Thesis’s narrative. After highlighting the growing recognition that galaxy groups deserve specific attention as distinct environments, this Chapter will summarise the multi-wavelength properties of galaxy clusters and groups and how these have been leveraged over the decades to identify them. A Section dedicated to the high-redshift Universe will conclude the Chapter, providing the foundation to interpret the results of this work at $z > 1.5 - 2$.

With the introductory background in place, **Chapter 2** will present AMICO, the detection algorithm employed throughout this work. It will outline the mathematical formalism, describe the detection workflow, and discuss the crucial role of input data and models, which underpin the method’s flexibility and performance. These two Chapters lay the groundwork for the analyses that follow, linking the motivation behind this Thesis to the methodology adopted for the detection and characterisation of these systems, described in **Part I** and **Part II**, respectively.

Part I: Detection

- **Chapter 3** is based on Toni et al. (2024). It describes the first cluster and group search performed with the AMICO algorithm in the COSMOS 2-deg² deep field, using the recent COSMOS2020 photometric galaxy catalogue and extending up to $z = 2$.
- **Chapter 4** is based on Toni et al. (2025a). It presents the largest deep catalogue of groups and protocluster cores to date, which leverages JWST's widest continuous observation (the COSMOS-Web survey) to push galaxy system detection up to $z = 3.7$ and down to more than 2 magnitudes fainter.

Part II: Characterisation

- **Chapter 5** is based on Toni et al. (2025b). It describes the analysis of the redshift evolution and group richness dependence of the quiescent galaxy fraction and red-sequence parameters in COSMOS galaxy groups (presented in **Part I**), across ~ 12 Gyrs of cosmic time.
- **Chapter 6** is based on Toni et al. in prep., a project based on our IRAM 30m telescope observations targeting CO line emission in three star-forming central brightest group galaxies found in COSMOS groups and caught in a rare stage of their lifetimes, showing clear signs of merger and interactions.

As an epilogue to this Thesis, **Chapter 7** and **Chapter 8** will respectively summarise the main results and broader impact of this work and attempt to sketch a non-exhaustive list of future (and ongoing) projects that this Thesis work has made possible.¹

¹In the work presented in this Thesis, we assume a standard concordance flat Λ CDM cosmology with matter density $\Omega_m = 0.3$, dark energy density $\Omega_\Lambda = 0.7$, and dimensionless Hubble constant $h = H_0 / (100 \text{ km/s/Mpc}) = 0.7$. Magnitudes are expressed in the AB system unless otherwise specified (Oke & Gunn, 1983).

1

Galaxy clusters and groups

Matter in the Universe, including galaxies, is distributed in a complex and interconnected structure known as the *cosmic web*. In this network, the densest regions (*nodes*) are connected by elongated bridges called *filaments* and flatter sheets called *walls*. Most of the galaxies reside in these structures, while the rest of space, vast and relatively empty, is occupied by the so-called *cosmic voids*. The cosmic web forms as a result of gravitational collapse, starting in the early Universe with initial small density fluctuations that then grow over time. As the Universe evolves, less extended and less massive structures form first and then merge and accrete to form more massive systems. In this way, the cosmic web grows and gets more complex, with filaments and walls establishing connectivity and enabling the flow and accretion of matter into nodes, where massive clusters of galaxies form, as visible in Figure 1.1 (see e.g. [Bond et al., 1996](#); [Kravtsov & Borgani, 2012](#); [Cautun et al., 2014](#); [Libeskind et al., 2018](#)). Despite being characterised by a complex composition of baryonic and non-baryonic matter in different forms, galaxy clusters are commonly identified as the most massive structures that have had the chance to fully form until now, and they trace the density peaks of the large-scale matter distribution.

The purpose of this Chapter is to outline the essential concepts and background that support the understanding of the subject addressed in this Thesis. However, before delving into it, attention should first be given to whether distinguishing between galaxy clusters and groups is meaningful in clarifying the focus of this Thesis.

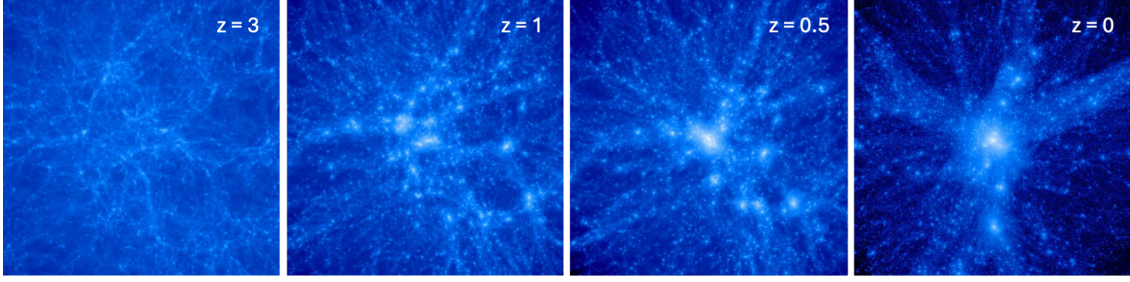


Figure 1.1. The emergence of a massive galaxy cluster in the simulated evolution of the dark matter density field. The forming cluster has a mass of $\sim 1.2 \times 10^{15} M_{\odot}/h$ at $z = 0$. The snapshots (from $z = 3$ to $z = 0$) illustrate the complex nature of the collapse in density peaks, with filamentary accretion, and substructure within the forming cluster-scale halo. Adapted from Kravtsov & Borgani (2012, Fig. 6).

1.1 Clusters or groups of galaxies?

The local environment we live in is a group of galaxies, the Local Group, which hosts the Milky Way together with other ~ 40 main galaxies (e.g. Grebel, 2005), making up a gravitationally bound assembly of a variety of galaxy types, with different matter content and stellar populations. But what exactly prevents it from being classified as a galaxy cluster? Is a galaxy group essentially a scaled-down version of a cluster? Despite there being no universally defined threshold to separate groups and clusters, these are generally distinguished in the literature by using limits, for instance, in mass at around $\sim 10^{14} M_{\odot}$ or in the number of galaxy members at around 50 galaxies (see e.g. Paul et al., 2017; Lovisari et al., 2021). Therefore, galaxy groups represent the transition range in halo masses from galaxies ($M \sim 10^9 - 10^{12} M_{\odot}$) to galaxy clusters ($M \sim 10^{14} - 10^{15} M_{\odot}$). Although the distinctive properties of galaxies and clusters have long been established, galaxy groups have historically received less attention, and for many years the literature lacked extensive studies dedicated to these intermediate structures (Paul et al., 2017). Indeed, recent observational studies and simulations have yielded results showing that galaxy groups exhibit substantial differences with respect to clusters and cannot be directly treated as smaller versions of them (see e.g. Oppenheimer et al., 2021, for a review). Groups are found to differ from clusters, which is reflected in the deviation from cluster scaling relations, with a breakaway point at $\sim 10^{14} M_{\odot}$ (Paul et al., 2017). Unlike massive clusters, which are relatively rare and emerge at the highest-density peaks of the cosmic web, groups are far more numerous and also form inside filaments and walls, tracing a much broader range of environments. Given the shallower potential wells, galaxy groups are also more sensitive to astrophysical phenomena such as galaxy mergers, shock heating, tidal interactions, galactic winds and feedback both from supernovae and active galactic nuclei (e.g. Giodini et al., 2009; McCarthy et al., 2010; Gaspari et al., 2011; Lovisari et al., 2015). As a result, they provide unique laboratories to study environmental effects on galaxy evolution that are not

as easily disentangled when studying only massive clusters.

Galaxy groups represent diverse environments and play an important role as building blocks of more massive systems. Over the past few decades, dedicated studies of groups have enabled their classification into three main categories based on optical and physical properties: loose groups, compact groups, and fossil groups. Loose groups (most typical galaxy groups) are associations of a few to a few tens of galaxies with sizes of the order of a megaparsec and relatively low velocity dispersions (e.g. Eke et al., 2004; Girardi & Giuricin, 2000; Einasto et al., 2003; Coenda et al., 2012; Martínez et al., 2013). They represent the majority of galaxy systems and are important laboratories for studying environmental effects on galaxy evolution. Compact groups, first systematically described by Hickson (1982), are small, compact systems containing a few galaxies in close proximity, with high galaxy number densities. This configuration promotes frequent interactions, mergers, and accelerated galaxy evolution (e.g. Mendes de Oliveira & Hickson, 1991; Coenda et al., 2012; Zandivarez et al., 2024; Tricottet et al., 2025). Fossil groups instead are defined by a dominant central galaxy, at least two magnitudes brighter than the next brightest member, accompanied by an extended X-ray halo indicative of a deep potential well (e.g. Jones et al. 2003, Eigenthaler & Zeilinger 2009, Aguerri & Zarattini 2021, Cossairt et al. 2022, the FOGO project series: e.g. Aguerri et al. 2011, Girardi et al. 2014, Zarattini et al. 2023). These systems are considered the aftermath of past mergers that have depleted the bright galaxy population, leaving a single luminous galaxy embedded in a massive halo. The study of groups and classification in these three categories highlights the diversity of galaxy group environments and their distinct role in shaping galaxy evolution and structure assembly.

Detecting and characterising galaxy groups is therefore essential and complementary to the study of clusters. Galaxy clusters serve as powerful probes for precision cosmology, complementing measurements from the cosmic microwave background (CMB) and large-scale structure surveys (see e.g. Allen et al., 2011; Clerc & Finoguenov, 2023, for reviews about the role of clusters in cosmology). The abundance and spatial distribution of galaxy clusters as a function of mass and redshift provide direct constraints on the growth of cosmic structure, making them sensitive probes of the matter density Ω_m and the amplitude of matter fluctuations σ_8 . Their clustering and weak lensing mass measurements further allow tests of the dark energy equation of state and the sum of neutrino masses. However, groups dominate the halo population by number, hosting the bulk of galaxies in the Universe. Therefore, understanding groups' properties, scaling relations, and abundance across cosmic time not only helps shed light on galaxy evolution processes but also impacts the constraints of cosmological parameters, possibly playing a role in reducing parameter tensions and degeneracies (e.g. Schellenberger & Reiprich, 2017; Sinha et al., 2018; Driver et al., 2022).

In this Thesis, the focus will primarily be on galaxy groups, as these are the dominant systems in the sky region we have targeted (although we have also detected a few small clusters, as

defined by the commonly adopted criteria). The following Sections present the multi-wavelength properties of clusters in general and the detection strategies that leverage them, with attention to groups whenever significant differences between the two are worth highlighting.

1.2 Galaxy population in the optical and near-infrared

Ever since galaxy clusters were first identified by visual inspection of photographic plates, they have been classified in terms of richness. In his pioneering work on galaxy cluster catalogues, [Abell \(1958\)](#) developed a richness criterion based on the number of galaxies brighter than $m_3 + 2$ (with m_3 being the magnitude of the third brightest member) within a characteristic radius later known as the Abell radius, laying the foundation for optical cluster detection and characterisation. In the optical and near-infrared (NIR), the signal from clusters arises mainly from the stellar emission of their member galaxies. Building on Abell’s approach, modern surveys allow for the detection of clusters and groups as overdensities of galaxies, now supported by accurate photometric or spectroscopic redshifts to disentangle projection effects. However, galaxy clusters and groups are not just simple overdensities of galaxies in space; they can be distinguished by characteristic properties of their member populations, compared to the field (a term that refers to the space occupied by isolated galaxies, characterised by low matter density). Cluster galaxies exhibit, for instance, clear differences in morphology, colour, star-formation activity, and stellar mass distribution, reflecting the strong influence that dense environments have on galaxy evolution.

One of the most striking features is the predominance of early-type, quiescent galaxies in clusters, especially in their central regions, in contrast to the star-forming galaxies (mostly spirals) that dominate lower-density environments. This colour and morphology bimodality is generally explained as the result of a strongly intertwined impact of intrinsic properties of galaxies, like stellar mass (*nature*) and factors related to the different environment galaxies experience in their lifetimes (*nurture*) (e.g. [De Lucia et al., 2012](#)). This *nature* vs *nurture* problem, or more specifically, the disentanglement of mass and environmental effect on quenching and transforming galaxies, has been widely studied, but it is still a matter of debate. Starting from the pioneering work of [Peng et al. \(2010\)](#), recent studies (e.g. [Chartab et al., 2020](#); [Taamoli et al., 2024a](#); [Zheng et al., 2025](#)) have shown that the relative efficiency of the two has a dependence on redshift and galaxy type (satellite or central). In fact, the environmental segregation, already known and first quantified by [Dressler \(1980\)](#), illustrates how environmental processes such as ram-pressure stripping, interactions, and strangulation act to quench star formation and transform spirals into lenticular and elliptical systems. As a consequence, dense environments typically host galaxy populations characterised by higher masses, lower star-formation rates (SFRs), and older and redder stellar populations compared to lower-density environments, at least for the more evolved

systems. A robust manifestation of this trend is the so-called red sequence: a tight grouping in colour–magnitude space formed by passively evolving galaxies with old stellar populations mainly assembled at early epochs. The slope and scatter of the red sequence provide insights into the ages and metallicities of cluster galaxies, while its presence (or absence) out to high redshifts can help trace the evolution of stellar populations over time in member galaxies. As a typical feature observed in clusters and groups, the red sequence reflects both the homogeneity of passive galaxies and their assembly and evolutionary history (e.g. [Kodama & Arimoto, 1997](#); [Stott et al., 2009](#)).

While massive clusters represent the densest regions where to study the effects of environmental processing, galaxy groups occupy an intermediate regime that bridges the field and the clusters, and their galaxy populations reveal distinct signatures of this specific environment. Groups tend to host a higher fraction of star-forming galaxies, consistent with a more diverse mixture of morphological types, with respect to clusters. However, environmental processes in groups do not act merely as a weaker version of those in clusters. Instead, the lower velocity dispersions and shallower potential wells of groups make some processing mechanisms even more efficient. For instance, galaxy–galaxy mergers and tidal interactions are enhanced in groups, leading to morphological transformations and the build-up of massive galaxies. Strangulation and starvation, related to the removal of a galaxy’s gas reservoir, are also impacting group haloes, gradually shutting down star formation over gigayear timescales (e.g. [McGee et al., 2009](#); [Peng et al., 2015](#); [Boselli & Gavazzi, 2006](#)). By contrast, processes such as ram-pressure stripping are generally less effective in groups, owing to their less dense intragroup medium. As a result, groups, which are the most common environment for galaxies, play a key role in the so-called ‘pre-processing’, that is, the transformation galaxies undergo before being accreted into clusters. Both simulations and observations have suggested that around half of cluster galaxies are brought in as part of infalling groups, and not directly from the field (e.g. [McGee et al., 2009](#); [Bianconi et al., 2018](#); [Benavides et al., 2020](#); [Lopes et al., 2024](#)). This makes the study of group galaxy populations essential both for understanding the hierarchical growth of structures and for disentangling the relative importance of different environmental mechanisms in shaping galaxy evolution.

To quantify the impact of the environment on cluster and group galaxies and to model them, the galaxy population can be characterised through statistical descriptors of its luminosity and spatial distribution. The luminosity function (LF) in galaxy clusters and groups quantifies the number of galaxies as a function of luminosity (or magnitude). In the literature, it is normally modelled by an analytical form known as the Schechter function ([Schechter, 1976](#)), which in terms of luminosity, L , reads

$$\phi(L) dL = \phi_{\star} \left(\frac{L}{L_{\star}} \right)^{\alpha} \exp \left(-\frac{L}{L_{\star}} \right) \frac{dL}{L_{\star}}, \quad (1.1)$$

and in terms of magnitude, m , can be expressed as

$$\phi(m) dm = 0.4 \ln(10) \phi_* 10^{0.4(\alpha+1)(m_*-m)} \exp(-10^{0.4(m_*-m)}) dm. \quad (1.2)$$

The LF parameters are the characteristic luminosity, L_* (or characteristic magnitude, m_*), the faint-end slope, α , and the normalisation factor, ϕ_* . The bright end of the LF is dominated by massive ellipticals, while the faint end traces the abundance of dwarf galaxies, which is sensitive to both feedback processes and efficiency of galaxy quenching. Variations in the LF, such as a steepening of the faint-end slope or an enhanced bright-end tail, trace therefore the cumulative effects of mergers, tidal interactions, and star-formation quenching processes in general. Field galaxies can also be obviously represented by an LF, but cluster and group LFs generally show an excess of massive bright galaxies (typically central galaxies) and tend to have a steeper faint-end slope, generated by the larger amount of quenched faint low-mass galaxies with respect to the field (e.g. De Propriis et al., 2003; Popesso et al., 2006; De Lucia & Blaizot, 2007; Lan et al., 2016; Negri et al., 2022). The LF can vary with cluster richness, redshift, and dynamical state. Thus, studying the LF parameters also offers a window into the co-evolution of galaxies and their host haloes.

In a complementary way to the luminosity function, the radial density profile describes the number density of galaxies as a function of cluster- or group-centric distance, revealing trends in morphological segregation, central concentration of passive galaxies, and therefore the spatial dependence of environmental processes. Together, LFs and density profiles provide a quantitative framework linking observed distributions of galaxies to the physical mechanisms driving their evolutionary lifetimes. The spatial distribution of galaxies inside a cluster or a group is a fundamental property of the member population, which relates to the interplay between baryonic and non-baryonic matter. Observational and simulation-based studies consistently find that galaxies in clusters/groups are distributed in a nearly universal, centrally concentrated way, with the number density of galaxies decreasing with increasing distance from the centre. The most widely adopted model for the radial galaxy member distribution is a Navarro-Frenk-White (NFW; Navarro et al., 1997) profile. The NFW profile is a widely used analytical model that describes the spherically averaged density distribution of dark matter haloes, $\rho(r)$, derived from N-body cosmological simulations. It provides for a cuspy central region where the density scales with the inverse of the radius and a steeper decline ($\rho \propto r^{-3}$) at large radii:

$$\rho(r) = \frac{\rho_s}{\left(\frac{r}{r_s}\right) \left(1 + \frac{r}{r_s}\right)^2}, \quad (1.3)$$

where ρ_s is the characteristic density and r_s is the scale radius. The concentration parameter, which indicates how matter is radially concentrated toward the centre, is defined as $c \equiv \frac{R_{200}}{r_s}$, where R_{200} is the radius of the sphere in which the mean density is 200 times the critical density

of the Universe at the redshift of that specific galaxy system¹. The concentration parameter is typically lower for the galaxy distribution than for dark matter, but both profiles exhibit a high degree of self-similarity (e.g. [Lin & Mohr, 2004](#); [Budzynski et al., 2012](#)). When observed, cluster/group galaxies are projected on the sky plane, so in observational studies and detection procedures, the projected formulation of the NFW is used, for instance, as proposed by [Bartelmann \(1996\)](#) and [Meneghetti et al. \(2003\)](#):

$$\rho_{2D}(x) = \frac{2\rho_s r_s}{x^2 - 1} f(x) \quad (1.4)$$

with x defined as $x \equiv r/r_s = c \frac{r}{R_{200}}$, and with

$$f(x) = \begin{cases} 1 - \ln \frac{1+a}{\sqrt{1-x^2}} & x < 1 \\ 1 - 2 \frac{\arctan a}{\sqrt{x^2-1}} & x > 1 \\ 0 & x = 1 \end{cases} \quad (1.5)$$

where

$$a = \begin{cases} \sqrt{\frac{1-x}{1+x}} & x < 1 \\ \sqrt{\frac{x-1}{x+1}} & x > 1 \end{cases}. \quad (1.6)$$

The NFW profile is consistent with predictions from cold dark matter simulations and is observed across a wide range of cluster masses and redshifts (e.g. [Pointecouteau et al., 2005](#); [Budzynski et al., 2012](#)). It should be noted that density profiles also exhibit a morphology-dependence, where redder, quiescent galaxies are more centrally concentrated and closely trace the overall mass profile, while star-forming ones have a shallower distribution and are more common in the outskirts of these systems. The same is expected for the mass-dependence, with more massive members favouring the central region. This is naturally a reflection of the evolutionary history and environmental effects within the galaxy system.

Clusters and large groups often contain substructures, especially in dynamically young or merging systems, leading to deviations from smooth, universal profiles. This is the case, for instance, for triaxial haloes and orientation effects like alignment with large-scale filaments (e.g. [Fu et al., 2024](#); [Gouin et al., 2020](#); [Despali et al., 2014](#)). The dynamical state and the large-scale environment of the system, therefore, influence the degree of central concentration and the overall density profile. Despite some studies showing density profiles can be better fitted by core-like and more flexible models like King ([King, 1966](#)) or Einasto ([Einasto, 1965](#)) profiles (especially in

¹The critical density of the Universe, ρ_c , is defined as $\rho_c = \frac{3H(z)^2}{8\pi G}$, where $G \approx 6.67 \times 10^{-8} \text{ cm}^3 \text{ g}^{-1} \text{ s}^{-2}$ is the gravitational constant. From now on, the subscript 200 will refer to quantities (e.g. mass, number counts) defined with respect to the sphere of radius R_{200} .

central regions and for specific classes of systems), the NFW profile remains a robust analytical description for the radial distribution of galaxies.

1.2.1 Detection methods in optical and NIR surveys

As discussed, the detection of galaxy clusters and groups in the optical and near-infrared relies on identifying overdensities of galaxies on the sky, which can be brought to three-dimensional space using accurate photometric redshifts (photo-zs) or spectroscopic redshifts (spec-zs). Although open questions remain regarding the connection between observed galaxy population properties and the underlying physical processes, well-established descriptors allow these systems to be characterised and modelled robustly. In particular, the distributions of galaxy magnitudes, the radial density profiles, the colour/morphology segregation, and the colour–magnitude relations provide powerful tools for the modelling and therefore for the identification of clusters and groups. Detection of galaxy clusters and groups in the optical and near-infrared has a long history and has steadily matured over the years. Ongoing and upcoming wide-field surveys aim to generate and exploit cluster and group catalogues to address a range of astrophysical processes and fundamental cosmological questions (e.g. [Allen et al., 2011](#); [Sartoris et al., 2016](#)). A variety of detection algorithms are currently employed in the literature, reflecting both the complexity of the data and the diversity of scientific goals.

Early developments focused on matched filter methods, which provided a statistical framework for identifying overdensities with a model-based approach. This technique, introduced by [Postman et al. \(1996\)](#) and extended by [Kepner et al. \(1999\)](#), allows for the detection of galaxy systems by enhancing their expected signal relative to noise, using prior knowledge of typical galaxy member distributions and noise statistics. This approach has been extended to the 3D space by incorporating redshift information, so it uses an a priori model for the expected cluster/group signal in position, redshift and luminosity. This model is therefore a spatial template (e.g. a projected radial profile) and a luminosity function, combined into a filter sliding across the catalogue (generally, across redshift slices). This filtering creates a map where candidates occupy peaks. The matched filtering method is highly flexible, multi-wavelength (it can also be applied to non-optical surveys; see below), and produces easily interpretable statistics and extensive characterisation of detected candidates, including proxies of mass like richness and therefore the possibility to quantify and improve the selection function (and so purity and completeness) of the sample. The main weaknesses of this family of methods reside in that same model dependence: wrong assumptions on the cluster model can bias the candidate selection and richness estimates, and the whole approach requires careful treatment of photo-z quality (like careful input data cleaning and validation, and proper redshift slicing). Iterative, deblending techniques that remove the imprint of a detection from the map and assignment of probabilistic memberships can significantly improve performance in crowded fields and permit systematic selection-function tests on mocks.

Examples of matched filters include the implementations of the 3D-MF cluster finder (Milkeraitis et al., 2010), and the optimal filtering method, AMICO (Bellagamba et al., 2018; Maturi et al., 2019). Several major optical and NIR surveys have successfully applied matched filtering techniques for galaxy cluster detection, notably SDSS (e.g. Banerjee et al., 2018), CFHTLS (e.g. Milkeraitis et al., 2010), J-PAS (Maturi et al., 2023), KiDS (Maturi et al., 2019, 2025), COSMOS (Toni et al. 2024, 2025a; this Thesis work) and Euclid (e.g. Euclid Collaboration: Adam et al., 2019; Euclid Collaboration: Bhargava et al., 2025, including GT).

Over time, a range of complementary techniques has been introduced. Among these, non-parametric density estimators based on Voronoi tessellation (e.g. Ramella et al., 2001) identify overdensities by partitioning the projected galaxy distribution into cells and flagging regions where cell areas are anomalously small (meaning the area is particularly dense with galaxies). Because the method does not impose an a priori profile, it is well-suited to detect systems with irregular or elongated morphologies, and it can be implemented in redshift slices to account for projection. The approach also yields a natural local density estimate and, in many implementations, a tentative probabilistic membership assignment. The Voronoi tessellation is a general approach that can also be used on non-optical data, and it is highly flexible, like the matched filters. Limitations are primarily practical: the choice of significance threshold and the handling of survey edges and masked regions, and the method becoming sensitive to shot noise in very sparse or shallow catalogues. A closely related approach is the Delaunay tessellation, the mathematical dual of the Voronoi scheme. In this framework, galaxies are connected to form triangles (2D) or tetrahedra (3D) such that no galaxy lies inside the circumcircle of any triangle. This network naturally defines the local surroundings properties for each galaxy, which can be used to estimate densities or to identify connected overdense regions. Compared to Voronoi tessellation, which provides a local density estimate, Delaunay tessellation focuses on connectivity and spatial relationships between galaxies. It forms the basis of algorithms such as the Delaunay Tessellation Field Estimator (DTFE; e.g. Cautun & van de Weygaert, 2011), which produces a continuous density field, and can also support percolation or clustering analyses. The main advantages of this method are its ability to capture the network structure of galaxy distributions and to interpolate densities across sparse regions. Limitations include sensitivity to survey edges, as for the Voronoi one, and increased computational costs relative to simple Voronoi density estimates. Delaunay-based methods have been successfully applied to deep imaging surveys for identification and for environmental studies up to high redshifts (e.g. Marinoni et al., 2002; Sarron & Conselice, 2021).

Another technique based on the location of galaxies is the family of Friends-of-Friends (FoF) algorithms. A FoF links galaxies into systems using distance (and, in redshift, line-of-sight velocity): any two galaxies separated by less than a chosen linking length are considered "friends" and belong to the same system. When redshifts are spectroscopic, this approach can map three-dimensional structures and allow for a relatively straightforward calibration with mocks by tuning

linking lengths to recover halo occupation. However, FoF results are sensitive to the chosen linking parameters, which can create under- or over-connections. Additionally, redshift-space distortions can elongate systems along the line of sight, and in photometric surveys, the necessary projection of the linking criterion severely degrades performance unless high-quality probabilistic photo- z information is incorporated. Well-known FoF-based catalogues have been constructed for major surveys using photo- and spec- z s (e.g. SDSS; [Farrens et al., 2011](#); [Wen et al., 2012](#); [Tempel et al., 2014](#)) and the approach has been shown to work well when combined with other complementary methods (e.g. [Knobel et al., 2012](#); [Stothert et al., 2019](#)).

Another statistical approach that has been applied to optical and near-infrared data is the Poisson Probability Method (PPM), proposed in [Castignani et al. \(2014a\)](#). This technique identifies candidate clusters and groups by estimating the probability that an observed overdensity of galaxies arises from random Poisson fluctuations of the background. Counts are measured within fixed or adaptive apertures placed across the field, and these are compared to the expectation value derived from the global or local background density. Overdensities with a sufficiently low probability of being due to chance are flagged as candidate detections. The method is interesting for its statistical simplicity and for its robustness in regimes with low galaxy counts, such as the search for distant clusters at $z \gtrsim 1$. However, it is also prone to projection effects and is sensitive to the accuracy of the background estimate, which can vary across a survey. Despite this, the PPM has proven effective in identifying intermediate- to high-redshift structures with systems in COSMOS, GOODS-N and SDSS successfully detected at $z \sim 1-2$ (e.g. [Castignani et al., 2014a, 2019](#); [Calvi et al., 2023](#)).

Another class of widely used techniques is that of wavelet filters, which work by decomposing the galaxy density field (usually in photo- z slices) into scale-dependent components and identifying significant peaks at scales corresponding to cluster/group sizes. The multi-scale nature of wavelets makes these methods particularly effective for simultaneously detecting systems of different physical sizes while suppressing large-scale background fluctuations; they are therefore well suited to heterogeneous data. Drawbacks include sensitivity to the choice of analysed scales and detection thresholds, a non-trivial fragmentation/merging behaviour that must be characterised on mocks, and non-negligible computational costs for very large datasets. Modern wavelet finders, like WaZP ([Benoist, 2014](#)) or PZWAV ([Gonzalez, 2014](#)), have been successfully applied to major surveys like the Dark Energy Survey (DES) and the S-PLUS surveys (e.g. [Aguena et al., 2021](#); [Werner et al., 2023b](#), respectively) and Euclid ([Euclid Collaboration: Adam et al., 2019](#); [Euclid Collaboration: Bhargava et al., 2025, including GT](#)).

Finally, a class of cluster finders with a long history and a wide employment in the literature is that based on the cluster red sequence (see e.g. [Gladders & Yee, 2000](#)). Red-sequence finders exploit the empirical fact that a large fraction of cluster members lie on a tight colour-magnitude locus of passively evolving galaxies. Operationally, this kind of finder scans colour-magnitude

(or multi-colour) space as a function of redshift to locate overdensities of galaxies consistent with a red sequence at a common distance, thereby strongly suppressing foreground contamination and enabling accurate photometric redshift estimates for the detections. The main advantages are the high purity for systems dominated by an evolved population and the ability to provide membership probabilities and therefore richness; the main limitations are selection bias against dynamically young or strongly star-forming systems that lack a prominent red sequence, sensitivity to photometric calibration, photometry quality and filter choice, and the need to model the red-sequence evolution with redshift. The red-sequence paradigm has set the base for the Red-Sequence Cluster Survey (e.g. [Gladders et al., 2007](#); [Gilbank et al., 2011](#)) and modern matched-red-sequence codes such as redMaPPer ([Rykoff et al., 2014](#)), which have been extensively used to compile large, well-calibrated catalogues for cosmological and astrophysical studies (e.g. [Rozo & Rykoff, 2014](#); [Rykoff et al., 2016](#); [McClintock et al., 2019](#)).

Recent years have also seen the emergence of machine-learning-based approaches to either integrate or replace classical methods. These include convolutional neural networks and object-detection architectures trained on images or on catalogue features, supervised classifier-s/regressors that predict cluster probability and richness, and methods that learn membership or mass proxies directly from multi-band imaging or from combinations of photometry and spatial information. The main advantage of these techniques is the flexibility in case of complex, non-linear patterns. The principal caveats are the dependence on the representativeness of the labels and the features of the training set, and the difficulty of producing an easily modelled selection function. Applications on real data and testing have shown that machine learning can help reach high completeness and purity (see e.g. [Grishin et al., 2023](#); [Tian et al., 2025](#)), but extensive validation on mocks and multi-wavelength cross-matches is required before such methods can be extensively adopted in cosmology pipelines.

Figure 1.2 shows an example of how cluster and group candidates emerge in different maps generated by four of the methods we have mentioned. Each of the approaches that have been developed for cluster/group detection has distinct strengths, limitations, and requirements in terms of data quality and completeness, resulting in varying ranges of reliability depending on redshift, richness, and survey depth. The above-mentioned methods have also been exploited in dedicated infrared surveys, which have contributed to the identification of clusters and groups up to high-redshift leveraging the emission of evolved or dust-obscured galaxies (e.g. [Eisenhardt et al., 2008](#); [Wen et al., 2018](#); [Wen & Han, 2018](#); [Gonzalez et al., 2019](#)). Optical and infrared detection methods complement selections at other wavelengths and form the basis of modern optical cluster and group studies, allowing the construction of large, well-characterised samples across a wide range of redshifts.

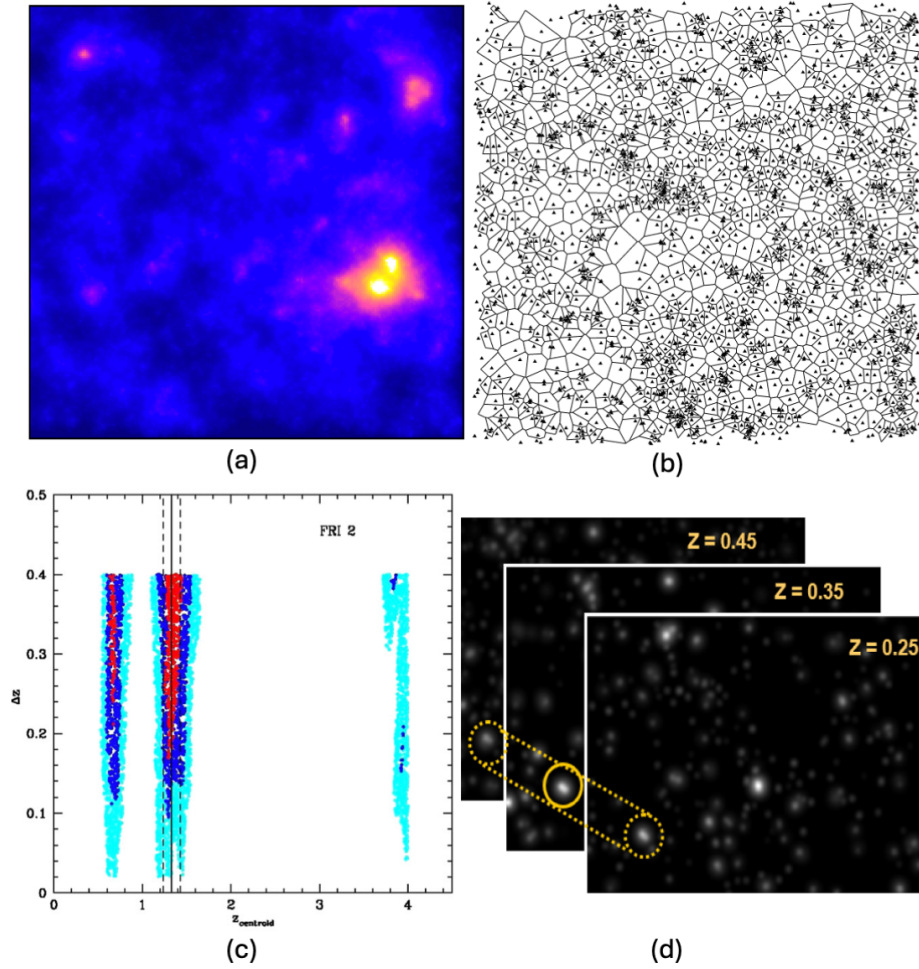


Figure 1.2. Examples of how cluster and group candidates emerge in (a) a matched filtered map (adapted from [Bellagamba et al., 2018](#)), (b) a Voronoi-tassellated map (adapted from [Ramella et al., 2001](#)), (c) a PPM diagram (adapted from [Castignani et al., 2014a](#)) and (d) three slices of a wavelet-filtered map (adapted from [Aguena et al., 2021](#)).

1.2.2 Peculiar galaxies and seeded detection

Besides the statistical description of the overall member galaxy population, clusters and groups host peculiar galaxies like the brightest cluster galaxy (BCG) or brightest group galaxy (BGG), respectively. These are typically located near the centre of the potential well and characterised by exceptional luminosity and mass. BCGs and BGGs are among the most luminous and massive galaxies in the Universe, showing properties that distinguish them from other galaxies (e.g. [Von Der Linden et al., 2006](#); [Kluge et al., 2020](#)). They are often giant ellipticals with extended stellar envelopes (the term *cD* or *central dominant* is also used to refer to these galaxies when they have a diffuse and extended envelope of stars), and their properties are tightly linked to the host’s assembly history, frequently shaped by mergers and accretion of satellite galaxies. The study of BCGs and BGGs, therefore, provides key insights into both galaxy evolution in extreme environments and the hierarchical growth of structures. Several campaigns have been carried out to study the properties of BCGs (e.g. [McDonald et al., 2016](#); [Webb et al., 2017](#); [Radovich et al., 2020](#); [Castignani et al., 2020b, 2022a](#); [Zenteno et al., 2025](#)). Observational studies targeting BGGs in the centre of groups have instead been carried out extensively in more recent time, to follow up on simulation predictions (e.g. [Pillepich et al., 2018](#); [Jackson et al., 2020](#); [Marini et al., 2021](#)), like the COSMOS BGG series ([George et al. 2012](#), [Gozaliasl et al. 2019](#), [Gozaliasl et al. 2024, including GT](#), [Gozaliasl et al. 2025, including GT](#)). Studies specifically targeting BGGs have revealed that these galaxies grow primarily through mergers and accretion, evolving gradually through a reddening of their stellar populations. Star formation in BGGs is generally low, although on average higher than in BCGs, with up to ~ 20 – 50% of them showing ongoing star formation and gas reservoirs, especially with increasing redshift (e.g. [Gozaliasl et al. 2024, including GT](#), [Gozaliasl et al. 2025, including GT](#)). The kinematics of stars and gas in BGGs also depends on the dynamical state of the host group, with relaxed groups showing better alignment while unrelaxed systems often display offsets and misalignments (e.g. [Raouf et al., 2021](#); [George et al., 2012](#); [Gozaliasl et al., 2019](#)). Observations of gas morphology reveal complex structures such as filaments and disks, and the stellar light profiles reflect the hierarchical assembly history of these peculiar galaxies (e.g. [Olivares et al., 2022](#); [Raj et al., 2020](#)). As BCGs and BGGs serve as signposts for their host systems, BCG/BGG-centred algorithms can be used to locate the central galaxy within a search radius and then characterise the surrounding galaxy overdensity; the BCG/BGG is often used as the system centre and as an anchor to estimate its richness. This strategy is conceptually simple and computationally efficient and frequently yields robust centres for relaxed systems, but it can fail when the central galaxy is not dominant, when miscentring occurs due to offsets, or in the case of line-of-sight projections. Such methods have produced rich observational catalogues, for example, the maxBCG catalogue derived from SDSS (e.g. [Koester et al., 2007](#)), which in turn have been used for statistical investigations of cluster scaling relations and cluster cosmology (e.g. [Zu et al., 2014](#); [Bauer et al., 2012](#)).

Previously in this Chapter, environmental effects and segregations have been discussed for galaxy members in clusters and groups. The environment can show its influence on galaxy evolution not only in optical colours and properties, but also in radio wavelengths. Radio galaxies powered by AGNs are frequently hosted by dense environments, often displaying features such as bent or distorted jets caused by interactions with the dense intracluster/intragroup medium. These characteristic radio-morphologies can be exploited as tracers of galaxy clusters and groups themselves (e.g. [Smolčić et al., 2007](#); [Castignani et al., 2014b](#)). In addition to AGN-related emission, clusters also exhibit diffuse radio emission—such as haloes and relics—which is not associated with compact sources, like radio galaxies. Such emission arises from synchrotron radiation, produced by relativistic particles spiralling around magnetic field lines, therefore requiring both large-scale magnetic fields and energetic particle populations. As a result, diffuse radio sources provide unique laboratories to study the physical processes shaping the intracluster medium (see [van Weeren et al., 2019](#), for a review).

1.3 The hot intracluster/group medium

Although galaxy clusters and groups can be identified in the optical and near-infrared thanks to their galaxy populations, the dominant baryonic component in terms of mass is the hot, highly ionised plasma in which galaxies are embedded. This plasma, known as the intracluster medium (ICM, or IGM in the case of groups), is gravitationally confined within the cluster potential well and typically exhibits particle densities between a few 10^{-2} and 10^{-4} cm^{-3} , with temperatures reaching 10^7 – 10^8 K in the richest systems, corresponding to 1–10 keV (see [Sarazin, 1986](#), for a review).

1.3.1 X-ray emission of the ICM/IGM

Under these conditions, the ICM makes clusters extremely bright X-ray sources, a property that has enabled their detection at these wavelengths ever since the first extragalactic X-ray surveys were carried out by missions such as Uhuru ([Gursky et al., 1971](#)) and Einstein ([Giacconi et al., 1979](#)). This diffuse X-ray emission originates mainly from thermal bremsstrahlung (free-free), produced by the deceleration of electrons in the Coulomb fields of ions in the hot ionised plasma, which produces a smooth continuum. The emissivity per unit volume and frequency is given by (e.g. [Rybicki & Lightman, 1979](#)):

$$\epsilon_\nu \propto n_e n_i T^{-1/2} e^{-h\nu/kT}, \quad (1.7)$$

where n_e and n_i are the electron and ion number densities respectively, T is the temperature in Kelvin, h is Planck constant, and k is Boltzmann constant. Integrating over frequency gives the bolometric emissivity, which scales as $\epsilon \propto n_e n_i T^{1/2}$, namely, the denser and hotter the gas, the

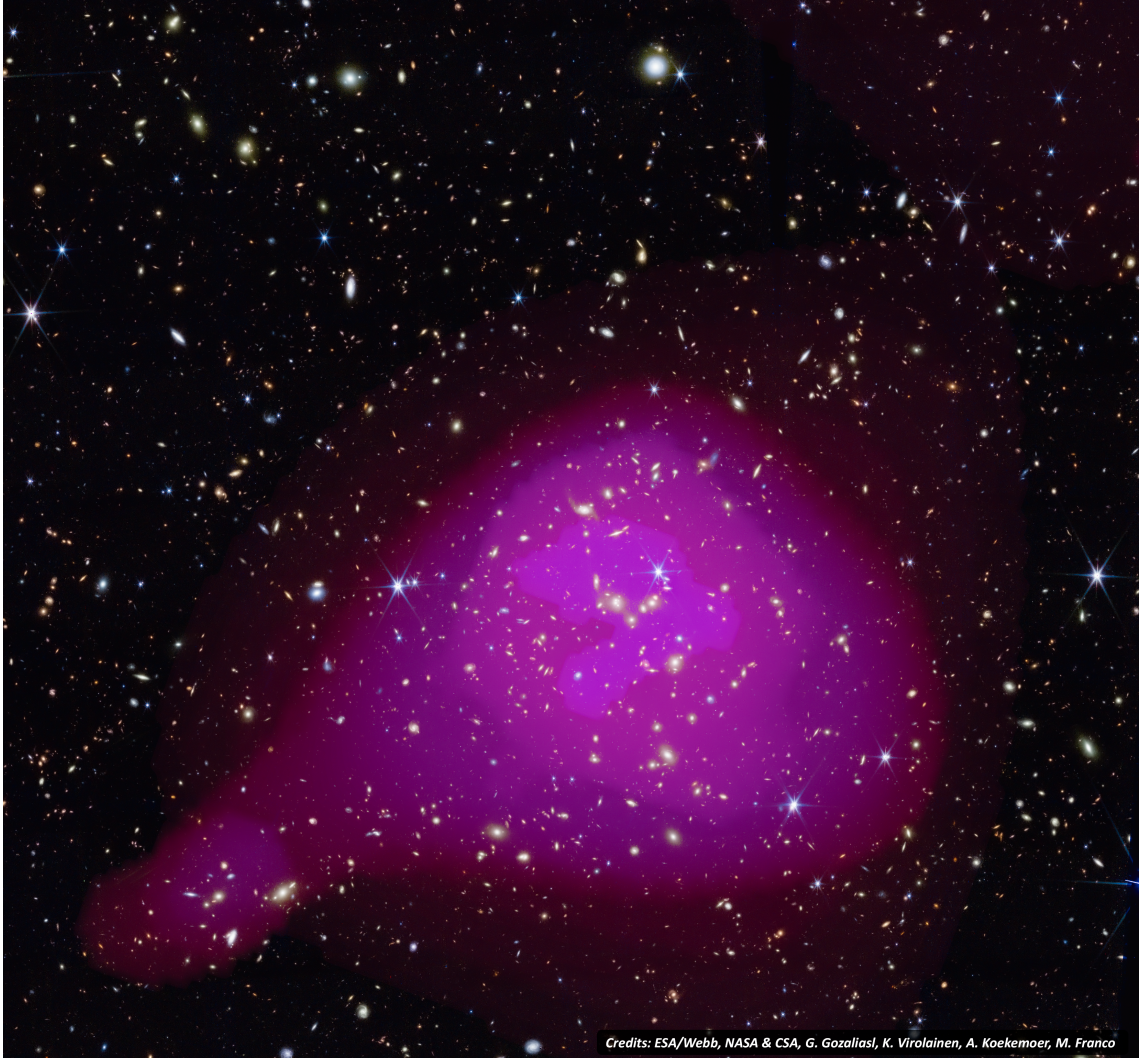


Figure 1.3. Optical and NIR emission of the galaxy component as captured by JWST and of the X-ray emission of the hot gas as seen by XMM-Newton and Chandra (pink cloud), for the most massive galaxy system in the COSMOS-Web field. JWST *rgb* (F444W as *r*, [F150W, F277W] as *g*, F115W as *b*) colour-composite image is overlapped with the X-ray emission from the combined XMM-Newton and Chandra 0.5–2 keV wavelet-filtered image. Credits: [Toni et al. \(2025a\)](#), see Chapter 4; [selected as ESA/Webb picture of the month](#) – April 2025; ESA/Webb, NASA & CSA; Gozaliasl, Virolainen, Koekemoer, Franco.

more luminous in X-rays. Typical observed X-ray bolometric luminosities yielded by this process fall in the range $10^{43} - 10^{45} \text{ erg s}^{-1}$. The heating of the gas to such extreme temperatures is thought to occur via shock heating and adiabatic compression during the process of hierarchical cluster assembly (e.g. Voit, 2005; Kravtsov & Borgani, 2012).

While thermal bremsstrahlung dominates the X-ray emission of the ICM, particularly at the high temperatures characteristic of rich clusters, additional processes can become relevant, especially in the soft X-ray regime. Bound-free emission is generated by recombination, when free electrons are captured by ions and release excess energy as photons, while bound-bound transitions give rise to discrete emission lines as electrons in ions move between quantised levels. Heavy elements, especially iron, silicon, and oxygen, dominate these line features, with the Fe XXV and Fe XXVI $K\alpha$ complexes at 6.7–6.9 keV being especially prominent (Mushotzky et al., 1996). The coexistence of continuum and line emission makes X-ray spectroscopy a powerful diagnostic tool. By fitting the bremsstrahlung continuum, robust estimates of the ICM temperature and density can be obtained, while the relative strengths of emission lines constrain ionisation states and reveal multi-phase gas components, providing direct insight into the metallicity of the ICM, thereby tracing the history of enrichment from supernovae and AGN feedback within the cluster galaxies (e.g. Renzini, 1997; Böhringer & Werner, 2010; Peterson et al., 2003). Furthermore, the width and shifts of X-ray lines can be used to study bulk motions and turbulence in the hot gas, a capability that has been demonstrated with high precision by the spectrometers onboard the Hitomi satellite (Hitomi Collaboration et al., 2016). Therefore, X-ray spectroscopy not only confirms the thermal origin of cluster emission but also opens a direct window into the thermodynamics, chemical history, and dynamical state of the ICM.

A substantial advancement in both imaging and spectroscopy was achieved at the turn of the millennium with the advent of Chandra (Weisskopf et al., 2000) and XMM-Newton (Jansen et al., 2001) space observatories. Chandra’s sub-arcsecond angular resolution revealed the complex thermodynamic structure of the ICM, including cavities inflated by AGN feedback and the intricate morphology of cooling cores (e.g. Fabian et al., 2000; McNamara & Nulsen, 2007). XMM-Newton, with its large effective area and high-throughput spectroscopy, provided complementary insights, enabling precise measurements of cluster temperature profiles and elemental abundances out to large radii (e.g. Kaastra et al., 2001; Pratt et al., 2007). Together, these observatories started a new era of X-ray cluster astrophysics, establishing the ICM as not only a tracer of cluster mass but also a laboratory for plasma physics, galaxy evolution, and feedback processes, complementary to the study of the galactic component. Figure 1.3 shows the richest and most massive system in the COSMOS-Web region, located at $z \sim 0.7$. The galactic component, imaged with the James Webb Space Telescope (JWST), is overlapped with the X-ray emission of the hot gas observed by XMM-Newton and Chandra (Toni et al., 2025a, this Thesis work).

The diffuse X-ray emission of the ICM is a robust and direct tracer of gravitationally bound

systems, which makes X-ray surveys a powerful tool for detecting clusters and groups, with low risk of projection effects, high purity and well-understood selection functions (e.g. [Allen et al., 2011](#); [Piffaretti et al., 2011](#); [Clerc & Finoguenov, 2023](#); [Bulbul et al., 2024](#)). Modern identification methods combine traditional image analysis, statistical modelling, and advanced machine learning to maximise detection efficiency and purity. These include wavelet and matched filter techniques to extract diffuse emission in contrast with point sources, using templates and a priori cluster profiles, combination with multi-wavelength data and association of members via cross-match with optical counterparts, providing redshift estimate (e.g. [Finoguenov et al., 2007](#); [George et al., 2011](#); [Lloyd-Davies et al., 2011](#)). Moreover, the strong correlation between X-ray luminosity and total mass, albeit with significant scatter, allows X-ray selection to provide a direct mass proxy (e.g. [Lovisari et al., 2015](#); [Mantz et al., 2010](#); [Leauthaud et al., 2010](#)). However, X-ray detection is limited by a reduced sensitivity to morphologically disturbed, low-mass, or high-redshift systems, due to surface brightness dimming, especially for shallower surveys. In fact, X-ray selection introduces biases by favouring cool-core clusters, namely those with bright, centrally peaked emission, massive systems and those with a relaxed dynamical state (e.g. [Rosati et al., 2002](#); [Käfer et al., 2019](#); [Kugel et al., 2024](#); [Popesso et al., 2024](#)). Groups, although far more numerous than massive clusters, pose additional challenges. Their IGM is cooler than the ICM ($kT \sim 0.5\text{--}2$ keV) and less dense, leading to fainter and softer emission that is more difficult to detect, requiring deeper observations (e.g. [Osmond & Ponman, 2004](#); [Eckmiller et al., 2011](#); [George et al., 2011](#)). In addition, scaling relations such as the luminosity–temperature relation steepen significantly in the group regime, reflecting the increasing role of non-gravitational processes like radiative cooling, galactic winds, and AGN feedback, which inject entropy into the gas and suppress its central density ([Pratt et al., 2007](#); [Sun, 2012](#)). Nonetheless, with deep pointed observations from Chandra and XMM-Newton (e.g. [Finoguenov et al., 2007](#); [Fassbender et al., 2011](#); [Pierre et al., 2016](#); [Koulouridis et al., 2021](#)), and, more recently, with the eROSITA All-sky Survey² ([Brunner et al., 2022](#)), the detection frontier has expanded toward both low-mass systems and higher redshift. eROSITA’s combination of sensitivity, spectral resolution, and wide-area coverage has revolutionised the detection of extended X-ray emission, delivering the largest ICM-selected sample of clusters and groups ever assembled. Its well-defined selection function and high purity enable statistically robust cosmological analyses and complement optical and NIR detection, thereby providing a cornerstone for cosmology but also for multi-wavelength astrophysical analyses in general ([Bahar et al., 2022](#); [Liu et al., 2022](#); [Seppi et al., 2023](#); [Bulbul et al., 2024](#)).

1.3.2 Sunyaev-Zeldovich effect in the microwaves

Galaxy clusters have long been detected in the optical and X-ray bands, but these are not the only available techniques. Over the past few decades, an equally powerful method has emerged

²<https://erosita.mpe.mpg.de/>

at microwave wavelengths, based on the Sunyaev–Zeldovich (SZ) effect (Sunyaev & Zeldovich, 1970). The SZ effect arises from inverse Compton scattering of CMB photons by the hot, thermal electrons of the ICM itself. Since the CMB is well described by a nearly perfect blackbody with a today temperature of $T_{\text{CMB}} \simeq 2.726$ K (Fixsen, 2009), the up-scattering of photons by cluster electrons imprints a measurable distortion on its spectrum at the cluster position. Specifically, the effect manifests as a shift, namely a decrement in intensity at frequencies $\nu \lesssim 217$ GHz and a corresponding increment at $\nu \gtrsim 217$ GHz, with the crossover frequency set by the physics of the scattering process. The amplitude of the signal is quantified by the Compton- y parameter, defined as

$$y = \frac{k_B \sigma_T}{m_e c^2} \int n_e T_e dl \quad (1.8)$$

where σ_T is the Thomson cross section, m_e the electron mass, c the speed of light, n_e the electron density, T_e the electron temperature, and the integral is along the line of sight. This parameter encapsulates the line-of-sight integrated thermal pressure of the ICM, making SZ measurements direct probes of the hot gas content of clusters. As for the X-ray selection, SZ-based detection of galaxy clusters relies on a combination of filtering techniques, profile fitting, often complemented by multi-frequency data (e.g. Tarrío et al., 2018). A major advantage of SZ detection is its nearly redshift-independent nature: since the CMB provides the background photons, the observable distortion depends only weakly on the cluster distance. As a result, SZ surveys are particularly efficient at detecting massive clusters out to high redshift, where X-ray surface brightness dimming becomes more severe (e.g. Allen et al., 2011). On the other hand, the method is intrinsically less sensitive to low-mass systems, such as galaxy groups. The signal amplitude scales roughly with the total thermal energy of the ICM, making the SZ signal of groups fainter than that of clusters. As a result, individual groups are generally below the detection threshold of current wide-field SZ surveys, although stacking analyses have successfully revealed their average signal (e.g. Planck Collaboration et al., 2013; Greco et al., 2015; Hill et al., 2018). This limitation highlights the complementarity of SZ, X-ray, and optical/NIR methods, since the latter remain more efficient in probing the low-mass end of the halo population.

SZ surveys have now delivered large catalogues of clusters with robust mass proxies, most notably from the Planck satellite (e.g. Planck Collaboration et al., 2016), the South Pole Telescope (SPT; Bleem et al. 2015) and the Atacama Cosmology Telescope (ACT; Hilton et al. 2021), which are complemented by targeted observations with instruments such as IRAM-NIKA2 (e.g. Kéruzoré et al., 2020) and ALMA (e.g. Kitayama et al., 2023). These catalogues have become crucial not only for studies of galaxy and cluster astrophysics but also as cosmological probes, since the abundance and spatial distribution of SZ-selected clusters provide strong constraints on the growth of structure across cosmic time (e.g. Aymerich et al., 2024; Salvati et al., 2022).

1.4 Gravitational lensing

Although galaxy clusters contain a significant baryonic component (about 15–20%, mainly in the form of the intracluster medium and galaxies), the vast majority of their mass resides in dark matter. Gravitational lensing provides a unique tool to probe this dominant, invisible component directly. Thanks to modern wide-field surveys and the development of dedicated algorithms, it has become possible to detect clusters and infer their total mass from the distortions imprinted on the images of background galaxies (e.g. [Schneider et al., 1998](#); [Maturi et al., 2005](#); [Shan et al., 2012](#); [Stapelberg et al., 2019](#); [Trobiani et al., 2025](#)). Photons emitted by distant sources are deflected according to general relativity by the gravitational potential of the intervening massive object, in this case, the cluster, which acts as a lens. The gravitational lensing effect is usually divided into strong lensing, producing prominent features such as giant arcs and multiple images (e.g. [Kneib et al., 1996](#); [Oguri & Marshall, 2010](#)), and weak lensing, which manifests as milder but systematic, statistical shape distortions of background sources, commonly quantified as *shear*, a measure of the stretching induced by the lensing mass distribution (see e.g. reviews by [Bartelmann & Schneider, 2001](#); [Bartelmann, 2010](#); [Hoekstra et al., 2013](#)). The great advantage of gravitational lensing is that it provides a way to infer the cluster total mass without depending on the physical state of the system, in contrast to X-ray or SZ techniques, which rely on assumptions about the equilibrium of the intracluster medium (e.g. [Allen et al., 2011](#)). Strong and weak lensing measurements, and the combination of the two, have been used extensively to obtain precise cluster mass estimates and to calibrate mass–observable scaling relations, thereby enabling the cosmological use of galaxy clusters (e.g. [Leauthaud et al., 2010](#); [Bellagamba et al., 2019](#)). Just to mention some, surveys such as CFHTLenS ([Heymans et al., 2012](#)), DES ([Abbott et al., 2020](#)) and HSC ([Miyazaki et al., 2018](#)) have shown the capability of weak lensing to deliver large cluster samples with robust mass calibration. Nevertheless, lensing-based detection has some limitations. Strong lensing is restricted to the most concentrated central regions of clusters and can be strongly affected by geometrical biases. Weak-lensing detection, although less biased, requires high-quality images and accurate control of systematics such as shape measurement errors, photo- z uncertainties, and intrinsic alignments of galaxies (e.g. [Köhlinger et al., 2015](#); [Mandelbaum et al., 2018](#)). Moreover, the signal-to-noise ratio of weak-lensing measurements scales with halo mass and redshift, making the detection of galaxy groups and high- z systems particularly challenging. Although the shallower potential wells of galaxy groups produce weaker shear signatures, stacking analyses of large group samples have successfully recovered their average lensing signal (e.g. [George et al., 2012](#); [Viola et al., 2015](#)), providing valuable insights into the transition regime between galaxy- and cluster-scale haloes. In summary, gravitational lensing offers a direct probe of the total mass distribution in clusters and groups, independent of baryonic tracers. While observational challenges can limit its effectiveness for individual low-mass systems, forthcoming deep and wide surveys such as

Euclid (Laureijs et al., 2011) and the Rubin Observatory LSST (Ivezić et al., 2019) are expected to significantly expand the role of lensing in cluster detection and cluster cosmology (e.g. Euclid Collaboration: Sereno et al., 2024; Euclid Collaboration: Ingoglia et al., 2025; Payerne et al., 2025).

Table 1.1 summarises the main cluster and group detection strategies discussed in this Chapter, with their main advantages and limitations.

1.5 High-redshift groups and protoclusters

Detecting and studying galaxy clusters and groups enables connections to be drawn between galaxy morphology, colour, star-formation activity, environment and assembly history. The dominance of red, early-type galaxies in clusters has been shown to characterise systems at least up to $z \sim 1$ (e.g. George et al., 2011; Fritz et al., 2014). This primarily quiescent, red, passively evolving population of galaxies is known to occupy a tight correlation in the colour-magnitude diagram, the red sequence. As previously described, the red sequence can be exploited as a tracer to create robust catalogues of clusters and groups at $z \lesssim 1$ (e.g. Gladders & Yee, 2000; Rykoff et al., 2014, 2016). However, while quiescent fractions and red-sequence properties have been well characterised in the local universe and up to intermediate redshifts ($z \sim 1 - 1.5$; e.g. Menci et al., 2008; Rudnick et al., 2009; Fritz et al., 2014; Hennig et al., 2017), the classical morphology-density relation observed in the local universe (Dressler, 1980) seems to gradually fade at $z \gtrsim 1.5$ and group and cluster galaxies exhibit properties more consistent with those of the field galaxies, in terms, for example, of SFRs and morphologies (e.g. Brodwin et al., 2013; Alberts et al., 2016), revealing more complicated and still relatively unexplored relations in this regime. Environmental effects in general, down to the group regime, have also been widely explored at $z \lesssim 1.5$ (e.g. Wilman et al., 2005; George et al., 2011; Salerno et al., 2019; Balogh et al., 2021; McNab et al., 2021; Reeves et al., 2021; Baxter et al., 2023; Kukstas et al., 2023). However, a better understanding of how the first galaxies evolved and quenched and build up the red sequence can be addressed by studying samples of groups and clusters at higher redshifts, up to the regime where clusters themselves were still forming ($z > 1.5 - 2$). This epoch has so far revealed a wide diversity in galaxy populations, with systems showing enhanced star formation, AGN, and merging activity (e.g. Brodwin et al., 2013; Alberts et al., 2016; Wang et al., 2016) and others (sometimes even coexisting) an early-phase red sequence and high red fractions, especially in the central regions of galaxy overdensities (e.g. Andreon & Huertas-Company, 2011; Spitler et al., 2012; Strazzullo et al., 2013, 2016; Zavala et al., 2019; Kiyota et al., 2025). In addition to this, the interval $z = 1.5 - 2.0$ marks the transition between virialised objects observed locally ($z \lesssim 1.5$) and the maturing phase of the so-called *protoclusters* ($2 \lesssim z \lesssim 3$), the progenitors of galaxy clusters (e.g. Muldrew et al., 2015; Overzier, 2016; Shimakawa et al., 2018). Even if there is no univocal observational definition of them in the literature, the term *protocluster* is used to refer to structures that are going to

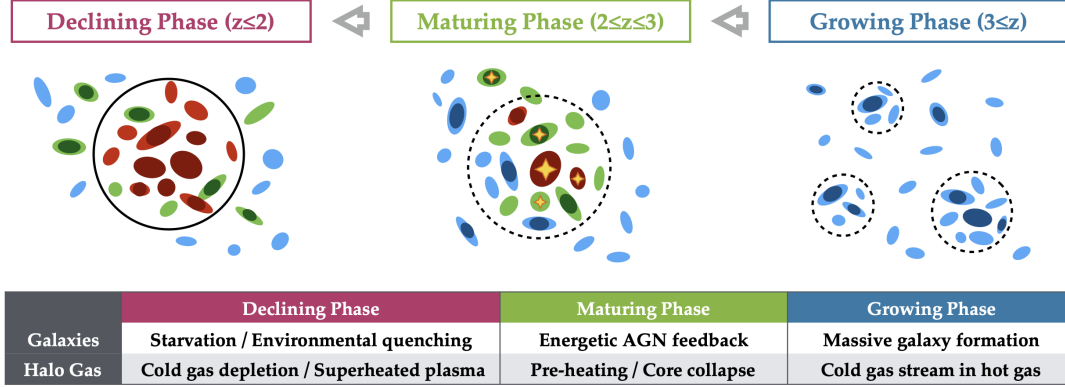


Figure 1.4. Schematic of how galaxy clusters (and their member galaxies) form and evolve over cosmic time. Protoclusters are extended high-redshift regions composed of multiple dense cores (right) associated with overdensity peaks. These group-sized protocluster cores evolve through gravitational collapse, AGN feedback (middle), driving the emergence by $z \sim 2$ of the virialised clusters we observe in the local Universe (left). Credits: Shimakawa et al. (2018).

evolve and eventually collapse into a galaxy cluster by $z = 0$ (Steidel et al., 1998; Overzier, 2016). Protoclusters generally exhibit extended, multi-halo and filamentary structure, and they exist in a range of different evolutionary states, since they are often observed while undergoing accretion and aggregation processes (Muldrew et al., 2015).

Given their diversity and relatively recent characterisation, definitions concerning protoclusters are often vague and inconsistent between different works. In this Thesis, we will adopt the following defining scheme, in line with that proposed by Shimakawa et al. (2018): (1) overdense fields on the scale of several tens of comoving Mpc³ fall under the more general category of large-scale structures; (2) overdense regions covering ~ 1 to a few tens of comoving Mpc are referred to as protoclusters; (3) single group-size ($\lesssim 1$ cMpc) overdensities or overdense cores/-peaks are called protocluster cores, or simply cores. In this scenario, protoclusters are extended objects (or better, regions) composed of several substructures (cores) corresponding to different overdensity peaks. These cores, which are found at high- z to have sizes and quasi-virialised conditions comparable to those of local galaxy groups, are thought to evolve and collapse during a maturing phase, starting indicatively from $z \sim 3$ (e.g. Shimakawa et al., 2018). The combined evolution of galaxies and ICM through energetic AGN feedback, environmental quenching and cold gas depletion culminates with the virialisation into the galaxy clusters we observe in the local Universe. This evolution scenario is well outlined in the sketch in Figure 1.4 (taken from Shimakawa et al. 2018; Figure 13). The study of high-redshift galaxy groups and protocluster

³a comoving Mpc (cMpc) is a Megaparsec ($\approx 3.086 \times 10^{24}$ cm) measured in the *comoving frame*, the reference system that follows the so-called *Hubble flow* driven by the Universe's expansion. Comoving distances are linked to proper distances by the scale factor.

regions is therefore key to understanding the evolution and star formation history of galaxies, their connection to the dark matter distribution, and the impact of physical processes such as gas heating and AGN feedback (e.g. Tanaka et al., 2012; Zhang et al., 2020).

The identification of galaxy clusters/groups and protoclusters at $z \gtrsim 1.5$ has been receiving growing attention only in recent times, thanks to modern deep multi-wavelength surveys and the development of dedicated search algorithms. This has revealed and proven the importance of studying these objects, as essential processes transforming galaxies in clusters take place already during early phases of cluster assembly. Detecting galaxy protoclusters and their cores includes identifying significant galaxy overdensities, confirming membership with spectroscopy and multi-wavelength data, and leveraging simulations to trace their evolutionary fate. Starting from the first serendipitous findings of overdensities of Ly-break galaxies (e.g. Steidel et al., 1998) and Ly α -emitters (e.g. Shimasaku et al., 2003), protoclusters and high- z groups have been found systematically in spectroscopic surveys (e.g. Cucciati et al., 2014, 2018; Lemaux et al., 2018; Diener et al., 2013), but also in photometric surveys (e.g. Spitler et al., 2012; Chiang et al., 2014). In the optical and infrared, overdensities of galaxies at high redshift can be detected using photometric redshifts or colour cuts targeting spectral features. While the red-sequence method loses efficiency beyond $z \sim 1.5$ due to higher levels of ongoing star-formation, a few interesting high- z candidates have been found to have an early passive galaxy population at $z \sim 2$ (e.g. Willis et al., 2013; Strazzullo et al., 2016; Kiyota et al., 2025). At higher redshifts, spectroscopic overdensity mapping and narrow-band imaging targeting Ly α (or H α) emission have been particularly powerful, revealing large protocluster regions at $z \gtrsim 2$ (Chiang et al., 2015; Calvi et al., 2023). Spectroscopic surveys such as VUDS (Lemaux et al., 2014, 2018), zCOSMOS (Diener et al., 2013, 2015), and merged archival data (e.g. Franck & McGaugh, 2016) have also mapped significant overdensities by tracing the three-dimensional distribution of galaxies. The advent of JWST is now unveiling dense protocluster cores at extremely high redshifts. To mention a few examples, the protocluster core revealed by Arribas et al. (2024) around two dusty-star-forming galaxies at $z \sim 6.9$, the one found behind SMACS0723-7327, at $z \sim 7.7$ (Laporte et al., 2022), and the A2744-z7p9OD at $z \sim 7.9$ (Morishita et al., 2023).

In the X-rays, extended emission from the ICM remains a robust tracer of virialised structures, though surface-brightness dimming as $(1+z)^{-4}$ strongly limits sensitivity at high redshifts. Nevertheless, Chandra and XMM-Newton have allowed the detection of the first high- z systems (e.g. Rosati et al., 2009; Wang et al., 2016), and now eROSITA is significantly broadening the census of massive intermediate- to high- z systems (e.g. Bulbul et al., 2024). On the other hand, detection using the SZ effect provides nearly redshift-independent sensitivity to hot, massive clusters and has enabled the production of catalogues of massive systems up to $z \sim 2$, though the method remains nearly insensitive to small and unvirialised systems.

Last but not least, signposts like AGNs (e.g. Hennawi et al., 2015; García-Vergara et al., 2022),

radio-galaxies (e.g. [Hatch et al., 2011](#); [Wylezalek et al., 2013](#)), and sub-millimetre galaxies (e.g. [Daddi et al., 2009](#); [Dannerbauer et al., 2014](#); [Casey et al., 2015](#); [Wang et al., 2021](#)) have uncovered a large number of overdense regions at high redshifts.

In summary, X-ray and SZ surveys are most efficient for finding mature, massive clusters in the declining/virialisation phase, while optical/infrared, spectroscopically-based methods, and sub-mm and radio detections for a complementary view are best suited for identifying groups and protoclusters at higher redshifts and studying their growing and maturing phase. Recently, the discovery of more protocluster candidates has also opened a window on the study of the formation processes involving BCGs and their connection to intracluster light (e.g. [Ito et al., 2019](#); [Werner et al., 2023a](#); [Chapman et al., 2024](#)). As we enter a new era of deep and wide surveys—like Euclid, the Rubin Observatory LSST, and JWST deep fields—it is now possible to get first statistical samples of high- z groups and protoclusters, allowing systematic exploration of their role in galaxy evolution and structure formation.

Table 1.1. Some of the techniques used in the literature for the detection of galaxy clusters and groups.

Method	Principle	Main Advantages / Limitations
Optical / NIR		
Matched Filter	fit galaxy density maps with an assumed radial/luminosity profile	well-defined selection function, flexible / sensitive to assumed templates and survey depth
Red Sequence	identify overdensities of galaxies on a tight colour–magnitude locus	high purity, robust at low-intermediate z / can miss groups with weak/absent red sequence, sensitive to bad photometry
galaxy-seeded	use BCGs/BGGs or radiogalaxies as seeds for targeted identification	simple, efficient for massive systems / less reliable for poor systems or systems without a central dominant
Friends-of-Friends (FoF)	link galaxies according to a maximum separation (linking length) in 2D/3D space	simple, flexible / subject to fragmentation or over-merging depending on parameter choice
Voronoi/Delaunay tessellation	detect overdensities using field tessellation based on galaxy density	non-parametric, sensitive to irregular morphologies, efficient at high- z / sensitive to masks and edges
Wavelets	identify peaks of a given size via multi-scale filtering	efficient for heterogeneous target detection / sensitive to scale choice, fragmentation and merging
X-ray extended emission		
	thermal bremsstrahlung from hot intracluster/intragroup medium	robust tracer of virialised systems / weak signal for groups or high- z systems, requires deep exposures
SZ effect		
Compton- γ signal	measure of CMB distortions from inverse Compton scattering by ICM/IGM	nearly redshift-independent, effective for massive clusters / weak signal for groups
Weak gravitational lensing		
Shear mapping	identify mass overdensities via background galaxy shape distortions	direct total-mass tracer / low S/N for groups, requires high-quality shape measurements
Specific methods for high-z groups and protoclusters		
galaxy overdensities or signposts	overdensities of tracers like Ly α emitters, sub-mm galaxies, radio-loud AGN or quasars; spectroscopy to study membership and 3D structure	trace structures at $z > 2$, spectroscopy provides robust confirmation / biased to specific tracers, sensitive to cosmic variance, requires deep surveys and spectral coverage

2

The AMICO algorithm

The Adaptive Matched Identifier of Clustered Objects (AMICO; [Bellagamba et al., 2011, 2018](#); [Maturi et al., 2019](#)) is an algorithm developed for the detection of galaxy clusters in photometric catalogues of galaxies. The algorithm is based on a technique known as linear optimal matched filtering. In general, this method allows the extraction of a specific signal from a dataset affected by a noisy background, maximising the signal-to-noise ratio (S/N). Further details about matched filters and their application can be found in the literature, e.g. in [Postman et al. \(1996\)](#) and [Maturi et al. \(2005\)](#). In this formalism, detection targets are defined by their resemblance to an a priori model describing the expected signal of the object of interest. In the context of galaxy cluster or group detection, such a model may, for example, be described by a radial profile and a luminosity function, both of which can vary with redshift. An important advantage of the optimal matched filtering formalism over other methods is its high flexibility. Rather than being designed for a specific dataset, it can be adapted to different kinds of targets and surveys across a wide redshift range, and an arbitrary number of properties can be added to the filter description. Moreover, it enables consistent comparison and combination of optical data with complementary observables, such as gravitational lensing, Sunyaev–Zeldovich, and X-ray observables.

In this Chapter, the AMICO detection algorithm will be presented and described, starting from its mathematical formalism to its recent successful applications to real data and its selection as one of the two official cluster detectors for the Euclid mission.

2.1 Mathematical formalism

The optimal matched filtering technique relies on the assumption that a dataset can be modelled as the sum of a signal component and a noise component. If we take for example a photometric catalogue of galaxies, the data component is represented by the galaxy density $D(\boldsymbol{\theta}, \mathbf{m}, z)$, which is a function of the angular position, $\boldsymbol{\theta}$, the redshift, z , and an array of n galaxy properties, $\mathbf{m}$. Despite AMICO being able to deal with an n number of galaxy properties, such as morphology, shape, and colour, the simplest and most commonly used configuration is with a single property, typically the photometric apparent magnitude. If we indicate the magnitude in a given filter as m , the galaxy density can be written as

$$D(\boldsymbol{\theta}, m, z) = A(\boldsymbol{\theta}_c, z_c) M_c(\boldsymbol{\theta} - \boldsymbol{\theta}_c, m, z) + N(m, z), \quad (2.1)$$

where the signal component accounting for the galaxies belonging to the cluster is the model, $M_c(\boldsymbol{\theta} - \boldsymbol{\theta}_c, m, z)$, that is the expected distribution of galaxies in the cluster centred in $(\boldsymbol{\theta}_c, z_c)$, multiplied by a normalisation factor, $A(\boldsymbol{\theta}_c, z_c)$, that is the *amplitude* of the cluster. The other term, $N(m, z)$, is the noise component associated with the contribution of field galaxies.

The signal amplitude A plays the role of a scaling factor representing the number of galaxies belonging to the cluster, and it is computed as a convolution of the data D with a *filter* defined via a constrained minimisation approach. The convolution with this filter, Ψ , guarantees an unbiased and minimum-variance estimate of the cluster amplitude:

$$A(\boldsymbol{\theta}_c, z_c) = \alpha^{-1}(z_c) \int \Psi(\boldsymbol{\theta} - \boldsymbol{\theta}_c, m, z) D(\boldsymbol{\theta}, m, z) d^2\theta dm dz - B(z_c), \quad (2.2)$$

where the last term is the background subtraction, with B being quantified as the mean contribution of the field galaxies to the signal amplitude. In case more than one galaxy property is being used, the integral is computed over the n dimensions of the given array of galaxy properties, $\mathbf{m}$. The normalisation constant, α , is defined as

$$\alpha(z_c) = \int \Psi^2(\boldsymbol{\theta} - \boldsymbol{\theta}_c, m, z) N(m, z) d^2\theta dm dz, \quad (2.3)$$

in such a way that it makes the amplitude, A , an indicator of the cluster signal in units of the model, M_c .

In case we assume the noise to be characterised by a white power spectrum, the optimal filter can be expressed as the ratio between the cluster model and the noise. In this way, when the

cluster signal dominates over the noise, the field galaxy contribution is suppressed by the filter:

$$\Psi(\boldsymbol{\theta} - \boldsymbol{\theta}_c, m, z) = \alpha^{-1}(z_c) \frac{M_c(\boldsymbol{\theta} - \boldsymbol{\theta}_c, m, z)}{N(m, z)}. \quad (2.4)$$

The AMICO filtering process, which takes place in the first phase of the detection process, is parametrised by three main filter constants characterising the filter function. These are defined in terms of the typical redshift probability distribution for a galaxy at a given redshift (e.g. the cluster candidate redshift, z_c).

The typical redshift distribution, $q(z_c, z)$, can be retrieved directly from the data as

$$q(z_c, z) = \frac{\sum_{i=1}^{N_{gal}} p_i(z - z_c + z_{peak,i}) p_i(z_c)}{\sum_{i=1}^{N_{gal}} p_i(z_c)}, \quad (2.5)$$

where $z_{peak,i}$ is the redshift at which $p_i(z)$ peaks, that is, namely the highest-probability redshift for the i -th galaxy.

Equation 2.5 is substantially a weighted mean where the redshift probability distribution of each galaxy is weighted by its value at the cluster candidate redshift and then shifted in order to peak there. The typical galaxy redshift distribution can be computed as an overall property of the whole dataset or differentially by dividing the sample into different sub-samples, for instance, to capture differences in the photometry quality across the dataset (see e.g. [Maturi et al., 2019](#)).

Once the typical redshift probability distribution is defined, the filter constants representing amplitude normalisation, average background contribution, and cluster variance can be written, respectively, as

$$\alpha(z_c) = \int \frac{M_c^2(\boldsymbol{\theta} - \boldsymbol{\theta}_c, m, z_c) q^2(z_c, z)}{N(m, z_c)} d^2 \theta dm dz. \quad (2.6)$$

$$\beta(z_c) = \int \Psi(\boldsymbol{\theta} - \boldsymbol{\theta}_c, m, z_c) q(z_c, z) N(m, z_c) d^2 \theta dm dz = \quad (2.7)$$

$$= \int M_c(\boldsymbol{\theta} - \boldsymbol{\theta}_c, m, z_c) q(z_c, z) d^2 \theta dm dz. \quad (2.8)$$

$$\gamma(z_c) = \int \Psi^2(\boldsymbol{\theta} - \boldsymbol{\theta}_c, m, z_c) q^3(z_c, z) M_c(\boldsymbol{\theta} - \boldsymbol{\theta}_c, m, z_c) d^2 \theta dm dz = \quad (2.9)$$

$$= \int \frac{M_c^3(\boldsymbol{\theta} - \boldsymbol{\theta}_c, m, z_c) q^3(z_c, z)}{N^2(m, z_c)} d^2 \theta dm dz. \quad (2.10)$$

The amplitude normalisation, $\alpha(z_c)$, which was already introduced in Equation 2.3 and enters in

the computation of the filter in Equation 2.4 and therefore in the amplitude (Equation 2.2) is also used, together with the above-defined constant $\beta(z_c)$ in the estimate of the average background to be subtracted in Equation 2.2:

$$B(z_c) = \alpha^{-1}(z_c)\beta(z_c). \quad (2.11)$$

In this way, whenever the galaxy density corresponds to that of the field galaxies, the amplitude is zero.

The third constant, $\gamma(z_c)$, is the one contributing to the estimate of the expected variance of the amplitude, given by background stochastic fluctuations and Poissonian fluctuations generated by cluster members, which is given by

$$\sigma_A^2(\theta_c, z_c) = \alpha^{-1}(z_c) + A(\theta_c, z_c) \frac{\gamma(z_c)}{\alpha^2(z_c)}, \quad (2.12)$$

where the two terms represent the stochastic background fluctuations and the Poissonian fluctuations produced by cluster members, respectively.

When dealing with catalogues of galaxies, the integrals reported above are discretised by running over the i -th elements of the catalogue. For instance, Equation 2.2 can be rewritten as:

$$A(\theta_c, z_c) = \alpha^{-1}(z_c) \sum_{i=1}^{N_{gal}} \frac{M_c(\theta_i - \theta_c, \mathbf{m}_i) p_i(z_c)}{N(\mathbf{m}_i, z_c)} - B(z_c), \quad (2.13)$$

where the redshift dependence of the model is absorbed by the photo- z distribution of each i -th galaxy, $p_i(z)$.

2.2 Input data and model requirements

The quality and accuracy of the results for such a detection algorithm are strongly dependent on a targeted and clean tuning of the input data and models required in this formalism.

This Section aims to briefly introduce the main input ingredients of a typical AMICO run and their importance.

Photometric galaxy catalogue. The primary and most trivial input is the target dataset. In the case of galaxy cluster or group detection, this is the photometric catalogue of galaxies with their locations and properties.

The AMICO algorithm reads each i -th galaxy in the catalogue as a data point $\mathbf{x}_i = [\theta_i, \mathbf{m}_i, p_i(z)]$ identified by sky coordinates, e.g. given in right ascension and declination, θ_i , the array containing all selected galaxy properties $\mathbf{m}_i$, and the photometric redshift probability distribution $p_i(z)$. As

already mentioned, AMICO is capable of dealing with more than one galaxy property at a time, so the property array $\mathbf{m}$ can contain a single magnitude or any other useful available observable one chooses to base the search on.

The redshift probability distribution function in photometric galaxy catalogues is generally represented in two ways. In some cases, the full probability distribution is reported across a set of redshift intervals, while in others, only the mode (or median) redshift and its 1σ uncertainties are provided. The AMICO algorithm includes the possibility to deal with both: either it reproduces the $p(z)$ as reported in the full distribution tables, or it analytically retrieves it in case the redshift information is given in the form: z_{peak} , z_{min} , z_{max} (i.e. the mode/median, and the values at $\pm 1\sigma$). In this latter case, the photo- z probability distribution is computed as a double Gaussian centred in z_{peak} and expressed by

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma_m} \exp\left(-\frac{(z - z_{peak})^2}{2\sigma^2}\right), \quad (2.14)$$

where σ and σ_m are defined as

$$\sigma = \begin{cases} z_{peak} - z_{min}, & \text{if } z < z_{peak} \\ z_{max} - z_{peak}, & \text{if } z > z_{peak} \end{cases} \quad \text{and} \quad (2.15)$$

$$\sigma_m = \frac{z_{max} - z_{min}}{2}. \quad (2.16)$$

In a detection algorithm like AMICO, which is purely based on photometric redshifts, the quality of the galaxy redshift distribution $p(z)$ plays a crucial role in the detection process. The galaxy $p(z)$, encapsulating uncertainties and photometry quality, enters all main quantities computed during the build-up of the optimal filter, presented in Section 2.1. In this formalism, photometric redshifts serve as the link between the analysis of density performed in redshift slices and the 3-dimensional space. As it will be presented below in Section 2.3, the assignment of galaxies to detected candidates is indeed based on a $p(z)$ -weighted contribution.

As in any data-based analysis, the pre-processing and cleaning of input data is crucial. Understanding the target photometric galaxy catalogue is the first step for a successful application of the detection algorithm. Depending on the characteristics of the dataset, pre-processing may include:

- the study of the uncertainties and the possible anomalies of the galaxy photo- z distribution shapes and parameters (e.g. identification and handling of unrealistically peaked or truncated distributions);
- the validation of the star-galaxy separation criteria (e.g. via magnitude vs full-width-half-

maximum (FWHM) diagram);

- the validation of the inaccessible or contaminated areas of the field (i.e. validation of the visibility mask, see the dedicated paragraph below);
- the analysis of the redshift distribution of galaxies;
- the computation of density maps in the target redshift range to spot anomalies and artefacts;
- the study of possible photometric redshift accumulation points (e.g. by computing a preliminary noise model, which traces the overall properties of the sample; see below);
- the inspection of anomalous-photometry sources and/or dropouts in specific filters.

Additionally, to guarantee uniform depth across the studied redshift range and limit depth inhomogeneities over the analysed survey area, a constant maximum magnitude can be imposed on the data. A commonly adopted magnitude cut is identified as the mode of the overall galaxy magnitude distribution. This helps reduce the contamination given by noisy or spurious sources: objects occupying the faint-end tail of the distribution are generally affected by large and unrealistic uncertainties.

Visibility mask. The second element needed by AMICO in input is the visibility mask, essential when working with real data. The visibility mask is a 2-dimensional pixelised map of the targeted field, which permits including in the cluster and group search the information about the footprint of the survey and the areas of the field that are inaccessible, or that were potentially contaminated during the observation phase. Using a visibility mask as input in the detection process has two crucial advantages: it can be used to filter out unsafe sources, excluding them from the input catalogue, and it allows the detection algorithm to account for masked areas during the detection procedure. Such areas include portions of the field outside of the observed field, haloes and spikes affecting areas in the proximity of bright foreground stars, where pixels saturate, and artefacts and patterns originate, but also ghosts connected to stars and reflections whose location is much less predictable.

In the framework of AMICO, the issue of masked areas can be handled by simply limiting the integrals of the filter constants α , β , and γ to the unmasked (safe and available) fraction of the surveyed area. The amplitude A is then computed in a consistent way, by considering the same galaxy distributions but by taking into account the accessible fraction of the area of coverage. In this way, the relative signal-to-noise ratio ends up being suppressed whenever the amplitude is derived in a partially masked region, compared to an unmasked one (see e.g. [Bellagamba et al., 2018](#)).

Normally, in a photometric galaxy catalogue, all sources are equally listed, but they are assigned a flag to help the catalogue user filter out affected sources. If a visibility mask is not

directly included in the target catalogue release or if it does not respond to the requirements for the cluster search, this can be generated directly from the data, for instance, making use of the provided flags. When this is not sufficient, inspection of density maps can help identify artefacts and consequently adjust mask boundaries to remove them.

Cluster model. In the formalism of optimal matched filters, the model is the template enclosing the information we have a priori about the signal we want to target, namely, the shape and properties of what we expect to detect. In the context of cluster or group searches, this corresponds to the expected galaxy distribution within a system as a function of the galaxy properties on which the search is based, and it represents one of the key input components of AMICO.

If we limit ourselves to the case of a single magnitude as galaxy property, m , a generic cluster model can be, for instance, analytically built with two different components, the radial and the magnitude distribution of galaxies:

$$M_c(r, m) = \Theta(r)\Phi(m), \quad (2.17)$$

where r is the distance from the cluster centre, $r = |\theta_i - \theta_c|$ and $\Theta(r)$ and $\Phi(m)$ are the radial and the magnitude distributions, respectively. Moreover, instead of having just one component, one can decide to split the sample into different galaxy populations, like central massive galaxies and satellites, quiescent and star-forming galaxies, and so forth. In this case, the different models representing each population are summed together. The expected distribution of galaxies within a cluster or group also depends on the system redshift z_c , which has so far been left implicit. In this framework, the redshift dependence is accounted for through the use of the galaxy redshift probability distribution, $p(z)$.

In practice, a cluster model template for AMICO is a cube that contains information about the galaxy density as a function of radius and galaxy properties, such as the magnitude, for each redshift slice, given a chosen redshift resolution. The units of the model are $\text{mag}^{-1} \text{deg}^{-2}$.

To select suitable functions for the radial and magnitude distributions, one can generally rely on existing analyses of cluster samples in which the properties of galaxy populations have been characterised, taking into account the chosen magnitude band and ensuring consistency with the relevant redshift range.

Noise model. As introduced in Equation 2.1, another element that needs to be modelled in the formalism of optimal filtering is the noise, $N(m, z)$. This component represents the properties of galaxies belonging to the field.

Within AMICO, the noise can be directly extracted from the dataset, and it is generally approximated to the overall galaxy properties of the target sample. This assumption is valid when the studied field is sufficiently large that the contribution of cluster galaxies to the noise is

negligible. Each galaxy is assigned to the corresponding magnitude–redshift bin, after which the counts are normalised by the bin volume and by the unmasked (i.e. effective) area.

The AMICO noise model is therefore 2-dimensional in the case of magnitude only and contains information about the galaxy density binned in redshift and magnitude. In this way, the noise is given by random counts, and therefore it has a constant mean over the area at a given redshift.

As shown in Equation 2.4, and as a consequence in all the filter constants, the noise enters at the denominator. To prevent numerical divergence, the imputation of a minimum threshold for the noise value is required in empty pixels, namely pixels with no galaxies. This concept, and the role of the noise in the computation of filter constants, plays a crucial role in the initialisation of an AMICO run in peculiar survey configurations such as that discussed in this Thesis, as it will be discussed in more detail.

2.3 Detection workflow

AMICO computes amplitude estimates on a 3-dimensional grid of sky positions and redshift, called *amplitude map*. After this phase, the AMICO cluster and group detection starts by selecting the candidates.

Given the mathematical formalism introduced in Section 2.1, the Gaussian likelihood for a candidate to be a real object can be computed as

$$\mathcal{L}(\theta_c, z_c) = - \int \frac{D_{obs}(\theta, m, z) - D(\theta, m, z)}{N(m, z_c)} = \quad (2.18)$$

$$= \mathcal{L}_0 + A^2(\theta_c, z_c)\alpha(z_c) \quad (2.19)$$

with $\mathcal{L}_0$ being a negative constant over the grid of coordinates. The two datasets D represent the observed galaxies and the galaxy data as modelled in Equation 2.1, respectively.

Intuitively, the largest structures produce the greatest increase in likelihood when the model is centred on them, so that amplitude maxima coincide with likelihood maxima at a given redshift. However, the quadratic dependence in Equation 2.19 removes information about the sign of the amplitude. As a result, high likelihood values are obtained for both galaxy overdensities and underdensities, preventing the likelihood alone from distinguishing between positive and negative amplitudes. Choosing to rely on the amplitude to map for the detection is what makes the optimal filter linear to the data (Bellagamba et al., 2011).

In AMICO, the signal-to-noise ratio (S/N), is simply defined as

$$S/N = \frac{A}{\sigma_A}, \quad (2.20)$$

where σ_A is the square root of the amplitude variance, σ_A^2 , defined in Equation 2.12. An alternative

definition of the signal-to-noise ratio, which will be referred to as S/N_{nocl} is the signal-to-noise ratio without shot-noise from cluster/group members in the amplitude variance ("nocl" stands for "no cluster contribution"). Therefore, the signal-to-noise ratio without cluster/group member contribution to the variance is expressed by

$$S/N_{nocl} = \frac{A}{\sqrt{\alpha^{-1}}}, \quad (2.21)$$

where α is the amplitude normalisation (Equation 2.6).

After computing the amplitude map for the analysed field, AMICO selects galaxy overdensities, namely cluster and group candidates, at the peaks of the map with the highest signal-to-noise ratio. The centres of the candidate detections are chosen on the amplitude grid and are not constrained to galaxy locations.

In order to generate a catalogue of detections starting from a 3D amplitude map, AMICO works in an iterative way, through a process of selection, probabilistic membership assignment, and cleaning. The number of detected clusters N_{det} depends on the chosen minimum signal-to-noise ratio, $(S/N)_{min}$, which stops the detection iteration and is set as an input parameter at the start of each run.

The iteration proceeds over the j -th detections (1 to N_{det}) as follows: pixels with S/N below the threshold are first discarded; the candidate is then identified at the peak pixel among the remaining ones, and its properties are stored. A membership probability is assigned to nearby galaxies, after which the map is cleaned to remove the detection imprint. Finally, the output quantities (like likelihood, amplitude, and variance) are updated, and the process continues until $(S/N)_{min}$ is reached. In the following paragraphs, we describe the main steps of the detection workflow in more detail.

Probabilistic membership assignment. As part of the iterative detection workflow, AMICO performs the assignment of a membership probability to galaxies surrounding the candidate detections. Given the j -th detection with amplitude A_j , located in (θ_j, z_j) , the probability $P_{i,j}$ of the i -th galaxy belonging to the j -th system is given by the ratio of the galaxy density of the cluster component alone to the total galaxy density, namely

$$P_{i,j} \equiv \frac{A_j M_j(\theta_i - \theta_j, m_i) p_i(z_j)}{A_j M_j(\theta_i - \theta_j, m_i) p_i(z_j) + N(m_i, z_j)}, \quad (2.22)$$

where the main quantities defined in Sections 2.1 and 2.2 enter, like the cluster model, the noise, the amplitude and the galaxy redshift probability distribution.

However, in reality, one galaxy can have a non-zero probability of being part of more systems. To make this possible, without knowing a priori the location of all detections, the field probability P_F is introduced. The field probability has an initial value of 1, which decreases at each association

by subtracting the specific contribution of each galaxy association. The field probability is exploited as a scaling factor to account for how much the galaxy is "still available" for further assignments. Thus, Equation 2.22 can be rewritten as

$$P_{i,j} = P_{F_i} \frac{A_j M_j(\boldsymbol{\theta}_i - \boldsymbol{\theta}_j, \mathbf{m}_i) p_i(z_j)}{A_j M_j(\boldsymbol{\theta}_i - \boldsymbol{\theta}_j, \mathbf{m}_i) p_i(z_j) + N(\mathbf{m}_i, z_j)}. \quad (2.23)$$

This always guarantees that for each galaxy $P_{i,j} \leq 1$, and that the sum of all association probabilities and the field probability of a given galaxy yields 1.

Map cleaning. The last step of the detection workflow of AMICO is the cleaning process. This step's goal is to remove the imprint of sequentially detected objects from the map, in order to allow the detection of smaller and blended structures.

The signal subtraction is performed by using the information about the probability of association described above as a weight for the contribution of the member galaxies. Thus, the new amplitude at the redshift of the k -th detection, resulting from the subtraction of the j -th one is recomputed as

$$A_{clean}(\boldsymbol{\theta}_j, z_k) = A(\boldsymbol{\theta}_j, z_k) - \sum_{i=1}^{N_{gal}} P_{i,j} \frac{M_c(\boldsymbol{\theta}_i - \boldsymbol{\theta}_j, \mathbf{m}_i) p_i(z_k)}{N(\mathbf{m}_i, z_k)}. \quad (2.24)$$

From a computational perspective, the cleaning step is the most resource-intensive stage of the workflow, as the map must be updated after each detection. Nevertheless, it substantially enhances detection performance, especially in crowded fields (Bellagamba et al., 2018). Figure 2.1 shows an example of how the cleaning works: the same slice of the original amplitude map (left) and the cleaned one (right) with some of the main structures which have been removed.

Output products and mass proxies. The principal output product of a typical AMICO run is the detection catalogue of candidate clusters or groups, accompanied by the list of candidate members for each detection. The produced catalogue includes detection properties. Among these properties, AMICO provides as output three different proxies of the mass of the cluster or group:

- the amplitude, A , the natural output of the filtering process given by Equation 2.2. As already mentioned, it scales with the galaxy content of the detection, giving a proxy of mass in units of the adopted cluster model;
- the apparent richness, λ , that represents an estimate of the visible number of galaxies;
- the intrinsic richness, λ_* , that represents an estimate of the visible number of galaxies inside R_{200} and with a z -dependent cut in magnitude.

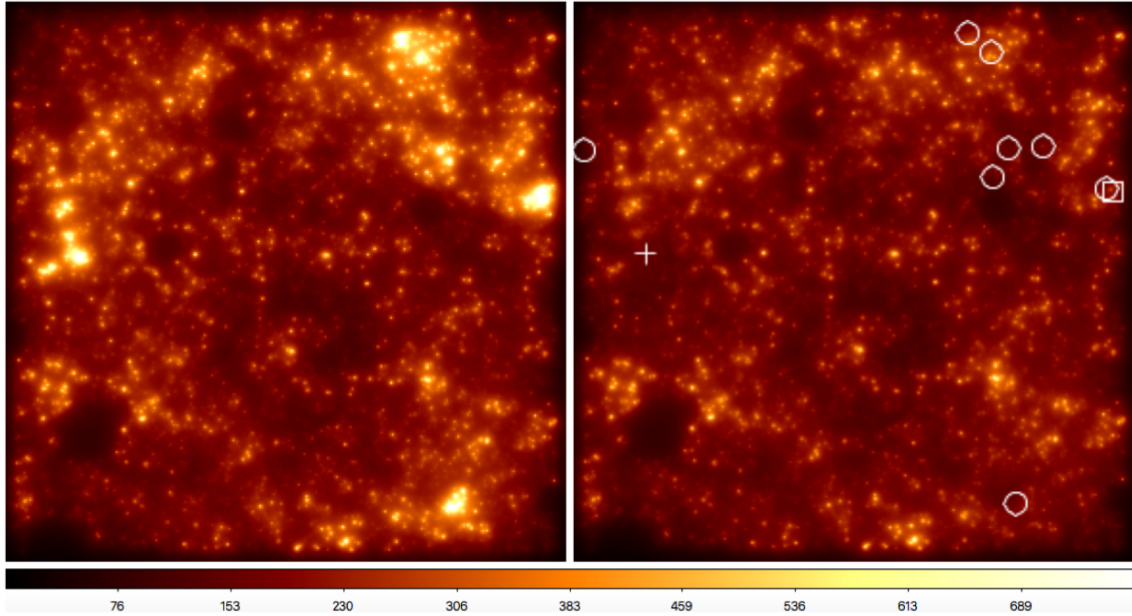


Figure 2.1. Slice of the amplitude map (left) and the corresponding cleaned amplitude map (right) for a test performed on COSMOS data at $z = 0.465$. In AMICO, the imprint of detected structures is removed from the map using the association probabilities to detect more easily blended systems. The white square in the right panel marks the first (highest-amplitude) removed detection, the cross indicates the second, and circles highlight the remaining main detections contributing to the amplitude in this slice. Both maps are shown with the same colour scale for direct comparison.

The apparent richness, λ , is defined as the sum of all probabilities of each i -th galaxy associated to a given j -th detection:

$$\lambda_j = \sum_{i=1}^{N_{gal}} P_{i,j}, \quad (2.25)$$

where $P_{i,j}$ is the membership probability expressed by Equation 2.23. This proxy is related to a direct observable, but it is redshift-dependent since the number of visible galaxies decreases with distance. The intrinsic richness follows the same approach, but with added constraints on magnitude and distance from the cluster/group centre:

$$\lambda_{*,j} = \sum_{i=1}^{N_{gal}} P_{i,j} \quad \text{with} \quad \begin{cases} m_i < m_*(z_j) + 1.5 \\ r_i < R_{200}(z_j) \end{cases}, \quad (2.26)$$

being z_j the redshift of the candidate. Here, $m_*(z)$ is the characteristic magnitude, one of the fundamental parameters of the luminosity function (see Equation 1.2), and it evolves with redshift. This function is part of the parameters to be chosen when building the cluster model as described in Section 2.2. Another parameter which is part of the model chosen for each run is R_{200} , which is the model radius. The definition of the cluster richness in Equation 2.26 is similar to the one implemented in redMaPPer (Rykoff et al., 2014), a widely used red-sequence-based detection algorithm.

Although all three quantities provide estimates of the galaxy content of a detected structure, key differences distinguish them from one another. The apparent richness, λ , is strongly dependent on redshift: more distant systems have fewer observable galaxies, given that the magnitude threshold is set by the survey depth limit. In contrast, the intrinsic richness is defined in a way that renders it nearly independent of redshift, providing a robust measure for inferring cluster and group masses across a wide range of redshifts. These trends remain valid as long as the magnitude cut $m_* + 1.5$ lies below the limiting magnitude of the dataset, which generally holds for deep and ultra-deep surveys within a wide redshift range, but can affect shallower wide survey configurations. Similarly, the amplitude offers a mass proxy that exhibits only weak redshift dependence. However, unlike λ_* , the amplitude is not directly tied to a direct observable, and it can be influenced by the presence of bright galaxies or very peaked redshift distributions, boosting the signal, due to the magnitude and redshift weighting in the filter.

2.4 Previous applications of AMICO

The AMICO algorithm has already been successfully applied to wide-field surveys, testing fields and mock catalogues. Thanks to its performances, it has been selected as one of the two official algorithms adopted for cluster detection by the Euclid mission (Laureijs et al., 2011; Euclid Collaboration: Adam et al., 2019), and it is planned to be exploited for other major future wide surveys like J-PAS and LSST.

One of the most successful past applications of the algorithm is the one to the Kilo-Degree Survey Data Release 3 (KiDS¹; de Jong et al., 2017), which gave origin to a cluster sample (AMICO-KiDS-DR3; Maturi et al., 2019) already used for several cosmological (e.g. Giocoli et al., 2021; Ingoglia et al., 2022; Lesci et al., 2022a; Romanello et al., 2024) and cluster galaxy population studies (e.g. Radovich et al., 2020; Puddu et al., 2021; Castignani et al., 2022a, 2023). The AMICO-KiDS-DR3 catalogue is the first cluster catalogue built with AMICO on the KiDS data and it exploited only the spatial and luminosity distribution of galaxies, without relying on colour information, to avoid red-sequence biases. Besides creating the catalogue, the aim was to provide a robust characterisation of its selection function, purity, completeness, and measurement uncertainties. To this end, the authors have generated realistic mock catalogues directly from KiDS data, incorporating observational effects such as photometric redshift errors, masks, and galaxy clustering. The resulting sample contains 7988 candidate clusters with $S/N > 3.5$ in the redshift range $0.1 < z < 0.8$ and over an effective area of 377 deg^2 . The purity is estimated to be $\sim 90\%$ across the full redshift interval. The analysis includes comparisons with GAMA spectroscopy (Driver et al., 2011) and with external catalogues (e.g. redMaPPer and Planck SZ), showing a clear correlation between different proxies of mass. The AMICO-KiDS-DR3 catalogue was then complemented by a weak lensing stacked analysis of 7,000 of the detected clusters up to $z = 0.6$, which allowed estimates of the mean halo masses and the construction of the relation with AMICO's mass proxies (Bellagamba et al., 2019; Sereno et al., 2020). Thus, this high-purity cluster catalogue enabled detailed studies of cluster galaxy populations as well as robust cosmological analyses.

Giocoli et al. (2021) analysed the weak-lensing signal beyond the virial radius, refining cluster mass estimates and estimating Ω_M . Competitive constraints on Ω_M and σ_8 were subsequently obtained from cluster abundances and stacked weak-lensing profiles (Lesci et al., 2022a) and from clustering (Lesci et al., 2022b; Romanello et al., 2024). Building on this, Ingoglia et al. (2022) measured the halo bias and constrained σ_8 , again via stacked weak-lensing analysis, while Busillo et al. (2023) used extreme-value statistics of the high-mass end to derive independent constraints on σ_8 . Additionally, Giocoli et al. (2024) measured the splashback radius from stacked lensing profiles, finding it consistent with R_{200} and potentially sensitive to optical selection effects.

¹<http://kids.strw.leidenuniv.nl/>

On the other hand, the properties of the cluster galaxy populations and their build-up and evolution in time were analysed by [Radovich et al. \(2020\)](#), while the extraction and evolution of the luminosity function were presented in [Puddu et al. \(2021\)](#). Moreover, a campaign of observations in the millimetric wavelengths to study the gas reservoir and environmental processing of gas in central star-forming BCGs was also carried out and presented by [Castignani et al. \(2022a, 2023\)](#).

Recently, the new AMICO analysis of the Fourth Release of the KiDS data ([Kuijken et al., 2019](#)) has been presented in [Maturi et al. \(2025\)](#). With respect to the KiDS-DR3 catalogue, the KiDS-DR4 (or KiDS-1000) AMICO cluster sample was retrieved on a significantly larger survey area ($\sim 839 \text{ deg}^2$), spanning $0.1 \leq z \leq 0.9$, with improved photometric redshifts thanks to additional near-infrared VIKING bands ([Edge et al., 2013](#)). The cluster sample comprises almost 24,000 detections ($S/N > 3.5$) and it is optimised for cosmological analysis over a wide tiled area. This time, the data-driven selection function was estimated within a blinding scheme, which serves as a protective measure against any subjective influences and therefore increases the robustness of the cosmological analysis. Upon unblinding, the cosmological constraints obtained from counts and weak-lensing were released ([Lesci et al., 2025](#)), revealing halved uncertainties, excellent agreement with recent independent measurements using KiDS-1000, and an intriguing $\sim 3\sigma$ tension with Planck CMB results.

The AMICO algorithm is also planned to be used for the $8,000\text{-deg}^2$ Javalambre Physics of the Accelerating Universe Astrophysical Survey (J-PAS; [Benitez et al., 2014](#)), for which performances have been assessed by applying the algorithm to the testing field of J-PAS, called miniJPAS ([Bonoli et al., 2021](#)). The results of this test have been presented in [Maturi et al. \(2023\)](#). By applying AMICO to $\sim 1 \text{ deg}^2$ (with redshift in the range $0.05 < z < 0.8$), the authors have detected systems with masses down to $\sim 10^{13} M_{\odot}/h$, finding 80, 30, and 11 clusters above S/N thresholds of 2.5, 3.0, and 3.5. By cross-matching with X-ray data, they calibrated mass–observable relations and showed that a stellar-mass–based proxy reduces scatter relative to richness. The narrow-band dataset yields redshift uncertainties as low as $\sigma = 0.0037(1 + z)$, placing miniJPAS between broadband and spectroscopic surveys. Extrapolation to the full $8,000 \text{ deg}^2$ J-PAS survey predicts $\sim 2 \times 10^5$ detections above $S/N = 3$, establishing J-PAS as a powerful resource for cluster cosmology.

2.5 Selection for the Euclid mission

As part of the preparation-to-launch phase of the Euclid mission, the Euclid Consortium² has proposed a challenge with the goal of comparing several candidate algorithms for cluster detection and selecting the best one(s), which would then be included in the official Euclid pipeline. The paper, [Euclid Collaboration: Adam et al. \(2019\)](#), reports the final results of this Cluster Finder Challenge (CFC), a blind, head-to-head evaluation of six optical/NIR cluster-finding algorithms on

²<https://www.euclid-ec.org/>

Euclid-like mocks. The goal is to select robust and complementary methods for the Euclid wide photometric survey ($\sim 15,000 \text{ deg}^2$, limited to $H \approx 24$), capable of delivering high completeness and purity in the range of interest (which is for masses above $10^{14} M_\odot$ and up to $z \sim 2$), with accurate redshifts, centring, and low-scatter mass proxies suitable for cosmology. The challenge used a 300 deg^2 Euclid light-cone (Merson et al., 2013; Ascaso et al., 2015) built on the Millennium simulation (Springel et al., 2005), reprocessed to impose Euclid+ground-based depths and realistic spectral energy distributions (SEDs). Photometric redshifts were derived with BPZ (Benítez, 2000), and a 20 deg^2 calibration field (with spec-zs) was provided to the participants. The mock is $H \leq 24$ limited and includes haloes down to $M_{200} \sim 10^{13.25} M_\odot$; performance is thus conservatively assessed for $M > 10^{14} M_\odot$ where Euclid will drive cosmology. The team also measures the mock’s intrinsic mass–richness relation to debias scatter estimates introduced by detections.

Six codes were participating: AMASCFI (an adaptive kernel with photo- z slicing, Adami & Mazure, 1999; Sarron et al., 2018), AMICO (described in this Chapter), HCFA (a hierarchical finder; Díaz-Sánchez, in prep.), PZWav (an adaptive wavelet; Gonzalez, 2014; Werner et al., 2023a; Thongkham et al., 2024), sFoF (a Friends-of-Friends; Farrens et al., 2011), and WaZP (a wavelet; Benoist, 2014). Each algorithm was able to return positions, redshifts, a ranking (usually S/N), and, for some, membership probabilities and richness/amplitude-like mass proxies. The evaluation consisted of matching detections to true haloes using multiple, tuned procedures (like cylindrical matching with membership-based cross-check when available) to quantify completeness, purity, redshift and centre accuracy, mass-proxy scatter, and fragmentation/over-merging. Importantly, by varying the matching parameters, it was possible to further constrain systematics in completeness to $< 3\%$, providing a robust baseline for algorithm selection.

For comparison, the catalogues were cut to fixed thresholds, mainly a mean purity $P = 80\%$ for $M > 10^{14} M_\odot$, $z < 2$ (and vice versa, completeness cuts $C = 80\%$ to evaluate the purity). At fixed purity: AMICO reaches $C = 91\%$, PZWav and WaZP $C = 83\%$. At fixed completeness: AMICO maintains $P = 93\%$, PZWav $P = 86\%$, WaZP $P = 83\%$. sFoF, AMASCFI, and HCFA follow, with lower purity and completeness. Redshift precision analysis yields good results across methods, with AMICO and WaZP having $\sigma_z/(1+z) = 0.015$, and PZWav $\sigma_z/(1+z) = 0.020$. Mass-proxy scatter is competitive: AMICO 0.19 dex, PZWav 0.16 dex, WaZP 0.18 dex (HCFA is 0.15 dex but at lower completeness). They also map completeness vs redshift in mass slices, showing that both AMICO and PZWav retain high performances for $10^{14} - 10^{14.5} M_\odot$ up to $z \sim 1.5 - 2$. Purity as a function of completeness for the six participating algorithms is shown in Figure 2.2.

The strengths of the AMICO algorithm, as demonstrated in the CFC, include: high completeness at fixed purity, best redshift precision, low centre offsets, and competitive mass-proxy scatter, all sustained across the whole mass-redshift range of interest. These characteristics, together with the probabilistic membership assignment, led to AMICO being officially chosen as a Euclid cluster finder. As part of the CFC criteria, catalogue complementarity was also taken into account.

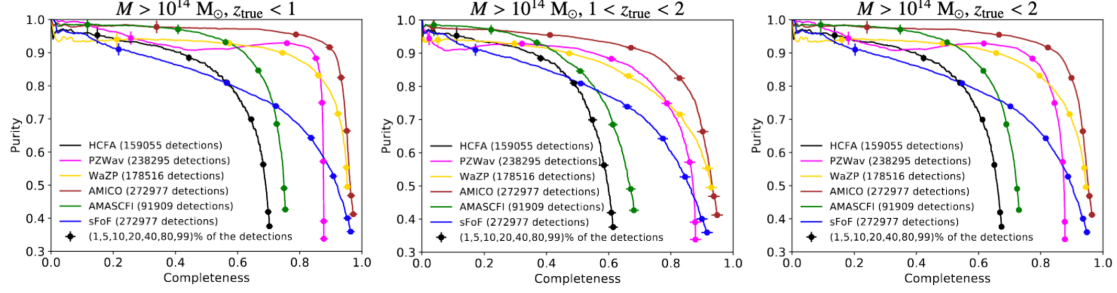


Figure 2.2. Purity–completeness curves for all CFC cluster finder catalogues. Dots mark the fraction of detections at each point, starting from the lowest ranked using the ranking property of each algorithm, with total counts given in the legend. The catalogues are trimmed at $z_{\text{true}} < 2$ and $M > 10^{14} M_{\odot}$, and the panels show results for three redshift bins as shown in the panel titles. Adapted from *Euclid Collaboration: Adam et al. (2019)*.

While AMICO delivers high overall performances alone, combining it with a second, structurally different method helps capture a larger variety of cases (e.g. morphology variations, photo- z tails). PZWav (followed by WaZP and sFoF) provides the largest incremental completeness over AMICO at high-mass/high- z , motivating its co-selection. PZWav is an adaptive wavelet method that enhances cluster-scale overdensities in redshift-sliced density maps using scale-space analysis and significance control. Its different signal model makes it complementary to matched filtering, and it performs particularly well in regimes where multi-scale structure or local backgrounds can complicate the use of filters. In the CFC, it achieved high completeness and purity, and good overall performance, especially competitive at the high- z , high-mass end. This balance of performance and complementarity with AMICO makes this the best combination for the scientific goals of the Euclid mission.

Recently, the official cluster detection pipeline of Euclid has been applied for the first time to the Euclid Quick Release 1 (Q1) real data³, mainly for validation purposes. The data pre-processing, the catalogue and the comparison with external data and spectroscopy are presented in *Euclid Collaboration: Bhargava et al. (2025, including GT)*, to which we have dedicated Appendix C.

2.6 Novelty of this Thesis work and newly implemented features

This Thesis project focuses on the application of the AMICO algorithm to the data available in the Cosmic Evolution Survey (COSMOS)⁴ 2-deg² field (Scoville et al., 2007). The main motivation, as anticipated, is to bridge the gap between large statistical cluster samples and in-depth environmental galaxy evolution and structure formation studies, by performing an innovative and challenging application of the detection algorithm up to high redshift. The COSMOS field offers

³<https://www.euclid-ec.org/science/q1/>

⁴<https://cosmos.astro.caltech.edu/>

the ideal opportunity to achieve this. Despite being a deep field, it covers a large area; therefore, it can probe a significantly larger volume with respect to standard "pencil-beam" deep fields, which is optimal for the study of large-scale structure and galaxy formation and evolution up to high redshift. Additionally, the COSMOS field boasts a unique multi-wavelength coverage, spanning the entire spectrum from radio to X-rays (e.g. [Civano et al., 2016](#); [Hasinger et al., 2007](#); [Smolčić et al., 2017](#)). It is among the most observed patches of the extragalactic sky and offers access to high-quality multi-band optical and near-infrared photometry (e.g. [Sawicki et al., 2019](#); [Koeke-moer et al., 2007](#); [Aihara et al., 2022](#); [McCracken et al., 2012](#)), including highly precise photometric redshifts, which are the basis of a robust group and cluster detection in three-dimensional space. COSMOS-based photometric galaxy catalogues, like the COSMOS2009 ([Ilbert et al., 2013](#)), the COSMOS2015 ([Laigle et al., 2016](#)) and the COSMOS2020 ([Weaver et al., 2022](#)) have been a valuable resource for extragalactic studies, particularly in the areas of galaxy formation and evolution, and galaxy environments. More recently, the central portion of the COSMOS field has been observed as part of the largest contiguous James Webb Space Telescope (JWST) program to date, a ~ 270 -hour Cycle 1 campaign covering approximately 0.54 deg^2 , known as the COSMOS-Web survey (PIs: J. Kartaltepe and C. Casey; [Casey et al., 2022, 2023](#)). The combination of JWST's deep near- and mid-infrared imaging with the extensive multi-band legacy of COSMOS has led to the construction of the new COSMOS-Web photometric galaxy catalogue, or COSMOS2025 ([Shuntov et al., 2025a](#)). The COSMOS field footprint showing the regions targeted by the main imaging instruments is displayed in Figure 2.3. Altogether, the availability of deep imaging, the unique multi-wavelength data-richness and the high-quality photometric redshifts provide an ideal opportunity to undertake this new and challenging application of AMICO. In this work, the AMICO algorithm has been applied for the first time to the most recent photometric data available for the COSMOS deep field. The novelty of this application, compared to past ones, resides, for instance, in the redshift extent, the depth, and the amount of area covered, and allowed the creation of two catalogues covering over ~ 12 Gyr of cosmic history and probing halo masses down to a few $10^{12} M_{\odot}$.

Therefore, this analysis was driven by a two-fold purpose. First, it aimed to push the limit of cluster detection into the regime of high-redshift groups and protoclusters and down to low-mass systems, providing the community with a valuable resource for studying structure formation and evolution of galaxies and AGNs within them. At the same time, being this the deepest and highest-redshift application of AMICO to date, it offered an ideal opportunity to benchmark, test, and adapt the algorithm to new and challenging survey configurations.

Compared to the surveys to which AMICO had previously been applied, the combination of small area, ultra-deep imaging, and wide redshift coverage in this work posed new challenges, providing a stringent test of the algorithm's flexibility and performance. With the implementation of algorithm enhancements, together with targeted data-cleaning procedures like the introduction

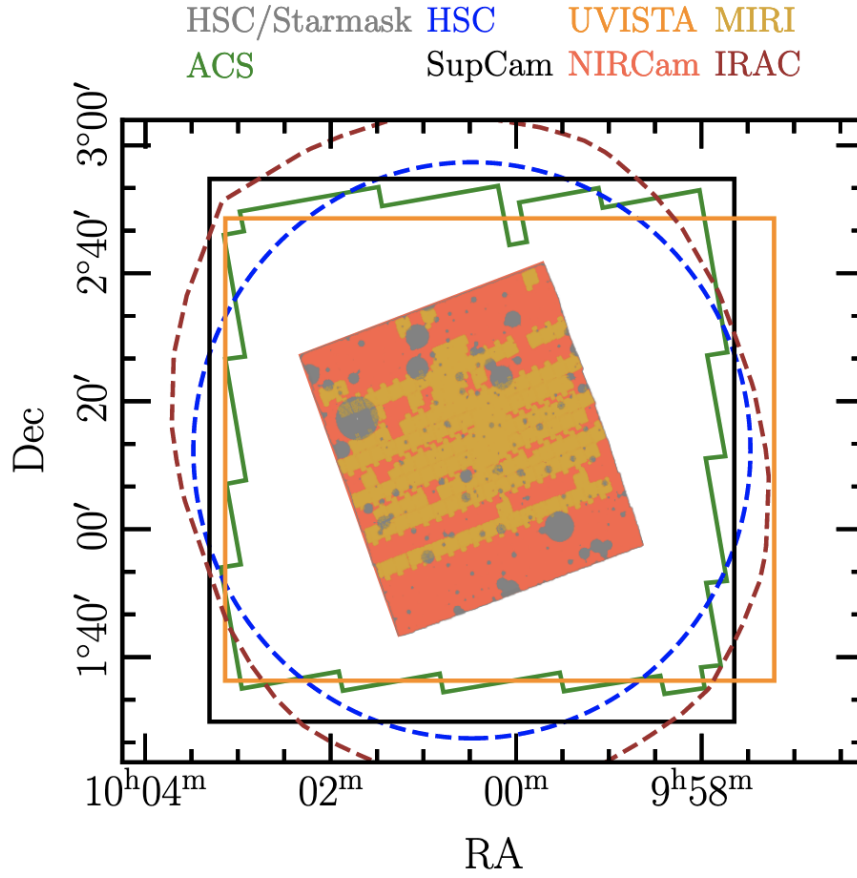


Figure 2.3. Survey footprint of the COSMOS field. Different regions show the footprints of different imaging instruments, colour-coded as in the legend. The JWST footprint from NIRCcam and MIRI is shown in the orange/yellow coloured areas. The grey regions mark the area affected by HSC bright star masks. Credits: [Shuntov et al. \(2025a\)](#)

of a maximum $p(z)$ value to discard too-peaked photo- z s, the algorithm was able to respond effectively to the data and produce robust results. This has opened new perspectives for the use of AMICO over small areas and deep fields. The new features developed and used in this work include:

- a cleaning routine for the cluster galaxy contamination to the noise model. Small areas may be affected by a non-negligible contribution of cluster galaxies when noise is extracted from the overall properties of all galaxies over the field. This results in peaks localised in redshift in the noise model, leading to biased results at those redshifts, where the noise is overestimated due to the presence of physical overdensities. To mitigate this effect, a cleaning method has been developed. This divides the field of interest into non-overlapping portions of roughly the same area (e.g. four quarters). Then, it computes the median of each corresponding pixel in the noise cube across the different portions using a $\kappa\sigma$ clipping; this effectively smooths out the imprint of physical overdensities that are located in a portion of the field but not in the others, without necessarily knowing the position of structures in the field a priori. Figure 2.4 shows different noise models (magnitude on the y-axis and redshift on the x-axis, lighter colours indicate higher galaxy counts in a given bin of magnitude and redshift). Panel (a) shows the noise extracted from the whole galaxy population in the full COSMOS field (from the catalogue used for the run described in Toni et al. 2024 and Chapter 3). In panels (b)–(f), we show the noise computed from 4 different datasets and configurations in the COSMOS-Web field: (b) full field, (c) north half of the field, (d) south half of the field, (e) noise using the cleaning method just introduced on four quarters of the field, (f) noise calculated on the galaxy catalogue by removing members as detected in Toni et al. (2024; Chapter 3). Reducing the analysed area makes the noise more sensitive to single physical overdensities of galaxies, which show up as visible peaks in the noise models. From panel (c), we can see, for instance, that overdensities might be present in the north half of the field at $z \sim 1.65$, while for the south half this seems to be the case for redshifts in the range $z \sim 2.25 - 2.45$. The median cleaning method we developed successfully smooths out these peaks, as visible in panel (e), even more efficiently and without a priori knowledge of the position of clusters and groups in the field, as shown in panel (f);
- the regularisation of the noise, particularly useful for applications of AMICO with low statistics, includes any tuning of the noise values, which helps account for empty pixels (magnitude-redshift bins without galaxies) and limits the integration when computing the filter constants. Starting from the existing imputation of a minimum value of the noise to avoid divergences, efficient regularisation can be achieved by attributing a very large value (easily distinguishable from noise values in the rest of the pixels) to empty pixels and/or to pixels at magnitude fainter than a redshift-dependent cut, based on m_* . The effect of this

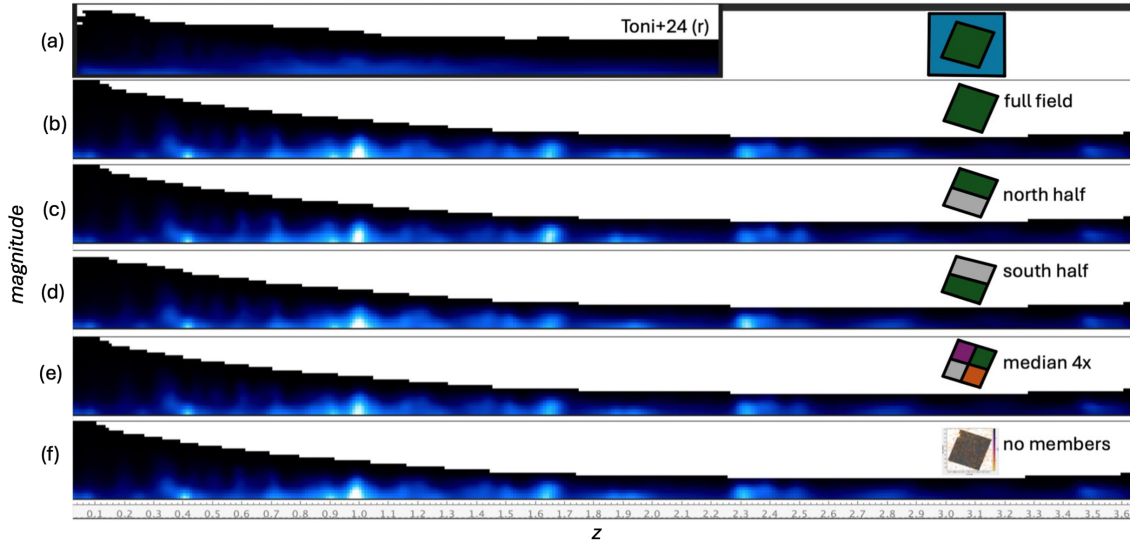


Figure 2.4. Noise as a function of magnitude (y-axis) and redshift (x-axis), where lighter colours indicate higher galaxy counts per bin. Panel (a) shows the noise extracted from the full COSMOS field (see Chapter 3), while panels (b)–(f) display results for different COSMOS-Web configurations: (b) full field, (c) northern half, (d) southern half, (e) median-cleaned quarters, and (f) catalogue with known group members removed. Reducing the field area increases sensitivity to local overdensities, visible as peaks in the noise distribution. The median cleaning method effectively smooths these peaks without prior knowledge of group positions. The figure is for illustrative purposes only and does not represent the final cleaned noise models used for the search described in Chapters 3 and 4.

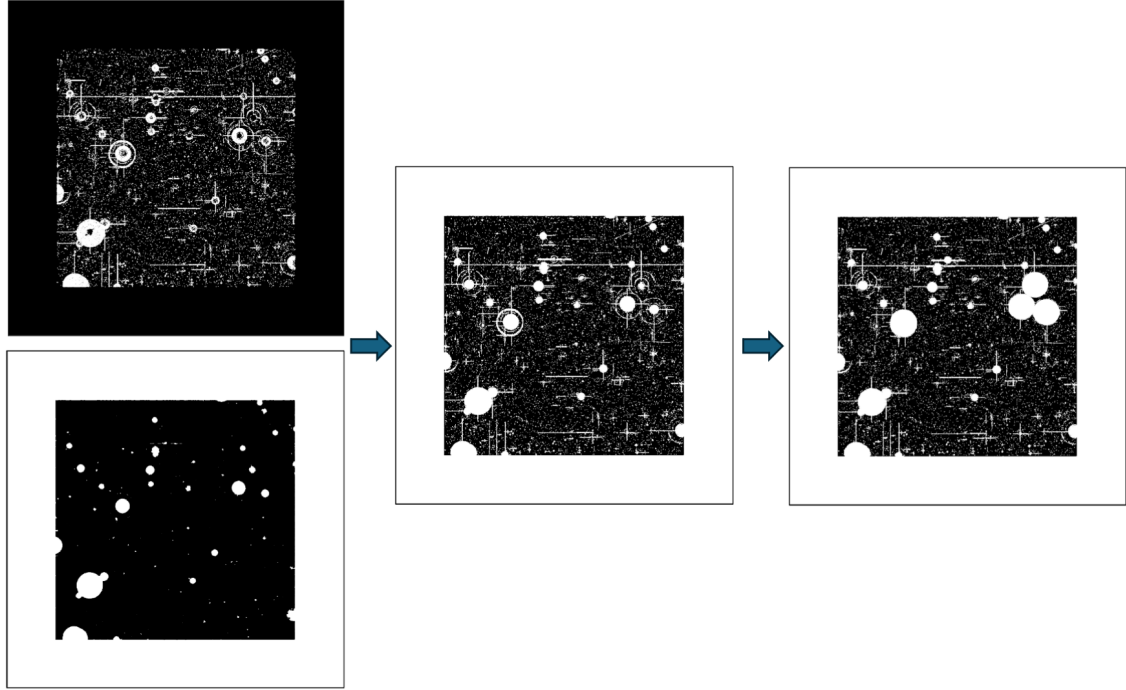


Figure 2.5. Example of constructing a composite visibility mask. The panels show: masking based on catalogue flags for masked objects (upper left) and for missing non-masked objects (lower left), their combination (middle), and the inclusion of additional circular masks around some bright stars to exclude potentially problematic areas (right). The sample used for this example is extracted from the COSMOS2015 catalogue.

regularisation is discussed in more detail in Chapter 3, Section 3.2;

- the possibility to generate a well-suited, adapted composite visibility mask. Rather than relying solely on the survey-provided mask contours, it allows the masking strategy to be fine-tuned to the requirements of a given field and scientific goal. The composite mask can incorporate information from flagged data, provided or custom-made mask contours, and star catalogues. An example of the construction of a composite visibility mask is shown in Figure 2.5. Starting from the catalogue flagging system, one can mask areas with masked objects (upper left panel) or lacking safe objects (lower left panel). The two merged (middle panel) can then be further integrated by polygons and/or circles around a list of stars or customised locations, to be more conservative and mask suspicious areas (right panel). This flexibility is particularly valuable in small, deep fields, where, given the high density of detections, the inclusion of spurious detections or the discard of safe areas can strongly affect the statistical properties of the sample.

Each of the features mentioned above and how they have been used for this specific application

will be discussed in more detail in the description of the detection process for the two catalogues produced in this work, in **Part I**, Chapter 3 and Chapter 4.

I

Detection

3

The AMICO-COSMOS catalogue

A CATALOGUE OF GALAXY CLUSTERS AND GROUPS

UP TO $z = 2$ WITH X-RAY COUNTERPARTS

This Chapter is dedicated to the first cluster and group search performed with the AMICO algorithm in the COSMOS 2-deg² field (Scoville et al., 2007), using the recent COSMOS2015 (Laigle et al., 2016) and COSMOS2020 (Weaver et al., 2022) photometric galaxy catalogues. The goal was to produce a new catalogue up to $z = 2$ by performing an innovative application of AMICO with respect to previous applications to wide-field surveys in terms of depth (down to $r < 26.7$), small area covered (~ 1.69 deg² of unmasked effective area), and redshift extent. This sample and the comparative analysis performed with the X-rays allowed for the calibration of mass-proxy scaling relations up to $z = 2$ and down to less than $10^{13} M_{\odot}$, and constitutes the basis for future applications of AMICO, such as the analysis of upcoming Euclid Deep Field data.

This Chapter is fully based on Toni et al. (2024), A&A, 687, A56^a

^aThe AMICO-COSMOS galaxy cluster and group catalogue with members is available at the CDS via anonymous ftp to [cdsarc.cds.unistra.fr](ftp://cdsarc.cds.unistra.fr) (130.79.128.5) or via <https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/687/A56>

3.1 COSMOS field data

The Cosmic Evolution Survey (COSMOS) (Scoville et al., 2007) has offered access over the years to a unique combination of deep data, with wavelength extension from radio to X-rays (e.g. Koekemoer et al., 2007; Zamojski et al., 2007; Cappelluti et al., 2009; Civano et al., 2016; Smolčić

et al., 2017) and large spectroscopic coverage (e.g. Lilly et al., 2007; Hasinger et al., 2018). This has made this 2 deg^2 field a source of large galaxy samples characterised by photometric redshifts of extremely good quality up to high redshift. The field properties and accessibility have made it ideal for studying the large-scale structure and the formation and evolution of structures in the Universe (e.g. Hung et al., 2016; Cucciati et al., 2018; Laigle et al., 2018).

The photometric galaxy catalogue used for this cluster and group search is the COSMOS2020 catalogue (Weaver et al., 2022), already widely used for galaxy evolution and environmental studies (e.g. Taamoli et al., 2024b; Gould et al., 2023). We nevertheless combined the catalogue with a sample from the previous release, COSMOS2015 (Laigle et al., 2016), due to some potential typical cluster galaxies missing in the COSMOS2020 selected catalogue. We visually inspected different classes of objects flagged as masked and selected a group of galaxies, labelled with `flag_peter = 6` or `4` in the catalogue by Laigle et al. (2016), which are bright extended galaxies and galaxies close to them, potentially located in dense environments and likely belonging to clusters or groups. These galaxies are misclassified, labelled with a general masking flag, are not identified, and have missing or unsafe photometry and/or missing redshift estimates in the newest release. Even if this galaxy sample from COSMOS2015 is statistically small, it contains typical cluster/group members, including several central bright galaxies that might affect the detection of some clusters and groups, especially at low and intermediate redshift.

The COSMOS2015 catalogue comprises more than half a million galaxies extracted from the $zYJHK_s$ image, with data from the Subaru (Taniguchi et al., 2015) and VISTA (McCracken et al., 2012) telescopes via SExtractor dual-image mode (Bertin & Arnouts, 1996). Photometric redshifts, estimated using LePhare software (Arnouts et al., 2002; Ilbert et al., 2006), reach an accuracy better than 1% at $z < 1.2$ when compared with spectroscopic samples.

The main improvement in the COSMOS2020 catalogue with respect to the COSMOS2015 catalogue is the addition of new ultra-deep optical and NIR data from the Hyper Suprime-Cam (HSC) Subaru Strategic Program PDR2 (Aihara et al., 2019) and from UltraVISTA survey Data Release 4 (McCracken et al., 2012; Moneti et al., 2023). The use of new data is reflected in the nearly doubled number of detected sources and in the ability to reach the same photometric redshift uncertainty at nearly one magnitude deeper than in COSMOS2015. Moreover, the addition of the deep HSC- i band to the detection process increases the completeness with respect to the previous release, especially for small blue galaxies (Weaver et al., 2022).

The COSMOS2020 release includes two different catalogues produced with independent extraction methods: (1) the CLASSIC catalogue, which follows the same approach used in COSMOS2015, namely it is created with SExtractor, after point spread function homogenisation, and (2) THE FARMER catalogue, produced by using purely parametric modelling with The Tractor (Lang et al., 2016; Weaver et al., 2023a). Both galaxy catalogues include in turn both photometric redshifts computed using LePhare and EAZY (Brammer et al., 2008). Weaver et al. (2022) presented a de-

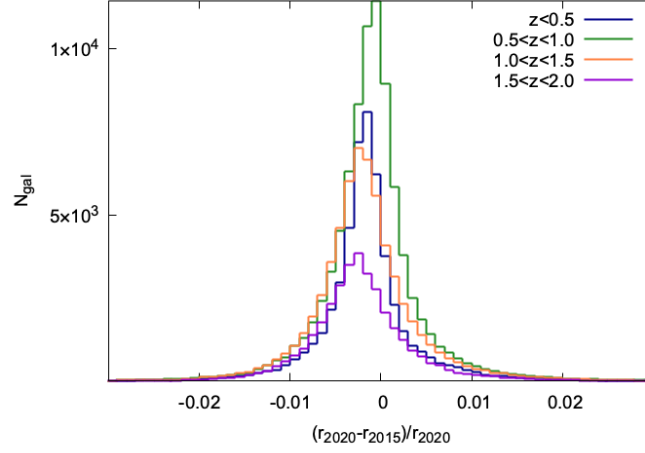


Figure 3.1. Distribution of the relative magnitude scatter in the r -band for the galaxies present both in the COSMOS2020 (r_{2020}) and in the COSMOS2015 (r_{2015}) catalogues. The different colours represent different redshift bins, as labelled in the plot.

tailed comparison between the different options, and showed that a very good level of consistency existed between the two redshift estimation methods and the two extraction methods.

For this work, we chose to use the CLASSIC catalogue in combination with photometric redshifts estimated with LePhare to maximise the number of available galaxies. The CLASSIC catalogue covers a slightly larger effective area than THE FARMER, which does not return photometry in the case of model failure in the extraction process (Weaver et al., 2022). Photometric redshifts estimated using LePhare, also chosen for better consistency with COSMOS2015, are estimated in COSMOS2020 with a configuration similar to that used in Ilbert et al. (2013) and were found to reach per cent accuracy at the brightest i magnitudes up to a maximum value of $\sim 0.02(1+z)$ for $i < 25$, with a catastrophic failure of just a few per cent. The consistency in the source extraction and the redshift computation between COSMOS2020 and COSMOS2015 became crucial since we integrated the selected galaxy catalogue of the newest release with a sample of galaxies from COSMOS2015. In addition to choosing the option with consistent photometric redshift estimation methods, we studied the consistency of the releases in terms of magnitude. We found a slight bias increasing with redshift, which is nevertheless negligible within the analysed redshift range. For most of the galaxies present in both catalogue releases, the relative magnitude difference in the two galaxy samples is less than one per cent, as shown in Figure 3.1 for the r -band. Within the studied redshift range, this magnitude difference is negligible with respect to the resolution of the cluster model, as will be described in Section 3.2.

We performed three independent AMICO runs using three different magnitudes as galaxy properties. Among the available bands, we chose to use HSC- r , UltraVISTA- Y , and UltraVISTA- H . We considered only the UltraVISTA portion of the COSMOS field (see the footprints in Figure

Table 3.1. Selected input galaxy catalogues for the three analyses with the total number of galaxies taken from both COSMOS releases ($N_{gal,TOT}$) and those from COSMOS2015 ($N_{gal,2015}$) in the range $0 < z < 2$.

<i>Instrument – band</i>	<i>depth</i>	$N_{gal,TOT}$	$N_{gal,2015}$
<i>HSC – r</i>	< 26.7	450984	1854
<i>UltraVISTA – Y</i>	< 26.1	442449	934
<i>UltraVISTA – H</i>	< 25.6	407067	883

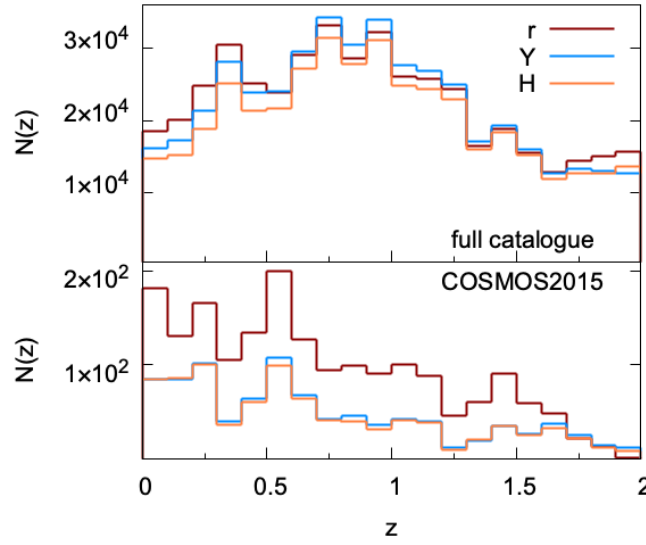


Figure 3.2. Distribution of selected galaxies as a function of redshift for the three runs in r -, Y -, and H -band magnitudes in the range $0 < z < 2$. *Top panel:* Full catalogues including selected galaxies of COSMOS2020 and the insert from COSMOS2015. *Bottom panel:* Only galaxies belonging to the insert from COSMOS2015.

2.3), namely with $RA[deg] \in [149.30, 150.79]$ and $Dec[deg] \in [1.60, 2.81]$, where the bands used for source extraction are available as well as the bands chosen for this analysis. We searched for clusters and groups in the unmasked fraction of this area (see Section 3.2), for a final effective area of 1.69 deg^2 . We selected three final galaxy catalogues, including all unmasked galaxies with available magnitude and photometric redshifts. We discarded galaxies with anomalies in the redshift probability distribution, $p(z)$, or with unrealistically peaked distributions that might bias the estimate of the candidate properties. We performed a magnitude cut on the mode of each magnitude distribution, defining the depth of the galaxy catalogues, as reported in Table 3.1. As a reference, the r -band galaxy catalogue we used extends to almost three magnitudes deeper than the catalogue used in the AMICO-KiDS-DR3 cluster search (Maturi et al., 2019). The distribution in redshift of the selected galaxies in the three input catalogues is shown in Figure

3.2 for $0 < z < 2$, where the top panel refers to the full input catalogues, while the bottom panel refers to the galaxies added from COSMOS2015 only. Table 3.1 summarises the properties of the input galaxy catalogues for $0 < z < 2$, with the total number of galaxies from both releases (total number of unmasked galaxies in the input catalogue) and the number of galaxies among these which are taken from the COSMOS2015 sample.

3.2 Applying AMICO to the COSMOS data

We applied the AMICO algorithm to the COSMOS field by running three different and independent analyses, each using one of the selected magnitudes. Even though AMICO can deal with an arbitrary number of quantities describing galaxy properties, we preferred to use one magnitude at a time, therefore without making explicit use of colour as a galaxy property. This reduces the possibility to bias the selection for the presence or absence of the cluster red-sequence. We used the following bands: HSC-*r*, UltraVISTA-*Y*, and UltraVISTA-*H*, down to a magnitude of 26.7, 26.1, and 25.6, respectively, as reported in Table 3.1. In all the analyses, each galaxy of the input catalogue is considered as a data point with ID, sky coordinates, a single magnitude, and an analytical photometric redshift probability distribution, $p(z)$, built with a Gaussian distribution characterised by z_{peak} and z_{min}, z_{max} , namely the 1σ errors, as given in the galaxy catalogue.

Cluster model. The cluster model describes the signal we expect to detect, namely the expected distribution of cluster or group galaxies in position and magnitude as a function of redshift. This can be constructed analytically from a radial profile and a luminosity function, as expressed by Equation 2.17. For the radial distribution, we used an NFW profile (Navarro et al., 1997) with parameters from the scaling relation presented by Hennig et al. (2017), where the properties of a sample of 74 SZ-selected massive clusters detected within the overlap between DES (The Dark Energy Survey Collaboration, 2005) and the SPT survey (Story et al., 2013) were studied. The sample extends up to $z \sim 1.1$. The estimate for the concentration parameter we adopted from this study is the mean full-population value $c = 3.59$, and the typical virial mass of the model was chosen to be $M_{200} = 10^{14} M_{\odot}/h$.

The luminosity function was assumed to follow a Schechter function (Schechter, 1976). The faint-end slope, α , was adopted as estimated by Zenteno et al. (2016) who analysed the 26 most massive clusters of the SPT survey sample and found a mean value for the full galaxy population of $\alpha = -1.06$.

The characteristic magnitude, m_* , and its redshift evolution were derived by evolving a typical massive elliptical galaxy with evolutionary synthesis models via the GALEV interface (Kotulla et al., 2009). For this model, we relied on a Kroupa initial mass function (Kroupa, 2002) and adopted a chemically consistent approach (Kotulla et al., 2009) for a massive elliptical formed at

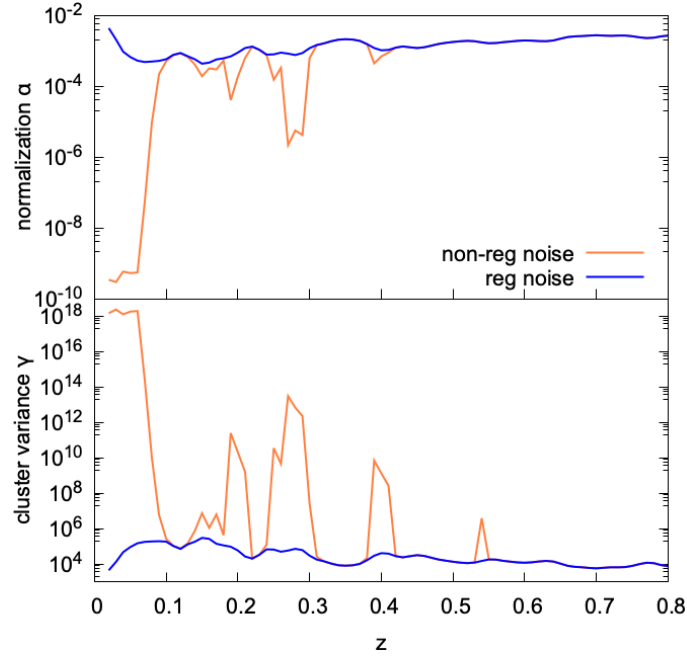


Figure 3.3. Effect of noise regularisation at $z < 0.8$ on the redshift trend of two filter constants: the amplitude normalisation, α (top panel; see Eq. 2.6), and the cluster variance, γ (bottom panel; see Eq. 2.9). The orange line represents a typical AMICO run on COSMOS data, using the Y -band magnitude, without noise regularisation, while the blue line represents the same run with the noise assessment described in the text.

$z = 8$ with a high-redshift, exponentially declining star formation burst. We did this following the same approach as in [Castignani et al. \(2022a, 2023\)](#), and consistently with formation scenarios and passively evolving stellar populations in massive elliptical galaxies (e.g. [Holden et al., 2004](#); [De Lucia et al., 2006](#); [Skelton et al., 2012](#); [Castignani & Benoist, 2016](#)). The cluster model we built for this search has a magnitude resolution of $\Delta m = 0.5$.

Noise model. The noise model, which accounts for field galaxies, can be approximated by the overall properties of the galaxy sample. This holds true under the assumption of negligible contribution of cluster/group galaxies. Although we observed overdensities localised in redshift in a noise model produced with the input COSMOS galaxy catalogues, we found that the cause of these peaks is not attributed to physical overdensities (i.e. cluster/group galaxies), whereas it is likely due to photometric redshift accumulation points. The negligibility of cluster/group contribution to the noise for the COSMOS field was proven in two independent ways: (1) by removing galaxies attributed to groups with probability $> 50\%$ up to $z = 1$ according to the membership catalogue produced by [George et al. \(2011\)](#) and (2) by taking the median of each noise pixel from non-overlapping portions of the field in order to attenuate the imprint of localised overdensities. Neither of these methods attenuated the observed peaks.

Given the small area covered by this search and the small statistics of the galaxy sample with respect to past applications of AMICO to wide-field surveys, a noise regularisation was needed in order to get robust results. To regularise the noise, we attributed an arbitrarily large value of noise to pixels without the contribution of any galaxy. This was necessary to make the integration space finite when computing the filter constants (see Eqs. 2.6, 2.7, and 2.9), and was empirically found to successfully make the filter constants more robust.

Figure 3.3 shows the influence that the performed regularisation has on two filter constants in comparison with a non-regularised noise model retrieved from COSMOS data sets. The regularisation of the noise successfully smooths out the largest fluctuations of the filter constants.

Masked areas. We created a visibility mask based on the input galaxy catalogue as follows. First, we masked the pixels where saturated stars fall, using the same sample [Weaver et al. \(2022\)](#) used for COSMOS2020. The sample is part of the Incremental Data Release of the HSC bright-star masks by [Coupon et al. \(2018\)](#), extracted from the Gaia DR2 ([Gaia Collaboration et al., 2018](#)), with magnitude threshold $G < 18$. We masked different areas depending on the magnitude of stars in this range by simply dividing them into two groups: brighter or fainter than the median value. Second, we masked pixels lacking HSC unmasked objects in the full CLASSIC catalogue in order to account for halos and diffraction spikes of bright foreground stars or other extended patterns. Finally, we masked pixels without any galaxy in the final joint COSMOS2020 and COSMOS2015 catalogue to account for the additional galaxies taken from the previous release. Masked pixels

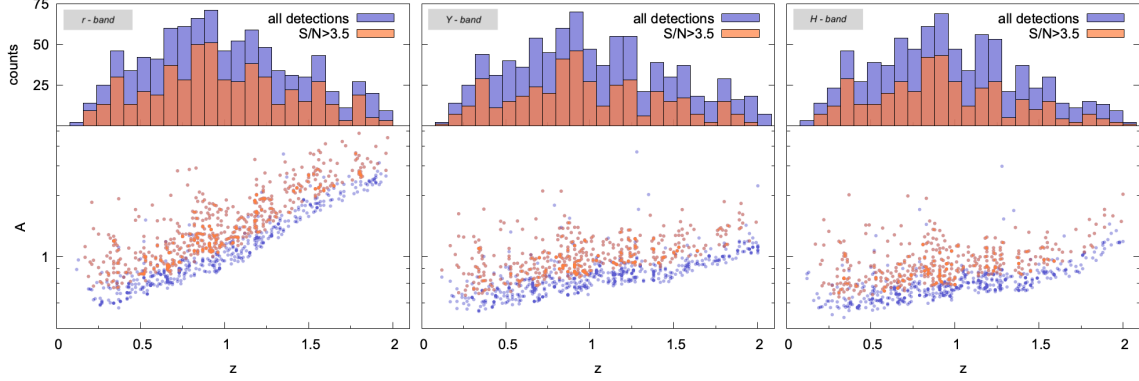


Figure 3.4. Number of detections (top panels) and amplitude (bottom panels) as a function of redshift for the initial cluster/group catalogues produced by analysis in r -band (left panels), Y -band (middle panels), and H -band magnitude (right panels). The orange histograms and points highlight detections with $S/N > 3.5$.

combining these criteria were used to create the final composite visibility mask. The final effective area on which the cluster and group search was performed is 1.69 deg^2 .

3.3 The cluster and group catalogues

In this section, we present the resulting catalogues of the cluster and group search performed in COSMOS with the AMICO algorithm. The candidate samples are then matched and compared, and a few examples of detections are discussed.

3.3.1 Results of the search

We performed a cluster and group search over the effective area of 1.69 deg^2 in the COSMOS-UltraVISTA field as previously described. We chose $(S/N)_{\min} = 3.0$ and cut the catalogues at $\lambda_* > 1$, which are typical values to minimise the number of spurious or extremely poor detections in terms of galaxy content. However, as we describe in Section 3.3.2, an a posteriori cut at $S/N = 3.5$ can be adopted to select a more robust subsample of the catalogue. We produced three initial catalogues from the three analyses performed by using different photometries:

r - band: We detected 893 candidates in the range $0.1 < z < 2.0$ by using the HSC- r magnitude as the galaxy property. Among these, 514 candidates ($\sim 58\%$) were detected with $S/N > 3.5$.

Y - band: We detected 845 candidates in the range $0.1 < z < 2.0$ by using the UltraVISTA- Y magnitude. Among these, 408 candidates ($\sim 48\%$) were detected with $S/N > 3.5$.

H - band: We detected 786 candidates in the range $0.1 < z < 2.0$ by using the UltraVISTA- H magnitude. Among these, 382 candidates ($\sim 49\%$) were detected with $S/N > 3.5$.

Figure 3.4 shows the distribution of the number of detections and their amplitude as a function of redshift. The minimum detectable amplitude increases as a function of redshift since the further away we observe, the harder it is to detect small systems. Differences between the number of detections and their properties between different runs can be due to the differences in depth and availability of galaxies in the input catalogues. This is also visible in the general trend of amplitude with redshift, which increases more steeply at $z > 1$ for the r -band catalogue with respect to the other catalogues. This is due to the larger number of available galaxies in the input catalogue.

The resulting catalogues include identification number, sky coordinates, redshift, S/N , amplitude A , apparent richness λ , intrinsic richness λ_* , likelihood, fraction of the detection that is masked, detection redshift probability distribution, and redshift uncertainty. In addition to the candidate cluster and group catalogue, AMICO creates the list of galaxy members for each detection with their membership and field probability.

3.3.2 Matching the r , Y , and H catalogues

We performed pairwise three-dimensional matching between the three catalogues adopting $dz=0.05(1+z)$ as the redshift separation and $drad = 0.5 \text{ Mpc}/h$ as the physical sky separation. These values were chosen according to the uncertainties of cluster/group redshift and position estimates and were tested by visual inspection of preliminary matching results. In the matching procedure, we used an a priori sorting of the input catalogues by amplitude, A .

We found correspondence between 561 candidate systems detected in the r and Y analyses, which is around 63% and 66% of the two catalogues, respectively. Between the Y and H analyses, 642 detections found a match, which is around 76% and 82% of the two catalogues. A total of 542 candidates were matched when comparing r and H analyses, which is around 61% and 69% of the two catalogues. The results of this matching analysis are displayed in Figure 3.5, where we show the redshift distribution and the S/N versus z of matched and unmatched detections for the three different pairs of catalogue combinations. As can be seen in Figure 3.5 and as highlighted by the S/N distribution of matched and unmatched objects for the r versus H comparison in Figure 3.6 (which is the comparison with the highest number of unmatched objects), the majority of detections without correspondence between runs have $S/N < 3.5$. This shows how this can be adopted as a reasonable and conservative S/N cut for the catalogue in order to have a more robust sample.

With two-way matching between different combinations of the three catalogues (i.e. matched whenever the detection is successfully paired in both directions of matching) and by adopting the same procedure described before, we found a total of 490 candidate clusters and groups detected independently in all three analyses. Given the results of the matching between catalogues resulting

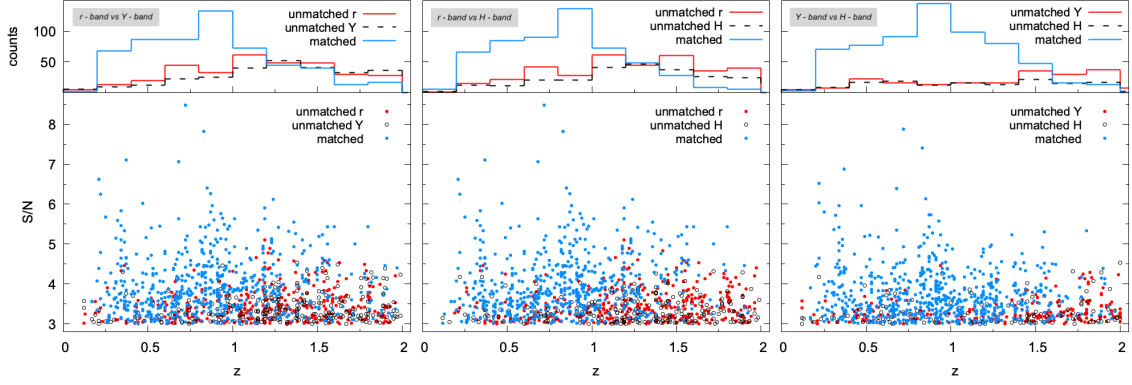


Figure 3.5. Number of detections (top panels) and their S/N (bottom panels) as a function of redshift for matched (blue) and unmatched (red) detections within AMICO runs: (from left to right) matching r vs Y , r vs H , and Y vs H . Filled red dots and solid red lines indicate unmatched detections in the first catalogue, and empty black dots and dashed lines indicate those in the second catalogue.

from the three runs, we created a final catalogue containing all candidates, both matched and unmatched (counting as one the detections successfully found in more than one run according to our matching). Our final catalogue contains a total of 1269 candidate clusters and groups, given that all runs were performed with $(S/N)_{\min} = 3.0$. Among these, 666 candidates were detected with $S/N > 3.5$ in at least one of the runs. When referring to properties of specific detections, hereafter, we report the mean of the values obtained in the three different runs, unless otherwise specified.

3.3.3 Examples of detections

We visually inspected the location of a sample of detected clusters and groups in the optical colour-composite images from HSC DR3 (Aihara et al., 2022). Figure 3.7 shows HSC images of four randomly chosen r -band detections, at different redshifts, with $S/N > 4.0$. Among the candidates, we successfully detected the main cluster of the system identified by Smolčić et al. (2007) via a wide-angle-tail radio galaxy. Smolčić et al. (2007) found that the radio galaxy coincided with an elliptical galaxy, which was identified as the BCG of a cluster at $z \sim 0.22$. We detected this cluster in all bands, and we identified 146 galaxy members (in the Y -band) with probability $> 50\%$, including the BCG (radio galaxy), which was assigned the highest probability and has a position consistent with the centre of the cluster as found by AMICO. The structure, as detected by AMICO, is shown in Figure 3.8, with associated members, amplitude, and X-ray Chandra+XMM-Newton (Gozaliasl et al. 2019; see Section 3.4) contours. We were not able to distinguish the assembly of more than one cluster as stated by Smolčić et al. (2007), probably because several saturated stars affected this region. Nevertheless, among the brightest galaxies associated with the cluster, we

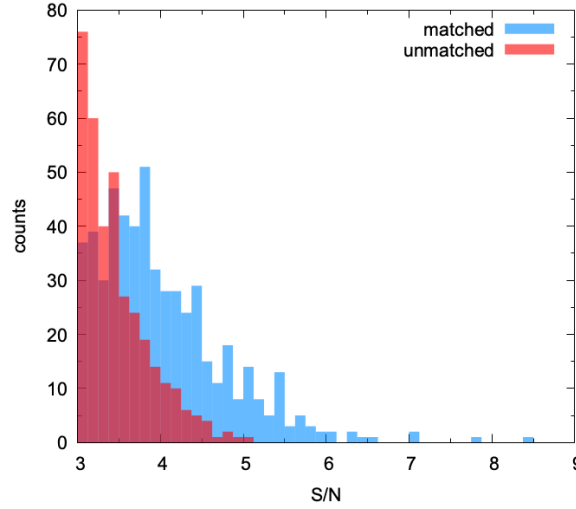


Figure 3.6. Distribution of matched (blue) and unmatched (red) detections in the comparison r vs H , as a function of signal-to-noise ratio. Most of the unmatched detections have $S/N < 3.5$.

found at least three galaxies coinciding with the position of the substructures and with the peaks of the diffuse X-ray emission. This structure has been identified by AMICO as a single cluster, but the presence of these bright galaxies far from the centre can indicate the grouping of several substructures.

We compared our list of candidates with other group and cluster catalogues available for the COSMOS field. Since the catalogue candidates were retrieved with different methods and therefore present different definitions of richness and quality of the detections, it is not straightforward to make consistent comparisons and considerations. We simply include in our final catalogue the corresponding identification numbers for the 10, 182, 307, 7, 11, and 585 detections matched within $dz=0.05(1+z)$ and $d_{rad} = 0.5$ Mpc/h, from the catalogues by [Zatloukal et al. \(2007\)](#), [Knobel et al. \(2009\)](#), [Knobel et al. \(2012\)](#), [Castignani et al. \(2014b\)](#), [Iovino et al. \(2016\)](#), and [Sarron & Conselice \(2021\)](#), respectively.

One of the goals of this study was to search for high-redshift clusters and groups. We detected 273 systems at $z > 1.5$, of which 125 were detected with $S/N > 3.5$ in at least one run. We found that 111 common identifications with the catalogue by [Sarron & Conselice \(2021\)](#) have $z > 1.5$, of which 3 were also detected by [Zatloukal et al. \(2007\)](#).

Detections overlapping along the line of sight. We performed two-dimensional matching on our final catalogue (without using redshift information), and we found 269 sets¹ of detections lying within 0.02 deg from each other. Some of these cases are in a configuration that makes it challenging to detect them, and therefore, the background objects are often not known in the

¹Pairs or groups of more than two detections

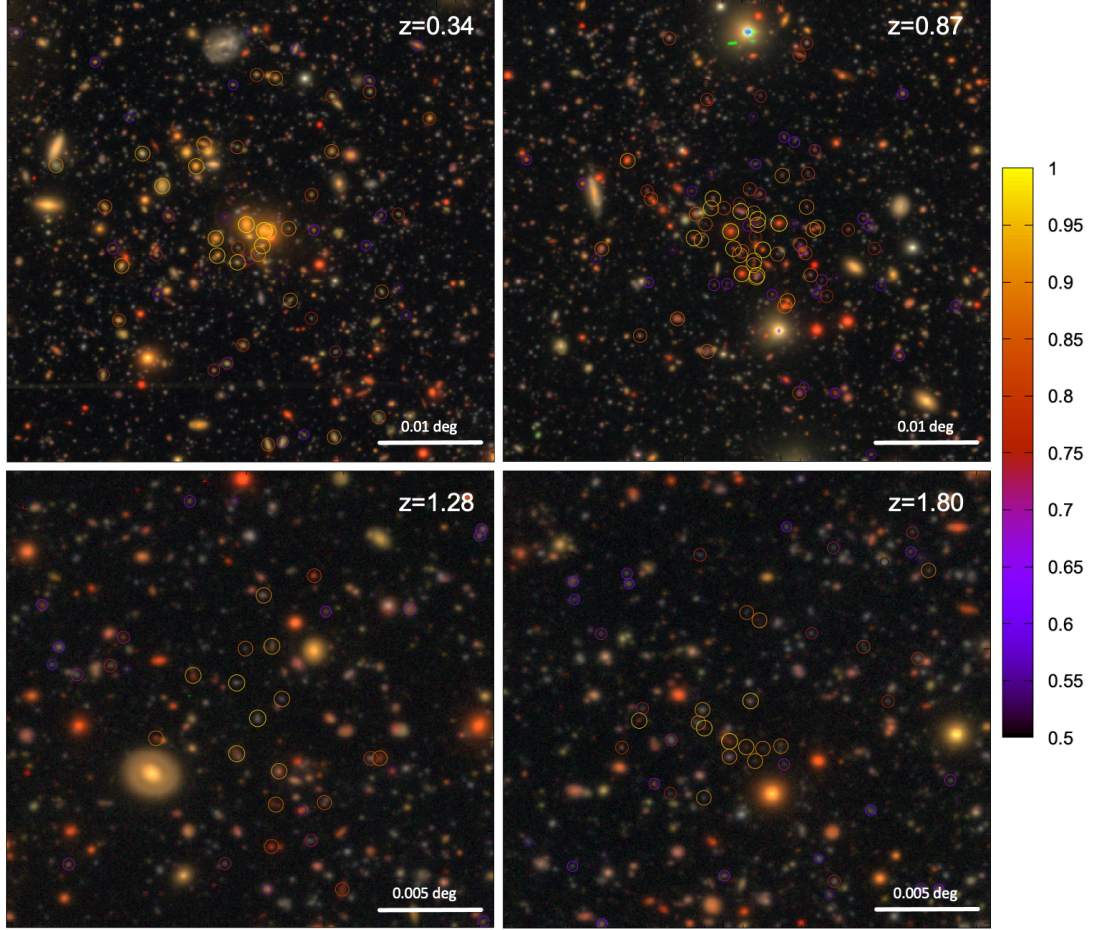


Figure 3.7. Four examples of detections at different redshifts as identified in the r -band run. Stamps are HSC g, r, i colour-composite images centred on the AMICO detection, with a side of 0.05 deg for the top panels and 0.025 deg for the bottom panels. The circles indicate the associated galaxies, sized and coloured according to their association probability (see colour bar at right). *Top left:* Candidate at $z = 0.34$ with $\lambda_{\star} \sim 17$; *top right:* Candidate at $z = 0.87$ with $\lambda_{\star} \sim 27$; *bottom left:* Candidate at $z = 1.28$ with $\lambda_{\star} \sim 26$; *bottom right:* Candidate at $z = 1.80$ with $\lambda_{\star} \sim 40$.

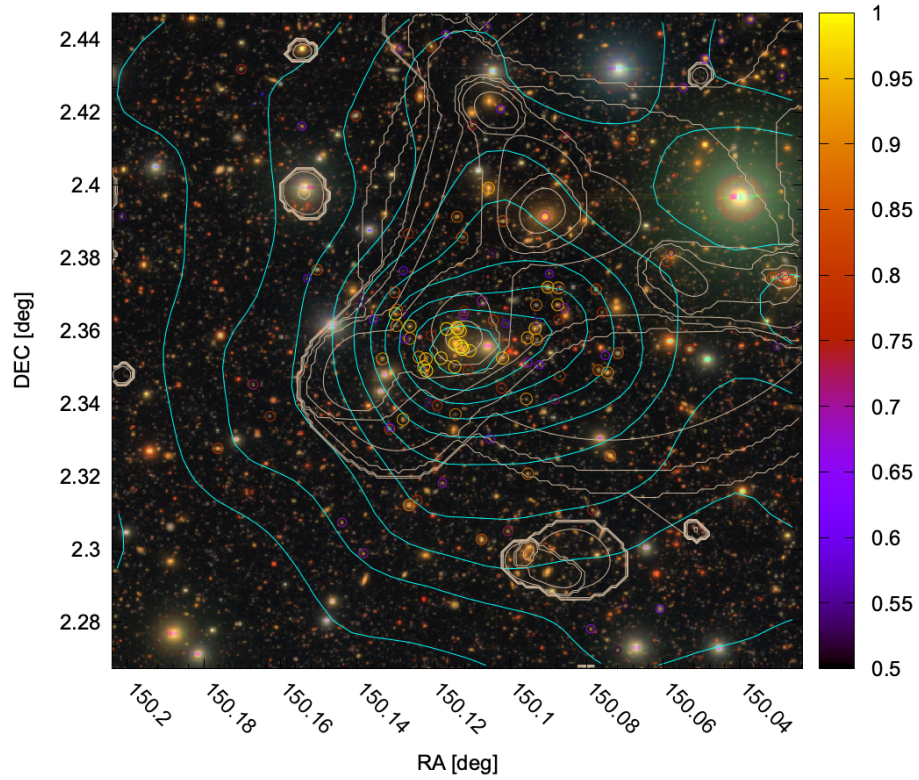


Figure 3.8. (Super-)group/cluster substructures at $z = 0.22$ found by Smolčić et al. (2007) via a radio-BGG, as detected by AMICO. The circles indicate assigned members with colour-coded probability (see colour bar at right) and are overlaid with AMICO amplitude and X-ray contours (cyan and white, respectively).

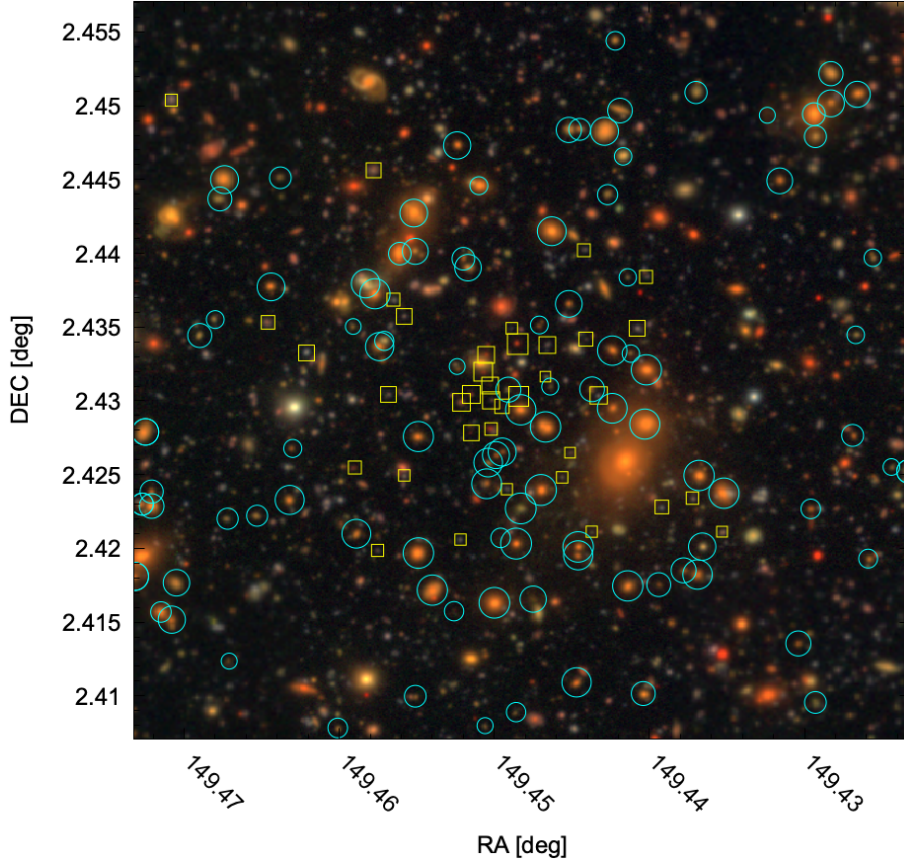


Figure 3.9. Pair of detections aligned along the line of sight, as detected in the Y-band run. Members of the foreground system, located at $z = 0.47$, are indicated by cyan circles. The yellow squares indicate background system galaxies ($z = 1.56$). The size of the circles and squares is proportional to the membership probability. The HSC g, r, i composite image has a size of 0.05 deg and is centred on the background system.

literature. For instance, we found that 116 sets of detections (of which 96 are pairs) have a sky separation ≤ 0.01 deg, involving a total of 253 objects that are almost in perfect alignment with one or more objects along the line of sight. An example of this kind of detection pair is shown in Fig. 3.9, where an already-known system at $z = 0.47$ (e.g. [Gozaliasl et al., 2019](#); [Knobel et al., 2012](#); [Sarron & Conselice, 2021](#)) is virtually aligned with a background system at $z = 1.56$, which is not included, for instance, in most optical and X-ray cluster and group catalogues. The identification of these sets of detections is important in this study, where the number of detected objects per square degree is high. In particular, having several detections close to each other on the sky plane can bias the estimation of X-ray properties. This is why we used the information about the position of these sets of detections to clean the catalogue for the calibration of the scaling relation described in Section 3.4.

3.3.4 Spectroscopic counterparts

We assigned spectroscopic counterparts to member galaxies associated by AMICO by making use of a sample of galaxies from 13 public spectroscopic surveys (zCOSMOS-b, Lilly et al. 2007; PRIMUS, Coil et al. 2011, Cool et al. 2013; GEEC2, Balogh et al. 2014; FORS2, Comparat et al. 2015; DEIMOS, Hasinger et al. 2018; VIS3COS, Paulino-Afonso et al. 2018; hCOSMOS, Damjanov et al. 2018; FMOS-COSMOS, Kashino et al. 2019; C3R2, Masters et al. 2019; MUSE, Rosani et al. 2020; LEGA-C, van der Wel et al. 2021; MAGIC, Epinat 2021, Abril-Melgarejo et al. 2021, Epinat et al. 2024; DESI, DESI Collaboration et al. 2024) collected in the COSMOS Spectroscopic Redshift Compilation (Khostovan et al., 2026, including GT)². We found that 612, 741, and 720 detections have at least one member with spectroscopic redshift associated with the cluster/group with a probability $> 50\%$, in the r -, Y -, and H -band analyses, respectively. Among these, 373, 449, and 454 (i.e. 61%, 61%, and 63%) have more than three members with a spectroscopic redshift. All of these candidates have a cluster/group redshift assigned by AMICO compatible with z_{spec} , which is defined as the mean spectroscopic redshift of the associated members, being $\Delta z / (1 + z_{spec}) < 0.03$. If we consider the final catalogue obtained by matching detections between the three different runs (as described in Sect. 3.3.2), the total effective number of detections with mean spectroscopic redshift based on more than three member galaxies is 567. All of these detections have a mean spectroscopic redshift compatible with that assigned by AMICO.

3.4 Comparison with X-ray group catalogues

We performed a comparative analysis of the AMICO detections with the X-ray selected groups presented in the work by Gozaliasl et al. (2019). In this work, a catalogue of galaxy groups was produced for the COSMOS field by using Chandra (Elvis et al., 2009; Civano et al., 2016) and XMM-Newton (Hasinger et al., 2007; Cappelluti et al., 2009) 0.5 – 2.0 keV combined data, cleaned via wavelet transform (Vikhlinin et al., 1998).

The catalogue contains 247 groups covering the range $0.08 \leq z < 1.53$ and with masses of $M_{200} = 8 \times 10^{12} - 3 \times 10^{14} M_{\odot}$. Virial masses were estimated via the scaling relation derived through stacked weak lensing analysis by Leauthaud et al. (2010). This X-ray selected catalogue is a revised and extended version of the COSMOS group catalogues presented in Finoguenov et al. (2007) and George et al. (2011). To optically validate the X-ray extended sources and estimate their redshifts, both spectroscopic (Hasinger et al., 2018) and photometric (Ilbert et al., 2009; Laigle et al., 2016; McCracken et al., 2012) galaxy samples were used. Data reduction and member identification, performed via a refined red-sequence approach, are described in detail in George

²We chose to rely on the 13 largest public surveys included in the compilation, since at the time this analysis was performed, the full Spec-z Compilation by Khostovan et al. (2026, including GT) was not publicly released yet. These 13 surveys constitute the bulk of the DR1.0 compilation release.

et al. (2011) and Gozaliasl et al. (2019).

We performed three-dimensional matching within $dz=0.05(1+z)$ and $drad = 0.5$ Mpc/h. We found 104 successful matches for the r -band run, 107 matches for the Y -band run, and 99 matches for the H -band run. If we consider AMICO detections matched within different runs, the total number of detections with correspondence in the X-ray catalogue by Gozaliasl et al. (2019) is 122, or 55% of the X-ray detections. For this comparison, we considered consistent volumes occupied by the cluster/group searches, namely accounting for different redshift extent and effective unmasked areas covered. This comparison displayed a general good matching quality with most of the successfully paired detections lying within 0.1 Mpc/h and within a redshift scatter of 0.02. We found a slight and negligible redshift bias: redshift estimates for AMICO are slightly smaller with respect to X-ray identifications, with an average bias of $\Delta z/(1+z) \approx -0.002 \pm 0.001$.

3.4.1 Quality flags

Gozaliasl et al. (2019) included four quality flags in their group catalogue. Flag 1 labels the safest detections, with spectroscopic members; flag 2 is for cases of foreground or background contamination; flag 3 signals a lack of spectroscopic counterparts; and flag 4 is assigned to the least safe sample, with an ambiguous optical counterpart association. In our comparative analysis, we found that 77% of the successful matches with AMICO catalogues are part of the safest X-ray sample. The safest sample of X-ray detections has more successful matches than unsuccessful matches, while for the detections flagged with 2, 3, and 4, the opposite is true: around 57% of the problematic detections in the X-rays (flag 2, 3, or 4) do not pair with any AMICO detection. This can be read as an indication of the reliability of the flagging system by Gozaliasl et al. (2019), and therefore the general reliability of the safest X-ray detections (flag 1).

3.4.2 The r , Y , and H analyses when compared with X-rays

A total of 86 clusters and groups have been found in all three AMICO catalogues and successfully matched with the catalogue by Gozaliasl et al. (2019). We can take the X-ray catalogue as a reference to make considerations on the different runs. We need to take into account only the common volume, so we restrict this comparison to the redshift range covered by the X-ray catalogue. The catalogue produced by the Y -band run and r -band run appears to be more complete with respect to the X-ray reference, having 107 and 104 matched detections, respectively. Then, we can consider the relative number of matches for those detections that are found by AMICO in only one of the runs, which we refer to as ‘unique’. Nevertheless, we found that the r -run matches only $\sim 4\%$ of its unique detections, whereas $\sim 7\%$ of unique Y and H detections are paired with X-ray sources.

Table 3.2. Mass-proxy scaling relation parameters and pivot values (see Eq. 3.1) based on detections matched with the group catalogue by Gozaliasl et al. (2019).

r-band			
	α	β	O_{piv}
λ_*	-0.320 ± 0.025	0.780 ± 0.129	21
A	-0.311 ± 0.025	0.917 ± 0.151	1.4
Y-band			
	α	β	O_{piv}
λ_*	-0.351 ± 0.026	0.961 ± 0.146	15
A	-0.330 ± 0.026	1.155 ± 0.198	1.0
H-band			
	α	β	O_{piv}
λ_*	-0.352 ± 0.027	1.027 ± 0.145	15
A	-0.341 ± 0.027	1.132 ± 0.180	0.9

3.4.3 Mass-observable scaling relations

We studied the relation between mass proxies provided by AMICO and X-ray-derived mass for the sample of successfully matched clusters and groups. The X-ray catalogue by Gozaliasl et al. (2019) comprises the 0.1-2.4 keV rest frame X-ray luminosity (L_X) measured within R_{500} and the virial mass M_{200} estimated through the scaling relation presented by Leauthaud et al. (2010). We studied the relation between X-ray virial mass (and luminosity) with AMICO amplitude (A) and with intrinsic richness (λ_*). For this analysis, we used only detections flagged as safe and with spectroscopic counterparts (flag 1 in Gozaliasl et al. 2019) that do not fall within 0.02 deg of another detection with $\lambda_* > 20$ and identified with $S/N > 3.5$. This reduces the possibility of introducing outliers due to projection effects and contamination of a nearby object.

Mass-observable calibration from AMICO mass-proxies was performed for the AMICO-KiDS-DR3 cluster sample (Maturi et al., 2019) via stacked weak-lensing analysis by Bellagamba et al. (2019). We used the scaling relation expression used by Bellagamba et al. (2018) to fit the data for the matched sample

$$\log \frac{M_{200}}{10^{14} M_\odot} = \alpha + \beta \log \frac{O}{O_{piv}} + \gamma \log \frac{E(z)}{E(z_{ref})}, \quad (3.1)$$

with O being the AMICO mass proxy (i.e. A or λ_*) and O_{piv} its median value in the considered sample. In this part of the analysis, we neglected the redshift dependency term of the relation

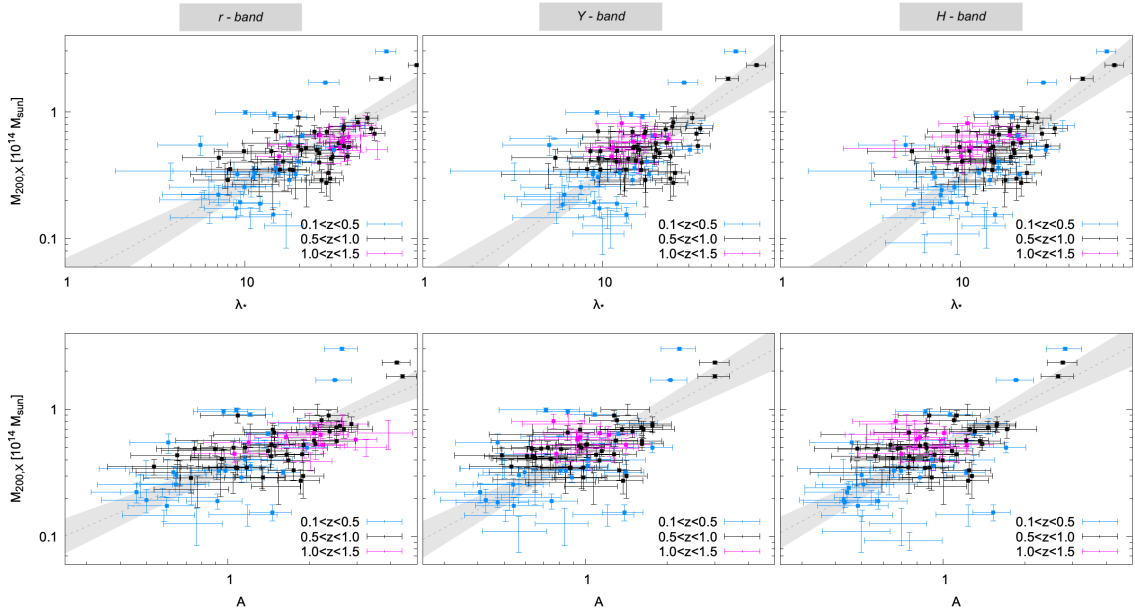


Figure 3.10. Relation between X-ray mass and richness (top panels) or amplitude (bottom panels) for the X-ray matched detections in the catalogue by [Gozaliasl et al. \(2019\)](#) in the range $z \in [0.1, 1.5]$. From left to right: Detections retrieved from analysis in r -, Y -, and H -band. The different colours refer to different redshift bins (see legend). The grey dashed line and shaded area indicate the best-fitting relation with the corresponding 1σ errors. Data points marked by squares are the detections belonging to the safest sample in the X-ray catalogue and without foreground, background, or nearby contamination. Error bar points without a central square are matched detections, not used for the scaling relation calibration.

($\gamma = 0$) since the limited size of our sample does not allow robust analyses of redshift evolution. The fitting analyses were performed by taking into account the 1σ errors for X-ray mass, the amplitude variance (Eq. 2.12), and the square root of λ_* . We made use of the R package for Linear Regression in Astronomy (LIRA) described in [Serenio \(2016\)](#). The best-fit parameters are reported in Table 3.2 and the corresponding relations shown in Figure 3.10. A defined relation is visible in almost all combinations among the catalogues, with a generally smaller scatter for the results of the r -band run with respect to the analyses using other photometries. The large scatter is expected for a sample extending over this wide range of masses. Detections indicated with squares in Fig. 3.10 are the detections used for the scaling relation fit (i.e. belonging to the safest sample detected in the X-rays) with reduced chances to have foreground or background contamination, and they all have spectroscopic members. The studied sample extends down to low masses, with the least massive matched group having $M_{200} \approx 9.3 \times 10^{12} M_{\odot}$ and $L_X \approx 3.3 \times 10^{41} \text{ erg/s}$.

3.5 X-ray counterparts for new detections

After matching AMICO detections and X-ray selected groups by [Gozaliasl et al. \(2019\)](#), we studied the sample of non-matched detections and looked for their possible X-ray counterparts. We measured X-ray properties for these detections directly at their locations as identified by AMICO.

We used the combined Chandra and XMM-Newton maps of emission residuals in the 0.5–2 keV band after removing instrumental and sky background, as well as unresolved X-ray emission. XMM-Newton provides approximately 70% of the sensitivity in the combined data. To test the effect of the total flux, including the unresolved emission, we only used the XMM-Newton data; compared to Chandra, these data have a larger reduction in sensitivity due to the removal of the contamination from nearby point-sources. In most cases, however, there is no contamination, and an upper limit using the XMM-Newton data without point-source cleaning is the most constraining. The data reduction in producing the maps is identical to the analysis published by [Gozaliasl et al. \(2019\)](#). Using the redshifts of the AMICO detections, we estimated the count rates inside the 200 kpc radius, and in obtaining the X-ray properties, we extrapolated these numbers to the iteratively estimated R_{500} radius, following the procedure outlined in [Finoguenov et al. \(2007\)](#). The size of the aperture closely matches the expected size of the emission zones of the AMICO detections found below the limit of the published X-ray catalogues in COSMOS, and so the flux extrapolations are minimal, while the overlap in the extraction zones between neighbouring detections is not significant at $z > 0.2$.

We kept only detections with significant X-ray emission by cutting at flux significance above the 1σ limit. The total number of new detections with X-ray flux estimates above the significance limit is 500 out of the 1147 analysed candidates ($\sim 44\%$). If we consider only objects detected by AMICO with $S/N > 3.5$, there are 267 with significant X-ray emission out of the total 577

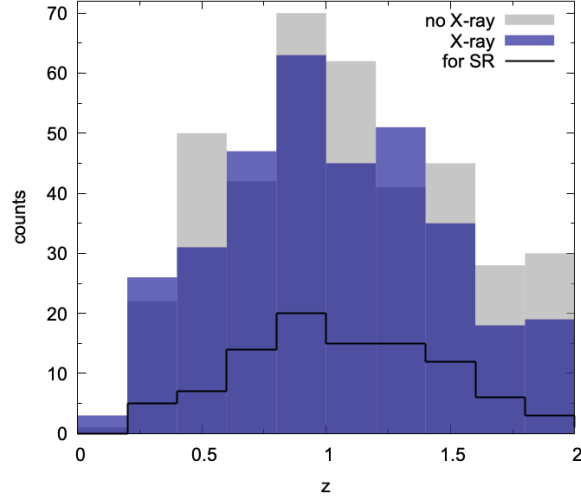


Figure 3.11. Distribution of the redshifts of new detections analysed in X-rays (from the Y-band run catalogue). The grey histogram represents detections having X-ray flux below the 1σ significance limit. The blue histogram refers to detections with significant X-ray emission ($\sigma > 1$). The solid line highlights the subsample selected to study the scaling relation (SR) between AMICO mass proxies and X-ray mass.

analysed candidates ($\sim 46\%$).

By adding together successful matches with the X-ray group catalogue and AMICO candidates with significant X-ray flux estimates, we created a sample of 622 candidate groups and clusters with optical and X-ray properties up to $z = 2$ and down to less than $10^{13} M_{\odot}$.

We repeated the calibration of the relation between AMICO mass proxies and X-ray mass, this time for the full sample of AMICO detections with X-ray estimates. In order not to include biased estimates due to foreground, background, or nearby contamination, we rejected all detections lying within 0.02 deg of the centre of another rich detection or lying on the extended X-ray emission of a bright cluster. This left us with 222 detections, which we used to calibrate the scaling relations. In Figure 3.11 we show the distribution in redshift of the new detections analysed for the Y-band run catalogue, displaying the detections with and without (blue and grey) significant X-ray flux and the ones eventually selected for the study of the mass-observable scaling relations (black solid line). In Figure 3.12, we show the consistency between detections matched with the X-ray catalogue (grey error bars) and new detections with X-ray estimates above the significance limit (coloured points), used for the scaling-relation calibration. Most of the new detections lie within the scatter of the catalogue-matched detections, consistently with the trends found by the best-fitting scaling relations. We found that at fixed amplitude or λ_* , new detections tend to be, on average, less luminous or less massive when compared to successful matches with the X-ray catalogue. This is due to the lower X-ray S/N level adopted in our analysis here, and it shows the

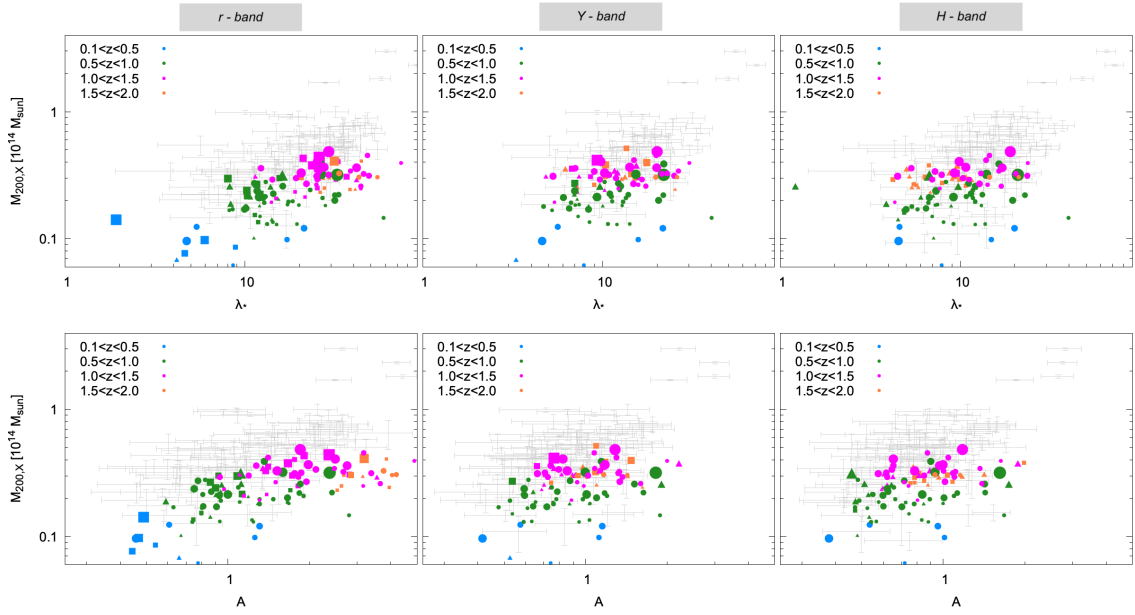


Figure 3.12. Same as in Fig. 3.10, but for the sample of new detections with X-ray flux estimates selected for the scaling relation study (coloured points) in the full redshift range, $z \in [0.1, 2.0]$. From left to right: Detections from analysis in the r -, Y -, and H -band. The different colours refer to the different redshift bins, as labelled in each panel. Circles are detections found in all three runs, triangles are found in two out of three runs, and squares indicate the unique detections from respective runs. The size of the symbols is proportional to the X-ray flux significance. As a reference, the grey error bars in the background indicate the detections used for the scaling relation derived from the direct matches with the X-ray catalogue (the same displayed in Fig. 3.10).

high completeness in the identification of bright X-ray extended sources in the [Gozaliasl et al. \(2019\)](#) catalogue.

3.5.1 Redshift dependence of the scaling relations

If we consider the whole sample of detections, including successful matches with the X-ray catalogue and the new detections with X-ray flux estimates, we have the opportunity to study the scaling relations for clusters and groups extending up to $z \sim 2$ and over a wide range of mass and richness. This also allows a study of the redshift-dependence of the scaling relations. We considered the scaling relation in Eq. 3.1, this time including the z -dependent term (i.e. $\gamma \neq 0$).

We chose $z_{ref} = 0.9$, namely the median redshift of the entire studied sample. The results of the fitting analysis, including the redshift dependence term, the scatter standard deviation and the chosen pivot values, are shown in Table 3.3. The best-fitting relations are also shown in Figure 3.13 where the data points are divided into four redshift bins with the relevant scaling relations computed at the typical redshifts of the bins ($z = [0.4, 0.8, 1.2, 1.6]$).

We found a negative redshift dependence for the r -band detected clusters and groups, for which, at fixed amplitude or λ_* , more distant objects tend to have lower masses. This trend seems to fade out for redder bands, although there is a negligible hint of an inverted trend for the relation $M_{200} - \lambda_*$ in the H -band catalogue.

3.5.2 New detections without X-ray counterparts

As mentioned above, nearly half of the detections analysed in the X-rays were found not to have flux significance above unity. The relative number of new detections with and without significant X-ray flux does not seem to be significantly affected by cuts in S/N or redshift (e.g. $\sim 43\%$ of analysed new detections still have $\sigma > 1$ and $S/N > 4.0$ at $z < 1$). Among these detections, we found and visually inspected several examples of clusters and groups with a significant number of assigned galaxy members. In particular, we selected 116 groups with $\lambda_* > 20$ (i.e. among our richest detections), detected with $S/N > 4.0$ at $z < 1.6$, that were found to have X-ray flux significance below 2σ . We chose this threshold in order to include both detections below the significance limit and with low significance ($1 < \sigma < 2$). We found a few cases of correspondence or vicinity to X-ray point sources. Thus, we investigated the impact of point-source removal in estimating the flux and luminosity of extended sources. We did so by repeating the measurements of X-ray properties for these 116 detections by considering the X-ray image without removing the point-sources. Undoubtedly, this represents a significant improvement for most of the analysed sources, as seen in Figure 3.14, where we show the relation between L_X and λ_* for this sample (compared to the scaling relation found for the full sample and the 2σ scatter) with (top panel) and without (bottom panel) point-source removal. Most of the selected detections become consistent with the rest of the sample by reintroducing the removed flux. Nevertheless, this does not seem to

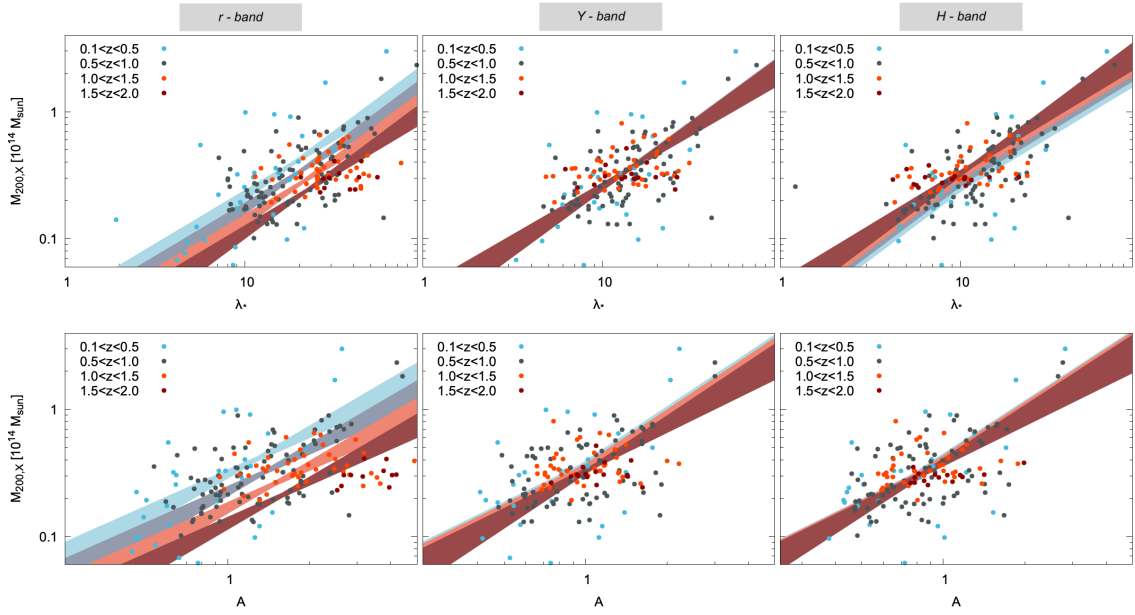


Figure 3.13. Relation between X-ray mass and richness (top panels) or amplitude (bottom panels) for the full sample including matched detections with the X-ray catalogue and new detections with X-ray flux estimate in the full redshift range, $z \in [0.1, 2.0]$. From left to right: Detections from analysis in r -, Y -, and H -band. The different colours refer to the different redshift bins, as labelled in the plot. The redshift-dependent best-fitting relation is shown for typical redshift values representing the four redshift bins, $z = [0.4, 0.8, 1.2, 1.6]$ (z increases from light blue to dark red) and with their 1σ error region. The best-fit parameters are also reported in Table 3.3. The redshift trend visible for r -band-based detections fades out for redder bands, where the scaling relations become consistent with no redshift dependence.

Table 3.3. Mass-proxy scaling relation parameters and pivot values (see Eq. 3.1) including redshift dependence based on the full selected sample with X-ray flux estimate. The standard deviation of the scatter σ is given in $\log_{10}$ space, being the scaling relations in the form $Y = \alpha + \beta X + \gamma Z \pm \sigma$, with $Y = \alpha + \beta X + \gamma Z$ referring to Eq. 3.1.

r-band						
	α	β	γ	σ	O_{piv}	z_{ref}
λ_*	-0.498 ± 0.019	0.954 ± 0.105	-0.966 ± 0.241	0.170 ± 0.016	20	0.9
A	-0.516 ± 0.021	1.098 ± 0.151	-1.309 ± 0.319	0.182 ± 0.017	1.4	
Y-band						
	α	β	γ	σ	O_{piv}	z_{ref}
λ_*	-0.482 ± 0.019	0.931 ± 0.110	-0.010 ± 0.172	0.160 ± 0.016	13	0.9
A	-0.452 ± 0.020	1.232 ± 0.187	-0.206 ± 0.194	0.171 ± 0.018	1.0	
H-band						
	α	β	γ	σ	O_{piv}	z_{ref}
λ_*	-0.516 ± 0.020	0.962 ± 0.103	0.424 ± 0.182	0.145 ± 0.017	11	0.9
A	-0.519 ± 0.022	1.220 ± 0.170	-0.094 ± 0.194	0.166 ± 0.018	0.8	

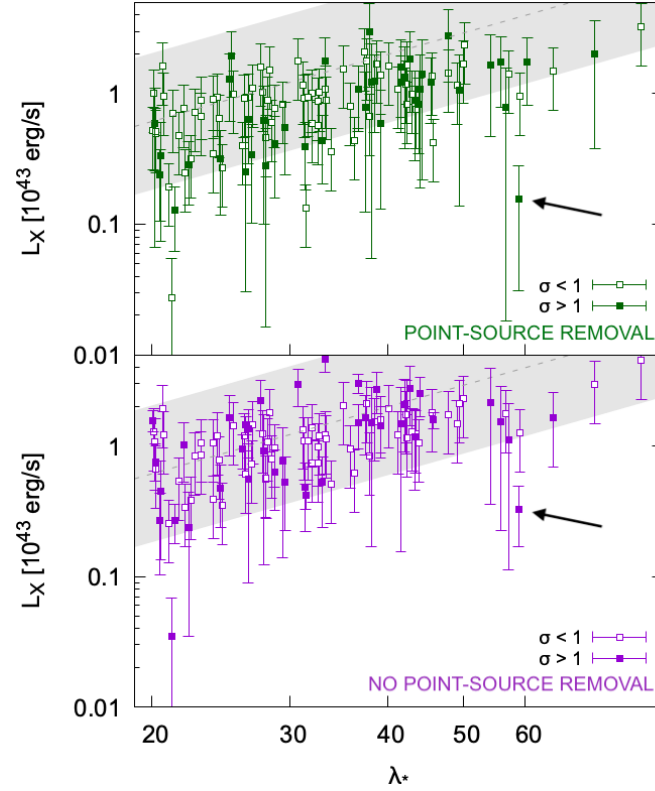


Figure 3.14. Relation between X-ray luminosity and intrinsic richness for the sample of richest detections ($\lambda_* > 20$, $S/N > 4.0$, $z < 1.6$) that were found to have non-significant X-ray emission, with X-ray measurements performed with (top panel) and without (bottom panel) point-source removal. Empty squares are detections with X-ray flux below the 1σ significance limit. The grey shaded area in the background is the 2σ scatter region relative to the $L_X - \lambda_*$ relation found at z_{ref} for the full r -band-detected sample.

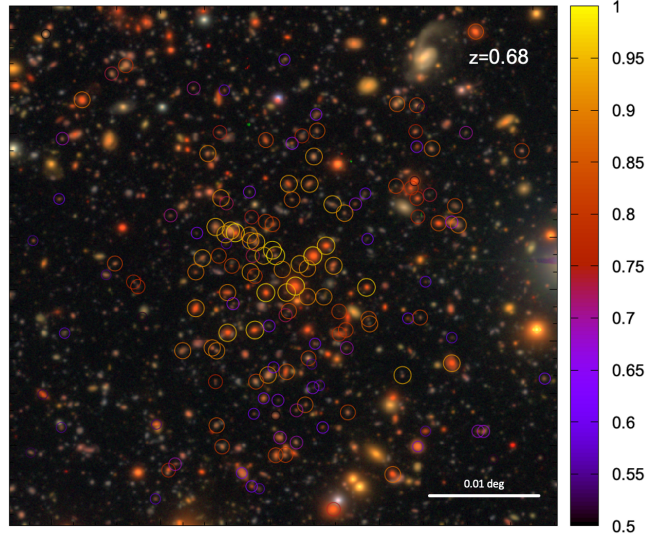


Figure 3.15. *Detection 5*: Candidate cluster detected at $z = 0.68$ with around 200 member galaxies ($\lambda_\star \sim 59$ in the r -band), without significant X-ray emission, even without point-source removal.

be the case for two of the detections. These two outliers do not display a significant improvement in their position with respect to the rest of the sample, even with increased flux significance. One of the two outliers, the object with the lowest richness, was found in all three runs at $z = 0.36$, with $S/N \sim 5$, $\lambda_\star \sim 19$, $A \sim 1$, but with very low X-ray luminosity and large errors. The second object was instead found at $z = 0.68$, and for simplicity we refer to it as *detection 5*.

Detection 5 is one of the richest detections without X-ray flux significance above 2.0, and it has been detected with a high signal-to-noise ratio in all AMICO runs ($S/N = 6.6$). Its intrinsic richness, for instance, in the r -band analysis is $\lambda_\star \sim 59$. It was found at $z = 0.68$, with ~ 200 galaxy members assigned with probability larger than 50% (among these, almost half are assigned with probability $> 75\%$). This is an example of a detection whose X-ray properties measured without point-source removal did not improve sufficiently for it to become consistent with the rest of the sample in terms of the relation between optical and X-ray properties. *Detection 5* is shown in Figure 3.15, and indicated by an arrow in Figure 3.14, with X-ray luminosity measured with and without point-source removal. The cluster/group redshift probability distribution, $P_{cl}(z)$, of *detection 5* shows no anomalies, and it is consistent with the $P_{cl}(z)$ of detections at similar redshift and with similar richness, which have also been found via X-ray selection.

This interesting detection is compatible with being part of the COSMOS wall (Iovino et al., 2016). Using the same data used by Abril-Melgarejo et al. (2021), retrieved with MUSE (Epinat 2021; Epinat et al. 2024; Bacon et al. 2010), we performed our dynamical analysis. The two groups mentioned in Abril-Melgarejo et al. (2021) (CG84 and CG84b) are responsible for the two peaks in the redshift distribution of AMICO members, when considering spectroscopic

redshifts, centred at $z = 0.6808$ (*group 1*) and $z = 0.6963$ (*group 2*). Running the `Clean` algorithm (Mamon et al., 2013) on these two peaks and considering ± 2000 km/s around each peak, results in the following characteristics for the two groups: *group 1* has 19 clean spectroscopic members, velocity dispersion of 263 ± 54 km/s, and a corresponding radius, $R_{200} = 397$ kpc; *group 2* has 7 spectroscopic members, velocity dispersion of 170 ± 62 km/s, and $R_{200} = 254$ kpc. We note that the mass estimates reported in Abril-Melgarejo et al. (2021) are associated with measuring higher values of velocity dispersion (of around 370 km/s) compared to our estimate, which we attribute to the contamination of interlopers, which in our analysis are more efficiently rejected. The number of member galaxies reported in Abril-Melgarejo et al. (2021) is indeed much higher: 35 and 31 (compared to our 19 and 7). The presence of interlopers, possibly arranged in a filamentary structure elongated along the line of sight, is reflected in the high richness measured by AMICO, given by the presence of a large number of galaxies within the redshift range accessible to the photometric redshifts, and in the relatively small X-ray flux produced by this structure, which is compatible with our spectroscopic mass estimate.

3.6 Conclusions

We robustly detected galaxy clusters and groups in the range $0.1 < z < 2$ by applying the AMICO algorithm to a photometric galaxy sample mostly based on the COSMOS2020 catalogue. We did this in three independent runs by using magnitudes in three different bands as galaxy properties: r -, Y -, and H -band. The final catalogue contains 1269 candidate clusters and groups, among which 666 were detected with $S/N > 3.5$ in at least one of the runs. Among the detected systems, 490 were detected in all three runs. All three runs provided us with new detections.

The main achievements of this study can be summarised as follows:

1. With the creation of a suitable cluster model, a composite mask, and the proper regularisation of the noise model, we found that the AMICO algorithm is efficient in detecting clusters and groups, even in peculiar survey configurations such as the deep 2-deg² COSMOS field.
2. We found 122 correspondences with the most recent X-ray selected group catalogue for the COSMOS field. Additionally, we confirmed X-ray emission with flux significance larger than 1.0 for another 500 AMICO detections, for a total sample of 622 AMICO candidate clusters and groups with associated X-ray properties up to $z = 2$.
3. The comparison with the X-rays allowed the calibration of the scaling relations between AMICO mass-proxies and X-ray mass up to $z = 2$ and down to less than $10^{13} M_{\odot}$. We found that the new detections with X-ray emission are consistent with the rest of the sample in terms of the relation between X-ray and optical properties.

4. The inclusion of the redshift dependence term in the scaling relation analysis showed that the Y -band and the H -band magnitudes used as galaxy properties for the search result in a more stable relation between X-ray mass and AMICO richness and amplitude (with respect to the r -band magnitude), which is an important result for setting the redshift dependence of the calibration.
5. We investigated the impact of point-source removal on X-ray estimates and found that in many cases, this can underestimate X-ray luminosity. Nevertheless, we found an interesting example of an underluminous X-ray object with respect to its galaxy content as detected by AMICO. This is possibly due to the presence of interloper galaxies arranged in a filamentary structure along the line of sight.

The creation of such a cluster and group catalogue for the deep COSMOS field, including robust mass proxies and lists of member galaxies, paves the way for different kinds of studies. This catalogue is a resource for the study of cluster and group galaxy populations and galaxy evolution, and for the study of the formation, evolution, and physics of clusters and groups themselves. Moreover, this work represents a key step towards the successful exploitation of the AMICO algorithm in other peculiar survey configurations and proves the importance of selecting suitable input parameters.

4

The COSMOS-Web catalogue

THE LARGEST DEEP SAMPLE OF GALAXY GROUPS AND PROTOCLUSTER CORES UP TO $z = 3.7$

This Chapter presents the creation of the largest deep catalogue of galaxy groups and protocluster cores to date. This achievement was made possible by JWST's largest contiguous observation in the COSMOS-Web field (Casey et al., 2023), which enabled the construction of the COSMOS2025 galaxy catalogue (Shuntov et al., 2025a). Using AMICO, we identified galaxy systems up to $z = 3.7$, producing a well-characterised catalogue of 1,678 detections over 0.45 deg^2 . We assessed the sample's purity and completeness through data-driven mock catalogues generated with the SinFoniA method, establishing the relation between detection S/N and sample purity. To validate high- z detections, we compared our catalogue with a compilation of known protoclusters in COSMOS. We further explored the spatial connections among detected cores via a 3D clustering analysis, revealing new candidate protoclusters. This comprehensive sample provides a unique resource for studying structure formation and galaxy evolution across diverse environments over ~ 12 Gyr of cosmic history.

This Chapter is fully based on Toni et al. (2025a), A&A, 697, A197^a

^aThe COSMOS-Web deep galaxy group catalogue with members is available at the CDS via anonymous ftp to [cdsarc.cds.unistra.fr](ftp://cdsarc.cds.unistra.fr) (130.79.128.5), via <https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/697/A197> or on GitHub <https://github.com/tonigreta/cosmos-web-groups>

4.1 COSMOS-Web data

The COSMOS-Web Survey (PIs: J. Kartaltepe and C. Casey; Casey et al., 2023) is a ~ 270 -hour observation program conducted using JWST. The survey spans 0.54 deg^2 and utilises four NIR-Cam filters (Rieke et al., 2023), achieving a 5σ point-source depth between 27.5 and 28.2 mag. Observations were carried out using the F115W, F150W, F277W, and F444W filters (Casey et al., 2023; Franco et al., 2025) and for a non-contiguous 0.19 deg^2 area imaged in the F770W MIRI filter (Wright et al., 2022; Harish et al., 2025), with a 5σ point-source depth between 25.3 and 26.0 mag.

The COSMOS-Web survey is a unique combination of the unprecedented depth and spatial resolution of JWST and the coverage and data availability of the COSMOS field (Scoville et al., 2007), making it a key resource for the study and definition of the large-scale structures in the cosmic web over around 50 million Mpc^3 (Casey et al., 2022).

As already introduced, the COSMOS field (Scoville et al., 2007; Capak et al., 2007) benefits from a rich legacy of multi-wavelength observations. Optical data include the u -band observations from CFHT’s MegaCam (Sawicki et al., 2019), the high-resolution data from the Hubble Space Telescope (ACS-HST) in the F814W band (Koekemoer et al., 2007), the Hyper-Suprime-Cam (HSC) imaging in the g , r , i , z , and y bands, in addition to the 13 intermediate and narrow bands of Subaru Suprime-Cam (SC) (Taniguchi et al., 2015). The UltraVISTA survey (McCracken et al., 2012; Moneti et al., 2023) completes the coverage at NIR-wavelengths with the Y , J , H , and K_s bands, which are complementary to JWST’s NIRCам and MIRI.

The new COSMOS-Web photometric catalogue (also known as COSMOS2025; Shuntov et al., 2025a) has been developed, targeting specifically the portion of the COSMOS field observed by JWST, succeeding previous COSMOS galaxy catalogues (e.g., Ilbert et al., 2013; Laigle et al., 2016; Weaver et al., 2022). More than 784,000 sources have been detected over the 0.54 deg^2 area, using a PSF-homogenised χ^2 detection image, which combines all NIRCам filters. Source extraction is challenging due to the variation in PSF sizes across space- and ground-based facilities, ranging from 0.05 to 1.0 arcseconds. To address this, SourceExtractor++ (Bertin et al., 2022) has been used to model Sérsic surface brightness profiles (Sérsic, 1963) at NIRCам resolution, followed by photometric extraction in each band. The resolution of JWST images enables the separation of previously blended sources (Arango-Toro et al., 2025). A more detailed description of the COSMOS-Web photometric catalogue used in this study is provided by Shuntov et al. (2025a).

4.1.1 Photometric redshifts

Accurate photometric redshift estimation is crucial for reliably detecting galaxy groups, as the precision of these redshifts directly affects the identification and characterisation of such structures. Photometric redshifts in the COSMOS-Web source catalogue are computed with the template-fitting code LePhare (Arnouts et al., 2002; Ilbert et al., 2006), with an expanded parameter space

for the template library allowed by the depth and coverage offered by this field. The template library consists of 12 templates based on stellar population synthesis models (Bruzual & Charlot, 2003), with 42 different ages and including various star formation histories (SFHs) and metallicities ($Z = 0.008Z_{\odot}, 0.02Z_{\odot}$) as outlined by Ilbert et al. (2015).

In order to assess photometric redshift accuracy, the COSMOS-Web redshifts have been compared to more than 11,000 spectroscopic redshifts (spec-zs, hereinafter), with confidence level $>97\%$, from the compilation created by Khostovan et al. (2026, including GT) (e.g., Lilly et al., 2007, 2009; Kartaltepe et al., 2010, 2015; Silverman et al., 2015; Kashino et al., 2019). The photometric redshift precision is around 0.01 for $m_{F444W} < 24$, with a 2% rate of catastrophic failures¹. Even for the faintest galaxies at $26 < m_{F444W} < 28$ and for high redshifts in the interval we are interested in ($z < 4$), the precision remains better than 0.03 with $\sim 10\%$ failure (Arango-Toro et al., 2025; Shuntov et al., 2025b,a).

4.1.2 Data cleaning and visibility mask

We cleaned the data set by keeping only galaxies (including active galaxies) based on LePhare classification. We removed masked objects, limiting the selection to both FLAG_STAR_JWST=0 and FLAG_STAR_HSC=0 to ensure not only the availability and reliability of JWST photometry but also of the external ground-based photometry, which is used to estimate photometric redshifts. Removing unsafe regions for ground-based data implies a loss of area of around 10% compared to the area selected only considering JWST photometry quality. We decided to sacrifice this area in exchange for the improvement in photo-z uncertainties and in physical properties resulting from SED fitting that including good-quality ground-based photometry, ensures (see Shuntov et al., 2025b,a, for further details). A cleaned and high-quality sample is crucial for the study of the clustering of galaxies in three dimensions, while, on the contrary, the inclusion of badly characterised galaxies with uncertain and inaccurate photo-zs can contaminate the sample with spurious detections, for instance, in correspondence with artefacts. For this reason, we additionally kept only the galaxies with the best photo-z quality flag, LP_warn_f1=0, to reject artefacts like snowballs and sources with inconsistent photometry between different bands. We then cut the catalogue at the mode of the magnitude distribution in the F150W band, which we chose as the reference band (see Sect. 4.2). This minimises the introduction of noise and spurious detections in the catalogue and defines the depth of the galaxy catalogue, that is $m_{F150W} = 27.3$. As a reference, this catalogue is almost 2 magnitudes deeper than the one we presented in Chapter 3 and more than 3 magnitudes deeper than the cluster search performed in KiDS (Maturi et al., 2019, 2025). To further reduce the contamination of the galaxy sample due to poor classification, we removed

¹Photometric redshift performance is estimated in terms of the σ_{NMAD} indicator, that is $\sigma_{NMAD} = 1.48 \times \text{median}[(|\Delta z - \text{median}(\Delta z)|) / (1 + z_{spec})]$, where $\Delta z = z_{phot} - z_{spec}$, the difference between photometric and spectroscopic redshift. Galaxies are defined as outliers if $|\Delta z| > 0.15(1 + z_{spec})$.

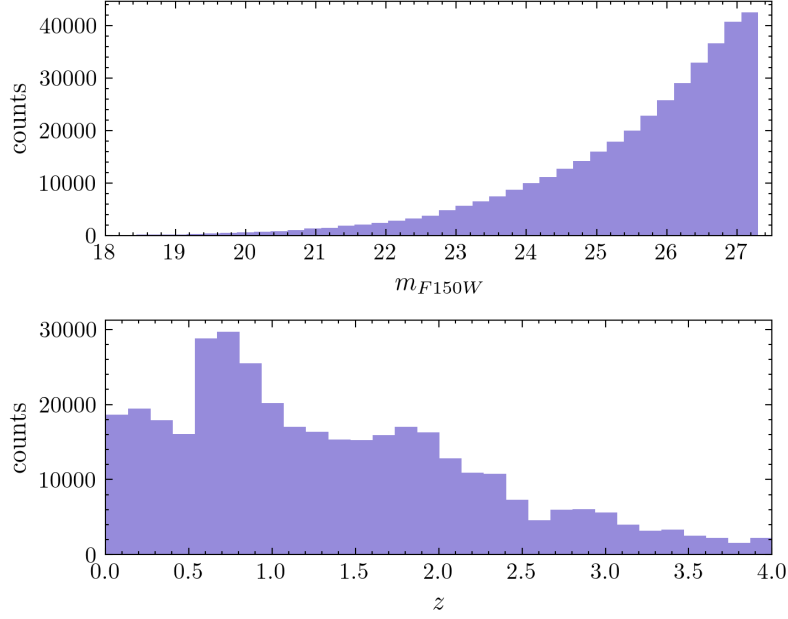


Figure 4.1. Distribution in magnitude (top panel) and redshift up to $z = 4$ (bottom panel) of the cleaned galaxy catalogue used as input for the group search.

all objects with an extremely small radius, choosing as a threshold radius $\sim 0.01''$, which is the value that divides into two distinct groups the sources in the m_{F150W} –radius plane. The total number of selected galaxies used for the group detection is 389248. The redshift and magnitude distribution of the cleaned input galaxy catalogue is shown in Fig. 4.1.

In order to include in our group detection the information about the footprint of the survey and the areas of the field that are inaccessible or potentially contaminated during the observation phase, we need to generate a so-called visibility mask. Using a visibility mask as input in the detection process has a twofold purpose: it filters out unsafe sources, excluding them from the input catalogue, and it allows the detection algorithm to account for inaccessible areas during the detection procedure. We based the visibility mask on the input galaxy catalogue, using an approach similar to that used in Chapter 3. In particular, we created the visibility mask starting by assigning the value 1 (masked) to all mask pixels devoid of galaxies in the selected catalogue described above. This ensures areas lying outside the field and areas affected by bright star halos are accounted for. Then, we built polygons to cover areas that may be affected by star spikes. To do this, we used the Incremental Data Release of the HSC bright-star masks by [Coupon et al. \(2018\)](#), extracted from Gaia DR2 ([Gaia Collaboration et al., 2018](#)), with magnitude $G < 18$. The spike and halo size follow an exponential relation with the magnitude of the star, with the same approach presented by [Coupon et al. \(2018\)](#), and the polygons were converted to binary masks using the `venice` code ([Coupon, 2018](#)). Besides these polygons, we visually inspected the galaxy density

maps in different redshift bins and manually masked areas with possible artefacts close to the star halos and/or field borders. During the visual inspection, due to suspiciously shaped overdensities at $z > 2$, we also discarded the area occupied by the central galaxies of a $z \sim 0.1$ known group in COSMOS, which is listed, for instance, by [Gozaliasl et al. \(2019\)](#) as ID20149. This makes this known bright object undetectable, but improves purity at higher redshifts. This masking procedure ensures our galaxy catalogue is as clean as possible from spurious or bad-photometry detections, especially for high- z detections.

The resulting composite visibility mask yielded an effective area in which the group search was performed of about 0.45 deg^2 .

4.2 Applying AMICO to the COSMOS-Web data

For this specific application, we adopted the same amplitude-map resolution used for the AMICO-COSMOS catalogue (see Chapter 3), that is $\Delta r = 0.3'$ (on the sky plane) and $\Delta z = 0.01$. We used AMICO with a single galaxy property, which we chose to be the magnitude in the JWST F150W band (simply m , hereinafter). This photometric band, in addition to having good resolution and low background noise with respect to other bands, was proven to be stable to the calibration of the mass-proxy scaling relation at high redshift observed in the H -band run described in Chapter 3. For what concerns the redshift probability distribution $p(z)$ of each galaxy, we decided to rely on an analytical Gaussian distribution that peaks at the redshift with the highest probability for each galaxy with its 1σ uncertainties.

Cluster model. To describe the expected distribution of cluster galaxies, we built an analytical model with a truncated Navarro-Frenk-White profile ([Navarro et al., 1997](#)) and a Schechter luminosity function ([Schechter, 1976](#)). We used the following parameters from the literature: normalisation as estimated by [Hennig et al. \(2017\)](#), concentration parameter from [Ragagnin et al. \(2021\)](#), and faint-end slope of the luminosity function from [Andreon et al. \(2014\)](#). The characteristic magnitude of the luminosity function, m_* , was estimated using evolutionary synthesis models with GALEV ([Kotulla et al., 2009](#)). We used the same GALEV configuration as in Chapter 3, evolving a massive ($\sim 10^{11} M_\odot$) elliptical galaxy. Figure 4.2 shows the redshift evolution of m_* overlapped with the selected galaxies (orange density contours) for a model with formation redshift, $z_f = 8$ and an exponentially declining star formation burst (purple line), and one with $z_f = 5$ and without burst (light blue line). The former marks the exponential cut-off in the number of galaxies and better describes the magnitude evolution trend of the data set up to at least $z \sim 3.7$, redshift to which we extended our search. This first m_* evolution trend was chosen to build the cluster model. We adopted a resolution in magnitude of $\Delta m = 0.5$ for the model and the noise describing group and field galaxies, respectively. This ensures sufficient galaxy statistics in each magnitude bin.

Noise model. We estimated the noise, which describes the distribution of field galaxies, by

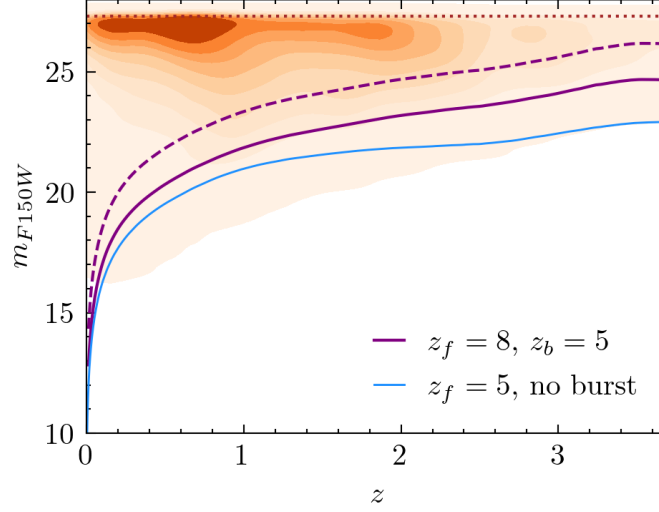


Figure 4.2. Magnitude and redshift of the galaxies of the input catalogue (orange density contours) and evolution with redshift of the characteristic magnitude of the luminosity function, m_* , for two different models with different formation redshifts and star formation burst, as shown in the legend. The model with $z_f = 8$ and past burst is the one used for the model in this work, marked by the solid purple line. This model better describes the trend of magnitude with respect to the one represented by the solid blue line. The dashed purple line indicates the same model, but for $m_* + 1.5$, which is the limit used in the definition of the intrinsic richness.

approximating it to the general distribution of the whole galaxy sample. Given the small area covered by this group search, we investigated the impact of group galaxies on the noise and observed visible peaks localised in redshift, which could be due to physical overdensities. We cleaned the noise model from the group galaxy contamination by dividing the field into four non-overlapping quadrants of roughly the same area and by taking the median noise value of each corresponding noise bin for the four quadrants as the final one. This cancels out the contribution of localised overdensities that are present in a specific quadrant and not in the others. Additionally, we performed a regularisation of the noise model similar to the one described in Chapter 3, attributing a large value to empty bins and those with $m < m_* - 3$.

4.3 Catalogue of galaxy groups

We performed a group search with AMICO over the effective area of 0.45 deg^2 in the COSMOS-Web field, as previously described. We ran the code down to $(S/N_{nocl})_{min} = 6.0$, where S/N_{nocl} is the AMICO signal-to-noise ratio, which does not include shot-noise from cluster/group members in the amplitude variance. When taken as a reference, this was proved to yield a more stable redshift dependence of the purity of the sample, with respect to the standard S/N , that includes both background and member contribution (Maturi et al., 2025). For further details on the definition

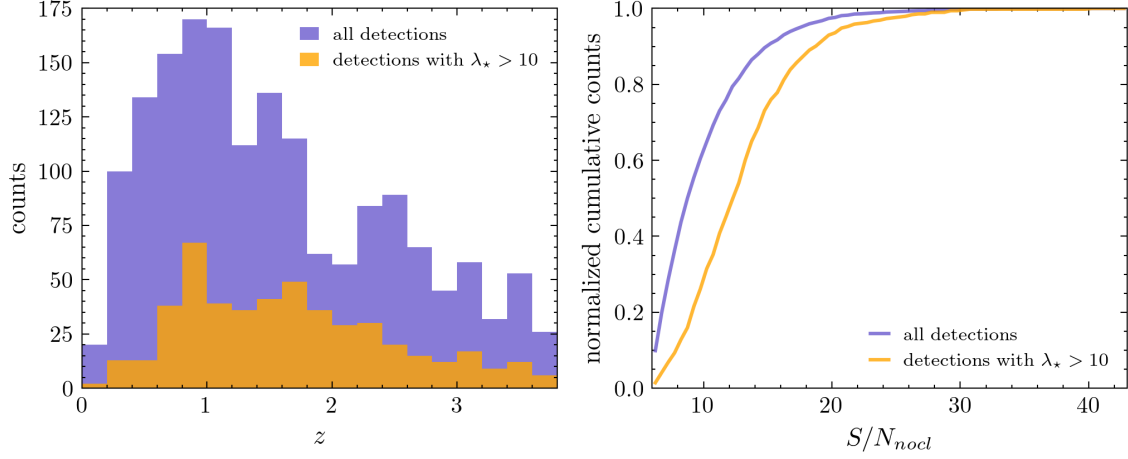


Figure 4.3. Redshift distribution (left panel) and normalised cumulative signal-to-noise distribution (right panel) for all the detections in the COSMOS-Web group catalogue. In orange, we show the richest detections, with $\lambda_\star > 10$.

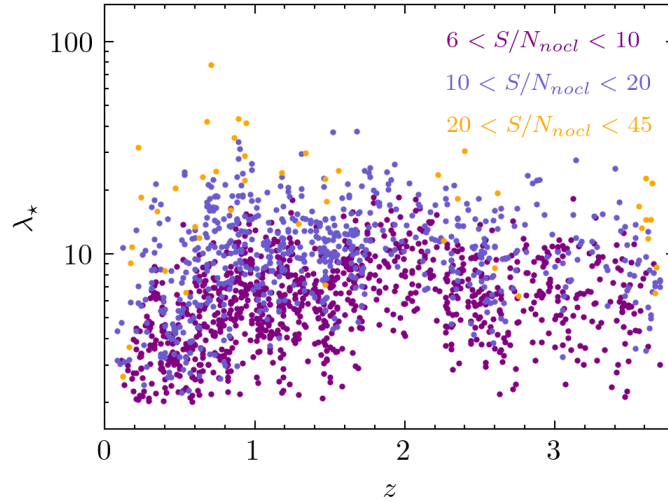


Figure 4.4. Intrinsic richness, $\lambda_\star$, for the sample of detected groups and its trend with redshift, in three different S/N_{nocl} bins as indicated in the plot.

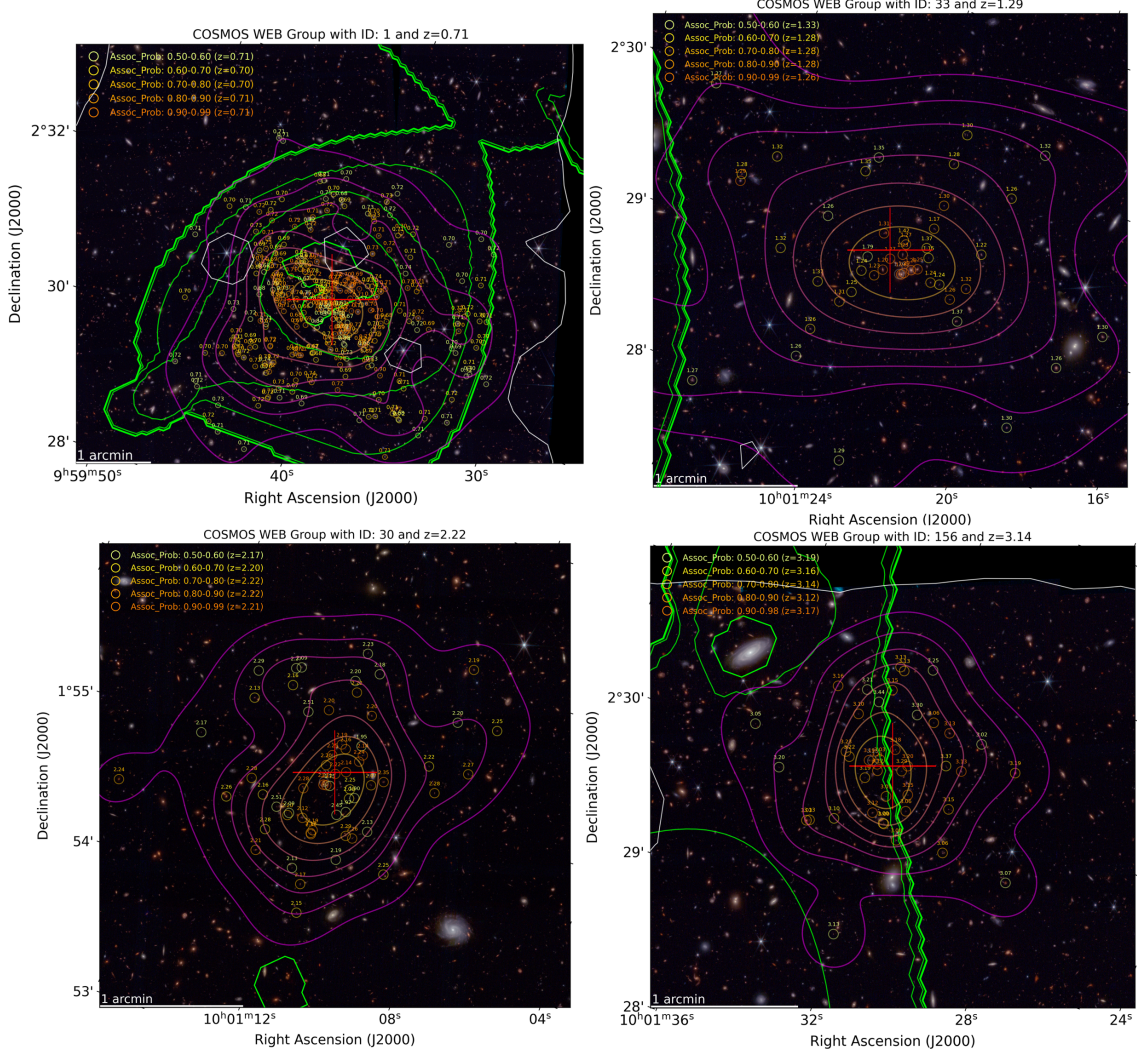


Figure 4.5. Four examples of detections present in the group catalogue, at different redshifts. Circles indicate member galaxies, colour-coded with membership probability and with their redshift printed next to each circle. Purple-to-yellow contours mark the density of galaxies and green contours the X-ray emission from the combined XMM-Newton and Chandra mosaic image in the 0.5–2 keV band. White lines delimit masked areas. The centre of the group is marked with a red cross. In the top left panel, a group at $z = 0.71$ with $\lambda_{\star} \sim 78$ and central AGN (the same group shown in Fig. 1.3). In the top right panel, a group at $z = 1.29$ with $\lambda_{\star} \sim 14$. In the bottom left panel, a group at $z = 2.22$ with $\lambda_{\star} \sim 24$ and on the bottom right panel, one at $z = 3.14$ with $\lambda_{\star} \sim 28$, flagged for being on the extended X-ray emission of a low- z group and for being at the edge of the field, as visible on the top of the stamp. JWST *rgb* images are in the same filters as in Fig. 1.3.

and on the difference between the two signal-to-noise ratios in AMICO, we refer the reader to Bellagamba et al. (2018) and Maturi et al. (2019). We then cut the catalogue at $\lambda_\star > 2$, which is a typical value to reject single and pairs of galaxies and minimise the number of unrealistic and spurious detections. Additionally, we rejected detections falling into the first and last redshift bin, which might be affected by border effects. The final catalogue we produced contains 1678 detections in the range $0.08 \leq z \leq 3.7$ and with $S/N_{nocl} > 6.0$. Despite AMICO has been widely used at $z < 1$ (Bellagamba et al., 2018; Maturi et al., 2019, 2023) and already tested (Euclid Collaboration: Adam et al., 2019), and applied to real data (see Chapter 3) up to $z = 2$, the detection of groups and clusters at even higher redshift, where they are still taking shape, is a quite unexplored regime, that we are interested in addressing with this work. The possibility of detecting objects at $z > 2$ with AMICO and the resulting sample derived from the application to the COSMOS-Web data are discussed in more detail in Sect. 4.4. Concerning the sample at $z < 2$, we compared the detections in the COSMOS-Web field with the previous COSMOS catalogue constructed with AMICO (described in Chapter 3). We performed a three-dimensional matching with maximum radial separation $dr = 1 \text{ Mpc}/h$ and maximum redshift separation $dz = 0.05(1+z)$, equivalent to $\sim 200 \text{ cMpc}$ at $z = 2$, which are typical values for group/cluster matching. In the matching, we allowed multiple associations in order not to exclude potential cases of fragmentation or over-merging of detections. For this comparison, we worked on the common volume covered by the two catalogues, which is defined by the interval $0 < z < 2$ and by the smallest effective area, namely the one of the COSMOS-Web field. Therefore, we discarded masked detections² from the previous AMICO-COSMOS catalogue according to the visibility mask described in Sect. 4.1.2. We found good correspondence between the two samples on the common volume, with a total number of 847 groups matched for the COSMOS-Web sample and 520 for the AMICO-COSMOS one. We explored the possibility for the difference in counts to signal fragmentation or over-merging. However, we found that the impact of fragmentation is marginal. Since we have used two different definitions of signal-to-noise as a lower limit to stop the detection process, we repeated the matching considering only COSMOS-Web detections with $S/N > 3.0$, which was the minimum threshold of the AMICO-COSMOS catalogue. This yielded a more consistent number of counts, with 400 matched detections for COSMOS-Web and 397 for AMICO-COSMOS. The percentage of matched detections for the AMICO-COSMOS catalogue is around $\sim 97\%$ and increases to $\sim 98\%$ when considering only detections with $S/N > 3.5$, which was identified as the threshold for selecting the most robust sample, based on the comparison between the detection performances in different photometric bands (see Toni et al., 2024, for details).

In Fig. 4.3, we show the redshift distribution of all the objects of this catalogue and the cumulative signal-to-noise distribution, marking with the orange line the distributions for the

²However, we allowed matching with masked detections if these fall within the tolerance of the matching.

detections with $\lambda_\star > 10$, a value to select the rich-end tail of the distribution in $\lambda_\star$ and identify the richest objects in the catalogue. In Fig. 4.4, we plot the intrinsic richness, $\lambda_\star$ vs z in three intervals of S/N_{nocl} . The trend of increasing intrinsic richness for increasing redshift observed up to $z \sim 2$ is expected, since the further we observe, the harder it is to detect poor and faint objects. At $z \gtrsim 2$, overdensities and numerous low- $\lambda_\star$ detections might be due to the fact that AMICO detects cores and substructures of clumpy extended protostructures rather than individual virialised clusters/groups. Additionally, the redshift interval $z \sim 2.4 - 2.6$ is occupied in COSMOS by several large overdensities of galaxies known, for example, as Hyperion and Colossus protoclusters (Cucciati et al., 2018; Lee et al., 2016, see Table 4.1 for further details and references). Figure 4.5 shows four examples of detections at different redshifts. The final group catalogue includes the columns described in Table 4.2, and it is accompanied by the list of member galaxies for each detection with membership and field probability as assigned by AMICO. As described in Section 2.3, this galaxy membership probability is calculated after AMICO detects candidates in the 3D amplitude map. In particular, the probability of the i -th galaxy belonging to the j -th detection is expressed by

$$P_{i,j} = F_i \frac{AM_c(\mathbf{x}_i)p(z_j)}{AM_c(\mathbf{x}_i)p(z_j) + N}. \quad (4.1)$$

Here, F_i is the probability of belonging to the field, which is used as a scaling factor to consider any previous associations of the galaxy. The field probability assumes a value of 1 for each galaxy when the iterations start and decreases at each new association. The vector containing the galaxy properties as described in Sect. 2.2, $\mathbf{x}_i$, contains here the sky position of the galaxy with respect to the group centre. The noise and the value of the galaxy redshift probability distribution at the group redshift, z_j , are here expressed by N and $p(z_j)$. The galaxy member catalogue and the membership probability for each galaxy associated with a detection are part of the output returned by AMICO. However, this membership association is by definition a model-dependent quantity. If one wants to study the properties of the clusters and groups themselves, such as the luminosity function or the density profile in a model-independent way, one needs to rely on a different method to identify possible members and field galaxies. For this reason, we included in our group catalogue the possibility of retrieving statistical membership, as it was done, for example, by Puddu et al. (2021) to study cluster luminosity functions. We proceeded as follows: in the membership catalog, we added a column named "FLAG_CYLINDER", which we assigned the value "1" for all galaxies located within a cylinder centered in the center of each detection, with a radius of $0.5 \text{ Mpc}/h$ and depth $0.01(1 + z)$; we then assigned the flag value "0" to all galaxies which do not fall in any of the cylinders. These flags help identify potential group and field galaxies without relying on the AMICO association probability. Besides this, we also created two three-dimensional maps of the volume covered by the group search, marking the effective area (taking into account only unmasked regions) inside or outside the cylinders, used to easily retrieve

the group and field volumes in a given redshift slice and perform background subtraction. This was done to make the membership catalogue usable for galaxy population and group property studies. However, when referring to group galaxies in this Chapter, we will make use of the membership probability directly returned by AMICO, selecting members as galaxies associated with $P > 50\%$, unless otherwise specified.

4.3.1 Purity and completeness of the sample with SinFoniA

We derived the purity and completeness of the sample by making use of data-driven mock catalogues, produced with the Selection Function extrActor (SinFoniA; [Maturi et al., 2019](#)). This method has already been applied to wide-field surveys such as KiDS ([Maturi et al., 2019, 2025](#)) and is currently part of the implementation of the Euclid pipeline. The idea is to generate mocks based on the input dataset used for the creation of our candidate catalogue, which is already divided into field and group galaxies via the association probability returned by AMICO. After this, the AMICO algorithm is applied to the mock galaxy catalogue in the same way as was done for the real data, and the list of generated mock groups is used as a "truth table" to study the resulting group catalogue. Therefore, the reference catalogue is matched with the results of the search and used to estimate the purity and completeness of the sample and the uncertainties on the retrieved properties of the groups. The SinFoniA algorithm exploits a Monte Carlo approach to create realisations of the universe and generate realistic mocks directly from the data. Therefore, the generated mocks are not based on any model or assumption, and they are able to reproduce the complexity of the real data. This reduces the possibility of introducing biases due to the cosmological assumptions, which are, for example, behind numerical simulations. Here, we briefly describe the main steps of the approach we adopted for this specific application and the results of our comparison against realistic mocks. For a complete description of SinFoniA, we refer the reader to [Maturi et al. \(2019\)](#).

Mock catalog. To create a data-driven mock catalogue, we proceeded as follows. First, we derived the cumulative distribution function (CDF) of a group sample identical to the original catalogue, but extending down to $S/N_{nocl} \sim 0$, and chose as reference ranking quantity the S/N_{nocl} returned by AMICO. We did this in redshift bins to account for a possible redshift dependence. The CDF of this sample is shown in Fig. 4.6, with different colours referring to different redshift bins. There is no significant redshift dependence for this application, proving once again the good quality of the photo- z estimation. Then, we extracted the group detections using the CDF at their signal-to-noise ratios as the probability to be extracted, since this can be interpreted as the likelihood of the detection to correspond to a real group. This was done to avoid a sharp cut in S/N_{nocl} when choosing which detections to consider. The field probability of galaxies is then recalculated by taking into account which detections have been rejected. Secondly, we extracted the field galaxies by using the membership information. The probability of the galaxy belonging

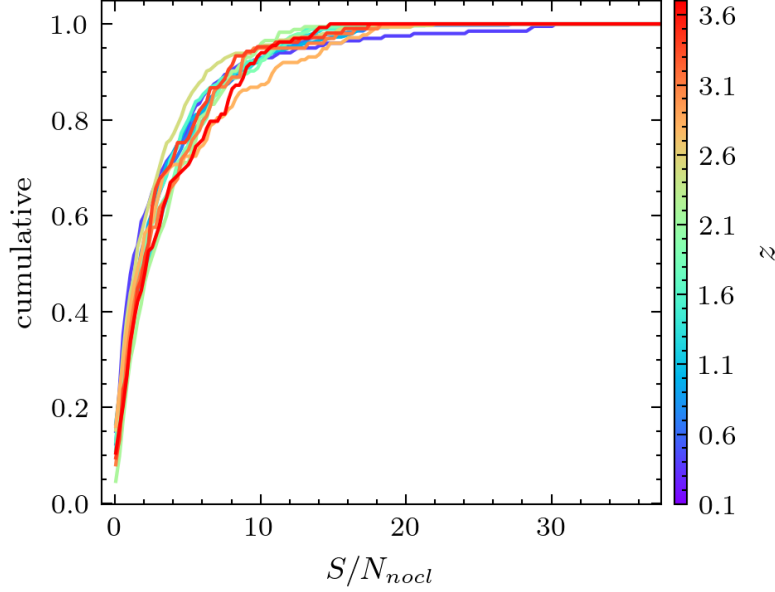


Figure 4.6. Cumulative distribution function (CDF) of S/N_{nocl} , in different bins of redshifts (colour bar on the right), used to select detections the mocks are based on, instead of using a sharp signal-to-noise cut.

to the field (F_i in Eq. 4.1) is used as the probability for the galaxy to be extracted as part of the mock field. For this application, we kept the original position and redshift of the field galaxies to maintain the noise spatial correlation and so the main features of the large-scale structure. Then, we generated the list of possible mock members. Once again, we extracted the possible members from the catalogue created via rejection based on the CDF.

All galaxies are then collected in bins of apparent richness and redshift so that the mock groups can be built by drawing from the full population in the corresponding bin. We used a bin resolution of $\Delta\lambda = 10$ and $\Delta z = 0.01$ for apparent richness and redshift, respectively. With this approach, mock groups are built with members extracted from stacks of different original detections, using the membership probability ($P_{i,j}$ in Eq. 4.1) as the probability of being extracted as members. However, this method tends to cancel out the intrinsic shape and orientation of clusters/groups. To restore their scatter in physical shape, we performed a coordinate transformation that does not affect density but introduces an ellipsoidal shape resembling that of dark matter halos in simulations (Despali et al., 2017). Mock group position and redshift are randomly shuffled, with maximum displacement values of $0.25 \text{ Mpc}/h$ and 0.01 in redshift. These values were chosen in order to introduce a perturbation that is not large enough to alter the features and structures on large scales, namely, to preserve the three-dimensional spatial correlation. Finally, we ran AMICO on the generated mock catalogue and used this as a reference to study the reliability of the original detections. Maturi et al. (2025) discuss the negligibility of the possible under-representation of

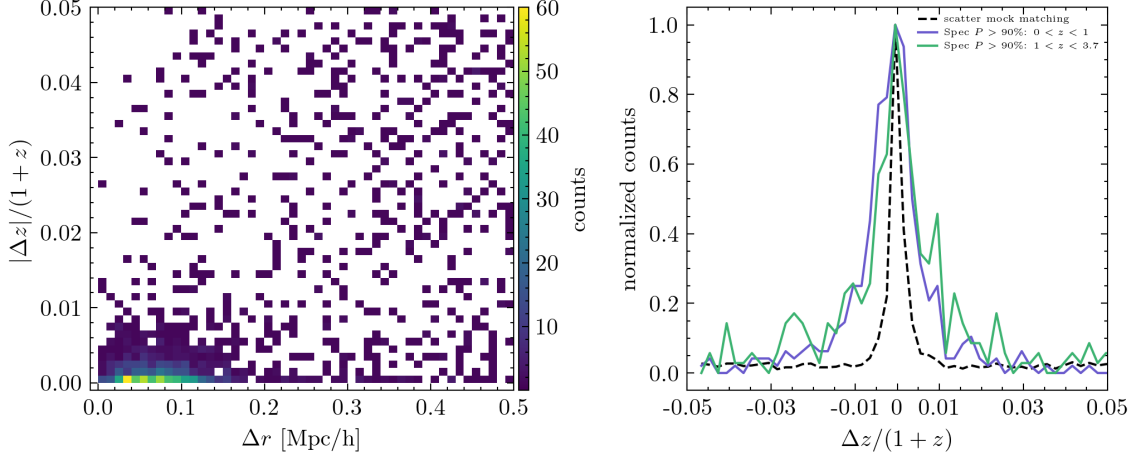


Figure 4.7. Left panel: Distribution of matched detections in the redshift scatter, $|\Delta z|/(1+z)$ - radial separation, Δr [Mpc/h] plane for an initial matching with maximum separation $dz = 0.05(1+z)$ and $dr = 0.5$ Mpc/h. Most of the matched detections are concentrated in the rectangular area at $\Delta r < 0.2$ Mpc/h and at $|\Delta z|/(1+z) < 0.01$. Right panel: Distribution of the redshift scatter for the mock matching (dashed black line) compared to the scatter between group photo- z and group spec- z estimated as the mean of spec-members with $P > 90\%$, in two intervals of redshift (solid purple and green lines). The distributions are normalised to 1.0 to better compare scatters.

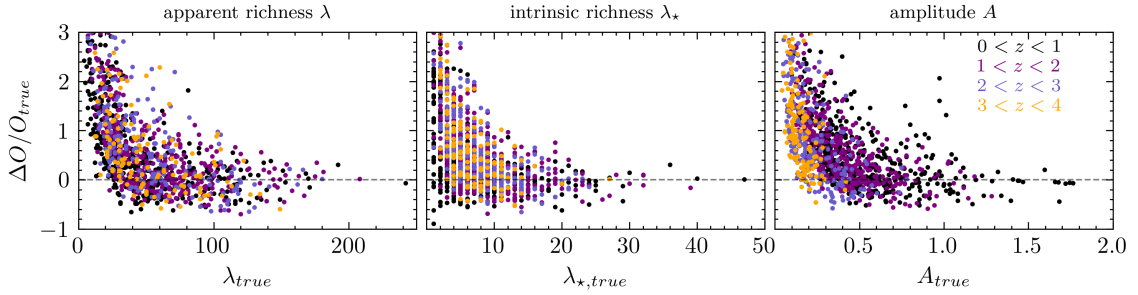


Figure 4.8. Relative scatter of the three observables (O , i.e., λ , $\lambda_\star$ and A , from left to right) between the detected observables, O_{det} , and the true observables as in the mocks, (O_{true}). The scatter is here expressed by $\Delta O = O_{det} - O_{true}$. Different colours mark different redshift bins as indicated in the plot on the right. Points with a scatter larger than 3.0 and the detection in Fig. 1.3 (which is much richer than the rest of the sample) are not shown for better visualisation of scatters and biases.

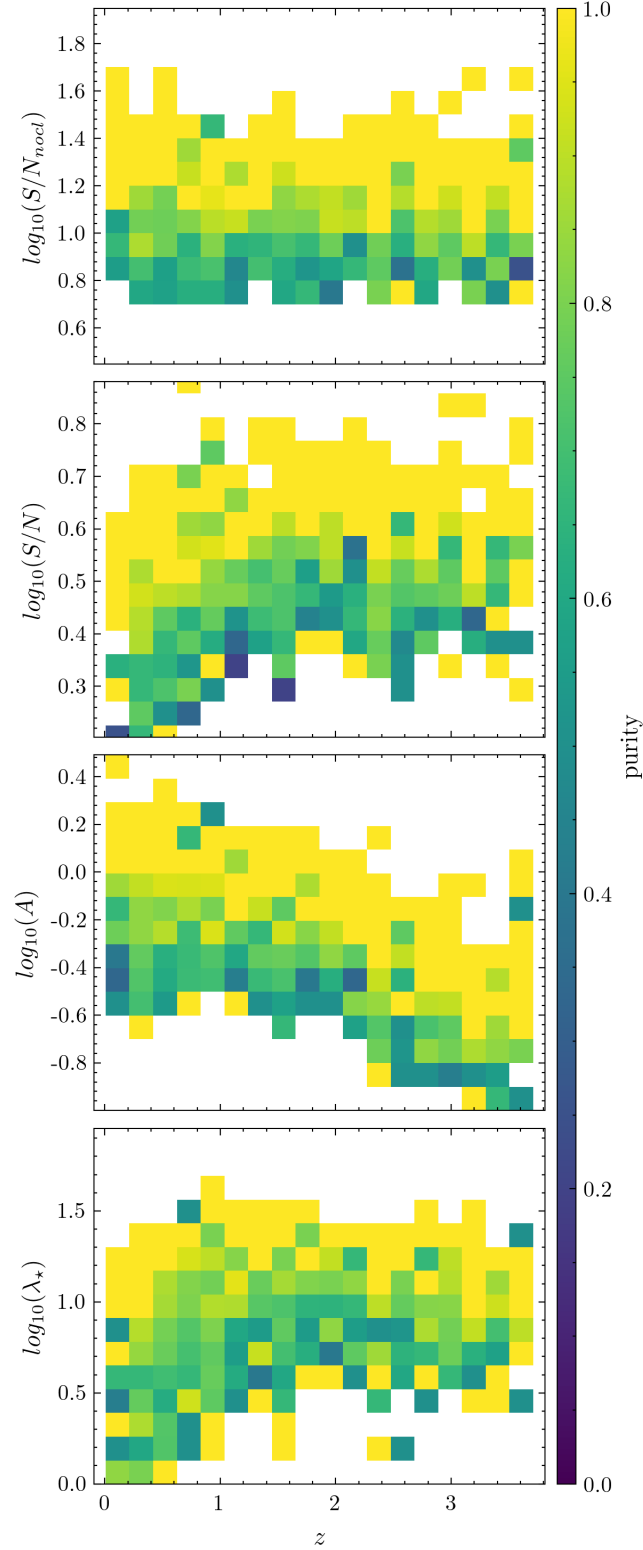


Figure 4.9. Purity of the group sample evaluated against the mock catalogue produced with SinFoniA. The four panels show the redshift dependence of (non-cumulative) purity referred to four different reference detection properties: signal-to-noise ratio without cluster contribution, with cluster contribution, amplitude, and intrinsic richness (from top to bottom).

small and blended objects in the mocks.

Matching procedure and group observables. When assessing the quality of detections through matching with a reference, it is important to ensure the good quality of the matching procedure itself. To do so, we performed an initial matching with relatively loose tolerance, namely $dz = 0.05(1+z)$ and $dr = 0.5 \text{ Mpc}/h$, and used an a priori sorting of the catalogues by richness to prioritise richer groups with respect to the more numerous poor groups. Then, we looked at the distribution of redshift and spatial separation between successfully matched detections, which is reported in the left panel of Fig. 4.7. The majority of matched detections are concentrated at $\Delta r < 0.2 \text{ Mpc}/h$ and at $|\Delta z|/(1+z) < 0.01$, suggesting these might be more appropriate tolerance values to be chosen for the matching in order to discard those likely to be random matches (see Fig. 4.7). We estimated the number of random matches by fitting the scatter distribution in redshift and position with a function given by the product of two Gaussian distributions, plus a constant that represents the value of the background consisting of random matches. We then repeated the same fitting with Cauchy distributions, which better represent the tails of the distribution. These tests showed that the background value is negligible with respect to the density inside the selected rectangle. We observed that typical scatter values in radial separation assume realistic values ($\Delta r = 0.2 \text{ Mpc}/h$) while for the redshift separation, the scatter distribution is very narrow, with most of the detections having their redshift matched within $\sim 0.005(1+z)$. In the right panel of Fig. 4.7, we show the normalised distribution of scatter in redshift for the mock matching (dashed black line) and for the comparison between photo- z and spec- z of the detected groups with spectroscopic counterparts (solid purple and green lines, for two redshift intervals). For details about the spectroscopic counterpart assignment, see Sect. 4.3.2. The shown spec- z values correspond to the mean spectroscopic redshift of members assigned with $P > 90\%$. We also tested lower probability thresholds, but the scatter did not significantly change. It is visible that the scatter in redshift in the mocks is significantly smaller when compared to the spectroscopic scatter.

The matching against the true catalogue allowed us to check for possible biases in the estimation of group observables, like the proxies of mass, amplitude, richness, and intrinsic richness. The relative scatter of the true and detected observable quantities is shown in Fig. 4.8 for apparent richness, λ , intrinsic richness, λ_* , and amplitude, A . No significant bias is visible in any of the observables and across the entire sample. As expected, the three proxies of mass are highly affected by the Malmquist bias, which is a selection effect and does not have to do with the detection method. This selection effect is significant for the smallest detections, as visible in Fig. 4.8, indicatively for $\lambda \lesssim 70$, for $\lambda_* \lesssim 10$, and for $A \lesssim 0.5$, with a slight redshift dependence for the latter.

Purity and completeness. The purity of a detected sample in reference to a true sample is defined as

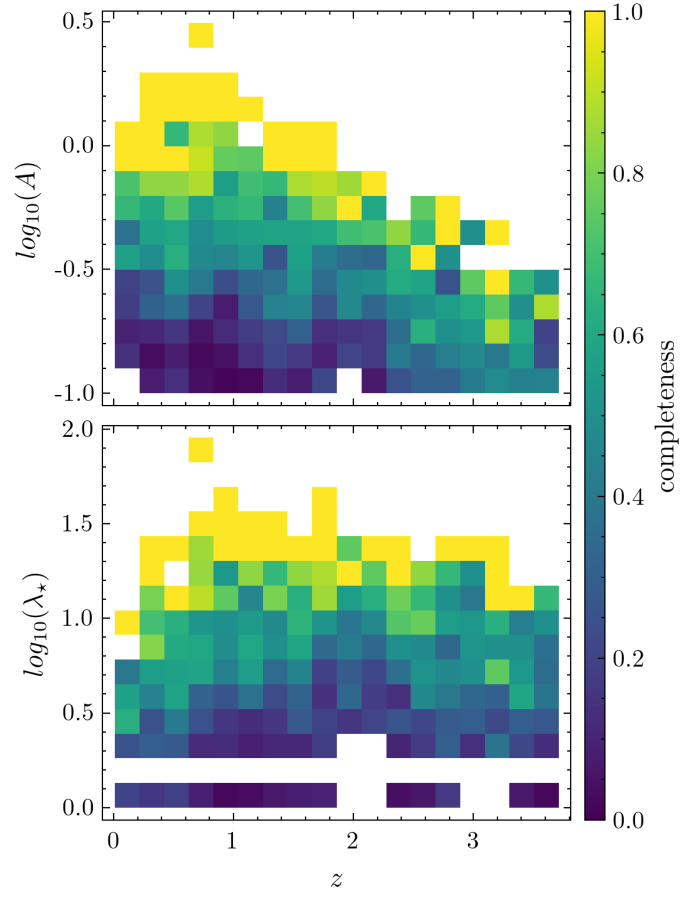


Figure 4.10. Completeness of the group sample evaluated against the mock catalogue produced with SinFoniA. The two panels show the redshift dependence of completeness referred to two different proxies of mass returned by AMICO: amplitude (top) and intrinsic richness, λ_* (bottom).

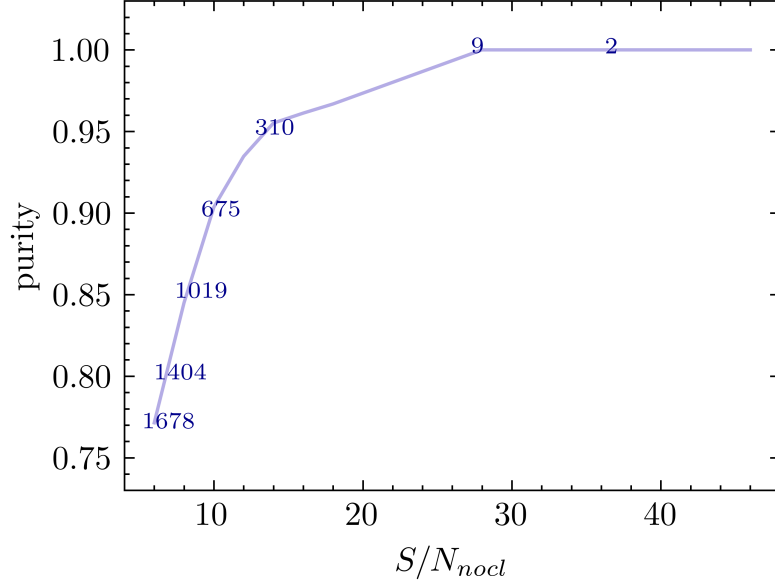


Figure 4.11. Relation between minimum S/N_{nocl} and purity of the sample. In correspondence with some reference values of purity, we report the number of selected detections.

the ratio of the successfully matched detections to the total number of detections; the completeness is computed as the ratio of the successfully matched detections to the total number of true objects. We estimated the purity and the completeness of our sample against the mock catalog produced with SinFoniA in bins of detected redshift and detected candidate properties, the two signal-to-noise ratios, S/N_{nocl} and S/N and the two main mass-proxies returned by AMICO, the amplitude A and the intrinsic richness, λ_* , for the purity and the two mass proxies for the completeness. The values of purity and completeness at different redshifts and for the different properties are shown in Figs. 4.9 and 4.10, respectively. As expected from the previous analyses of AMICO cluster and group samples from real data and simulations, the S/N_{nocl} value is the detection property with the best stability in redshift when analysing the purity of the catalogue. This means that this indication of signal-to-noise can be easily linked to a desired purity threshold. We found that, for this application, a cut at $S/N_{nocl} \sim 10$ (selecting around 670 detections) would identify a 90%-pure group sample, while a cut at $S/N_{nocl} \sim 7$ (selecting around 1400 detections) would set purity at 80%. Figure 4.11 shows the signal-to-noise cut (horizontal axes) to be applied to this group sample for each desired level of purity (vertical axes). Concerning completeness, for this sample it resulted in being less redshift-dependent for the intrinsic richness (bottom panel of Fig. 4.10), with an 80% completeness reached at around $\lambda_* \sim 16$.

4.3.2 Spectroscopic counterparts

We associated spectroscopic redshifts with our detections, using the compilation of spectroscopic catalogues for the COSMOS field by [Khostovan et al. \(2026, including GT\)](#). Public and private surveys included are, for instance, [Lilly et al. \(2007, 2009\)](#); [Kartaltepe et al. \(2015\)](#); [Hasinger et al. \(2018\)](#); [Kashino et al. \(2019\)](#) and many others. Additional references can be found in [Khostovan et al. \(2026, including GT\)](#). We selected sources with high-quality spectra, namely with a confidence level larger than 80%. This yields a spectroscopic sample with more than 66,000 galaxies with spectroscopic redshift in the COSMOS field. We assigned spectroscopic counterparts to our galaxy members identified by AMICO, via positional matching within $1''$, which is the same separation used to assign counterparts from different surveys in the compilation ([Khostovan et al., 2026, including GT](#)). Thanks to the large spectroscopic coverage offered by the COSMOS field, we were able to assign a spec- z to around 20% of the group members assigned by AMICO (4300 members assigned with $P > 50\%$ have a high-quality spectroscopic redshift). We defined the quality of the spectroscopic association to the group as Q , which is the ratio of the sum of the membership probabilities of the spectroscopic members to the number of spectroscopic members, $Q = \sum_i P_i / N_{gal}$. We found that 535 detections in total have at least three members with a spec- z . 948 detections have at least one galaxy member with spec- z and $Q > 70\%$. A total of 1075 detections have at least 3 spectroscopic members or an association with $Q > 70\%$. Among these, around 80% of the group candidates have the redshift assigned by AMICO that is consistent with z_{spec} . Here, z_{spec} is defined as the mean spectroscopic redshift of the associated members. The consistency criterion is met when the relative difference between the two redshifts is less than 3%, that is, $|\Delta z| / (1 + z_{spec}) < 0.03$, which is around the largest σ of the photo- z uncertainty. By assigning spectroscopic counterparts to our group members, we were able to exclude the presence of any significant general z -dependent bias between photo- z s and spec- z s, namely between AMICO redshift (based on galaxy photo- z s) and spectroscopic redshift of the group, estimated as the average redshift of the available spectroscopic members. This kind of bias was, for instance, found in the KiDS sample at low redshift and consequently corrected ([Maturi et al., 2025](#)). The mean bias of the sample with more than 3 spectroscopic members is $\Delta z / (1 + z_{spec}) = -0.011 \pm 0.051$.

However, even if a significant mean bias characterising the sample is not visible, individual detections may show an inconsistency between the redshift assigned by AMICO and the one based on the available spectroscopic members. Therefore, we studied and marked with a dedicated flag all the detections that have a spectroscopic inconsistency of redshifts. For this analysis, we used only groups with at least 4 spectroscopic members or with spectroscopic quality $Q > 95\%$. The flag is assigned to all detections that have relative redshift scatter, $|\Delta z| / (1 + z_{spec}) > 0.15$, which is the same limit chosen to compute outlier fractions for galaxies in the COSMOS-Web galaxy catalogue (see e.g., [Casey et al., 2023](#); [Shuntov et al., 2025b](#)). Only 30 detections are affected

by spectroscopic mismatch as just described. This confirms the already assessed good photo- z quality of COSMOS-Web data used for this group search (Shuntov et al., 2025b; Arango-Toro et al., 2025). In this work, we also leveraged the availability of a large compilation of spec- z s to test the AMICO algorithm on a hybrid galaxy sample, using also the spectroscopic redshifts, when available. This application is described in Appendix A.

4.3.3 Detection flags

Besides flagging the detections in our catalogue without spectroscopic members and those with spectroscopic members with redshift not compatible with the photometric one, we added useful flags to help filter the sample depending on the study that has to be performed with this catalogue. Our flagging system is based on a list of base-2 flag bits referring to the following properties, ordered starting from the least severe:

1. lack of spectroscopic members
2. less than 3 arcmin from a border edge
4. X-ray projection or proximity flag ³
8. masked fraction larger than 25%
16. central X-ray selected AGN (obtained via matching with the X-ray sources in COSMOS-Web or COSMOS2020)
32. spectroscopic mismatch (see Sect. 4.3.2)
64. low intrinsic richness ($\lambda_\star < 5$)

In our detection catalogue, we introduced the flag "DETECTION_FLAG" which represents the sum of all these detection flag bits. Thus, if more than one criterion is present, the flags are summed together; for example, a group without spectroscopic members and $\lambda_\star = 3$ will be flagged with 65. According to this flagging system, around 67% of the entire sample down to $S/N = 6.0$ is flagged with a value smaller than 20, which is an example of how to select robust detections based on their individual properties.

4.4 Candidates at $z \geq 2$

The depth of this catalogue and the availability of high-quality photometric redshifts make this application ideal for testing the AMICO algorithm in the relatively unexplored regime of high-redshift groups, protogroups, and protoclusters.

³In the same way as it was done in Chapter 3 to clean the sample for the calibration of scaling relations, this flag marks detections within 0.02 deg from the centre of a rich detection or lying on their extended X-ray emission.

4.4.1 COSMOS protocluster compilation

In order to benchmark our results in this challenging regime, we first collected literature about structures already known in the COSMOS field. We created a compilation of all currently known clusters, groups, protoclusters, or protocluster cores in the range $2.0 \lesssim z \lesssim 4.0$, which is the interval we are interested in for this comparison.

The collected detections with relative references and sample properties, like position and redshift, are summarised in Table 4.1. It should be noted that for this application of the AMICO algorithm to the COSMOS-Web field, we made use of a cluster model, as described in Sect. 4.2, so we are relying on our knowledge of how galaxies are distributed in clusters and groups at low redshift, which might not be what we actually see at $z > 1.5$. According to currently favoured scenarios for cluster formation, present-day galaxy clusters and groups are the aftermath of the assembling, growing, and maturing of protocluster cores (see e.g., Shimakawa et al., 2018). In this high- z group search, we expect the algorithm to detect cores or possibly virialised substructures of protoclusters which may extend for several tens of Mpc. For this reason, to perform a consistent comparison of our high- z sample with the known structures in the literature, we considered peaks and substructures inside extended protoclusters as individual objects. Just as an example, the 10 density peaks of Elentari (Forrest et al., 2023) and the 7 peaks of Hyperion (Cucciati et al., 2018) are counted as 17 different objects in our protocluster compilation (see Table 4.1). In this compilation, we included not only objects detected as overdensities of photometric or spectroscopic redshifts (e.g., Chiang et al., 2014; Diener et al., 2013; Sarron & Conselice, 2021), but also discovered with several other approaches and methods, using, for example, the emission from distant galaxies or radio-galaxies (e.g., Geach et al., 2012; Castignani et al., 2014b; Daddi et al., 2022), the mapping of Lyman α forest (e.g., Lee et al., 2014) and the group X-ray emission (e.g., Wang et al., 2016, Gozaliasl et al. in prep.). Since this protocluster compilation might be used as a reference also in other studies beyond the comparison in this work, we included all detections in the COSMOS field, and not only in the COSMOS-Web portion of it. In case the object falls outside the COSMOS-Web field, we make that explicit in Table 4.1. Additionally, we did not perform internal matching to discard possible double detections, since it is not straightforward to define the spatial limits of protoclusters. However, if the discovery of a structure is generally attributed to more than one work, we report the corresponding references (see Col. 6 of Table 4.1).

4.4.2 Our candidates in the COSMOS-Web field

We ran a three-dimensional matching between our catalogue and the compilation, in which we allow an AMICO detection to have more correspondences in the compilation, given that we did not discard objects potentially detected multiple times, as previously mentioned. The matching was run as follows: for each AMICO detection, we attributed a successful match with every other

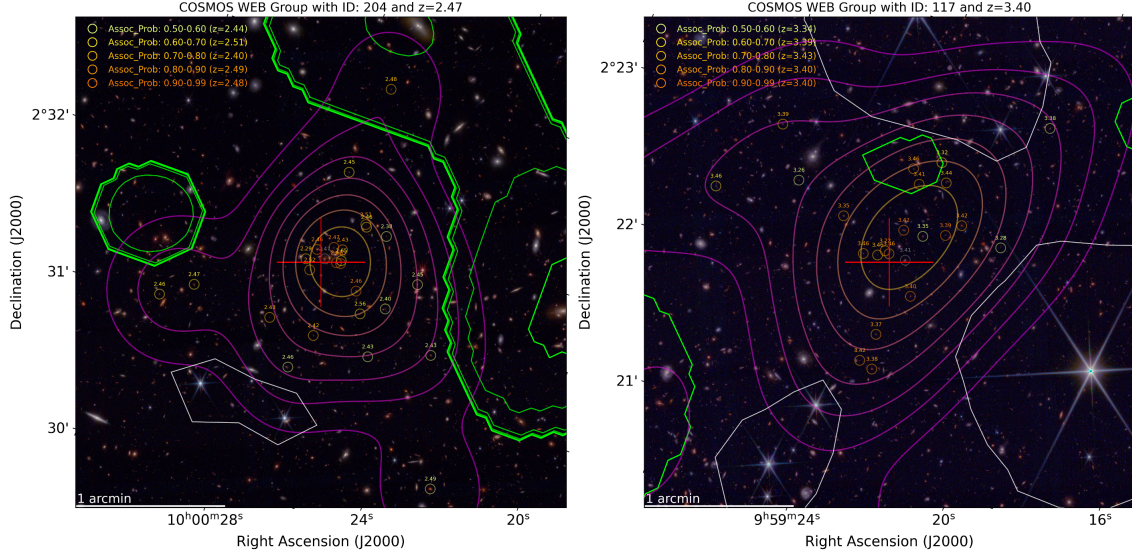


Figure 4.12. Two examples of high- z detections not known in the literature. Stamps are JWST colour-composite images, annotations are the same as described in the caption of Fig. 4.5. Left: Candidate at $z = 2.47$ with $\lambda_\star \sim 11$. Right: Candidate at $z = 3.40$ with $\lambda_\star \sim 25$.

known object that lies within a cylinder of radius $dr = 1.0 \text{ Mpc}/h$ and redshift depth $dz = 0.05(1+z)$. These values are similar to scatter values chosen in literature for matching clusters/groups at very high redshift (see e.g., [Sarron & Conselice, 2021](#)). In this procedure, we used an a priori sorting of the AMICO catalogue by $(S/N)_{\text{no cl}}$. In the last column of Table 4.1, the matched protoclusters are indicated with a checkmark. Whenever the reference detection is not a single object but a catalogue or a structure with multiple peaks, we indicate next to the checkmark the percentage of matched peaks or detections over the total number of objects falling in the available area according to our visibility mask (described in Sect. 4.1.2). We successfully matched 205 of our groups to be compatible with detected protoclusters or substructures of protoclusters at $z \geq 2$ already known in literature, like the Hyperion protocluster structure ([Cucciati et al., 2018](#); [Casey et al., 2015](#); [Wang et al., 2016](#)), Elentari ([Forrest et al., 2023](#)), CC2.2 ([Darvish et al., 2020](#)), and others. We identified a total number of 316 new high- z objects⁴ with $2 \leq z \leq 3.7$. These are interesting candidate protocluster cores and protogroups that do not match any of the structures we collected from the literature in the field. Two examples of new high- z detections are shown in Fig. 4.12, one object is at $z = 2.47$ and one at $z = 3.40$.

As mentioned above, at such high redshifts, we detected with AMICO groups that might be cores or density peaks of more extended protoclusters. For this reason, we explored the possibility of identifying protoclusters made of multiple AMICO detections by applying a 3D clustering algo-

⁴During the revision phase of the paper this Chapter is based on, a catalogue of overdensity peaks was published by [Hung et al. \(2025\)](#). This catalogue includes objects compatible with $\sim 49\%$ of our new detections. The overall matching rate is $\sim 93\%$.

rithm. We treated our high- z detections as points with 3D coordinates and performed a clustering analysis with the `sklearn` algorithm for Density-Based Spatial Clustering of Applications with Noise (DBSCAN; Ester et al., 1996). We used a minimum of two members per cluster, to allow any cluster size down to pairs of detections, and a clustering scale of 0.05 deg as angular separation and $0.02(1+z)$ as maximum redshift separation. This analysis resulted in 111 potential large-scale protoclusters, with a maximum size of 14 detections in the same object. We show the largest of these structures in Fig. 4.13, to which we gave the name of AmicOne (pronounced [a'mikone] or [a'mikwʌn]), the first and largest protocluster candidate detected with AMICO. This structure is formed by the clustering of 14 different cores, detected in the range $2.5 \leq z \leq 3.09$, and it includes two groups consistent with those (IDs 37 and 47 in our compilation, in Table 4.1) detected by Diener et al. (2013). The diametral size of the structure is about 20 cMpc on the plane of the sky at $z = 2.65$, and it extends over a wide range of redshifts, equivalent to $\sim 600 \text{ cMpc}$.

4.5 Conclusions

We produced a deep galaxy group catalogue based on the new COSMOS-Web photometric galaxy catalogue down to $F150W < 27.3$. The detection procedure was performed with the AMICO algorithm, a widely tested cluster detector currently used in Euclid, KiDS, and other major surveys. The group search was performed over an effective area of 0.45 deg^2 and up to $z = 3.7$, covering the transition from protoclusters to virialised objects and their maturing phase. We detected 1678 groups with a signal-to-noise ratio (S/N_{nocl}) larger than 6.0. For each detection in our catalogue, we provided information about position, redshift, signal-to-noise ratio, mass proxies (like amplitude A and intrinsic richness, λ_*), masked fraction, flags indicating detection features, like spectroscopic counterparts, presence of AGNs, and spectroscopic confirmation of redshift. In addition to the detection catalogue, we also provide a list of members with their association probability and the information needed to estimate statistical membership. The structure of the released catalogue with the description of each column is summarised in Table 4.2.

We evaluated the purity of our sample against realistic mocks generated with the SinFoniA algorithm. This method allows for the creation of mock catalogues, capturing the complexity of real data, without relying on any a priori assumption. We found the relationship between purity and S/N_{nocl} , which is the detection property that yields the most redshift-independent relation with purity. This makes it possible to establish cuts of the original catalogue based on the desired purity level. For example, we found that a cut at $S/N_{\text{nocl}} \sim 10$ would correspond to selecting a sample of around 670 objects with purity above 90%.

We leveraged the successful application to COSMOS data up to $z = 2$ presented in Chapter 3 and the good quality of the photometric redshift and depth of the COSMOS-Web survey to create this deep catalogue of galaxy groups and to explore the possibility of detecting protocluster cores

with AMICO at $z \geq 2$. We successfully detected 316 new objects at $z \geq 2$ and 205 compatible with being part of protoclusters and high- z groups known in the literature. To perform this comparative analysis, we created a compilation of known objects at $z > 2$ in COSMOS and matched them with our detections. In total, our catalogue contains 509 candidate groups and protocluster cores in the range $2 \leq z \leq 3.7$, of which around 400 are not isolated but lie within 111 Mpc-scale structures we found by performing a 3D clustering analysis with a minimum cluster size of 2, with the DBSCAN algorithm. Among these structures, the one including the highest number of substructures (14 cores) is a large protocluster candidate found at redshift $z \sim 2.5 - 3.0$, and we have assigned it the name AmicOne.

Such a deep and well-characterised sample of groups, extending over a wide range of richness and redshift, is an important resource for the study of group and cluster assembly and evolution, from high- z cores to the massive objects we observe today. Additionally, it makes it possible to study several aspects of galaxy populations in different environments. For instance, it provides information about the numerous but relatively unexplored population of low-mass groups, and about galaxies in the outskirts of groups and clusters.

Table 4.1. Compilation of protoclusters and high-redshift clusters and groups known in the literature for the COSMOS field.

ID	Name	RA [deg]	DEC [deg]	z	Reference(s)	Matched
1	ZFOURGE/ZFIRE	150.094	2.251	2.095	(1), (2)	✓
2	G237	150.507	2.312	2.160	(3)	
311	-	150.474	2.328	2.160	(4)	✓
315	-	150.208	2.015	2.230	(5)	✓
3	CC2.2	150.197	2.003	2.232	(6)	✓
314	-	150.060	2.200	2.300	(7)	
12-16	COSTCO	[149.706, 150.129]	[2.024, 2.275]	[2.047, 2.391]	(8), (9), (10)	✓ (100%)
312	Colossus	150.060	2.310	2.440	(11), (7), (12), (13)	
313	-	150.130	2.373	2.470	(14)	✓
310	CL J1001+0220	150.250	2.333	2.506	(15)	✓
5-11	Hyperion	[149.958, 150.331]	[2.112, 2.404]	[2.423, 2.507]	(16)	✓ (100%)
341-345	LATIS1-D2	[149.532, 150.347]	[1.988, 2.274]	[2.457, 2.685]	(17)	✓ (100%)
91-266	-	[149.409, 150.749]	[1.628, 2.793]	[2.000, 2.840]	(18)	✓ (82%)
4	QO-1000	150.043	1.694	2.770	(19)	∅
335-340	-	[149.601, 150.620]	[1.593, 2.289]	[2.010, 2.900]	(20), (21)	✓ (100%) *
316	VPC-1000	150.068	1.967	2.900	(22)	✓
17-55	-	[149.870, 150.588]	[1.766, 2.603]	[2.013, 2.957]	(23)	✓ (82%)
267-297	-	[149.509, 150.693]	[1.707, 2.674]	[2.070, 3.080]	(24)	✓ (92%)
334	MAGAZ3NE	150.429	2.506	3.120	(25)	✓
321-326	-	[149.583, 150.346]	[1.967, 2.603]	[2.194, 3.295]	(26), (27)	✓ (80%)
317-320	CCPC	[149.980, 150.070]	[2.110, 2.280]	[2.098, 3.303]	(28)	✓ (100%)
308-309	MAGAZ3NE	[149.850, 150.117]	[2.333, 2.564]	[3.367, 3.380]	(29)	✓ (100%)
298-307	Elentari	[149.732, 150.424]	[2.275, 2.583]	[3.248, 3.409]	(30)	✓ (100%)
56-90	-	[149.410, 150.616]	[1.543, 2.794]	[2.000, 3.500]	(31)	✓ (82%)
327-333	-	[149.706, 150.751]	[1.798, 2.762]	[2.323, 3.613]	(32)	*

(1) Spitler et al. (2012); (2) Yuan et al. (2014); (3) Polletta et al. (2021); (4) Koyama et al. (2021); (5) Geach et al. (2012); (6) Darvish et al. (2020); (7) Lee et al. (2016); (8) Ata et al. (2022); (9) Dong et al. (2023); (10) Edward et al. (2024); (11) Lee et al. (2014); (12) Diener et al. (2015); (13) Chiang et al. (2015); (14) Casey et al. (2015); (15) Wang et al. (2016); (16) Cucciati et al. (2018); (17) Newman et al. (2020); (18) Sarron & Conselice (2021); (19) Ito et al. (2023); (20) Castignani et al. (2014b); (21) Castignani et al. (2019); (22) Cucciati et al. (2014); (23) Diener et al. (2013); (24) Chiang et al. (2014); (25) McConachie et al. (2025); (26) Daddi et al. (2021); (27) Daddi et al. (2022); (28) Franck & McGaugh (2016); (29) McConachie et al. (2022); (30) Forrest et al. (2023); (31) Gozaliasl et al., in prep.; (32) Sillassen et al. (2024);

The table shows arbitrary identification numbers in our compilation (ID), name of the protocluster or the survey, if available (Name), coordinate in degrees or range of coordinates (RA, DEC), redshift or range of redshifts (z), reference(s) and whether or not the objects are matched with one of our detection (and in which fraction if the reference is a catalog or list of objects). (∅) Outside the COSMOS-Web field; (*) All outside the COSMOS-Web field except for one object, which is unmatched; (*) All outside the COSMOS-Web field except for two objects, both matched.

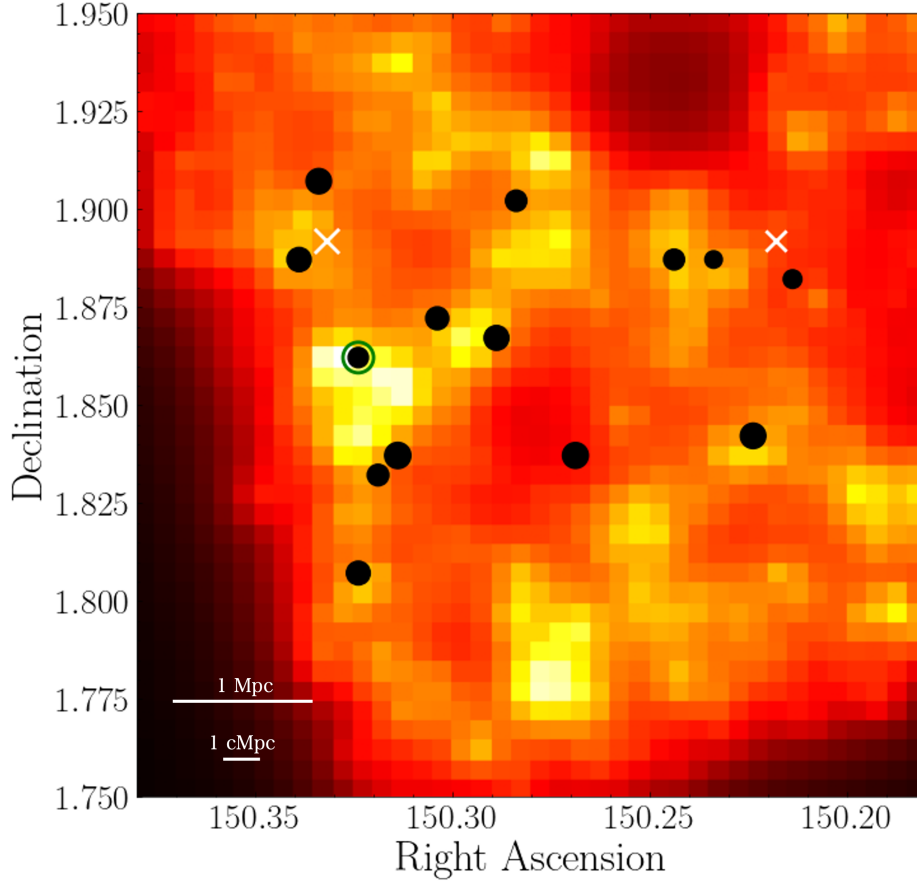


Figure 4.13. AmicOne protocluster candidate, consisting of fourteen cores (AMICO detections) with $z \in [2.5, 3.09]$ (black points) clustered together in the same large-scale structure, according to our clustering analysis. In the background, the amplitude map returned by AMICO at the redshift slice of the detection marked by the green circle, namely at $z = 2.65$. Detections (black points) have different redshifts in the range; this is why they are not all lying on an amplitude peak. The amplitude peaks at their individual redshifts. The white crosses mark the position of the protogroups found by [Diener et al. \(2013\)](#). The size of the points is proportional to the redshift.

Catalog	Column Name	Description	Range
Group catalog	ID (1)	Group identification number	1–1816
	RA (2)	Right Ascension (R.A.) in degrees	149.66–150.56
	DEC (3)	Declination (Dec.) in degrees	1.73–2.67
	Z (4)	AMICO group redshift	0.08–3.70
	SN (5)	S/N (incl. background + cluster)	1.28–6.65
	SN_NOCL (6)	S/N_{nocl} (S/N incl. only background)	6.00–43.11
	AMP (7)	Signal amplitude	0.12–2.51
	MSKFRC (8)	Masked fraction	0.03–0.79
	LAMBDA (9)	(Apparent) richness	14–456
	LAMBDA_STAR (10)	(Intrinsic) richness ($m < m_* + 1.5$, $r < R_{200}$)	2.0–77.6
	DETECTION_FLAG (11)	Detection flag (see Sect. 4.3.3)	0–120
	Z_SPEC (12)	Mean spectroscopic redshift (if available)	0.08–3.80
	N_SPEC (13)	Number of members with spectroscopic redshift	0–103
	ZPHYS_SIGM (14)	1σ lower uncertainty on redshift	0.004–0.281
	ZPHYS_SIGP (15)	1σ upper uncertainty on redshift	0.000–0.192
Member catalog	GALID (1)	COSMOS-Web galaxy identification number	0–783999
	FIELD_PROB (2)	Probability to belong to the field	0–0.995
	ID (3)	Group identification number	1–1816
	ASSOC_PROB (4)	Probability to belong to the group indicated in "ID"	0.005–1

Table 4.2. Descriptions of columns for the group and member catalogue.

II

Characterisation

5

Quiescent population build-up

INVESTIGATION OF RED-SEQUENCE AND QUIESCENT FRACTION EVOLUTION IN COSMOS GROUPS UP TO $z = 3.7$

In this Chapter, we investigate the redshift evolution and group richness dependence of the quiescent galaxy fraction and red-sequence parameters in COSMOS galaxy groups (detected in **Part I** of this Thesis), spanning a wide redshift range, from $z = 0$ to $z = 3.7$. We relied on a machine-learning method to estimate quiescence probability. We computed quiescent fractions via two approaches: one using AMICO membership probabilities and one model-independent. The red sequence (RS) was identified by estimating the ridgeline with probability-weighted photometry and $\kappa\sigma$ clipping, applied to both rest- and observed-frame magnitudes. We find that the quiescent galaxy population steadily build up from $z = 1.5 - 2$, with faster growth in the richest systems. We report a rare overdensity of quiescent galaxies at $z = 3.4$, among the most distant early RS observed. Using COSMOS2025's broad wavelength coverage, we trace RS evolution over ~ 12 Gyr, finding no significant trends in slope or scatter.

This Chapter is fully based on Toni et al. (2025b), A&A, accepted & in press.

5.1 Catalogue of galaxies and galaxy groups

For this analysis, we used the COSMOS-Web photometric galaxy catalogue (also known as COSMOS2025; Shuntov et al., 2025a), which specifically targets the COSMOS-Web Survey (Casey et al., 2023). The survey covers 0.54 deg^2 with four filters F115W, F150W, F277W, and F444W in NIRCcam (Rieke et al., 2023; Franco et al., 2025), achieving 5σ point-source depths of 27.5–28.2

mag. These filters offer a coverage of the near-infrared regime that is highly complementary to the UltraVISTA survey filters. An additional non-contiguous area is observed with the F770W MIRI filter (Wright et al., 2022; Harish et al., 2025).

The COSMOS2025 catalogue sources were detected using a PSF-homogenised χ^2 image combining all NIRCам bands, using SourceExtractor++ (Bertin et al., 2022) to address PSF variations across datasets. This approach boasts high spatial resolution reflected in the deblending power of the source extraction. In the COSMOS2025 catalogue, redshifts (and all physical properties) are derived using the LePhare template-fitting code (Arnouts et al., 1999; Arnouts et al., 2002; Ilbert et al., 2006) with an expanded set of templates based on Bruzual & Charlot (2003) models (Ilbert et al., 2015). When compared to high-confidence spectroscopic samples, the photometric redshifts achieve a precision of ~ 0.01 for bright galaxies ($m_{F444W} < 23$) with $< 2\%$ catastrophic outliers, and maintain < 0.03 precision with $\sim 10\%$ outliers even for $26 < m_{F444W} < 28$. When divided between quiescent and star-forming, according to the NUVrJ criterion (Ilbert et al., 2013), galaxy photo-zs achieve 0.008 and 0.013 precision, and 2.26% and 1.95% outliers, respectively (see Table 3 in Shuntov et al., 2025a).

We based our analysis on the COSMOS-Web group catalogue, presented in Chapter 4 extending up to $z = 3.7$. We also include an analysis of our first COSMOS catalogue, AMICO-COSMOS (Chapter 3), to explore the relation with group X-ray emission. This catalogue boasts the availability of associated X-ray properties, like luminosity, flux, and mass, for over 600 candidate groups and clusters.

5.2 Galaxy classification with machine learning

Since the AMICO galaxy group catalogues are based on input photometric galaxy catalogues in the COSMOS field, our goal is to classify COSMOS galaxies. Several methods have been developed over the past years to select quiescent and star-forming galaxies (see e.g. Pearson et al., 2023, for an overview), spanning from a classic sharp cut in sSFR (Salim et al., 2018), to colour selections (e.g. Whitaker et al., 2011; Ilbert et al., 2013) and the use of spectroscopic features (e.g. Baldwin et al., 1981; Gallazzi et al., 2014). In this work, in order to classify galaxies into quiescent (red) and star-forming (blue), we used a new approach that incorporates the information from several "classical" methods to train a machine learning (ML) model. We chose to test two algorithms, one non-linear and one linear: the extremeGradientBoosting classifier (XGB; Chen & Guestrin, 2016) and the LinearDiscriminantAnalysis classifier (LDA; Fisher, 1936), both as implemented in scikit-learn Python package (Pedregosa et al., 2011).

5.2.1 Machine learning algorithms

The XGB algorithm belongs to the class of gradient boosted trees (Friedman, 2000). It classifies data by building decision trees sequentially, where each new tree corrects the errors of the previous ones using gradient descent optimisation. This algorithm is well-suited for classification tasks on structured data, offering high predictive accuracy, built-in handling of missing values, and regularisation to mitigate overfitting. It also includes parameters to effectively handle imbalanced class distributions. Compared to linear classification methods, it can model complex, non-linear relationships more effectively, despite being more computationally expensive and based on more complex modelling.

The LDA algorithm classifies data using linear combinations of features, assuming that targets in classes are normally distributed. This makes it a simple and efficient classification method, especially for small datasets for which these assumptions hold. However, it is less flexible than non-linear methods like XGB, less efficient with complex patterns, and does not include a default handling of imbalanced classes or missing values. For this application, we are dealing with highly imbalanced data, since photometric galaxy catalogues are expected to contain many more examples of blue star-forming galaxies rather than red elliptical galaxies. In the classification problem, these two are therefore referred to as the majority and minority classes, respectively. Class imbalance can be mitigated through the generation of synthetic data. One of the most used methods for data generation is the Synthetic Minority Over-sampling Technique (SMOTE; Chawla et al., 2011), which creates additional examples for the minority class via feature interpolation between existing data. This generally helps improve decision boundaries and reduces bias towards the majority class. We tested the classification with LDA both with and without the generation of synthetic data. The results of this test are shown in Table 5.1 and discussed in Sect. 5.2.4. The LDA method also requires the data to be scaled and normalised before training, which was done with the `StandardScaler` method from `scikit-learn`.

5.2.2 Training and testing set

We chose to train and test the ML models on a galaxy sample based on the COSMOS2015 galaxy catalogue (Laigle et al., 2016). Choosing this sample ensures a realistic representation of the features and properties of the galaxies that we want to classify. We chose not to train the algorithm on a subsample of COSMOS2025 since we aim to explore the full sample without compromising statistics and without introducing the risk of repeated data points in the training and target set, which may lead to overfitting. By choosing COSMOS2015 as the training set, we have no data-point overlap with the target dataset. Additional details about the training/testing set choice can be found in Appendix D.

The first step to create a training/testing set is to clean the catalogue to reduce the possibility

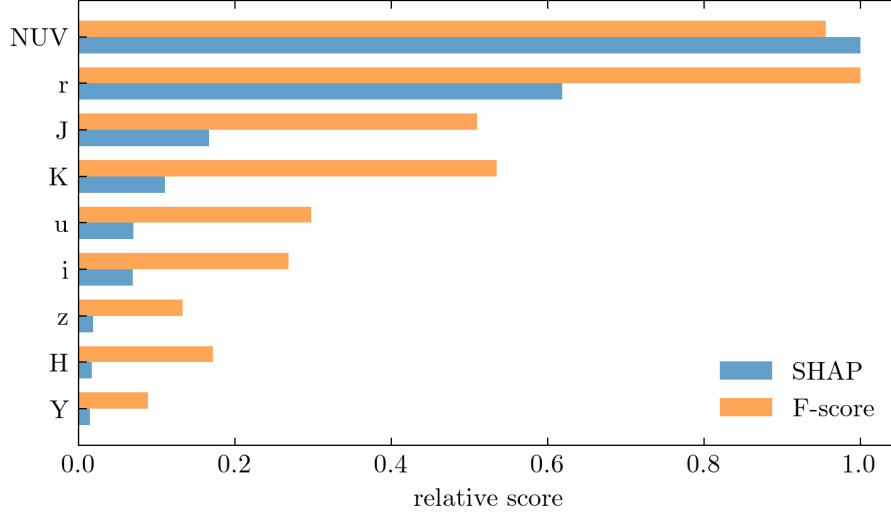


Figure 5.1. Feature importance for an initial XGB training, measured by the F-score (relative split frequency; orange bars), highlights bands directly used to define ground-truth labels are the most frequent in decision splits. To better assess feature impact, we evaluate SHAP scores (Lundberg & Lee, 2017), which quantify the relative feature contribution to predictions (blue bars). Both approaches confirm that rest-frame magnitudes in r , NUV , K and J provide the strongest constraining power.

of training the models on misclassified or spurious objects. For this reason, we performed a conservative cleaning applying the following cuts:

- we removed stars ($\text{type}=1$) and unclassified objects due to SED fitting failure;
- we selected only data with high quality photometry according to the SExtractor flagging system in optical broad bands and UltraVISTA bands: $[\text{BAND}]_{\text{FLAGS}} < 4$;
- we removed masked objects, by applying the AMICO-COSMOS visibility mask and by filtering for FLAG_PETER in $(0, 4, 6)$ (for details on this selection, see Chapter 3);
- we removed galaxies located within 0.05 degrees from pairs of bright star halo (or halo-border pair), resulting in a more conservative mask where the field is more fragmented;
- finally, we cut the dataset at $H < 25$ (corresponding to the mode of the magnitude distribution) to discard faint sources with large magnitude errors and homogenise completeness.

This cleaning yielded a sample of approximately 300,000 galaxies in our training and testing set, across the range $0 < z < 6$. In order to establish the "ground truth" in the classification of galaxies for the training/testing set, we used the classification label yielded by several classical methods. We proceeded as follows. First, we selected four classical methods widely used in the literature to discriminate between the two galaxy classes:

- NUVrJ (Ilbert et al., 2013): a rest-frame colour-colour cut at $NUV - r = 3(r - J) + 1$ and $NUV - r = 3.1$;
- sSFR threshold (Salim et al., 2018): a fixed cut in sSFR at $10^{-10.5} \text{ yr}^{-1}$, as adopted for instance by Euclid Collaboration: Humphrey et al. (2023) and Bisigello et al. (2020);
- Sa-colour evolution (Andreon, 2006; Radovich et al., 2020): a cut given by the evolution in z of the observed colour of a modeled Sa-type galaxy. Colours bracket the $4000 \text{ \AA}$ break;
- NUVrK (Davidzon et al., 2016): a rest-frame colour-colour cut at $NUV - r = 1.37(r - K) + 2.6$ and $NUV - r = 3.15$ and $r - K = 1.3$.

We classified galaxies into three categories according to these four methods: the galaxy is flagged as quiescent if it satisfies most of the four criteria (7% of the sample). The same procedure is followed to classify star-forming galaxies (92% of the sample). If the classification is ambiguous (e.g. 2 vs 2), the galaxy is stored in the third class. This third class allows us to identify galaxies with the following characteristics: those with uncertain classifications, because of missing UV photometry and/or ambiguous SED fitting, as well as those in a physical transition phase between the red sequence and the blue cloud, a regime commonly referred to as the green valley. This category of objects represents $\sim 1\%$ of the full sample. Given the small statistics, it could not be introduced as a third class in the learning process and the classification. Therefore, we decided to remove these objects from the training and testing sets. This reduces ambiguity in the classification and results in an improvement of the performances we describe in Section 5.2.4 of about 2 – 5% in all evaluation metrics for the minority class. The final selected sample is then randomly divided into training (2/3) and testing set (1/3) using the `scikit-learn train_test_split` function. Additional information about the classical method thresholds, their properties, and the distribution of labeled galaxies in redshift and mass can be found in Appendix D.

5.2.3 Feature engineering

In machine learning, *features* are the measurable properties available in the data that we select to make predictions on the objects we want to target. Feature engineering is the ML pre-processing step of selecting, evaluating, and adjusting features, including the handling of missing data, so that the ML model can better capture patterns in the data to make predictions. First of all, we selected all rest-frame magnitudes available in both COSMOS2015 and COSMOS2020. These are the magnitudes in the following bands: NUV , u , r , i , z , Y , J , H and K . To get an estimate of the impact of features on the decision-making process, we performed an initial training with the XGB algorithm, which easily allows the evaluation of feature importance. The results are shown in Fig. 5.1. The importance is measured in terms of the F-score (or weight; orange bars), which is the number of times a feature appears in the decision tree splitting. The r , NUV , J , and K

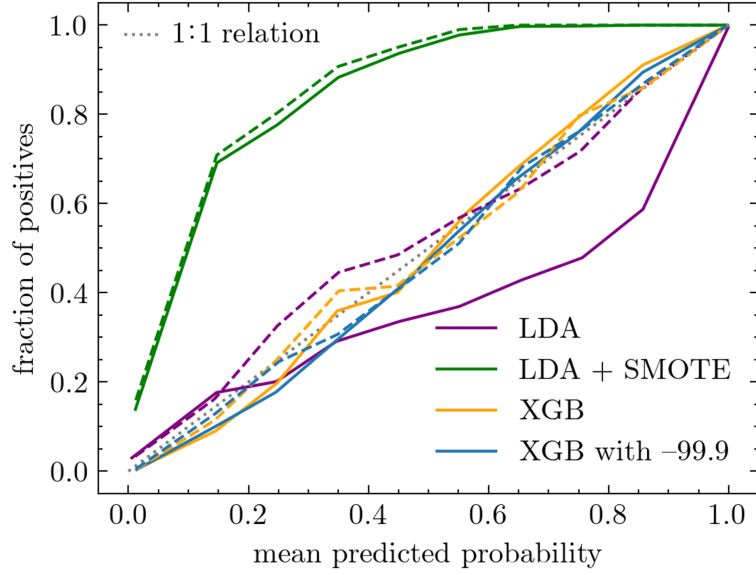


Figure 5.2. Calibration curves showing the bias between predicted probability and fraction of actual positives labels (star-forming galaxies) using all rest-frame magnitude bands (solid lines) and the four most impacting in the decision (dashed line). Different colours represent different methods and configurations, as in the legend, and the dotted grey line represents the ideal 1-to-1 relation.

bands are the most frequent features used in the decision-making process. This result is expected, as these bands are directly used in assigning the ground-truth labels, as described in Section 5.2.2. To gain deeper insight into feature impact, we also evaluated the SHAP score (Lundberg & Lee, 2017), which indicates how much each feature contributes to the model’s predictions (blue bars). The rest-frame magnitudes in r , NUV , J , and K are confirmed as the features with the greatest constraining power in the model predictions. These 4 magnitude bands are also the only rest-frame magnitudes available in the current release of the COSMOS2025 galaxy catalogue at the moment of writing. Since the selection of features is not trivial, and often a few features with high importance simplify the model and lead to better performance than a large number of features, we chose to test and compare results using all magnitudes and with only the best-ranked 4. In Sect. 5.2.4, we describe the differences in performance depending on the set of chosen features in more detail.

The XGB algorithm is sparsity-aware, meaning it can naturally handle missing data. This can be particularly useful in regions with reduced photometric coverage, as it allows us to include galaxies that are missing some rest-frame magnitudes rather than discarding them. Missing values are imputed with an arbitrary constant (-99.9), chosen to be clearly distinguishable from real values. Because XGB does not create artefacts from such artificial flags, this strategy is effective; indeed, Euclid Collaboration: Humphrey et al. (2023) showed that using a flagging number, like

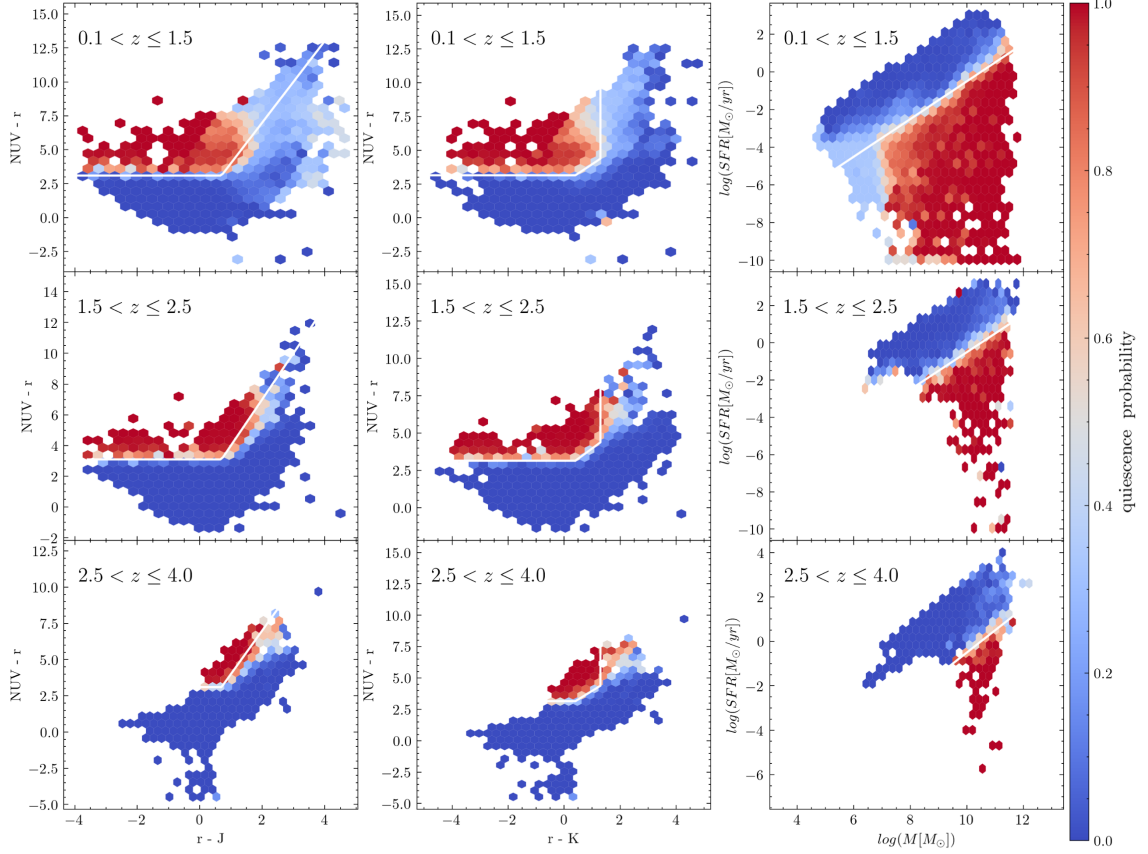


Figure 5.3. Results of our probabilistic classification of COSMOS-Web galaxies with XGB with imputation, as in the NUVrJ (left column), NUVrK (middle column), and SFR- M (right column) planes. Each hexagon is colour-coded by the mean quiescent probability (scale bar on the right), with rows corresponding to different redshift bins. As expected from our training, the method consistently reproduces classical cuts (white lines), while also providing additional insight near the classification boundaries. Classification is ambiguous for some low-mass galaxies at low redshift (light blue region in the top panels).

−99.9, outperforms other imputation methods such as mean, median, or minimum. This approach enables the algorithm to learn from both fully and partially characterised galaxies. We tested the XGB algorithm with and without imputation of a constant value to recover missing data.

5.2.4 Classification performance and results

To evaluate the algorithms and configurations, we rely on the standard metric used to evaluate classification performance against the testing set: precision, recall, and F1-score. Precision (P) is the fraction of correctly identified objects among all objects assigned to a class, analogous to purity in astronomical samples. Recall (R) is the fraction of true objects correctly identified, analogous to completeness. The F1-score combines both precision and recall via their harmonic mean (i.e. $2PR/P + R$). These three quantities range between 0 and 1.

Table 5.1. Performance of the classification methods in terms of precision, recall, and F1-score. The upper block refers to the training with all available rest-frame magnitude bands, the bottom block to the one using only the four with the highest decision impact. Testing set sample size and red/blue ratio: 94236 galaxies with 7.1% red/blue ratio; 98659 with 7.2% red/blue ratio, when missing values are imputed.

Method	Precision (red/blue)	Recall (red/blue)	F1-score (red/blue)
NUV,u, r, i, z, Y,J,H,K			
LDA	0.94/0.99	0.84/1.00	0.89/0.99
LDA + SMOTE	0.69/1.00	1.00/0.97	0.81/0.98
XGB	0.96/1.00	0.95/1.00	0.95/1.00
XGB with −99.9	0.96/1.00	0.93/1.00	0.94/1.00
NUV,r, J,K			
LDA	0.90/0.99	0.90/0.99	0.90/0.99
LDA + SMOTE	0.66/1.00	1.00/0.96	0.80/0.98
XGB	0.96/1.00	0.95/1.00	0.95/1.00
XGB with −99.9	0.95/0.99	0.91/1.00	0.93/1.00

The performances evaluated against the testing set are summarised in Table 5.1, in terms of the metrics just introduced, both for the minority (quiescent galaxies, red) and for the majority class (star-forming galaxies, blue). The different methods yield comparable results for the classification of star-forming galaxies, whereas for the identification of quiescent galaxies, the XGB algorithm tends to perform better overall. Interestingly, applying SMOTE to help LDA with class imbalance increases recall for the minority class but significantly reduces precision. This may be due to an overlap between the two class distributions in the feature space, happening when classes share some similar properties and boundaries are not clearly separable, which is expected to cause

overconfidence toward the minority class when synthetic data are injected (e.g. Chawla et al., 2011; Lu et al., 2024).

However, in this application, we are concerned not only with the quality of the binary classification of our target sample, but particularly with the reliability of the probability assigned to each object of belonging to a given class, since these probabilities are used to estimate quiescent fractions and identify the red sequence. Therefore, in addition to recall, precision, and F1-score, we analysed how well the probabilities are estimated. Figure 5.2 shows the so-called reliability curve (also known as calibration curve), which plots the fraction of actual positives (in this case, "true" star-forming galaxies) against the mean predicted probability for that class, for all four methods, in bins of predicted probability. This diagnostic illustrates how closely the predicted probabilities match the true outcomes. A reliable, unbiased model produces a curve that follows the 1-to-1 diagonal in the plot: for example, if the mean predicted probability in a given bin is 0.7, then roughly 70% of the objects in that bin should belong to that class.

One case in which good classification performance does not imply good calibration is the model trained using data augmentation to balance the classes. Addressing the imbalance between star-forming and quiescent galaxies can lead the model to underestimate the probabilities of the majority class (positives), producing a reliability curve well above the 1- to-1 line (see the green curves in Fig. 5.2). Another behaviour revealed by the calibration curves is the sensitivity of LDA to the number of features selected. In simple linear models, a small number of strong features can yield better results than a more complex model based on more galaxy properties.

By implementing the imputation of the missing values instead of complete discard, we observed a slight decrease in the performance for minority class recall, but a more reliable probability estimate. For this application, given the importance of an unbiased probability and the stability of the algorithm with respect to the number of features used, we decided to opt for an XGB-based classification with the imputation of a constant value for missing data. In Figure 5.2, this configuration is represented by the blue dashed line, which is the most stable and closest to the 1-to-1 relation.

The results of the application of our chosen model for the probabilistic classification, to the COSMOS2025 dataset used for the COSMOS-Web group search, are shown in Fig. 5.3: each hexagon is colour-coded according to the mean probability to be quiescent (bar on the right) inside that hexagonal bin; each column shows a different diagram used in classical methods, while rows refer to different redshift intervals. As expected from the labelling of our training set, our method is consistently mapping classical empirical cuts (displayed as white lines), while providing additional insight, particularly in regions near the classification boundaries. The algorithm struggles in the classification of galaxies with masses below $10^7 M_\odot$ in the first redshift bin, populating the light blue region in the top panels. These galaxies mostly have $z < 0.2$, and are expected to have a small impact on our analysis, as we are particularly interested in evolution at earlier cosmic times.

In a recent parallel work, [Asadi et al. \(2025\)](#) classified Farmer COSMOS2020 galaxies ([Weaver et al., 2022](#)), using a gradient boosted tree algorithm trained on a mock galaxy catalogue based on semi-analytical models and using 28 colours. Their approach differs from ours in several key aspects: it relies on simulated rather than observed data for training, uses apparent colours instead of rest-frame magnitudes, and labels the training set solely through an sSFR threshold. Since their trained model is publicly available, we applied it to the new COSMOS-Web data. The outcomes align with their findings for COSMOS2020: specifically, their model returns a higher fraction of quiescent galaxies at all redshifts compared to classical thresholds such as NUVrJ. Because our method is based on the consensus among multiple classical thresholds (see Sect. 5.2.2 and Appendix D), the model of [Asadi et al. \(2025\)](#) also predicts a consistently higher quiescent fraction relative to the approach presented in this Chapter.

5.3 Quiescent galaxy fractions

We studied the fraction of quiescent (red) galaxies in groups and its dependencies on group richness and redshift. The quiescent fractions are computed with two different methods: pure membership (with testing of secondary association subtraction) and cylinder background subtraction.

Following common practice to reduce biases due to colour uncertainties at the faint end and fully sample typical ranges occupied by RS galaxies, we introduced a redshift-dependent luminosity limit to define the quiescent fraction. This limit is often defined at 2 or 3 magnitudes fainter than an evolving characteristic magnitude or the magnitude of a reference galaxy, such as the brightest one (e.g., [Stott et al., 2009](#); [Cerulo et al., 2016](#); [Radovich et al., 2020](#)). The depth of COSMOS-Web allows us to push this limit deeper than in previous studies and to adopt a threshold at $m_*(z_j) + 4$, where z_j is the redshift of the group. In other words, for each group, we analyzed all galaxies up to 4 magnitudes fainter than the magnitude of a typical elliptical galaxy at the knee of the group luminosity function, $m_*(z_j)$, in our reference band (NIRCam F150W), which was estimated by evolving a massive passively evolving elliptical galaxy (see Chapter 4 for details) with evolutionary synthesis models ([Kotulla et al., 2009](#)). This limit allows for a deep sampling of sub- m_* galaxies (galaxies fainter than the luminosity function knee), which include the strongest signatures of RS build-up (e.g., [Stott et al., 2007, 2009](#)). It should be noted that the $m_* + 4$ limit is not a sharp cut for the galaxy sample, because it depends on the redshift of the group the galaxies are associated with, so galaxies are limited in general by the survey’s completeness limit (see [Shuntov et al., 2025a](#); [Toni et al., 2025a](#)). However, at $z \sim 2.5$, the evolving $m_*(z_j)$ reaches the sample magnitude limit. Beyond this redshift, the $m_* + 4$ cut is no longer significant, because the faint-end is determined by the survey’s limit. This implies that some of the faintest galaxies in our last two redshift bins – in particular, low-luminosity star-forming galaxies – might be undetected because beyond the survey limit, which could lead to a slight

overestimation of the quiescent fraction in those bins with respect to lower redshifts. However, as discussed later, this effect appears not to be significant in our sample, indicating that this selection yields a consistent population across cosmic time while fully leveraging the depth and statistical power of COSMOS-Web. In this section, we briefly introduce the different approaches for computing the quiescent fraction and present the results for this group sample.

5.3.1 Pure membership and secondary association subtraction

Our sample includes the probability for each galaxy to belong to a given group, as provided by the AMICO algorithm. Moreover, we have the probability for each galaxy to be quiescent as we estimated in Section 5.2. The simplest way to compute quiescent fractions is by summing the probabilities to be quiescent, conditional on the probability of belonging to a group, and dividing it by the total, regardless of the class attributed in Sect. 5.2. Given that $P_{i,j}$ is the AMICO probability of the i -th galaxy of belonging to the j -th group (see [Maturi et al., 2019](#), for details) and that $P_{red,i}$ and $P_{blue,i}$ are the (complementary) probabilities of the galaxy to be quiescent or star-forming, respectively, the quiescent fraction of the j -th group, $f_{q,j}$, can be estimated as:

$$f_{q,j} = \frac{\sum P_{red,i} P_{i,j}}{\sum P_{red,i} P_{i,j} + \sum P_{blue,i} P_{i,j}} = \frac{\sum P_{red,i} P_{i,j}}{\sum P_{i,j}}. \quad (5.1)$$

This computation relies on AMICO membership probabilities, which are model-dependent, computed during the detection process and based on the chosen model of that specific group search, that is built on an NFW profile ([Navarro et al., 1997](#)) and a Schechter luminosity function ([Schechter, 1976](#)), for typical low-mass clusters or large groups ([Toni et al., 2024, 2025a](#)). The AMICO association probability assignment also takes into account the possibility for a galaxy to be associated with multiple group detections. Thus, we introduced a small correction in the computation, which consists of considering only the main (highest probability) association for each galaxy and subtracting from it the probabilities to be associated with other detections (secondary associations). The probability $P_{i,j}$ in Eq. 5.1 is therefore substituted by

$$P_{i,j-new} = P_{i,j-main} - \sum_{N_{assoc}-1} P_{i,j-secondary} \quad (5.2)$$

However, this correction resulted to be minimal, and the results with and without correction, statistically compatible.

The results are reported in Fig. 5.4, where we show the weighted quiescent fraction, that is, the quiescent fraction expressed by Eq. 5.1, weighted by the purity of the sample in the corresponding redshift and richness bin, as estimated against realistic mocks produced with SinFoniA ([Maturi et al., 2019](#)) and presented in Chapter 4. Figure 5.4 shows that the population of quiescent galaxies seems to start assembling consistently shortly before $z = 2$, with earlier and accelerated build-up

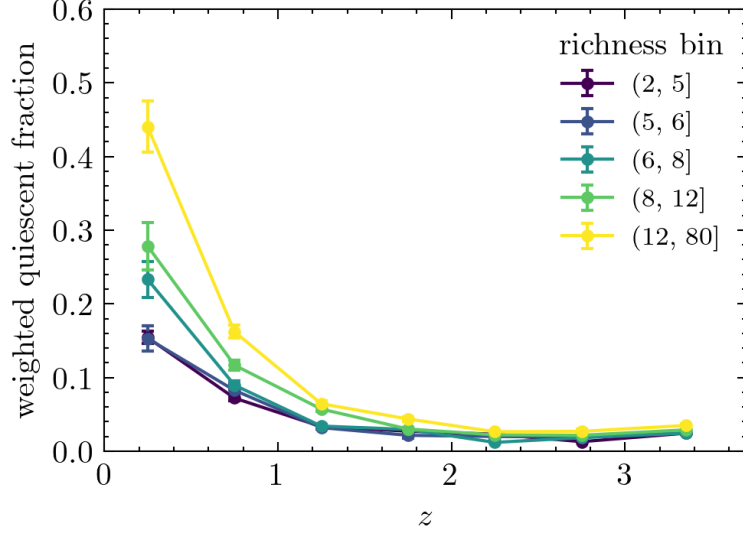


Figure 5.4. Purity-weighted quiescent fraction vs. redshift in different richness (λ_*) bins, as in the legend.

for the richest systems. The slight increase in the quiescent fraction observed at $z \gtrsim 3$ may be influenced by the survey’s magnitude limit, as previously discussed. However, this rise is minimal and also consistent with the expected contribution from a few rare, early RS galaxies that may already be emerging at these high redshifts (see Sect. 5.4.1). For this and the subsequent analyses, we chose richness bin sizes to homogenise the number counts whenever possible, and we estimated uncertainties using bootstrap resampling with 1000 realisations.

5.3.2 Cylinder background subtraction

The AMICO algorithm produces model-dependent membership probabilities for the group member galaxies. In some cases, when quantities used to model the distribution of galaxies in the template are studied, this might lead to biased considerations due to the underlying model chosen for the detection. This is the case for radial profiles and luminosity functions (see e.g. Puddu et al., 2021). In Toni et al. (2025a), we proposed a new method based on a cylinder-based background subtraction technique to retrieve model-independent results for individual and stacked analyses. For the study of the quiescent fraction, we tested the cylinder method, performing background subtraction and estimating the quiescent fraction in a model-independent way. The implementation proceeds as follows:

1. The first step is the definition of cylindrical volumes around each group. We chose to make the cylinder extend $\pm 0.01(1+z_j)$ from the group centre in the redshift direction, where z_j is the group redshift. On the sky-plane, we chose a variable radius given by the root mean square (RMS) of the group-centric galaxy distribution. The choice of the cylinder

parameters is described in detail in Appendix D. We call the volumes inside the cylinders “group regions.”

2. Then, we must consider that not all volume is filled with galaxies, as areas can be masked due to the presence of bright stars and/or image artefacts and patterns. We estimated effective areas on the entire field (i.e. subtracting masked areas), using the same mask adopted in the detection process in each redshift slice (using the AMICO detection redshift resolution, $\Delta z = 0.01$). We accounted for masked areas both inside and outside the cylinders, namely for “group regions” and “field regions,” respectively.
3. We estimated quiescent and star-forming galaxy counts as the sum of the probability to be quiescent (P_{red}) or star-forming (P_{blue}), weighted by the value of the galaxy redshift probability distribution at the group redshift. Thus, the red and blue counts are expressed by

$$N_{\text{red}} = \sum_i P_{\text{red},i} p_i(z_j), \text{ and } N_{\text{blue}} = \sum_i P_{\text{blue},i} p_i(z_j), \quad (5.3)$$

where the $p_i(z_j)$ is the redshift probability of the i -th galaxy at the redshift of the host group detection (z_j);

4. We did this computation for galaxies in both group and field regions and normalised counts by the volumes ($V_{(1)}$ and $V_{(0)}$, respectively), in the redshift slices the cylinder crosses;
5. The final galaxy density within groups is the density in the group region minus the density in the field region at that redshift. Given that the subscripts (1) and (0) refer to galaxies and volumes in the group and field regions, respectively, the final density of quiescent (red) galaxies (equivalent for the star-forming, or blue) in the j -th group is given by

$$\rho_{\text{red},j} = \sum_{(1)} P_{\text{red},i} p_i(z_j) V_{(1)}^{-1} - \sum_{(0)} P_{\text{red},i} p_i(z_j) V_{(0)}^{-1}; \quad (5.4)$$

6. Finally, the quiescent galaxy fraction is defined as the background-subtracted density of quiescent galaxies (Eq. 5.4) over the total density. Galaxies and respective volumes located in the cylinder redshift slices, outside the cylinder volume, and along the line of sight of it, are rejected and not included in the calculation to avoid contamination. Thus, the final expression for the cylinder quiescent galaxy fraction, $f'_{\text{q},j}$, is given by

$$f'_{\text{q},j} = \frac{\rho_{\text{red},j}}{\rho_{\text{red},j} + \rho_{\text{blue},j}} \quad (5.5)$$

$$= \frac{\rho_{\text{red},j}}{\sum_{(1)} p_i(z_j) V_{(1)}^{-1} - \sum_{(0)} p_i(z_j) V_{(0)}^{-1}}. \quad (5.6)$$

Cylinder-based quiescent fractions were computed only for groups with a fraction of masked area not larger than 40% to reject extremely masked detections. We also removed groups in the extreme redshift bins to avoid border effects when considering the cylinders. Additional discussion about this computation and the method we used to deal with possible negative densities can be found in Appendix D. To retrieve these results, as well as those using the membership probability, we weighted the quiescent fraction obtained as just described by the value of the purity in the corresponding bins of redshift and richness. Results for the cylinder-extracted quiescent fractions are shown in Fig. 5.5 for four different bins of richness. For comparison, we report in grayscale the results for the pure-membership quiescent fractions in the corresponding richness bins (Fig. 5.4). Apart from having larger uncertainties due also to lower statistics and to the scatter given by the background, the trend with redshift of the cylinder-based fractions is in line with a rapid build-up of the quiescent population starting at $z \sim 2$, and $z \sim 2.5$ for the richest groups, as seen when using pure memberships. However, when present, the quiescent population seems to be more dominant in rich groups when estimated through the cylinder-based method, which may be interpreted as a consequence of an effective field-galaxy contamination removal or an effect of background fluctuation, since the cylinder method retrieves fractions in comparison to the field population at that redshift. The discrepancy is less prominent in the low-richness bins. In conclusion, the model-dependent pure-membership method and the nearly model-independent¹ cylinder method, despite their different assumptions and computation procedures, produce strikingly consistent trends over nearly 12 Gyr of cosmic time.

5.3.3 X-ray luminosity of groups and environmental effects

In recent years, growing attention has been dedicated to the observed differences and expected biases between X-ray- and optically selected samples of galaxy clusters and groups. It is now increasingly recognised that X-ray-bright and X-ray-faint (or X-ray-undetected) systems often display distinct physical properties, indicating that different selection methods can preferentially identify different underlying populations. These discrepancies trace physical signatures of processes such as AGN feedback and halo assembly history, commonly referred to as halo assembly bias, and their impact becomes especially significant when transitioning to the galaxy group regime (e.g., [Lovisari et al., 2021](#); [Eckert et al., 2021](#); [Popesso et al., 2024](#); [Marini et al., 2025](#)).

In this Section, we consider the AMICO-COSMOS cluster sample ([Toni et al., 2024](#)) comprising 622 candidate groups with X-ray flux above the significance level. Among these, 222 detections have a robust estimate of X-ray properties that were used to estimate group mass and to calibrate

¹It should be noted that, although AMICO does not explicitly use colour selections like RS-based detection algorithms, biases may still arise from the intrinsic properties of the data. In particular, quiescent galaxies tend to be massive and systematically brighter than star-forming ones. Even after removing model (and therefore magnitude) dependence via the cylinder method, the analysis may remain affected by the generally higher photo- z accuracy of quiescent galaxies compared to star-forming ones.

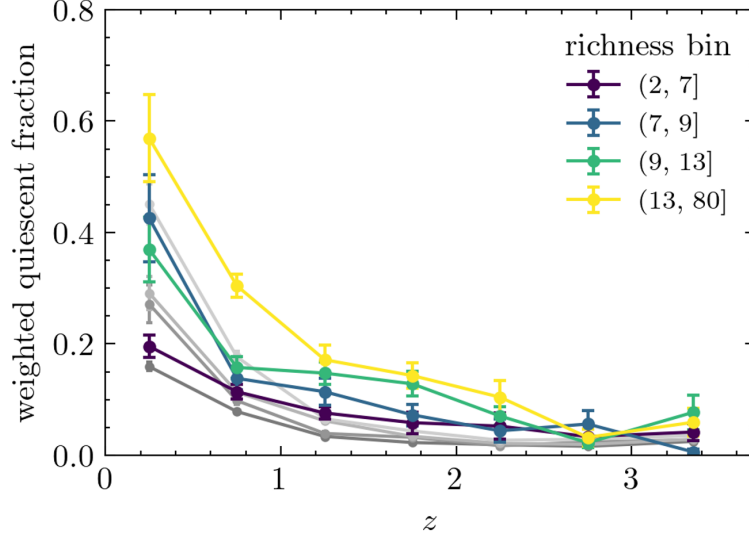


Figure 5.5. Purity-weighted quiescent fraction as estimated with the cylinder background subtraction, as a function of redshift for groups with $\lambda_* > 10$. Different colours show different richness bins, as in the legend. The greyscale trend lines refer to the fractions obtained with the pure membership method in corresponding richness bins (see Fig. 5.4).

mass-proxy scaling relation in [Toni et al. \(2024\)](#). In this sample, around half of the detections have no significant X-ray emission, and this is fairly independent of S/N or redshift cuts. We leverage the availability of X-ray flux and X-ray flux significance (i.e. σ_X) in this sample to study the quiescent fraction and its evolution for the population of X-ray bright groups and those for which a significant X-ray emission was not detected. We calculated the quiescent fraction with the pure membership method, in order not to limit the statistics of the sample, dividing it into the population of robust X-ray emitters ($\sigma_X > 1$ and good quality flag from [Toni et al. 2024](#); 164 groups) and the population with X-ray flux under the significance limit ($\sigma_X \leq 1$; 359 groups). The difference between quiescent fraction for the two populations is shown in Fig. 5.6 as a function of redshift (top panel) and AMICO signal-to-noise ratio, S/N_{ocl} (bottom panel) for two different richness bins (blue and orange lines). The result is that groups with clear X-ray emission tend to have a quiescent fraction consistently higher than or at most equal to that of groups without significant X-ray emission. In particular, rich groups detected with high S/N and groups at $z \lesssim 0.5$ seem to have significantly more dominant population of quiescent galaxies when they are bright in the X-rays.

In a recent study, [Popesso et al. \(2024\)](#) have compared the GAMA *vs* eROSITA samples analysing the properties of X-ray-undetected clusters and groups. Among the considered properties, they have studied the position of the clusters with respect to the large-scale structure (LSS), finding that $\sim 90\%$ of low-emission clusters and groups, often undetected in the X-rays, are located in filaments (and sheets) rather than in the nodes of the cosmic web. Interestingly, this might be

correlating with the properties of the hosted galaxy population, like the pre-accretion processing and therefore the star-formation activity. Groups residing in nodes are more likely to accrete galaxies pre-processed in filaments and, therefore, are more quenched on average. On the other hand, groups hosted by filaments accrete galaxies from the field that are therefore less processed. To check whether this is the case for our sample as well, we compared the position of our groups with that of galaxies that have been found to be located in filaments or in clusters based on local density in LSS maps. For this, we used data from the COSMOS LSS study by [Darvish et al. \(2014\)](#), which covers the redshift interval $0.1 \leq z \leq 1.4$ and the more recent data based on the COSMOS2020 galaxy catalogue, which extends from $z = 0.5$ to beyond the higher end of our group sample (Taamoli et al, in prep.).

The results of this test are shown in Fig. 5.7 for the COSMOS2020-based LSS; results for the LSS map by [Darvish et al. \(2014\)](#) show very similar trends. In Fig. 5.7, the horizontal axis represents the filament signal, defined as the ratio of the number of filament galaxies in a cylinder centred on the group to the number of filament + cluster galaxies, while the vertical axis indicates the probability density distribution smoothed and normalised using Kernel Density Estimation (KDE). To compute the filament signal, we counted galaxies in the central $0.25 \text{ Mpc}/h$ of each group and defined the filament signal as the ratio of the number of galaxies identified as part of a filament to the total (number of filament galaxies + number of cluster/node galaxies). In this way, groups with a filament signal close to 0 are likely hosted by a node (they are rich in cluster/node galaxies), while groups with filament signals close to 1 sit further away from nodes, likely found in filamentary environments connecting nodes. Groups without significant X-ray emission (pink curve) have around 2 times higher density in filaments (filament signal > 0.8) with respect to X-ray bright ones (green curve). The latter are instead more likely to be found in denser environments in correspondence or close to cosmic-web nodes (filament signal < 0.2).

Interesting insights can be gained from the investigation of group properties through a comparison between optically selected and X-ray selected clusters and groups. A comprehensive analysis of this aspect, for what concerns the AMICO-COSMOS groups and that incorporates the X-ray characterisation of the new COSMOS-Web group catalogue, will be presented in a forthcoming dedicated work.

5.4 Red sequence

We now turn our attention to how the red sequence in our group sample evolves over a wide range of redshifts, from $z = 0$ to $z = 3.7$. To study this, we need to identify the red sequence and characterise it through the main parameters of the RS ridgeline.

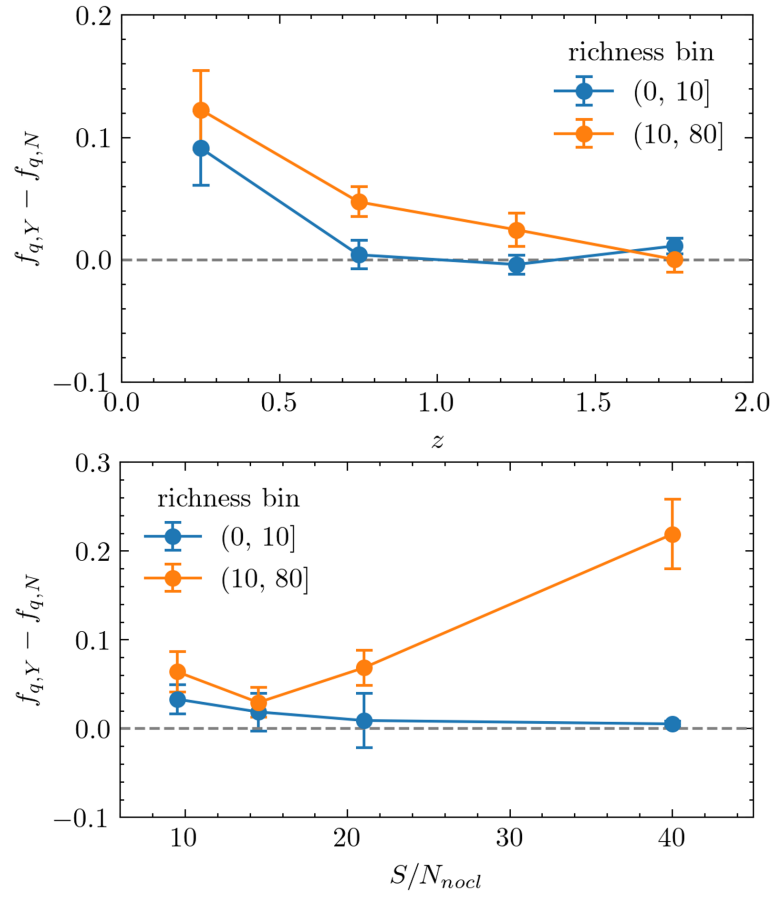


Figure 5.6. Quiescent fraction difference between X-ray bright (Y) and X-ray faint (N) groups as a function of redshift (top panel) and AMICO S/N_{nocl} (bottom panel), in two bins of richness as in the legend.

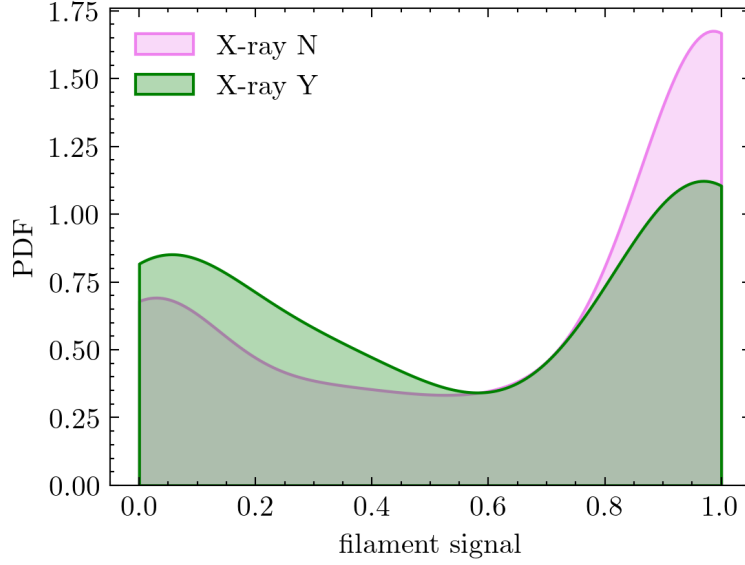


Figure 5.7. Filament signal (ratio of number of filament galaxies to number of filament and cluster galaxies) probability density function for X-ray bright (Y) and X-ray faint (N) groups, using the LSS data and galaxy classification by Taamoli et al. in prep. Consistent results are obtained using the LSS information provided by Darvish et al. (2014)

5.4.1 Observed red sequence with rest-frame matching

The first step to robustly detect the observed red sequence is to identify the appropriate combination of colours and magnitude that, at a given redshift, allows a clean identification of the sequence in the CMD. This can be done following the matched rest-frame photometry method (Blakeslee et al., 2006; Stott et al., 2009), namely by choosing two magnitude bands that bracket a characterising spectral feature, like the $4000 \text{ \AA}$ break, and the reddest of the two as the reference magnitude for the CMD. This reduces the impact of the k -correction and the redshift dependence of the colours, straddling a feature that strongly traces different stellar populations in RS and non-RS galaxies.

The COSMOS-Web galaxy catalogue features a comprehensive set of magnitude bands spanning from optical to mid-infrared wavelengths. This coverage is uniform and continuous thanks to the complementarity of the JWST NIRCcam filters and the UltraVISTA filters, making this the perfect opportunity to study observed red sequence evolution over a wide interval of redshifts, using the matched rest-frame photometry technique. This can be seen in Fig. 5.8, which shows the observed wavelength of the rest-frame $4000 \text{ \AA}$ break (black solid line) at different redshifts over the studied interval, and which filters are sampling and bracketing it (vertical dashed lines indicate when the break exists and enters adjacent bands). As clearly visible, the filter coverage of this spectral feature is continuous until at least $z \sim 3.5$, namely virtually covering the full COSMOS-Web group sample. Our selection of the $4000 \text{ \AA}$ -bracketing band pairs and the relative

intervals are shown in Table 5.2. Once the red sequence is sampled in consistent observed colours, the ridgeline can be identified as the best-fitting line representing the locus of the quiescent member galaxies in the CMD.

Table 5.2. Red-sequence selection method: observed colour and magnitude choices by redshift interval.

Redshift Range	Colour	Magnitude
$0.01 < z < 0.38$	$g - r$	r
$0.38 < z < 0.70$	$r - i$	i
$0.70 < z < 1.10$	$i - z$	z
$1.10 < z < 1.40$	$z - Y$	Y
$1.40 < z < 1.70$	$Y - J$	J
$1.70 < z < 2.35$	$F115W - F150W$	$F150W$
$2.35 < z < 3.25$	$F150W - H$	H
$3.25 < z < 3.70$	$H - Ks$	Ks

As in Sect. 5.3, we performed this RS analysis over 4 magnitudes from the m_* in the F150W band, which, given the depth of the galaxy catalogue, guarantees redshift-dependent completeness until at least $z \sim 2.5$, without compromising the statistics. We computed the RS parameters by performing a weighted least squares (WLS) fitting analysis of the group CMDs on all galaxies attributed to a group with a probability larger than the minimum threshold imposed by AMICO, that is, 0.005. In a similar way to what we have done for the red fraction in Section 5.3.1, we weighted each galaxy with the probability of being quiescent, conditional on the probability of being a group member. This represents the first guess for the position of the red sequence, which is nevertheless still potentially contaminated by central and/or bright star-forming galaxies that might draw the ridgeline guess towards bluer colours. To refine the position guess, we performed an iterative cleaning of outliers by fitting with a 3σ clipping. Once convergence is reached (over maximum 5 iterations), the following RS best-fit parameters are stored for each group detection: the scatter computed on final retained galaxies, the average colour, which is the colour of the RS ridgeline computed at m_* and the number of RS galaxies, which we defined as the number of quiescent member galaxies ($P_{red} > 50\%$ and $P > 50\%$) lying within a typical ± 0.3 scatter (e.g. De Lucia et al., 2007; Martinet et al., 2015) from the best-fit line.

Besides the direct count of RS galaxies, we also computed λ_{RS} , which is the sum of the membership probability of RS galaxies, in analogy to AMICO apparent richness, λ . Figure 5.9 shows the fraction of λ_{RS} to the λ (sum of all membership probabilities of galaxies used for the initial guess) as a function of redshift, for the sample of groups with an RS identified with $\lambda_{RS} > 2$. This is an arbitrary limit that roughly corresponds to having at least 3 quiescent galaxies on the RS ridgeline. The yellow line and points indicate the average trend of the RS fraction with redshift. The size of the red points is proportional to the S/N of the group detection, which is used as a

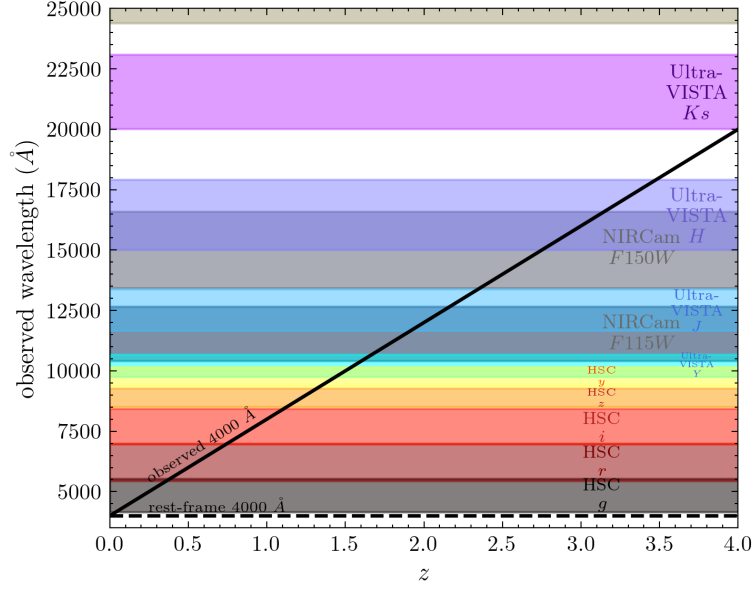


Figure 5.8. The sampling in the COSMOS-Web filters of the observed 4000 Å, the spectral feature used for RS detection (black solid line). Different colours indicate different bands, as in the legend, while vertical lines are the exit/enter point of the break in adjacent bands. Pairs of bands bracketing the break are reported in Table 5.2. The complementarity of COSMOS-Web dataset filters allows for a contiguous coverage up to $z \sim 3.5$.

weight for the average. The grey trend line is reported from the quiescent fraction analysis for the central richness bin (see Fig. 5.4), while the blue trend line shows the same as done with the apparent magnitude but using rest-frame J and K_s , which we describe in Sect. 5.4.2. Quiescent fractions and RS fractions differ in terms of how they are defined and computed. The former is based on the probability of a galaxy being a member and being quiescent, while the latter also leverages the colour homogeneity of the quiescent population (extracted by fitting a tight relation in the CMD). We find consistent trends between the two independent computations.

In the bottom plot of Fig. 5.9, we show the fraction of groups with an identified RS with $\lambda_{RS} > 2$ over the total number of detections per redshift bin, for S/N -cut at 6, 7, and 10 (i.e. purity level of 77%, 80%, and 90%). The overall trend follows that of the quiescent fraction and RS fraction, with the build-up of the RS settling between $z = 2$ and $z = 2.5$. The peak at $z \sim 0.7 - 0.8$ may be due to the presence of the COSMOS Wall (Iovino et al., 2016), which hosts more mature systems with established RSs. The total number of groups with an identified RS in this way is 214.

The highest-redshift RS detected in this study is located at $z = 3.4$ and it contains 3 quiescent member galaxies with very similar colours within around 250 kpc/ h from each other. This high- z object is among the new protocluster cores not known in the literature before, which we reported in Chapter 4, detected independently in a parallel work presented by Hung et al. (2025). The protocluster core is rich ($\lambda_* \sim 25$, $\lambda \sim 119$) and contains 5 galaxies with spectroscopic redshift,

according to our spectroscopic counterpart assignment in the COSMOS spec- z compilation (Khostovan et al., 2026, including GT). The spectroscopic redshifts are compatible with that assigned by AMICO, solely based on photometry, which confirms the presence of an overdensity of galaxies at this location. However, none of the quiescent galaxies of the RS we found have a public spec- z available. If confirmed, this would represent one of the highest-redshift examples of early red sequence detected to date (see the $z \sim 4$ protocluster reported by Tanaka et al. 2024, and the overdensity discovered at $z = 3.44$ by Jin et al. 2024). The CMD of this $z = 3.4$ detection (ID CW117) is shown in Fig. 5.10, where we plot AMICO member galaxies with size proportional to their membership probability and their colour referring to the probability of being quiescent and field galaxies as grey points. In the plot, we also report the main parameters of the RS ridgeline and the RS richness (λ_{RS}), while the RS galaxies are marked by black circles around the points. The orange dashed line indicates the expected colour of a population passively evolving since $z = 5$, which is our reference model. This represents the colour of a typical m_* -galaxy, according to the same configuration chosen for the AMICO model used for the group search. The model was computed with the evolutionary synthesis model generator GALEV (Kotulla et al., 2009), relying on a Kroupa initial mass function (Kroupa, 2002) and adopting a chemically consistent approach (see Kotulla et al., 2009, for further details) for a massive elliptical formed at $z_f = 8$ with a more recent medium-intensity (0.2) star-formation burst at $z = 5$. This approach is similar to the one already used by Castignani et al. (2022a, 2023). In the bottom left of the panel, we report the sky distribution of the galaxies in the central region of this high- z overdensity, with the same notation as in the CMD plot.

5.4.2 Results and comparison with rest-frame red sequence

We performed RS fitting with the same procedure described above (3σ -clipped fit on WLS initial guess) based on the LePhare rest-frame magnitudes in the UltraVISTA- J and - K_s bands available in the COSMOS2025 release. We confirmed all RS detected with the observed magnitudes, mostly with compatible parameters. We additionally detected RS (mainly with low richness, $\lambda_{RS} < 3$) for another 11 groups.

The RS parameters as a function of redshift for the RS ridgelines detected in this study are shown in Fig. 5.11. When computing the fitted relations and mean trends, each group candidate is weighted by its S/N to suppress the impact of the smallest and least reliable detections on the overall fit. The size of the points is proportional to the S/N of the groups. The top panel shows the average observed colour for all the detections in the COSMOS-Web group catalogue, compared to the reference model we described above in Sect 5.4.1 (orange lines). Colours are computed in the associated redshift-dependent bands, as summarised in Table 5.2, which explains the jumps in average colour when the $4000 \text{ \AA}$ switches between adjacent bands. Besides this effect, the trend of the model is consistent with the evolution of the observed colours in our sample

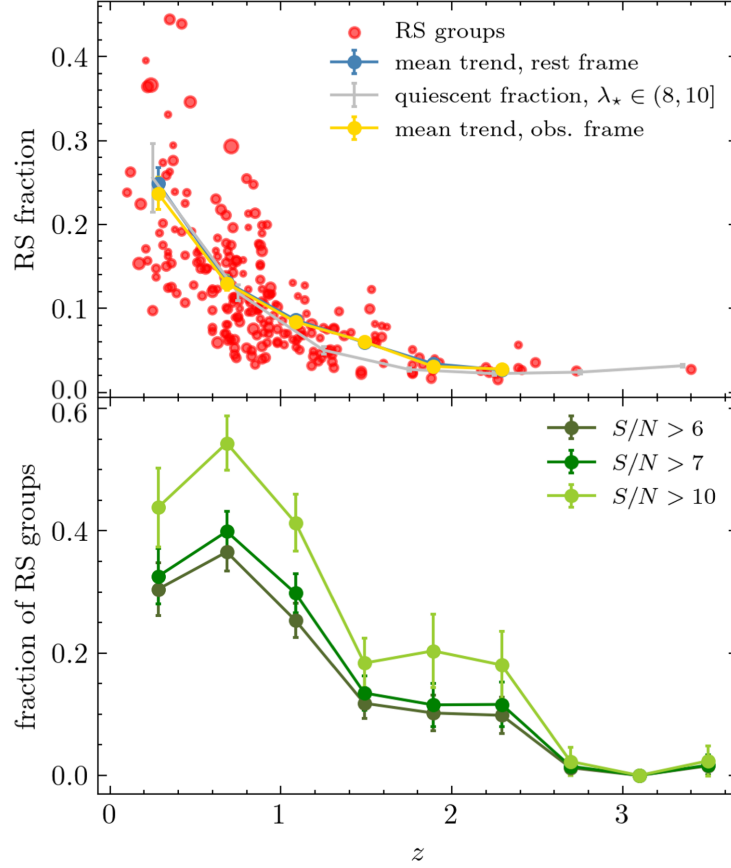


Figure 5.9. Top panel: fraction of the selected RS galaxies over the total (λ_{RS}/λ) in each group, as a function of redshift for the sample with identified red sequence. The yellow and blue lines indicate the mean fraction in each z -bin weighted by the group S/N (size of red points is proportional to the group S/N). Bottom panel: the fraction of groups with identified red sequence at different S/N (or purity) levels, as a function of redshift.

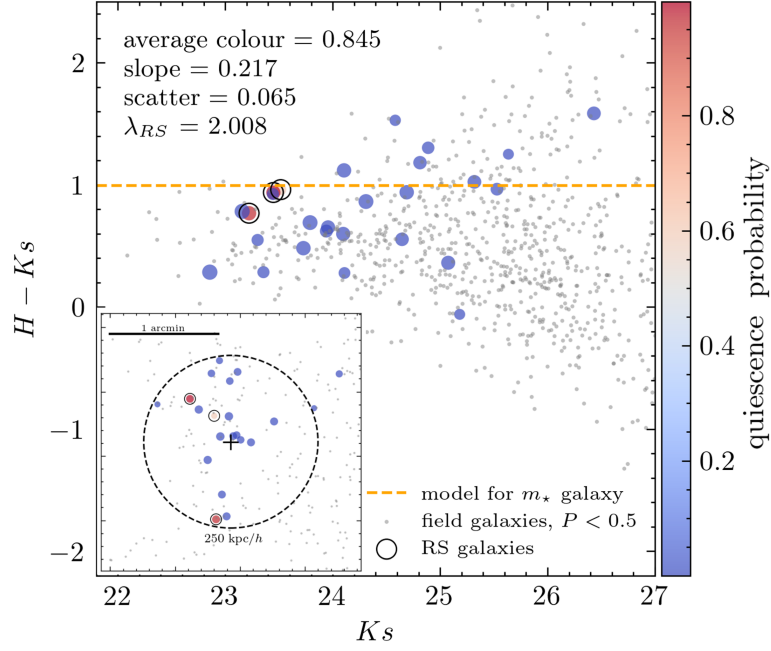


Figure 5.10. Observed-frame CMD of an overdensity of galaxies at $z = 3.4$ (CW117) with 3 quiescent galaxies (red points with black contours) with colours consistent with RS evolutionary synthesis models for a typical m_* elliptical (orange line). This is the highest- z RS we detected.

until at least $z \sim 2.4$. For the last two redshift intervals, the observed colour of the ridgeline is not consistent with a prototypical quiescent galaxy, which is expected since at this redshift we do not fit any red sequence for most of the groups, and the ridgeline parameters are tentative and mainly driven by bright star-forming galaxies populating groups and cores at such high redshifts. The location in the plot of the highest- z RS presented above (CW117) is indicated by a yellow star. As already mentioned, the average observed colour of this RS is consistent with the model. The second and third panels show the ridgeline slope. In the second plot, we include all groups in the sample (grey points) with the relative total fitted trend (black line) and its 1σ error (shaded area). Red points mark groups for which it was possible to identify a red sequence with $\lambda_{RS} > 2$. This latter sub-sample is shown in the third plot from the top, where we compare the slope computed with observed rest-frame matched colours (red points and orange line and shaded area) with that based on rest-frame J and K_s magnitudes (blue points and blue line and shaded area). The two trends are consistent within confidence intervals, with the rest-frame colours offering a stronger constraining power. Under the assumption of negligible k -correction effect, the RS slope trend reflects metallicity evolution (e.g., [Stott et al., 2009](#)). We found a slight increase in the absolute value of the slope with redshift. However, the steepening of the RS over the studied redshift range is very mild, and there is no clear evidence of a significant evolution. The average RS slope evolves of only $\Delta\text{slope} \sim 0.05$ across 12 Gyrs. This weak trend is consistent with little or no

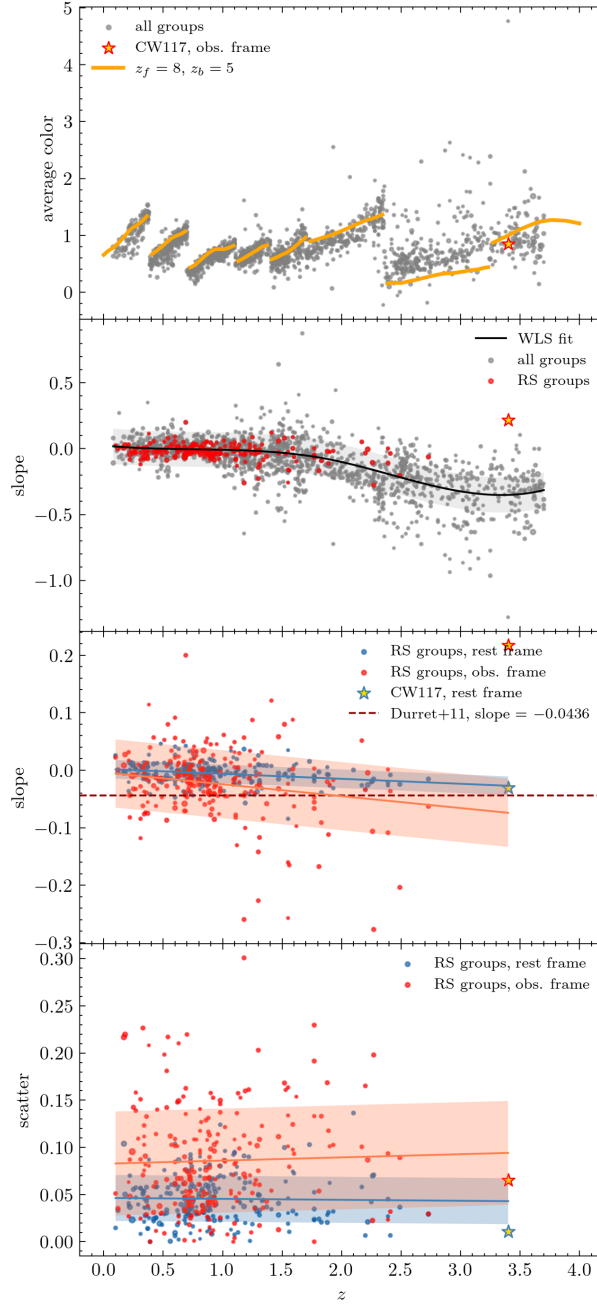


Figure 5.11. Evolution with redshift of the main RS parameters for our group sample. From top to bottom: the first panel shows the average observed colour as in Table 5.2, compared to our reference $m_*(z)$ model (orange lines). The star shows the CW117 detection, which has RS color compatible with that expected at that redshift for typical RS galaxies; the second panel shows the slope, with all groups in gray and those with RS in red; the third panel is a zoom on the slope of RS-groups only, with the parameters as retrieved using observer-frame (red points and orange fit line) and rest-frame magnitudes (blue point and fit line); the fourth panel shows the same but for the scatter of RS-groups. The yellow star shows the RS parameters of CW117, in the observer frame (red contours) and the rest frame (blue contours).

evolution in the mass–metallicity relation of quiescent galaxies, in agreement with, for instance, what was found for clusters up to $z \sim 1.5$ by [Mei et al. \(2009\)](#) and [Cerulo et al. \(2016\)](#). Therefore, our results support this picture, indicating that the mass–metallicity relation was already in place and has remained largely preserved since even earlier cosmic times. A slope for the RS typically adopted in the literature ([Durret et al., 2011](#); [Martinet et al., 2015](#); [Takey et al., 2019](#)) is shown as a dashed dark red line for comparison. Once again, the location of CW117 is indicated by a yellow star, with red contours for observed colours and blue contours for the rest frame. The slope of the RS in CW117 is a clear outlier in the observed colours (red star in the second and third panels of Fig. 5.11), because the quiescent galaxies lie very close together in the CMD (see Fig. 5.10), which drives the fitted slope to become large and positive.

As for the slope, our results for the RS scatter (fourth panel of Fig. 5.11) show little to no evolution over the studied redshift range. The RS scatter is found to be below 0.1 mag for most of the detections ($\sim 64\%$), especially when measured in the rest frame ($\sim 98\%$). The persistence of small scatters might suggest that the bulk of stars in RS galaxies formed early and over relatively short timescales, with limited diversity in their subsequent star formation histories. However, a more quantitative interpretation, such as estimating age dispersion, depends sensitively on photometric uncertainties, colour choice, RS definition, group assembly history, luminosity thresholds, and survey depth; therefore, it is not straightforward to draw firm conclusions about the absolute age spread of the population from the observed RS scatter. However, the average scatter values and trend behaviour are consistent with previous studies based on both simulations and semi-analytical models (e.g., [Menci et al., 2008](#); [Romeo et al., 2015](#)) and on observations (e.g., [Hao et al., 2009](#); [Mei et al., 2009](#); [Cerulo et al., 2016](#)). We found that the scatter little-to-no evolution reported by [Cerulo et al. \(2016\)](#) for clusters up to $z = 1.5$ persists at higher redshifts, with an average rest-frame value of $\approx 0.045 \pm 0.025$ mag, consistent with that found by [Mei et al. \(2009\)](#) for elliptical (0.042 ± 0.021 mag) and for elliptical + lenticular galaxies (0.062 ± 0.015 mag). When considering observed colours, average RS scatter and its spread increase with respect to rest-frame measurements ($\approx 0.084 \pm 0.054$), consistent with the low-redshift observations by [Hao et al. \(2009\)](#) and semi-analytical model predictions ([Menci et al., 2008](#)).

5.5 Conclusions

We conducted a study of the evolution in redshift and the dependence on group richness of the quiescent galaxy fraction and of the red-sequence parameters for COSMOS groups detected with the AMICO algorithm, across around 12 Gyrs of cosmic history, from $z = 0$ to $z = 3.7$. We have used multiple complementary methods to estimate the quiescent fraction and detect the red sequence, finding consistent results across different approaches. In total, we identified 225 red-sequence ridgelines using rest-frame colours, with 214 also consistently characterised using

rest-frame-matched observed colours. The main results of our study are as follows:

- Machine learning classifiers (both linear and non-linear) provide a powerful framework for leveraging photometric multi-band survey data to robustly distinguish between quiescent and star-forming galaxies. In particular, gradient boosting decision trees, such as XGB, offer efficient handling of missing data and class imbalance, while delivering reliable probabilistic classifications. In our application, when tested against a testing dataset, this approach achieved an F1-score exceeding 93%, even in the case of limited availability of rest-frame magnitudes.
- The AMICO algorithm identifies galaxy groups without relying explicitly on colour information. As a result, the AMICO catalogues enable a robust analysis of member galaxy properties, such as colours and luminosities, in both model-dependent and model-independent ways. Our results on the build-up of the quiescent galaxy population are found to be consistent across these methods.
- We found that the quiescent galaxy population in groups builds up steadily starting from $z = 1.5 - 2$ in all richness bins, with evidence of an earlier and accelerated growth for the richest systems, consistent with the scenario proposed, for instance, by [Cerulo et al. \(2016\)](#).
- By analysing X-ray characterised AMICO detection, we found an indication that X-ray faint groups show, on average, lower quiescent fractions than X-ray bright ones. This might be connected to their location in the cosmic web, as X-ray faint groups are more likely to be found in filaments, where the environmental pre-processing of galaxies is lower compared to cosmic nodes, as well as the local density.
- We detected and characterised 225 RS ridgelines in COSMOS-Web groups, providing an extensive redshift coverage. We report the discovery of a rare compact (~ 250 kpc/ h scale) overdensity of quiescent galaxies at $z = 3.4$, which, if confirmed, may represent one of the most distant early red sequences detected to date, not known in the literature before our recent group search in COSMOS-Web (presented in Chapter 4).
- We found that the observed average colours of the red-sequence ridgelines in our sample are consistent with those predicted by a passively evolving elliptical galaxy model. Regarding the slope and scatter of the red sequence, we observed no significant evolution with redshift, suggesting that the absence of significant trends found in previous studies of more massive galaxy clusters and at lower redshift may persist in the group regime and at earlier epochs.

This study extends the detection and characterisation of the red sequence and quiescent galaxy fractions beyond $z > 1.5$ and into the regime of galaxy groups. Such an analysis is a natural application of a deep and well-characterised dataset like the COSMOS-Web group catalogue.

Especially for high- z candidates and RS evolution trends, we stress the necessity of spectroscopic confirmation, which we plan to carry out in a dedicated campaign. In forthcoming work, we will further investigate the correlation between galaxy properties and X-ray emission (i.e. intracluster/-group medium properties) and its connection with the LSS. Additionally, we plan to explore the properties and role of the central brightest group galaxies (BGGs), examining how their evolution relates to the environment and the assembly history of their hosts.

6

Brightest Group Galaxies

ENVIRONMENTAL PROCESSING OF MOLECULAR GAS IN COSMOS BGGs

This Chapter presents the pilot program of our planned CO follow-up campaign on star-forming Brightest Group Galaxies (BGGs). The aim is to observe, characterise, and compare the molecular gas reservoirs and their depletion timescales in central galaxies of the COSMOS group catalogues developed in this work. Our current observations target three star-forming BGGs caught in a rare transitional phase, likely undergoing transformation driven by interactions with close companions. We report CO(1→0) line emission in one system, and two non-detections. By combining molecular gas constraints with fiducial star formation rates derived from total infrared emission, we infer gas depletion timescales indicating that, despite their active star formation and interaction signatures, some BGGs are already experiencing efficient gas exhaustion or suppressed gas replenishment. Our findings hint that environmental processes in galaxy groups can strongly regulate the availability of molecular gas and drive rapid evolutionary phases in central galaxies, bridging the gap between gas-rich cluster centrals and passively evolving systems.

This Chapter is fully based on Toni et al. 2026, to be submitted to A&A ^a

^aObservations have been carried out under project numbers 079-25 (PIs: Toni & Castignani) with the IRAM 30m telescope. IRAM is supported by INSU/CNRS (France), MPG (Germany) and IGN (Spain).

Brightest cluster galaxies (BCGs) and brightest group galaxies (BGGs) are remarkable laboratories for exploring the effect of dense environments on the transformation and evolution of galaxies. The observed large masses and luminosities of present-day BCGs and BGGs are believed to be the aftermath of a complex and still-debated evolutionary history. In galaxy clusters, BCGs

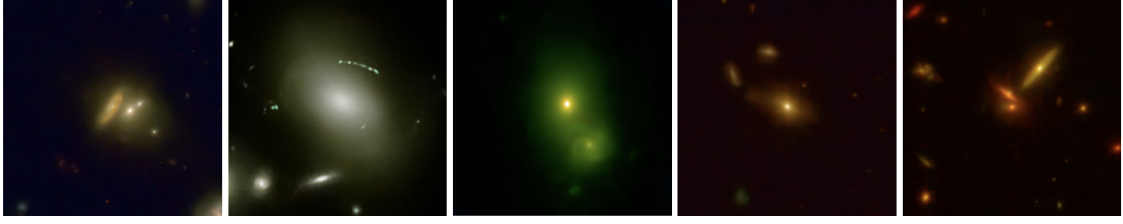


Figure 6.1. JWST NIRCam colour-composite images of five BGGs in COSMOS in the range $z \sim 0.2 - 3.1$. Adapted from [Gozaliasl et al. \(2025, including GT\)](#).

have long been subject to detailed observational and theoretical study, specifically focusing on the assembly of their extreme masses, their quenching histories and their connection with host halo and intracluster light (e.g. [De Lucia & Blaizot, 2007](#); [Lin & Mohr, 2004](#); [Kravtsov et al., 2018](#)). These galaxies serve as important benchmarks, challenging models of hierarchical formation: currently favoured paradigms propose that while the bulk of their stars are formed at early stages ($z > 3$), a consistent assemble of mass takes place via minor dry mergers at lower redshift, significantly increasing the stellar content of the galaxy, although with potentially strong dependence on the host halo and large-scale environment (e.g. [De Lucia & Blaizot, 2007](#); [Stott et al., 2011](#); [Lidman et al., 2012](#); [Einasto et al., 2024](#)).

Observations and cosmological simulations consistently indicate that the assembly histories of BCGs lead to the formation of massive elliptical galaxies, which consequently evolve passively, showing quenched star formation, at least in the local Universe (e.g. [Martizzi et al., 2012](#); [Zhao et al., 2017](#); [Chu et al., 2022](#)). For instance, [Chu et al. \(2022\)](#) analysed a large sample of more than a thousand BCGs, finding that only $\sim 7-9\%$ exhibit indications of residual star formation at $z < 0.7$.

Strongly related to the stellar content and the evolution of BCGs is the rest of the baryonic budget, encompassing a multiphase gaseous reservoir of molecular gas, dust, and ionised components. Numerous studies in the local Universe have revealed that some BCGs host substantial reservoirs of molecular gas ([Edge, 2001](#); [Salomé & Combes, 2003](#); [Hamer et al., 2012](#); [McNamara et al., 2014](#); [Russell et al., 2014](#)), and in several cases extended cold or ionised filaments ([Olivares et al., 2019](#); [Russell et al., 2019](#)). In cool-core clusters, molecular gas masses of the order of $10^{10} - 10^{11} M_{\odot}$ are common (e.g. [Fogarty et al., 2019](#); [Castignani et al., 2025](#)), often co-spatial with ionised and dusty filaments (e.g. [Olivares et al., 2019](#)). Such configurations likely trace intracluster medium (ICM) condensation, which can fuel residual star formation or active galactic nuclei (AGN) activity (e.g. [Fogarty et al., 2019](#); [Castignani et al., 2025](#); [Ubertosi et al., 2025](#)). Although rare at higher redshift, some BCGs were found to exhibit gas-rich phases, such as the Phoenix A BCG, which hosts a $\sim 2 \times 10^{10} M_{\odot}$ molecular reservoir at $z = 0.597$ ([McDonald et al., 2014](#)) with an extreme star-formation rate (SFR) of $\sim 800 M_{\odot} \text{ yr}^{-1}$ ([McDonald et al., 2013](#)), implying rapid assembly of $\sim 3\%$ of its stellar mass within ~ 30 Myr.

While BCGs are overall well-represented in the literature (e.g. McDonald et al., 2016; Webb et al., 2017; Radovich et al., 2020; Zenteno et al., 2025), BGGs, their analogues in galaxy groups, have historically received less attention. Nonetheless, BGGs are of fundamental importance: groups are the most common environment for galaxies, and baryonic physics such as gas cooling, feedback, and mergers can play a strong role in low-velocity-dispersion systems. Recent efforts have indeed emphasised that BGGs may not simply scale down BCG behaviour but can follow distinct evolutionary paths and with different time scales. In the COSMOS field, the campaign to study the properties of BGGs in the X-ray-selected group sample by Gozaliasl et al. (2019) has shown that $\sim 20\%$ of local BGGs are star forming, rising to $\sim 50\%$ at $z \sim 1$. Moreover, it was found that their stellar ages show a spread somewhat younger than semi-analytic expectations (Gozaliasl et al., 2024, including GT), particularly when offset from the X-ray centres, which may indicate recent group merging history (Gozaliasl et al., 2019). Earlier work in the same series (Gozaliasl et al., 2018) has shown that BGGs present a bimodal distribution between quiescent and star-forming systems, with mean SFRs ~ 2 dex above the median SFR of the overall galaxy population at comparable redshift, and that a considerable amount of the stellar mass growth in star-forming BGGs may derive from *in situ* star formation. Indeed, a non-negligible fraction of BGGs at low redshift ($z \lesssim 0.2 - 0.3$) still exhibit signs of active star formation (Gozaliasl et al., 2024, including GT), a result that is particularly noteworthy given the expectation that central galaxies in mature systems should be generally quiescent at recent epochs. The presence of star-forming BGGs in the nearby Universe suggests that residual or accreted cold gas—possibly condensing from the IGM (e.g. Temi et al., 2018; Olivares et al., 2022)—can contribute to sustaining residual star formation or AGN fueling within the group core environment. Recent studies have also highlighted the diversity in central galaxy growth within group haloes, in terms of, for instance, star formation histories, kinematics, and structural properties (e.g. Jung et al., 2022). In parallel, studies have shown that, in group-scale environments, the diversity and properties of BGGs are closely linked to those of their host haloes and even to their Mpc-scale environment (e.g. Raouf et al., 2021; Einasto et al., 2024). Altogether, these findings underscore the importance of group-scale haloes in the investigation of the ongoing interplay between cooling, feedback, and star formation in dense environments. Some examples of BGGs in COSMOS, as observed by JWST, are shown in the panels of Figure 6.1.

In this context, an in-depth investigation of BGG formation and evolution is timely and compelling. This Chapter explores the environmental processing of central galaxies in COSMOS groups at intermediate redshifts ($z \sim 0.3$) by studying the properties of their molecular gas reservoirs. This is achieved thanks to targeted observations carried out with the IRAM 30m telescope (obs. prop. ID 079-25, PIs: Toni & Castignani).

6.1 Motivations and goals

Carbon monoxide (CO) line emission is the primary observable tracer of molecular hydrogen (H_2) in galaxies. The H_2 molecule lacks a permanent dipole moment, making its detection in the cold molecular gas difficult. In contrast, CO is abundant, has a relatively low excitation temperature ($T_{\text{ex}} \sim 5$ K for the CO(1 $\rightarrow$ 0) transition), and its rotational transitions fall conveniently in the millimetre and submillimetre regimes, where they can be more easily observed. Thus, CO emission provides a reliable proxy for tracing the spatial distribution, kinematics, and total mass of the molecular gas reservoir of a galaxy, given an assumed CO-to- H_2 conversion factor (α_{CO}). Detecting CO line emission in BCGs provides a direct probe of the gas that can fuel star formation and AGN activity (e.g. [Edge, 2001](#); [Salomé & Combes, 2003](#)). Conversely, non-detections of CO can indicate effective and rapid quenching, providing stringent tests of theoretical models of galaxy–host coevolution.

In the work described in this Chapter, we aimed to detect CO emission in three star-forming galaxies at $z \sim 0.3$ located in the cores of galaxy groups within the COSMOS field, in order to confirm (or put upper limits on the total mass, in case of non-detections) the presence of substantial molecular gas reservoirs. Building upon the previous CO-detection campaigns targeting intermediate to high- z BCGs with the IRAM 30m telescope located at Pico Veleta, Spain and presented in [Castignani et al. \(2019, 2020b,c, 2022a, 2023\)](#), our new observations extend this analysis to the group regime. This is the pilot program of an observation campaign which will ultimately enable the construction of the first statistically significant sample of group and cluster central galaxies at intermediate redshift with robust CO measurements and therefore characterisation of molecular gas properties.

Our program is designed to constrain the molecular gas content and star formation efficiency of central galaxies in group environments, shedding light on the environmental processing of gas at intermediate redshift ($z \sim 0.3$, corresponding to a lookback time of ~ 3 Gyr). Despite growing efforts, the literature still substantially lacks large samples of distant central ellipticals in dense environments with observations of their molecular gas feeding the star formation (see e.g. [Dunne et al., 2021](#)). The lack of such observations is even more significant when we consider group centrals (e.g. [Olivares et al., 2022](#)), which trace intermediate environments. By comparing the derived molecular hydrogen mass (M_{H_2}), gas-to-stellar mass fraction, depletion time, excitation conditions, and SFRs with those of both local and high-redshift analogous in galaxy clusters and field galaxies (e.g. [Tacconi et al., 2018](#); [Freundlich et al., 2019](#); [Fogarty et al., 2019](#); [Dunne et al., 2021](#)), we aim to characterise the role of group environment, also in comparison to that of clusters, in regulating gas accretion and star-formation fueling or potential rapid exhaustion in massive central galaxies.

6.2 Star-forming brightest group galaxies

We selected our targets among the central galaxies of the clusters and groups detected with AMICO in the COSMOS field described in Chapter 3, the AMICO-COSMOS catalogue. This catalogue covers an effective area of $\sim 1.69 \text{ deg}^2$, allowing us to have good statistics of central galaxies in COSMOS, in the redshift interval we are interested in.

6.2.1 AMICO-COSMOS BGG catalogue

We created the AMICO-COSMOS BGG catalogue by considering a combination of vicinity to the group centre as estimated by AMICO, galaxy luminosity, and membership probability, as it was also done by Radovich et al. (2020) for the analysis of AMICO-KiDS galaxy population properties. However, despite these galaxy properties are related to each other in how the detection and member assignment are performed by AMICO, the membership assignment is sensitive to the shape and peakiness of galaxy photo- z probability distribution (see e.g. Toni et al. 2024; Euclid Collaboration: Bhargava et al., 2025, including GT), which enters Equation 2.23 as a Gaussian distribution with mode and 1σ errors taken from the photometric galaxy catalogue (as expressed by Eq. 2.14). Spectral Energy Distribution (SED) fitting and therefore accurate photo- z estimation for bright extended galaxies can be challenging at intermediate/low redshifts, where extended galaxies' morphology and substructures are resolved (e.g. Soo et al., 2018; Wilson et al., 2020; Weaver et al., 2022). Indeed, we found a few examples of bright extended central galaxies (especially at $z \lesssim 0.5$) which, despite clearly exhibiting the properties of a BCG/BGG, were attributed with low probability to their host system due to their redshift probability distribution being, for instance, very peaked at a value slightly different from the system centre (namely, of the other member galaxies on average). An example can be appreciated in Fig. 3.9: the centre of the foreground system, a cluster located at $z = 0.47$, corresponds to a very bright extended galaxy with consistent photo- z (albeit peaked and slightly off the average) that is, however, not marked by a cyan circle as the other members. The galaxy has a membership probability $P < 50\%$, despite likely being the BCG of this detected system.

This effect impacts only a small number of systems, but it can influence BGG selection¹, particularly within the redshift range relevant to this study. Therefore, we have developed a new method to safely select BGGs without relying solely on membership probability. We proceeded by considering the following criteria:

- vicinity to the system centre (sky distance), Δx : we filtered the galaxy sample by keeping only galaxies with $\Delta x < 0.3R_{200}$, where R_{200} is the AMICO model radius, the same used for

¹we simply refer to the brightest central galaxies as BGG, as most of the galaxy systems in COSMOS are expected to have group-halo properties, as discussed in this Thesis.

the definition of the intrinsic richness, λ_* (Eq. 2.26), and 0.3 is a typical value used to limit the search to the central area of each candidate cluster/group;

- vicinity to the system redshift, Δz : a constraint in redshift is needed, since we do not use any a priori limitation in membership probability. To limit the selection of the BGG to the vicinity of the cluster/group candidate, also in redshift space, we imposed $\Delta z < 0.05(1 + z)$, where z is the system redshift. This choice is in line with the average photometric redshift probability uncertainty of COSMOS2015 and COSMOS2020 galaxies;
- galaxy magnitude, m : among the obtained galaxy sub-sample, we selected the most likely BGG as the galaxy with the lowest apparent magnitude m ;
- membership probability, $P_{i,j}$: if the first and the second brightest have a magnitude difference < 0.5 , we selected the one with the highest membership probability among the two.

For all the detections in the cluster/group catalogue, we stored the first and second choices according to these criteria. We repeated this selection for all three catalogues, resulting from taking each of the three photometric bands as the reference band, using the magnitude in the same photometric band used for detection in each run. The final catalogue of central galaxies is the result of the merging of the three catalogues, keeping any selected candidate, in a way that each detection can have multiple options for its chosen BGG. As we have used the input photometric galaxy catalogue chosen for the AMICO-COSMOS cluster/group search, galaxies here selected and their photometric properties are taken from COSMOS2020 (Weaver et al., 2022) or, in a few cases (see Chapter 3 of this Thesis for details), COSMOS2015 (Laigle et al., 2016). Each BGG in the catalogue is assigned two scores, one related to its association probability and one related to its distance on the sky plane to the centre of the host system. The first score is the ratio of the BGG membership probability to the highest probability assigned by AMICO in that cluster/group. The second score is the distance to the system centre in units of R_{200} , the AMICO model radius. The combination of these two gives an indication of the probability of the galaxy being the BGG of the system. We have expanded and utilised this BGG selection method also for the COSMOS-Web data. The resulting BGG catalogue was already exploited for the investigation of size–mass evolution of central group galaxies up to $z = 3.7$. This work has been presented in Gozaliasl et al. (2025, including GT) to which we have dedicated Appendix B.

6.2.2 Selection of targets

For our specific goal here, to ensure the feasibility of CO-line detection, we cross-matched the central galaxies with high-confidence measurements from public spectroscopic redshift catalogues (the same listed in Chapter 3, Section 3.3.4). In the case of aforementioned scores being low, we used spec-zs to rule out ambiguous cases. The spectroscopic counterpart assignment yielded

~ 800 sources with spec- z consistent with the redshifts of the associated parent cluster or group. Among the selected galaxies with spectroscopic redshifts, we further limited ourselves to star-forming sources, selecting only those with $\text{SFR} > 20 M_{\odot} \text{ yr}^{-1}$ according to COSMOS2020 physical properties derived via SED fitting with LePhare (Arnouts et al., 1999; Ilbert et al., 2006), at intermediate redshifts, namely in the range $0.2 < z < 0.4$. This redshift interval was chosen to target non-local galaxies without compromising the feasibility of the observations with the IRAM 30m telescope.

This selection procedure yielded three central galaxies, all belonging to groups with masses in the range $M_{200} \simeq (1 - 3) \times 10^{13} M_{\odot}$, thus lower than those typically associated with galaxy clusters. As reported in Table 6.1, our targets have intermediate redshifts, $z \sim 0.25 - 0.35^2$, high stellar masses, $M_{\star} \simeq (1 - 7) \times 10^{10} M_{\odot}$, typical of ellipticals at the centre of clusters and groups, and are clearly star-forming (SFR multi-wavelength diagnostics will be discussed below in Sect. 6.2.3). These galaxies may therefore represent the intermediate- z analogues of local star-forming central galaxies such as Perseus A and Cygnus A (e.g. Salomé et al., 2011; Privon et al., 2012; Fraser-McKelvie et al., 2014).

Table 6.1 summarises the main physical properties of the three targets, including the expected SFR for a galaxy of the same mass at the same redshift on the star-forming main sequence (MS) according to Speagle et al. (2014). Star formation rates reported in Table 6.1 are the fiducial values we assume for this study, which are extracted by integrating dust emission after performing SED fitting. Multi-wavelength indicators and SFR estimates are described in Sect. 6.2.3.

Our three selected targets are shown in Figure 6.2, where the upper panels show colour-composite Hyper Suprime Cam (HSC) images (Aihara et al., 2022), while the bottom panels show UltraVISTA Ks (McCracken et al., 2012) and Hubble Space Telescope’s ACS (Koekemoer et al., 2007) single-band images. The images display hints of spiral arms and a clearly disturbed morphology for the three targets, possibly induced by the interaction with the surrounding companions, suggesting strong processing of gas and influence of the environment.

6.2.3 Star formation rates

Our three target BGGs have a probability of being star-forming larger than 99% according to the machine learning based classification method developed by Toni et al. (2025b) and presented in Chapter 5. Rest-frame diagnostic diagrams like NUVrJ (Ilbert et al., 2013) further confirm that the sources are star-forming and clearly separate themselves from the population of quiescent galaxies. In Fig. 6.3, we report the position of our three selected targets in the NUVrJ diagram of the AMICO-COSMOS member galaxies with probabilities from Toni et al. (2025b).

A key component of our analysis is the connection between star-formation activity and

²spectroscopic redshifts of these three targets are all DESI-EDR and DESI-DR1 redshifts (DESI Collaboration et al., 2024, 2025), also reported in the COSMOS Spectroscopic Redshift Compilation (Khostovan et al., 2026, including GT).

Table 6.1. Properties of our selected targets: three star-forming BGGs in COSMOS.

Galaxy name (ID COSMOS2020) (1)	RA [hh:mm:ss.s] (2)	Dec [dd:mm:ss.s] (3)	z_{spec} (4)	M_* [$10^{10} M_\odot$] (5)	SFR [$M_\odot/\text{yr}$] (6)	sSFR [Gyr^{-1}] (7)	sSFR _{MS} [Gyr^{-1}] (8)
BGG 0957+0204 (683413)	09:57:47.570	02:04:15.880	0.25288	$1.2^{+0.1}_{-0.1}$	$6.7^{+1.1}_{-1.1}$	$0.55^{+0.11}_{-0.11}$	0.16
BGG 1000+0207 (745502)	10:00:10.940	02:07:24.760	0.33904	$6.3^{+1.1}_{-1.0}$	$29.8^{+3.6}_{-3.6}$	$0.48^{+0.10}_{-0.10}$	0.11
BGG 0958+0223 (1041562)	09:58:38.830	02:23:48.750	0.35510	$6.9^{+0.8}_{-0.7}$	$45.8^{+5.1}_{-5.1}$	$0.67^{+0.11}_{-0.10}$	0.11

Notes. Column description: (1) BGG identification; (2-3) J2000 equatorial coordinates; (4) spectroscopic redshift²; (5) stellar mass (Weaver et al., 2022); (6) SFR derived from spectral energy distribution modelling of FIR emission and using the Kennicutt (1998) relation, which we take as fiducial value (see Sect. 6.2.3); (7) specific SFR, i.e. $\text{sSFR} = \text{SFR}/M_*$; (8) sSFR for main sequence field galaxies with redshift and stellar mass of our targets estimated using the relation found by Speagle et al. (2014).

properties of the molecular gas reservoirs, for which we need reliable and justified estimates of the SFR. The COSMOS2020 catalogue (Weaver et al., 2022) includes extensive multi-wavelength photometry from the ultraviolet through the near- and mid-infrared. Specifically, it combines deep imaging from GALEX, HSC, UltraVISTA, and Spitzer-IRAC over the wavelength range $\sim 0.15 - 5 \mu\text{m}$ (Weaver et al., 2022). However, this wavelength coverage does not include the full mid-infrared regime nor extend into the far-infrared (FIR; $\gtrsim 20 \mu\text{m}$) range, where the bulk of dust-reprocessed emission associated with star formation emerges, as typically found in massive and dusty galaxies (e.g. Kennicutt & Evans, 2012; Casey et al., 2014). Additionally, COSMOS2020 SED fitting is solely based on the optical and near-infrared part of the spectrum, which often fails to fully and reliably trace star formation in systems with significant dust content (e.g. Casey et al., 2014; Leja et al., 2017; Malek et al., 2018).

Independent tracers of star formation, such as emission lines ($\text{H}\alpha$ and $[\text{OII}]$) and total FIR luminosities, therefore play a crucial role in building a comprehensive view of star formation activity (see e.g. Calzetti, 2013, for a review). The emission lines provide information on recent star formation on ~ 10 Myr timescales, while FIR luminosities trace the reprocessed light from

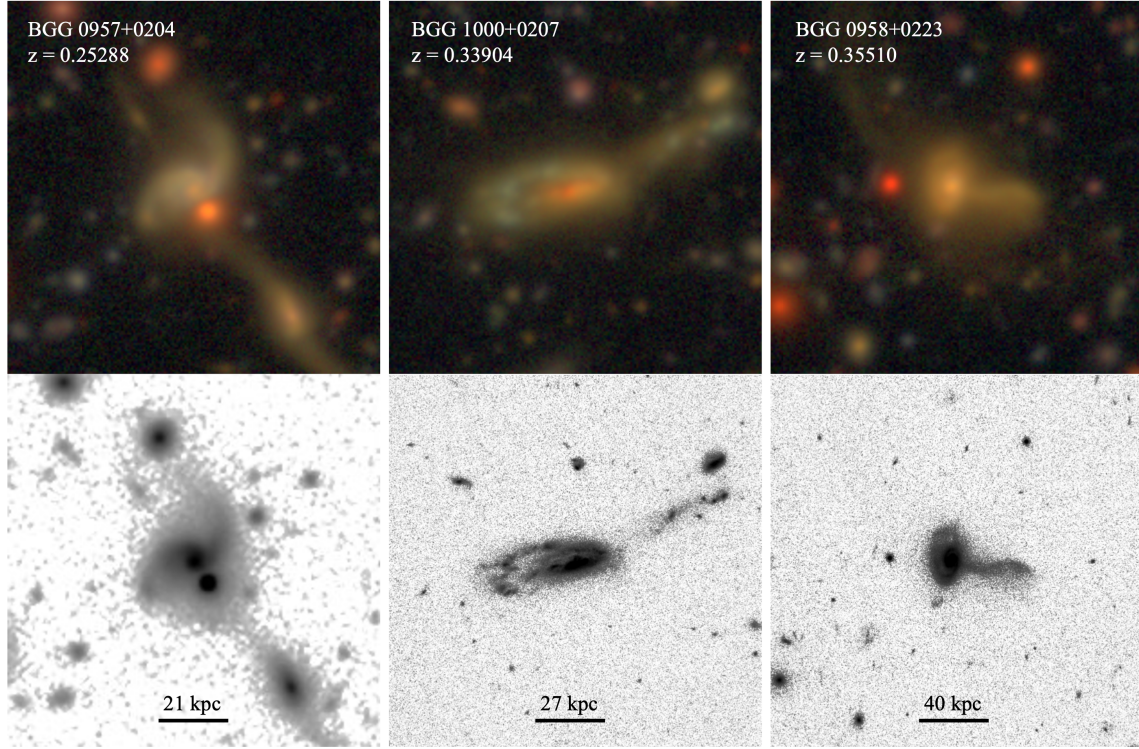


Figure 6.2. From left to right: Target BGG 0957+0204, BGG 1000+0207 and BGG 0958+0223. Top panels: Colour composite $30'' \times 30''$ images of the three targets in HSC g, r, i bands; bottom panels: $30'' \times 30''$ UltraVISTA Ks image (for BGG 0957+0204) and HST-ACS I mosaic (for BGG 1000+0207 and BGG 0958+0223). Physical scales are reported at the bottom of the panels. North is up, east is left.

young stars embedded in dust and longer timescales, of the order of ~ 100 Myr (e.g. [Calzetti et al., 2025](#)). By comparing SFRs obtained from different estimators, we can assess their agreement and interpret them in light of the timescales they probe. Table 6.2 reports the results in terms of SFR and corresponding sSFR for the multi-wavelength indicators we analyse in this Section.

Line diagnostic from DESI spectra. We investigate the star formation properties of the BGGs using DESI DR1 spectroscopy ([DESI Collaboration et al., 2025](#)), following the general approach adopted in previous studies (e.g. [Castignani et al., 2022a, 2023](#)). Star-formation rates are estimated from emission-line diagnostics, exploiting the proportionality between the SFR and the luminosities of the $H\alpha$ recombination line and the [OII] forbidden-line doublet at $3726 \text{ \AA}$ and $3729 \text{ \AA}$ ([Kennicutt, 1998](#)). Throughout this work, all SFRs are calibrated to a [Chabrier \(2003\)](#) initial mass function (IMF).

Emission-line fluxes are retrieved from the DESI DR1 Galaxy and Quasar Value-Added Catalogues (VAC; [Myers et al., 2023](#)), which provide measurements derived from the FASTSPECFIT

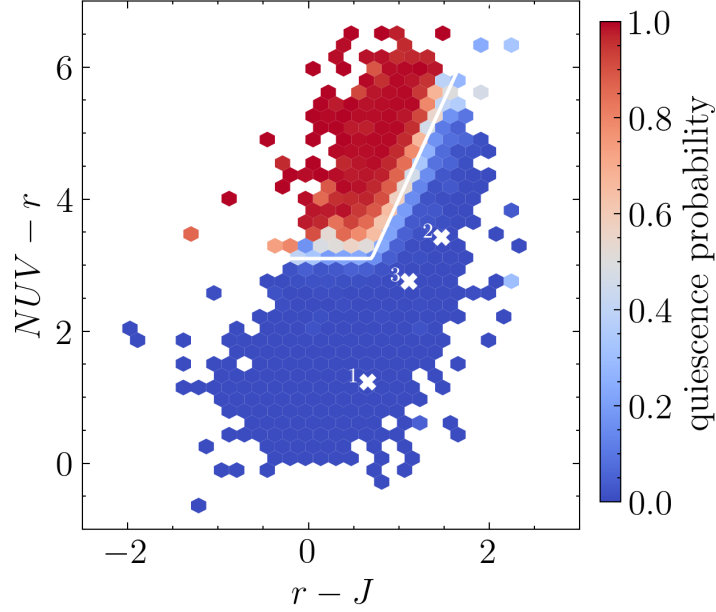


Figure 6.3. Location of our three targets (white crosses: 1. BGG 0957+0204; 2. BGG 1000+0207; 3. BGG 0958+0223) in the NUVrJ diagnostic diagram of the AMICO-COSMOS member galaxies. Colours indicate the mean probability to be quiescent in each hexagonal bin, as in the bar on the right. Rest-frame colours of the target galaxies are clearly separated from the passive population (red hexagons).

spectral fitting pipeline (Moustakas et al., 2023). These fluxes correspond to the narrow emission-line components measured after subtraction of the stellar continuum.

Because DESI spectra are obtained through fibres with a diameter of 1.5 arcsec, aperture corrections are required for extended galaxies. We correct all emission-line fluxes using the aperture correction parameter provided by the FASTSPECFIT Emission-Line Catalog³, which estimates the ratio between the total galaxy flux and the flux captured within the fibre based on the best-fit morphological model from Legacy Survey imaging (Dey et al., 2019; DESI Collaboration et al., 2025). When multiple measurements are available for a given galaxy, we adopt the spectrum with the largest number of coadded exposures to maximise signal-to-noise⁴.

Aperture-corrected fluxes are converted to luminosities using the spectroscopic redshift of each galaxy. Dust attenuation is incorporated into the SFR estimates through an attenuation term $A_{H\alpha}$. For galaxies with high signal-to-noise $H\beta$ measurements, we estimate nebular extinction directly from the Balmer decrement ($H\alpha/H\beta$), adopting a Calzetti attenuation law (Calzetti et al.,

³<https://data.desi.lbl.gov/doc/releases/dr1/vac/fastspecfit/>

⁴This part of the work uses data obtained with the Dark Energy Spectroscopic Instrument (DESI). DESI construction and operations is managed by the Lawrence Berkeley National Laboratory. To this end, we acknowledge the use of services/data provided by the Astro Data Lab at NSF’s NOIRLab. NOIRLab is operated by the Association of Universities for Research in Astronomy (AURA), Inc. under a cooperative agreement with the National Science Foundation.

Table 6.2. SFRs and relative sSFR of our target BGGs.

star formation indicator	BGG 0957+0204	BGG 1000+0207	BGG 0958+0223
FIR	$6.7^{+1.1}_{-1.1}$	$29.8^{+3.6}_{-3.6}$	$45.8^{+5.1}_{-5.1}$
(this work)	$0.55^{+0.11}_{-0.11}$	$0.48^{+0.10}_{-0.10}$	$0.67^{+0.11}_{-0.10}$
FIR	—	$49.6^{+3.2}_{-3.2}$	$30.3^{+2.8}_{-2.8}$
(Kartaltepe+10)	—	$0.79^{+0.14}_{-0.15}$	$0.44^{+0.06}_{-0.07}$
H α	$13.3^{+1.3}_{-1.3}$	$11.8^{+3.6}_{-3.6}$	$61.9^{+9.4}_{-9.4}$
	$1.11^{+0.15}_{-0.17}$	$0.19^{+0.07}_{-0.07}$	$0.90^{+0.16}_{-0.17}$
[OII]	$21.9^{+2.3}_{-2.3}$	$22.0^{+2.1}_{-2.1}$	$61.3^{+11.1}_{-11.1}$
	$1.82^{+0.26}_{-0.30}$	$0.35^{+0.07}_{-0.07}$	$0.89^{+0.18}_{-0.19}$
SED	$54.0^{+10.2}_{-10.2}$	$56.8^{+22.3}_{-13.8}$	$70.3^{+16.8}_{-13.2}$
(COSMOS2020)	$4.49^{+0.95}_{-1.00}$	$0.90^{+0.38}_{-0.27}$	$1.02^{+0.26}_{-0.23}$

Notes. For each star-formation indicator, we specify the SFR in the upper row and the relative sSFR in the bottom row, in units of $M_{\odot}/\text{yr}$ and Gyr^{-1} , respectively. Star formation indicators from top to bottom are: the FIR flux we estimated by integrating the dust component of the SED; the integrated FIR flux but with the values reported by Kartaltepe et al. (2010); the H α emission line from DESI spectra; the [OII] emission line from DESI spectra; the UV-optical SED fitting from Weaver et al. (2022).

2000). For BGG 1000+0207 where the $H\beta$ measurement is of lower significance and the Balmer decrement yields highly uncertain attenuation values, we adopt an alternative approach based on the infrared-excess–stellar-mass (IRX– M_*) relation by [McLure et al. \(2018\)](#). Empirical studies have shown that IRX correlates strongly with stellar mass and exhibits only weak redshift evolution out to at least $z \sim 1$, providing stable population-averaged attenuation estimates for star-forming galaxies (e.g. [Pannella et al., 2015](#); [Álvarez Márquez et al., 2016](#)). The inferred infrared excess is converted into an effective $H\alpha$ attenuation assuming a Calzetti attenuation curve.

With these conventions, the $H\alpha$ - and [OII]-based SFRs can be expressed as (see, e.g. [Gilbank et al., 2010](#); [Zeimann et al., 2013](#); [Old et al., 2020](#)):

$$\frac{\text{SFR}_{H\alpha}}{M_{\odot} \text{ yr}^{-1}} = 5.0 \times 10^{-42} \frac{L_{H\alpha}}{\text{erg s}^{-1}} 10^{0.4 A_{H\alpha}}, \quad (6.1)$$

$$\frac{\text{SFR}_{[\text{O II}]}}{M_{\odot} \text{ yr}^{-1}} = 5.0 \times 10^{-42} \frac{L_{[\text{O II}]}}{\text{erg s}^{-1}} 10^{0.4 A_{H\alpha}} r_{\text{lines}}^{-1}, \quad (6.2)$$

where r_{lines} accounts for the relative attenuation between the [OII] and $H\alpha$ lines under the adopted extinction law.

Following [Gilbank et al. \(2010\)](#), we adopted $r_{\text{lines}} = 0.5$, and when considering the [OII] doublet, we have taken the total flux of the doublet components and summed their uncertainties in quadrature. The $H\alpha$ and [OII] doublet fluxes before aperture correction for BGG 0957+0204, BGG 1000+0207, and BGG 0958+0223 are $F_{H\alpha} = (146.3 \pm 1.6)$, (46.0 ± 1.5) , and (145.0 ± 1.5) and $F_{[\text{O II}]} = (88.0 \pm 2.9)$, (12.3 ± 3.6) , and (41.1 ± 4.1) , respectively, in units of $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$. Thus, the resulting SFRs for BGG 0957+0204, BGG 1000+0207, and BGG 0958+0223 are $\text{SFR}_{H\alpha} = (13.3 \pm 1.3)$, (11.8 ± 3.6) , and (61.9 ± 9.4) and $\text{SFR}_{[\text{O II}]} = (21.9 \pm 2.3)$, (22.0 ± 2.1) , and (61.3 ± 11.1) , respectively, as reported in Table 6.2.

SED fitting and IR luminosity with LePhare. As an independent estimate of the total star-formation activity, we compute SFRs based on the total infrared luminosities. The far-infrared luminosity (L_{IR}), integrating the dust emission over $\sim 8\text{--}1000 \mu\text{m}$, provides a direct tracer of star formation that is largely insensitive to the detailed dust geometry affecting optical diagnostics ([Kennicutt, 1998](#)).

In order to independently estimate SFRs from infrared luminosities, we performed SED fitting for our three target BGGs. In line with previously adopted approaches in the literature (e.g. [Castignani et al., 2019](#); [Poitevineau et al., 2023](#)), we performed galaxy SEDs modelling using LePhare ([Arnouts et al., 1999](#); [Ilbert et al., 2006](#)). The far-infrared photometry was fitted separately in order to capture potential dust-related emission, adopting the set of 105 infrared templates from [Chary & Elbaz \(2001\)](#). Photometric observations at shorter wavelengths were instead modelled with the CE_NEW_MOD template set, which includes a suite of galaxy spectra

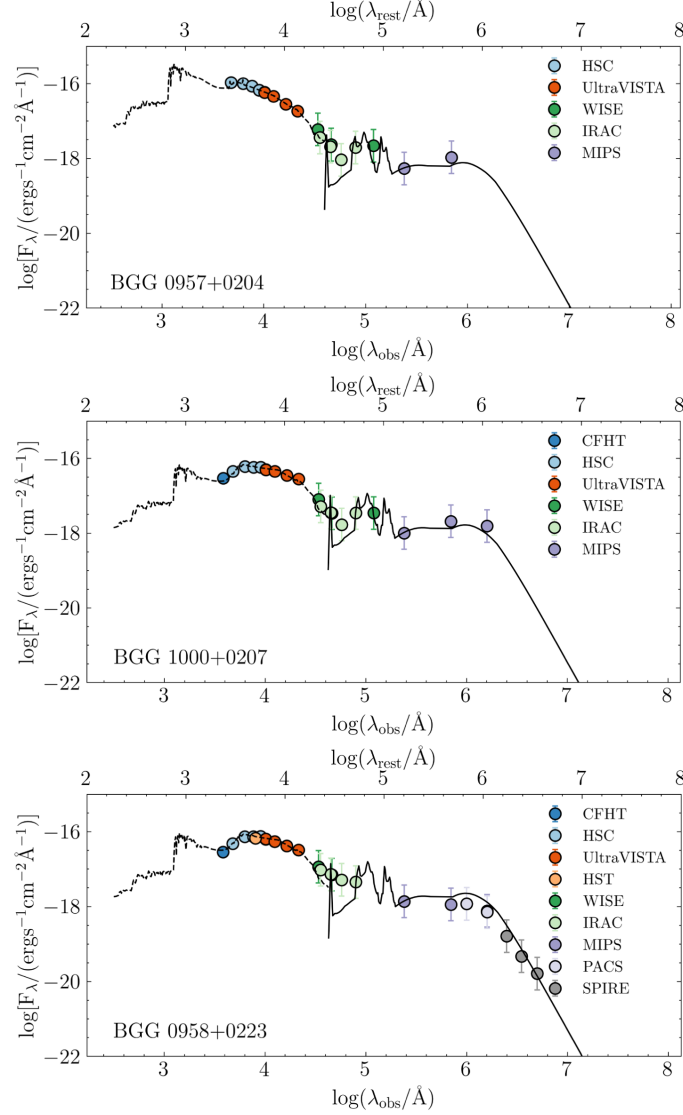


Figure 6.4. SEDs and respective LePhare best fitting modelling for the three target BGGs from u -band to FIR. Different colours refer to different instruments as in the legend. Dashed and solid lines indicate best fits for the stellar and dust components, respectively (see the text for details about instruments, relative bands and adopted templates).

generated through interpolation (Arnouts et al., 1999)⁵.

For the SED fitting, we used photometric archival data from u -band to FIR from the COSMOS2020 galaxy catalogue and from the NASA/IPAC Extragalactic Database (NED). Archival photometry includes optical and near-infrared data from CFHT (Sawicki et al., 2019), HSC (Aihara et al., 2019), HST ACS (Koekemoer et al., 2007) and UltraVISTA (McCracken et al., 2012; Moneti

⁵We also double checked the results of our SED fitting by making use of the GAZPAR tools, operated by CeSAM-LAM and IAP.

et al., 2023), and infrared from Spitzer/IRAC and MIPS (Frayser et al., 2009; Kartaltepe et al., 2010; Capak et al., 2013), WISE (Cutri et al., 2021) and Herschel/PACS and SPIRE (Ma & Yan, 2015). Our results for the best-fitting SED modelling using separate handling of stellar and dust components are shown in Fig. 6.4 for the three target BGGs.

From the resulting fits, we derived the total infrared luminosity (L_{IR}) by integrating the best-fit model of the dust component in the rest-frame 8–1000 μm range. This luminosity was then translated into an SFR using the Kennicutt (1998) relation, adjusted for a Chabrier (2003) initial mass function:

$$\frac{\text{SFR}_{\text{FIR}}}{M_{\odot} \text{ yr}^{-1}} = 1.075 \times 10^{-10} \left(\frac{L_{\text{IR}}}{L_{\odot}} \right). \quad (6.3)$$

Dust emission integration yielded the following IR luminosities: $\log(L_{\text{IR}}/L_{\odot}) = 10.8, 11.4$, and 11.6 for BGG 0957+0204, BGG 1000+0207, and BGG 0958+0223, respectively, translating into $\text{SFR}_{\text{FIR}} = (6.7 \pm 1.1), (29.8 \pm 3.6)$, and $(45.8 \pm 5.1) M_{\odot}/\text{yr}$. We assume these as fiducial values for the rest of the analysis as reported in Table 6.1. Based on their high IR luminosities, the three BGGs can be considered luminous infrared galaxies (LIRGs; see e.g. Sanders & Mirabel, 1996), which are often associated with significant star formation activity (e.g. Pereira-Santaella et al., 2015; Herrero-Illana et al., 2017; Castignani et al., 2020d). Although BGG 0957+0204 has an IR luminosity slightly below the formal LIRG threshold of $L_{\text{IR}} \simeq 10^{11} L_{\odot}$, it lies very close to this limit and is therefore considered comparable to the rest of the sample.

This conversion of IR luminosity into an SFR estimate is valid provided that the FIR luminosity predominantly arises from reprocessed stellar light associated with young massive stars. We assumed this condition holds for our three targets based on the following considerations. We performed WISE colour-based classification (Stern et al., 2012; Jarrett et al., 2017) using AllWISE photometry (Cutri et al., 2021), according to which BGG 0957+0204 is classified as an intermediate disk, BGG 1000+0207 as a starburst galaxy and only BGG 0958+0223 as an AGN, in line with the emission-line-based AGN classification reported in the DESI VAC (Myers et al., 2023). It is worth mentioning that BGG 0958+0223 has WISE colours $W1-W2 = 0.86$ mag (Vega) and $W2-W3 = 2.88$ mag (Vega), which are in between the AGN and intermediate disk classes.

However, for all our targets, the FIR emission is well fitted by the templates of Chary & Elbaz (2001). This suggests that the FIR emission primarily originates from dust heated by star formation, with minimal AGN contamination, which would instead result in a steep-spectrum SED. The absence of a strong power-law component in the SED of BGG 0958+0223 (see Fig. 6.4) suggests the assumption of dominant dust emission holds for this target as well.

Finally, we compared our estimates of the total infrared luminosities with the ones inferred by Kartaltepe et al. (2010). However, only luminosities for BGG 1000+0207 and BGG 0958+0223 are provided in this study. Reported FIR luminosities for BGG 1000+0207 and BGG 0958+0223 are $\log(L_{\text{IR}}/L_{\odot}) = 11.64 \pm 0.03$ and 11.45 ± 0.04 , respectively, yielding $\text{SFR}_{\text{FIR}} = (49.6 \pm 3.2)$ and $(30.3 \pm 2.8) M_{\odot} \text{ yr}^{-1}$, as reported for comparison in Table 6.2. Overall, the resulting SFRs agree

well with those derived from our FIR-based estimates.

Table 6.2 summarises and compares the results from different multi-wavelength SFR indicators we considered in this analysis, including the FIR emission we estimated by integrating the dust component of our best-fit SED modelling, the integrated FIR-based SFR inferred from fluxes reported by Kartaltepe et al. (2010), the H α and [OII] emission line from DESI spectra, and the UV-optical SED fitting from Weaver et al. (2022).

For each BGG, the different SFR estimates are broadly consistent, agreeing within a factor of ~ 3 , which is in line with expectations when comparing different SFR tracers (e.g. Kennicutt & Evans, 2012; Figueira et al., 2022). However, the comparison of star formation rates derived from multiple diagnostics also highlights some variations among the different tracers, possibly reflecting their sensitivity to distinct physical timescales. In general, UV-optical SED fitting without IR constraints tends to yield higher SFRs, while optical emission-line indicators provide lower or intermediate values that may be affected by extinction and non-star-forming excitation mechanisms (see e.g. Kennicutt & Evans, 2012; Conroy, 2013; Old et al., 2020). As they trace the dust emission heated by star formation, infrared-based estimates typically offer a more robust measure of the star formation rates and are therefore adopted as fiducial values in the following analysis. Therefore, we used the estimated values obtained via our SED fitting that we performed with LePhare (see Table 6.2). The diversity of SFR estimates underscores the complex and heterogeneous nature of star formation in these systems. A more detailed interpretation of the individual galaxies, and of the possible physical origin of the discrepancies among the various SFR indicators, is presented in Sect. 6.4.

6.3 IRAM observations and data reduction

We observed the three sources using the Eight Mixer Receiver (EMIR) mounted onto the IRAM 30m telescope at Pico Veleta in Spain. The observations were carried out in July, September, October, and November 2025 (ID: 079-25; PI: G. Toni & G. Castignani). Altogether, each source was observed for about 2.3 to 3.3 hours (double polar).

Each of the E090, E150, and E230 receivers, operating between 1 and 3 mm, offers 4 $\times$ 4 GHz instantaneous bandwidth covered by the correlators. For all three sources we used the E090 and E230 receivers to simultaneously target the CO(1 $\rightarrow$ 0) and CO(3 $\rightarrow$ 2) lines, redshifted to $\sim (3.3 - 3.5)$ mm and $\sim (1.1 - 1.2)$ mm, respectively.

The IRAM 30m half-power beam width (HPBW) is ~ 16 arcsec $\frac{\lambda_{\text{obs}}}{2 \text{ mm}}$ (Kramer et al., 2013), where λ_{obs} is the observer frame wavelength. All sources are thus unresolved by our IRAM 30m observations. Indeed, the typical CO-to-optical size ratio is ~ 0.5 (Young et al., 1995), and in general of the order of unity. The wobbler-switching mode was used for all observations to minimise the impact of atmospheric variability. The Wideband Line Multiple Autocorrelator (WILMA)

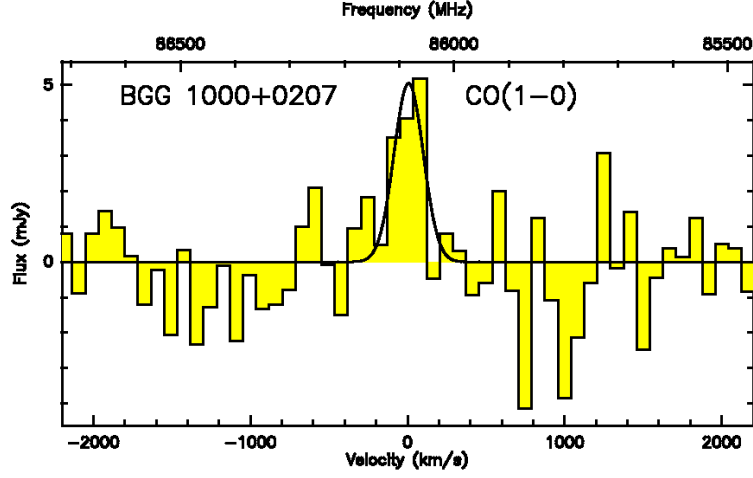


Figure 6.5. Baseline-subtracted CO(1→0) spectrum of BGG 1000+0207. The solid line shows the Gaussian best fit. The flux (y-axis) in units of mJy is plotted against the relative velocity with respect to the BGG redshift.

was used to cover the 4×4 GHz bandwidth in each linear polarisation. We also simultaneously recorded the data with the fast Fourier Transform Spectrometers (FTS) as a backup at 200 kHz resolution.

We encountered overall unfavourable weather conditions in July and September, often characterised by atmospheric instability, which is common, considering that the observations were carried out during the local summer season. However, in late October and late November, we had optimal weather conditions. Altogether, we had system temperatures ranging between $T_{\text{sys}} \simeq 115\text{--}124\text{ K}$ at $\sim 3\text{ mm}$. At $\sim 1\text{ mm}$ system temperature ranged $T_{\text{sys}} \simeq 343\text{--}424\text{ K}$.

The data reduction and analysis were performed using the CLASS software of the GILDAS package⁶. We applied standard efficiency corrections to convert 1) the antenna temperature T_{a}^* into the main beam temperature T_{mb} , and then 2) T_{mb} into the corresponding CO line flux, with a conversion factor of 5 Jy/K. In particular, we adopted the following efficiency corrections: $T_{\text{mb}}/T_{\text{a}}^* = 1.17$ and 1.70, at $\sim 3\text{ mm}$ and $\sim 1\text{ mm}$, respectively.⁷

We detected the CO(1→0) line for BGG 1000+0207. The baseline-subtracted spectrum is shown in Fig. 6.5. Conversely, we were able to set upper limits for the other BGGs analysed in this work. Concerning CO(1→0) lines of BGG 0957+0204 and 0958+0223, we reached a root mean square (rms) of 1.35 and 1.32 mJy, for velocity channels of 97.75 and 105.7 km/s, respectively, after removing the baseline with a linear fit. Similarly, for the CO(3→2) line of BGG 0957+0204, 1000+0207, and 0958+0223, we reached rms values of 6.66, 3.95, and 4.57, with velocity channels of 65.58, 69.65, and 70.49, respectively. In the cases of non-detection, we were then able to set

⁶<https://www.iram.fr/IRAMFR/GILDAS/>

⁷<https://www.iram.es/IRAMES/mainWiki/Iram30mEfficiencies>

3σ upper limits to the velocity integrated flux of $S_{\text{CO}(J \rightarrow J-1)} \Delta v$, where we assumed a velocity resolution and width of 300 km s^{-1} , which is typical of massive galaxies such as those in our samples (e.g. Edge, 2001; Castignani et al., 2020c; Dunne et al., 2021; Castignani et al., 2025).

6.3.1 Molecular gas reservoirs

For our three sources, we inferred measurements and 3σ upper limits to the $\text{CO}(J \rightarrow J-1)$ velocity integrated luminosities, $L'_{\text{CO}(J \rightarrow J-1)}$, in units of $\text{K km s}^{-1} \text{ pc}^2$, and the velocity-integrated $\text{CO}(J \rightarrow J-1)$ fluxes, $S_{\text{CO}(J \rightarrow J-1)} \Delta v$, in units of Jy km s^{-1} . To this aim, we used the following relation from Solomon & Vanden Bout (2005):

$$L'_{\text{CO}(J \rightarrow J-1)} = 3.25 \times 10^7 S_{\text{CO}(J \rightarrow J-1)} \Delta v \nu_{\text{obs}}^{-2} D_L^2 (1+z)^{-3}, \quad (6.4)$$

where ν_{obs} is the observed frequency in GHz of the $\text{CO}(J \rightarrow J-1)$ transition, D_L is the luminosity distance in Mpc, and z is the redshift of the target galaxy.

We then estimated the total H_2 gas mass, M_{H_2} , of the BGGs or its upper limit directly from the velocity-integrated CO line luminosities, given that $M_{H_2} = \alpha_{\text{CO}} L'_{\text{CO}(1 \rightarrow 0)} = \alpha_{\text{CO}} L'_{\text{CO}(J \rightarrow J-1)} / r_{J1}$, where α_{CO} is the CO-to- H_2 conversion factor and $r_{J1} = L'_{\text{CO}(J \rightarrow J-1)} / L'_{\text{CO}(1 \rightarrow 0)}$ denotes the excitation ratio. For the two galaxies without clear CO detection, we used $\text{CO}(1 \rightarrow 0)$ flux to estimate M_{H_2} upper limits. We adopted a Galactic CO-to- H_2 conversion factor $X_{\text{CO}} \approx 2 \times 10^{20} \text{ cm}^{-1} (\text{K km s}^{-1})^{-1}$, namely, $\alpha_{\text{CO}} = 4.36 M_{\odot} (\text{K km s}^{-1} \text{ pc}^2)^{-1}$ (Solomon et al., 1997; Bolatto et al., 2013). The results for the molecular gas properties are reported in Table 6.3. The choice of a typical Galactic conversion factor is in line with that adopted in previous studies and thus enabled us to perform a homogeneous comparison, presented in Sect. 6.4.2 (e.g. Geach et al., 2011; Castignani et al., 2019, 2020c,b; Dunne et al., 2021).

However, it should be noted that our selected galaxies have FIR-based $\text{SFR} \sim 10 - 50 M_{\odot}/\text{yr}$, with specific star-formation rates (sSFR) in the range between $\sim 3.4 - 6.1 \times \text{sSFR}_{\text{MS}}$ (see Table 6.1), hence occupying the regime above the star-forming main sequence (MS), which is typically characterised by sSFR values larger than $3 \times \text{SFR}_{\text{MS}}$. Therefore, a lower conversion factor, typical of starburst galaxies, $\alpha_{\text{CO}} \sim 0.8 M_{\odot} (\text{K km s}^{-1} \text{ pc}^2)^{-1}$, could also be an appropriate choice in our specific case. We discuss the effect of adopting a starburst α_{CO} conversion factor below in Sect. 6.4.1.

6.4 Discussion

Given the star-forming nature of these galaxies, as confirmed by their colours and by multiple independent SFR indicators, and considering the presence of disturbed morphologies indicative of interactions, one would generally expect significant CO emission tracing large molecular gas

Table 6.3. Results of our CO observations.

Galaxy	CO(J→J-1)	ν_{obs} [GHz]	int. time [h]	$S_{\text{CO}} \Delta v$ [Jy km/s]	L'_{CO} [$10^9 \text{ K km pc}^2/\text{s}$]	M_{H_2} [$10^{10} M_{\odot}$]	τ_{dep} [Gyr]	$\frac{M_{\text{H}_2}}{M_{\star}}$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
BGG 0957+0204	1→0 3→2	92.012 276.023	2.28 2.28	< 0.694 < 2.80	< 2.21 < 0.99	< 0.96 (< 0.18)	< 1.43 (< 0.26)	< 0.80 (< 0.15)
BGG 1000+0207	1→0 3→2	86.085 258.242	3.33 3.33	$1.18^{+0.31}_{-0.31}$ < 1.71	$6.90^{+1.81}_{-1.81}$ < 1.11	$3.01^{+0.79}_{-0.79}$ ($0.55^{+0.15}_{-0.15}$)	$1.01^{+0.29}_{-0.29}$ ($0.19^{+0.05}_{-0.05}$)	$0.48^{+0.15}_{-0.15}$ ($0.09^{+0.03}_{-0.03}$)
BGG 0958+0223	1→0 3→2	85.065 255.181	3.25 3.25	< 0.705 < 1.99	< 4.54 < 1.42	< 1.98 (< 0.36)	< 0.43 (< 0.08)	< 0.29 (< 0.05)

Notes. Column description: (1) BGG name; (2-3) CO(J→J-1) transition and observer frame frequency; (4) on-source integration time (double polar); (5) CO(J→J-1) velocity integrated flux; (6) CO(J→J-1) velocity integrated luminosity; (7) molecular gas mass $M_{\text{H}_2} = \alpha_{\text{CO}} L'_{\text{CO}(1\rightarrow0)}$, if $\alpha_{\text{CO}} = 4.36 M_{\odot} (\text{K km s}^{-1} \text{ pc}^2)^{-1}$ is assumed, and in parentheses if a starburst value is assumed, namely $\alpha_{\text{CO}} = 0.8 M_{\odot} (\text{K km s}^{-1} \text{ pc}^2)^{-1}$; (8) depletion timescale $\tau_{\text{dep}} = M_{\text{H}_2}/\text{SFR}$; (9) molecular gas-to-stellar mass ratio. Columns (5-9) show upper limits at 3σ , with the exception of the CO(1→0) detection value for BGG 1000+0207.

reservoirs (e.g. [Violino et al., 2018](#); [Lisenfeld et al., 2019](#); [Castignani et al., 2022a](#)). However, as presented in Sect. 6.3 and summarised in Table 6.3, two out of the three BGGs are not detected in CO, yielding upper limits on their molecular gas masses being below $\sim 10^{10} M_{\odot}$. These limits imply molecular-gas-to-stellar-mass ratios in the range of $M_{\text{H}_2}/M_{\star} \lesssim 0.3 - 0.8$. By combining the molecular gas estimates (or their upper limits) with the relatively high SFR values, as those inferred in particular for BGG 1000+0207 and BGG 0958+0223, we obtain a variety of depletion timescales $\tau_{\text{dep}} = M_{\text{H}_2}/\text{SFR}$, ranging from the shortest value of 400 Myr (for BGG 0958+0223) up to ~ 1 Gyr or possibly above, for BGG 0957+0204 and BGG 1000+0207.

Although all are star-forming galaxies, a possible interpretation is that these BGGs are observed at different phases of their evolution. As further discussed below, environment-driven processes are possibly causing the exhaustion of the gas reservoirs and driving the galaxies towards rapid quenching of their star formation.

BGG 0957+0204. BGG 0957+0204 displays a disturbed morphology and is likely undergoing interaction with a nearby companion galaxy, which can be clearly seen in Fig. 6.2. Both galaxies are unresolved by our CO observations, within the IRAM 30m beam. Major interactions and mergers are generally expected to be gas-rich systems, as illustrated by nearby interacting galaxies such as

the Antennae system and by statistical studies at low and intermediate redshift (e.g. [Wilson et al., 2000](#); [Saintonge et al., 2012](#)). In this context, the absence of a CO detection in BGG 0957+0204 is particularly noteworthy.

The molecular gas mass is constrained from the rms of the CO observations, yielding an upper limit of $M_{\text{H}_2} < 9.6 \times 10^9 M_\odot$. Adopting our FIR-based SFR estimate, the resulting depletion time is $\tau_{\text{dep}} \lesssim 1.4$ Gyr, consistent with typical values for main-sequence galaxies at similar redshift. However, shorter values are not excluded, given the upper limit.

The combination of an interaction-driven morphology and the upper limit to the molecular gas content, despite the ongoing star formation activity, suggests a rapidly evolving system. Rather than being driven solely by interaction-induced star formation, the lack of detectable molecular gas likely points to environmentally driven processes, such as starvation due to suppressed gas accretion within the group environment. In this scenario, the interaction may have accelerated gas consumption, while the surrounding environment prevents the replenishment of the cold gas reservoir, leading to a non-detection in CO.

To explain the observed phenomenology, another interpretation may be that the galaxy experienced a recent episode of enhanced star formation. The IR-based SFR estimate, $\sim 7 M_\odot \text{ yr}^{-1}$, is indeed smaller than line-based SFR estimates ($\text{H}\alpha$, [OII]), in the range of $\sim 13\text{--}22 M_\odot \text{ yr}^{-1}$. This difference, given that the former estimate generally probes star formation on a larger timescale (~ 100 Myr) than the latter (~ 10 Myr), suggests that this BGG may have recently experienced an enhancement of the star formation activity, possibly induced by galaxy-galaxy interaction and gas infall, followed by a rapid consumption of the molecular gas.

BGG 1000+0207. BGG 1000+0207 exhibits a red central component surrounded by blue substructures or satellite-like features and an elongated structure towards north-west (see Fig. 6.2). These features indicated a complex morphology and, possibly, spatially distinct stellar populations. The BGG is also gas rich, given our detection of a substantial molecular gas reservoir $M_{\text{H}_2} \simeq 3 \times 10^{10} M_\odot$, yielding a large molecular-gas-to-stellar-mass ratio of $M_{\text{H}_2}/M_\star \simeq 0.5$, even higher than the values of $M_{\text{H}_2}/M_\star \simeq 0.1 - 0.4$, previously found in highly star forming (SFR $\gtrsim 100 M_\odot/\text{yr}$) intermediate-redshift BCGs such as RX1532 ([Castignani et al., 2020c](#)) and M1932 ([Fogarty et al., 2019](#)) from the CLASH survey ([Postman et al., 2012](#)). Such a large molecular gas reservoir for BGG 1000+0207, together with a substantial IR-based SFR of $\sim 30 M_\odot \text{ yr}^{-1}$, yields a depletion time $\tau_{\text{dep}} \simeq 1.0$ Gyr, as typically found in main sequence star-forming galaxies. Altogether, these features suggest that the gas fuelling sustaining the ongoing star formation is effective, while the molecular gas reservoir is also efficiently consumed.

However, our IR-based SFR estimate of $\sim 30 M_\odot \text{ yr}^{-1}$ is higher than the line-based SFR values inferred from $\text{H}\alpha$ and [OII], in the range $\sim (12 - 22) M_\odot \text{ yr}^{-1}$. Based on the different timescales probed by these two estimators, these values suggest a possible decline of the star formation

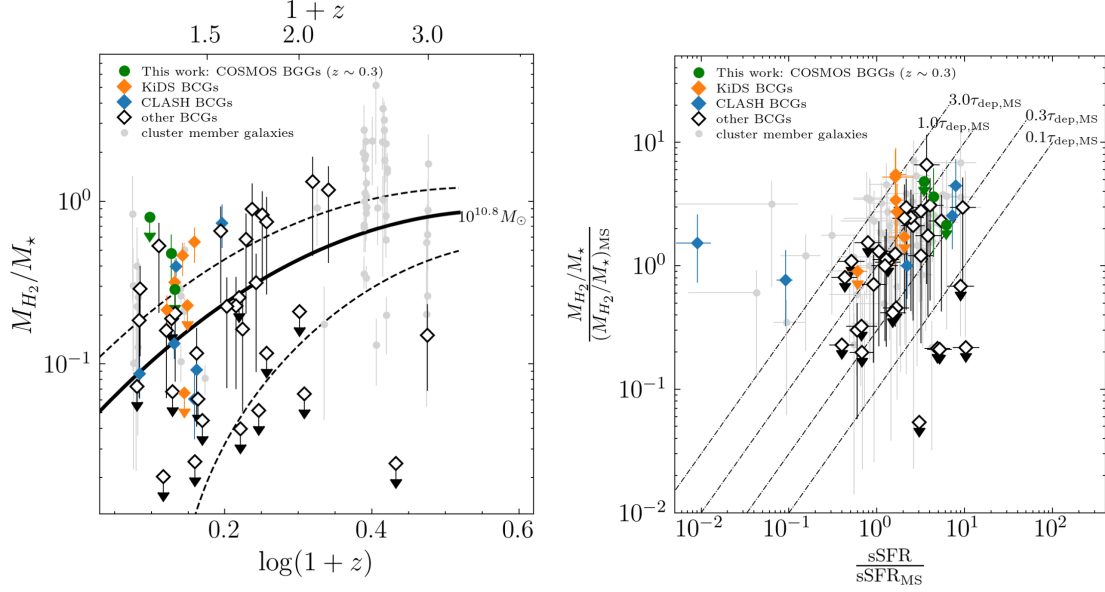


Figure 6.6. Molecular gas properties of BCGs (diamonds) and cluster galaxies (grey points) detected in CO from the literature, compared to the BGGs studied in this work (filled green circles). BCGs marked in orange are KiDS BCGs (Castignani et al., 2022a, 2023) and those marked in blue are CLASH BCGs (Castignani et al., 2020c; Fogarty et al., 2019). Left: Molecular gas-to-stellar mass ratio as a function of redshift. The solid line shows the main-sequence scaling relation for field galaxies with $\log(M_*/M_\odot) = 10.8$ from Tacconi et al. (2018), with dashed lines indicating the 1σ uncertainty. Right: Molecular gas fraction versus specific star formation rate, both normalised to their main-sequence values following Tacconi et al. (2018) and Speagle et al. (2014). Dot-dashed lines indicate constant molecular gas depletion times in units of the main-sequence depletion time. All values are calibrated to the same α_{CO} conversion factor for a consistent comparison.

activity over the past ~ 100 Myr. Therefore, BGG 1000+0207 may be observed in a phase preceding the gas exhaustion and the onset of quenching.

BGG 0958+0223. BGG 0958+0223 has a disturbed morphology, suggesting recent or ongoing dynamical perturbations. It also hosts an AGN, as indicated by its optical emission-line properties and as confirmed by the WISE colour classification described in Sect. 6.2.3. This galaxy is the only BGG in our sample to exhibit a broad $H\alpha$ emission-line component in the DESI spectrum, characterised by a line width $\sigma = 2300$ km/s and a monochromatic luminosity of 1.23×10^{42} erg s $^{-1}$, as typically found in the broad line region of Type 1 AGN (e.g. Stern & Laor, 2012; Liu et al., 2019).⁸ Interestingly, these values yield a single-epoch black hole mass of $\log(M_{BH}/M_\odot) \approx 8.26$, obtained using the calibration provided by Shen et al. (2011). High black hole masses $\log(M_{BH}/M_\odot) \gtrsim 8$ are indeed commonly found in giant ellipticals, as those hosted by central cluster galaxies and BCGs

⁸As previously mentioned in Sect. 6.2.3, we stress that the $H\alpha$ emission line was decomposed into narrow and broad emission line components, and only the former was used to estimate the $H\alpha$ -based SFR.

in particular (e.g. [Lauer et al., 2007](#); [Phipps et al., 2019](#)). Despite the evidence of AGN emission, the IR SED lacks a strong power-law component, indicating that the AGN does not dominate the far-infrared emission (see Fig. 6.4). Moreover, the overall agreement among SFR estimates from different tracers (Table 6.2) further strengthens the reliability of these estimates.

Concerning the molecular gas properties, our CO observations yield an upper limit of $M_{\text{H}_2} < 1.98 \times 10^{10} M_{\odot}$, corresponding to a short depletion time of ~ 400 Myr, given its $\text{SFR} \approx 46 M_{\odot}/\text{yr}$. Furthermore, BGG 0958+0223 has an IR-based SFR that is smaller than that derived from $\text{H}\alpha$ and $[\text{OII}]$ emission lines (see Table 6.2), similarly to BGG 0957+0204 (see Sect 6.4). This may suggest a recent enhancement of the star formation activity, on time scales of ~ 10 Myr, as probed by the line-based SFR estimators.

Taken together, these results point to a highly efficient but potentially short-lived episode of star formation, where its fuel, the molecular gas reservoir, is being consumed rapidly. BGG 0958+0223 may therefore be observed during a transient phase of enhanced star formation, possibly triggered by galaxy interactions, while AGN feedback has not yet led to the suppression of star formation.

6.4.1 Molecular gas consumption and environment

Interestingly, our three target galaxies appear to sample different phases of star formation and molecular gas consumption within a group environment. Despite the diversity of morphologies and star formation properties, the molecular gas constraints play a key role in shaping the interpretation of their evolutionary state. In all cases, the CO observations provide meaningful constraints on the molecular gas mass and possibly their star formation properties.

As previously discussed, BGG 1000+0207 is the only one that is detected in CO. It hosts a substantial molecular gas reservoir, which sustains its ongoing star formation, well above typical MS levels. However, despite having similar star formation rates, the other two BGGs have not been detected in CO. This may suggest that in these galaxies, gas depletion precedes star formation quenching, a result that has also been proposed in previous studies on massive galaxies (e.g. [Castignani et al., 2022b](#)).

Altogether, the comparison between BGG 1000+0207 and the two non-detected systems indicates that, once the molecular gas reservoir starts to be consumed or its replenishment is suppressed, the star-formation decline occurs on different timescales, depending on the specific galaxy properties. As BGG 1000+0207 has an IR-based SFR that is larger than line-based SFR, it is possible that BGG 1000+0207 has been experiencing a star formation decline over the past 100 Myr, which will possibly lead to quenching, unless a gas replenishment occurs.

Additionally, in this work, we adopted a Galactic CO-to- H_2 conversion factor, which is typical of MS galaxies. However, our targets have inferred SFRs above the main sequence (see Table 6.2). Therefore, we consider the possibility of adopting instead a lower conversion

factor, more appropriate for starbursts, $\alpha_{\text{CO}} \sim 0.8 M_{\odot} (\text{K km s}^{-1} \text{ pc}^2)^{-1}$ (Bolatto et al., 2013). For comparison, we have reported molecular gas properties for both cases, in Table 6.3. Overall, assuming a starburst conversion factor would imply a decrease of more than a factor of 5 of M_{H_2} and as a consequence of τ_{dep} and $\frac{M_{\text{H}_2}}{M_{\star}}$ reinforcing the proposed scenario of a rapid exhaustion of molecular gas, possibly induced by efficient environmental-driven processing of gas.

While interaction-driven star formation may contribute to rapid gas consumption, the environment may also play a significant role by limiting the replenishment of the cold gas supply. In such a scenario, starvation caused by the suppression of gas accretion, combined with environmental processing such as tidal interactions or stripping, may accelerate the transition toward lower star formation activity. Visual inspection of Fig. 6.2 suggests a complex morphology for our BGGs, elongated structures, blue (possibly star-forming) components, and nearby companions. As an example, the southern companion of BGG 0957+0204 appears elongated along the north-east to south-west direction, with possible tidal features.

Similar cases of gas-poor or rapidly evolving systems have been discussed in previous studies (e.g. Castignani et al., 2019, 2020a), particularly in relation to galaxy compactness and clumpy morphologies. In this context, the galaxies presented here provide additional evidence that environmental effects can strongly influence the balance between gas fuelling and consumption, leading to short star formation timescales even in systems that show signatures of recent or ongoing interactions.

6.4.2 Comparison with the literature

The molecular gas properties derived for our sample of BGGs can be further interpreted by comparing them with existing CO observations of massive galaxies in denser environments, in particular BCGs and cluster galaxies in general. Comparing our results with these studies allows us to investigate whether central galaxies in groups follow similar trends to their cluster counterparts, or whether the group environment leads to distinct molecular gas exhaustion.

Our comparison sample includes 103 galaxies, from 61 different clusters in the interval $0.2 \leq z \leq 2.0$. This sample is extracted from the compilation described in Tables A.1 and A.2 of Castignani et al. (2020b), further updated with more recent observations in Castignani et al. (2023).

Figure 6.6 compares the molecular gas properties of the BGGs studied in this work with those collected from the literature. Cluster galaxies are shown as grey points, while BCGs are indicated by diamonds, and our BGGs by filled green circles. Orange diamonds highlight KiDS BCGs (Castignani et al., 2022a, 2023) while blue ones mark CLASH BCGs, including the gas-rich and highly star-forming RX1532 (Castignani et al., 2020c) and M1932 (Fogarty et al., 2019), located in cool-core clusters at similar redshifts to our targets. In the left-hand panel, the molecular gas-to-stellar mass ratio is shown as a function of redshift. The solid line represents the scaling

relation for main-sequence field galaxies with $\log(M_\star/M_\odot) = 10.8$, corresponding to the median stellar mass of our sample, as derived by Tacconi et al. (2018), while the dashed lines indicate the associated 1σ intervals. This comparison highlights how the BGGs studied in this work span a range of gas fractions that overlaps with, or is slightly higher than, those of gas-rich cluster galaxies at similar redshifts.

The right-hand panel shows the molecular gas fraction as a function of the specific star formation rate, both normalised to their respective main-sequence values following the prescriptions of Tacconi et al. (2018) and Speagle et al. (2014). The dot-dashed lines indicate loci of constant molecular gas depletion time in units of the main-sequence depletion time. In this plane, our target BGGs occupy an intermediate region between extreme, cooling-flow-fed BCGs with intense star formation, such as RX1532 and M1932 in the CLASH sample, and the less star-forming and less gas-rich BCGs identified in the KiDS survey and hosted by galaxy clusters.

Specifically, the BGGs studied here seem to exhibit depletion times that are comparable or shorter than those of typical star-forming cluster galaxies (from KiDS and CLASH samples). However, our BGGs also have less extreme star-formation activity than that observed in strongly star-bursting, cooling-flow systems. This suggests that central galaxies in the group regime may experience efficient gas consumption without the continuous replenishment provided by cluster-scale cooling flows. In this scenario, star formation is sustained only transiently and is regulated by a combination of limited gas supply and environmental processing, placing BGGs in a transitional phase between gas-rich, highly star-forming cluster core galaxies and fully quenched central galaxies in more massive halos, typical of clusters.

6.5 Conclusions

We presented IRAM 30m CO observations of three BGGs extracted from the AMICO-COSMOS group sample (Chapter 3; Toni et al. 2024), aimed at constraining their molecular gas content and the efficiency of star formation in group environments. The three systems exhibit diverse morphologies, including disturbed structures, blue sub-components, and signatures of interactions, indicative of ongoing or recent dynamical activity.

Among the targets, BGG 1000+0207 is detected in CO, revealing a substantial molecular gas reservoir of $M_{\text{H}_2} \sim 3 \times 10^{10} M_\odot$, while BGG 0957+0204 and BGG 0958+0223 remain undetected, with upper limits of $M_{\text{H}_2} \lesssim 2 \times 10^{10} M_\odot$. When combined with multi-wavelength SFR estimates, these measurements point to distinct star formation regimes: BGG 1000+0207 is characterised by sustained, efficiently fuelled star formation with a depletion time of order ~ 1 Gyr. On the basis of star formation diagnostic and its disturbed morphology, BGG 0957+0204 shows some evidence in favour of a recent episode of enhanced star formation, possibly followed by gas exhaustion. BGG 0958+0223 exhibits a particularly short depletion time, indicative of a transient and highly efficient

episode of star formation, possibly followed by gas exhaustion, similarly to BGG 0957+0204.

Altogether, our results indicate that intermediate- z star-forming BGGs in group environments occupy an intermediate position, in terms of their molecular gas content and star formation values, between strongly star-forming, cooling-flow-fed BCGs in massive clusters and less star-forming BCGs. While the molecular gas reservoirs and star formation activities of the BGGs in this work are less extreme than those observed in classical cooling-flow systems, they are nonetheless sufficient to sustain short-lived phases of enhanced star formation. This may suggest that environmental processes in groups can regulate the availability and consumption of cold gas, driving rapid evolutionary phases even in massive central galaxies.

As an IRAM pilot project, we plan to expand this study by enlarging our sample both within the COSMOS field and in other extragalactic surveys, an effort that will be facilitated by the steadily increasing spectroscopic coverage. A larger statistical sample will allow us to investigate possible dependencies between star formation activity and molecular gas excitation, and to assess whether the trends suggested by our pilot study persist across different evolutionary stages and environments. In this context, the new public catalogue of galaxy groups from the COSMOS-Web survey, extending out to $z = 3.7$ (Chapter 4; [Toni et al. 2025a](#)), provides a unique opportunity to probe the molecular gas properties of central galaxies at $z > 1$, where current CO observational coverage remains sparse (see Fig. 6.6).

Finally, confirming and disentangling the relative impact of different quenching mechanisms will require high-resolution interferometric observations—such as those achievable with ALMA or NOEMA—to spatially resolve the molecular gas distribution and excitation conditions in central group galaxies (e.g. [Castignani et al., 2025](#)). Such observations will be crucial to link the global gas properties to the underlying physical processes driving star formation and quenching in dense environments.

III

Epilogue

7

Conclusions

A conclusion, by definition, marks the end of a discussion. Yet, as this Thesis comes to a conclusion, it also represents a beginning. The work and the results presented in the first part of this Thesis lay the foundation for future scientific investigations, providing the community with resources and tools to explore their compelling questions.

In this Thesis, we robustly detected galaxy clusters, groups, and protocluster cores up to $z = 3.7$ in the COSMOS field using the AMICO algorithm, and we have characterised them leveraging multi-wavelength data. The algorithm, previously tested on wide-field surveys, was successfully enhanced and adapted to the peculiar survey configuration of deep fields, such as the COSMOS, and even more challenging, the COSMOS-Web field. This application has resulted in two robust catalogues, covering up to $z = 3.7$ and down to a few $10^{12} M_{\odot}$, and has enabled the identification of hundreds of systems that remained undetected until now. As part of the validation of the produced catalogues, X-ray counterparts have been assigned, enabling the calibration of mass proxy scaling relations; the purity and the completeness have also been estimated through the creation of data-driven, realistic mock catalogues. The subsequent scientific analyses based on these catalogues have provided new insights into the build-up of the quiescent galaxy population over ~ 12 Gyrs, allowed the identification and investigation of rare systems and the study of the formation and environmental processing of bright central galaxies.

The main results of the work presented in this Thesis can be summarised as follows, by grouping them into four thematic areas.

1. Creation of well-characterised publicly available catalogues

The AMICO algorithm has proven efficient in detecting clusters and groups, even in challenging survey configurations such as the deep 2-deg² COSMOS field. Its first application to the COSMOS(2020) data produced a robust catalogue of galaxy clusters and groups up to $z = 2$ and over an effective area of 1.69 deg², **the AMICO-COSMOS catalogue**. The final sample contains 1269 candidates with $S/N > 3$, resulting from three independent runs performed using magnitudes in three different bands as galaxy property. Although all three runs yielded new detections, most of the unmatched detections occupy the low-end of the S/N distribution, demonstrating the reliability of the method. Cross-matching with the latest COSMOS X-ray group catalogue (Gozaliasl et al., 2019) revealed 122 correspondences. Additionally, X-ray emission with significant flux was confirmed for another 500 AMICO detections, yielding a total of 622 AMICO candidate clusters and groups with associated X-ray properties up to $z = 2$. This matching analysis enabled the calibration of the scaling relations between AMICO mass-proxies and X-ray mass up to $z = 2$ and down to less than $10^{13} M_{\odot}$. The inclusion of a redshift-dependent term revealed the stability of these relations when using near-infrared magnitudes, and the analysis of X-ray point-source removal highlighted instances of underluminous systems due to line-of-sight interlopers. All properties, including X-ray measurements, were published in the catalogue.

Building on the insights and experience gained from the creation of the AMICO-COSMOS sample, we attempted to push the limit even further: the AMICO algorithm was applied to the new COSMOS-Web photometric galaxy catalogue down to a magnitude of 27.3 (in the F150W band), over an effective area of 0.45 deg² and up to $z = 3.7$, covering the transition from protoclusters to virialised objects and their maturing phase. This represents the deepest application of the AMICO algorithm to date and has yielded **the COSMOS-Web catalogue**, a large, deep sample comprising 1678 candidate galaxy groups and even protocluster cores at $z \geq 2$. Among these latter, 316 were previously unknown, and 205 matched known high-redshift systems, for a total of 509 candidates in the redshift range $2 \leq z \leq 3.7$. The high- z regime analysis was performed by compiling a list of all known structures in COSMOS in the redshift range covered by our sample and by comparing them with our COSMOS-Web detections. The protocluster compilation with the relevant references was also released. 400 among the high- z candidates were found to be embedded in large-scale structures identified through 3D clustering with the DBSCAN algorithm. The purity and completeness of this sample were evaluated against realistic mocks generated with SinFoniA, a data-driven approach which allowed for the creation of mock samples, capturing the complexity of real data. This enabled us to find the relationship between purity and S/N , to establish cuts of the original catalogue based on the desired purity level. For example, a cut at $S/N \sim 10$ corresponds to selecting a sample of around 670 objects with purity above 90%.

2. Methodological and technical insights on structure detection with AMICO

This work has demonstrated that AMICO, an algorithm proven to be particularly effective in creating large cosmological samples in wide-field surveys, can be successfully adapted and optimised for deep, small-area surveys, and for extensive searches for structures at high redshift. Several enhancements were introduced to improve AMICO’s performance in these peculiar survey configurations. These include a procedure to mitigate the impact of physical overdensities on the noise model, techniques to stabilise the algorithm in low-statistics regimes, and a flexible masking approach that integrates multiple sources of information to optimise survey coverage. Together, these enhancements increase the robustness and flexibility of the algorithm in peculiar conditions, allowing for the reliable detection of low-mass and high-redshift structures, and providing a solid foundation for the creation of well-characterised catalogues that can be used in future studies. The careful preparation and pre-processing of the input data, the construction of suitable cluster models, a robust masking strategy with particular attention to the integration of ground-based and space-based data, and the regularisation of the noise model were all essential to unlock the power of the detection algorithm and achieve reliable results when pushing the limit to the lowest-mass systems and the highest redshifts. For these reasons, the methodologies we developed and tested in this Thesis will be of great use for the successful exploitation of the algorithm in other deep surveys, such as the Euclid Deep Fields, drilling fields, calibration and testing fields (e.g. [Maturi et al., 2023](#)), as well as in other peculiar survey configurations.

Another methodological aspect explored in this Thesis is the capability of the algorithm to detect overdensities and large-scale structures in a regime where regular and virialised systems are not expected to have formed yet. The excellent photometric redshift quality, the depth of the COSMOS-Web survey, together with the availability of large-scale structure catalogues derived with complementary methods in the field, provided an ideal opportunity to explore the possibility for AMICO to detect protocluster cores at $z \geq 2$. The methodological approach we adopted consists of a new strategy which has proven to be effective: instead of attempting to reproduce protocluster shapes and properties with complex models, we relied on the exact same model used for the detection of virialised clusters and groups at lower redshifts. This idea is based on the similarities (in terms of size and dynamical state) between overdensity peaks or protocluster cores and low- z groups, for which we can expect these to be detected as small systems when using a standard cluster model for the AMICO detection. And this is what happens when AMICO is applied to deep high- z data in the COSMOS-Web field. The larger overdense regions can then be reconstructed by analysing the clustering properties of the detected cores, namely by identifying tens-of-Mpc-scale structures made by the clustering of several overdensities.

3. Galaxy evolution and environmental processing

A natural application of a deep and well-characterised dataset, such as the COSMOS-Web group catalogue, is the study of the statistical properties of the group galaxy population. Using this catalogue, we conducted an extensive analysis of the evolution with redshift and the dependence on group richness of the quiescent galaxy fraction and of the red-sequence parameters, over roughly 12 billion years of cosmic history, from $z = 0$ to $z = 3.7$. Multiple complementary methods were employed to estimate the quiescent fraction and to identify and characterise the red sequence, yielding consistent results. In total, 225 red-sequence ridgelines were identified using rest-frame colours, with 214 also robustly detected using matched observed colours. This work extends the detection and characterisation of the red sequence and quiescent fractions well beyond $z > 1.5$ and down to low-richness systems, a regime previously lacking extensive investigation.

To classify galaxies and distinguish between quiescent and star-forming populations, we adopted a probabilistic machine-learning approach based on gradient boosting decision trees, which proved to be highly efficient even with class imbalance, achieving an F1-score of over 93%. Importantly, the AMICO algorithm identifies galaxy groups without explicitly relying on colour information, enabling a nearly unbiased analysis of member galaxy properties. We found that the quiescent galaxy population in groups builds up steadily from $z = 1.5 - 2$ across all richnesses, with faster and earlier growth in the richest groups. The first galaxies settle onto the red-sequence ridgeline by $z \sim 2$, consistent with currently favoured evolutionary scenarios. Overall, observed average colours of the red-sequence ridgelines are consistent with those predicted by passively evolving elliptical galaxy models, and we found no significant evolution in the slope or scatter of the red sequence with redshift. By extending this analysis to the AMICO-COSMOS catalogue, which boasts the availability of X-ray properties, we found an indication that X-ray faint groups exhibit, on average, lower quiescent fractions than X-ray bright ones.

Moving from population-level analysis to individual peculiar galaxies, we detected for each of our candidates the most probable brightest cluster/group galaxy (BCG/BGG), including alternative candidates when the selection was ambiguous. During the visual inspection of the AMICO-COSMOS BCG/BGG catalogue, we identified three remarkable star-forming central galaxies caught in a rare stage of evolution, clearly undergoing mergers or interactions with close companions. To probe the role of molecular gas in sustaining their star formation, we obtained observing time at the IRAM 30m telescope to search for CO emission lines at millimetric wavelengths. One of the targets was successfully detected in CO, revealing a substantial molecular gas reservoir of $M_{\text{H}_2} \sim 3 \times 10^{10} M_{\odot}$, while for the undetected targets, we were able to set upper limits of $M_{\text{H}_2} \lesssim 2 \times 10^{10} M_{\odot}$. When combined with multiwavelength SFR estimates, these measurements point to distinct star formation regimes, including sustained, efficiently fuelled star formation, or recent episodes of transitional enhanced star formation, possibly followed by gas exhaustion.

Our results, when compared with the literature, indicate that intermediate- z star-forming BCGs in group environments may occupy an intermediate position between strongly star-forming, cooling-flow-fed BCGs in massive clusters and more quenching BCGs.

Together, these results outline a coherent picture in which both environmental effects and internal gas-regulation processes shape galaxy evolution within groups—from the collective build-up of the quiescent population to the transformation of individual central galaxies. The analyses presented in this Thesis demonstrate the power of deep, well-characterised group catalogues to address long-standing questions about how galaxies form and evolve in dense environments. Yet, on the galaxy-evolution front, we have only begun to scratch the surface: the potential of these new catalogues extends far beyond the applications explored here, offering a wealth of opportunities for future investigation. Some of these promising directions are discussed later in Chapter 8.

4. High- z Universe and rare systems

The new application of the AMICO algorithm to the deep COSMOS-Web survey has opened a new window onto the high-redshift Universe. This work has demonstrated that the same detection formalism successfully employed in wide-field surveys at low-to-intermediate redshift can be adapted to deep, small-area observations, enabling the identification of structures at the earliest stages of cluster and group formation. Using a standard cluster model, the application of AMICO to the COSMOS-Web catalogue has yielded a large sample of high- z groups and protocluster core candidates, including over 300 new detections at $z \geq 2$. To assess their reliability, we compiled and cross-matched all known protocluster and group candidates in the COSMOS field in this redshift range, finding 205 associations with previously reported systems and confirming the robustness of our detections.

Among the new detections, we identified a particularly remarkable structure: a large, 20-cMpc complex of 14 protocluster cores clustered within the redshift range $z \sim 2.5\text{--}3.0$. This overdensity, which we named AmicOne, represents the first and largest protocluster candidate detected with AMICO in real data, offering a compelling target for future spectroscopic follow-up and multi-wavelength characterisation. In total, around 400 of our high-redshift detections are found within such extended, interconnected structures, further illustrating the capability of AMICO to trace the early stages of cosmic web assembly.

In parallel, during our galaxy colour investigation, we discovered a compact (~ 250 kpc/ h) overdensity of quiescent galaxies at $z = 3.4$, which, if confirmed, may represent one of the most distant red sequences identified so far. The system, previously unknown in the literature, provides a rare glimpse into the early emergence of environmental quenching. Identifying passive galaxies in such young and immature systems offers a unique opportunity to pinpoint the onset of quenching, explore its connection to the assembly of the densest structures and study some of

the Universe’s oldest evolved galaxies. To follow up on this interesting discovery, we recently submitted a JWST Cycle 5 proposal (PI: Toni) to spectroscopically confirm the nature and internal structure of this protocluster core. The goal is to characterise the stellar populations and ages of its central galaxies, to address key open questions on the timing of galaxy evolution, and to test whether the assembly of dense structures can accelerate quenching and produce among the oldest quiescent galaxies in the Universe—potential progenitors of early brightest cluster galaxies.

Our analysis also revealed several projection effects not known in the literature. One of these was identified as a rich detection without significant X-ray emission, even after the re-inclusion of flux subtracted during point-source removal. This was found to be a very rare apparent cluster that, upon analysis of galaxy spectroscopic redshifts, revealed a double-group structure aligned along the line of sight, possibly inside a filament. Such systems, although very uncommon, underscore the importance of combining multi-wavelength data and spectroscopic information to confirm true physical shapes and orientations of such objects.

Overall, this work highlights how deep, high-quality photometric data—combined with a flexible, robust detection algorithm—can efficiently unveil both the earliest structures and rare configurations in the Universe. Even in a field as thoroughly explored as COSMOS, new and intriguing systems continue to emerge. As one might say, *the COSMOS field is one of the most observed patches of the sky—yet it still holds surprises*. These findings establish a foundation for further future studies of structure formation and galaxy evolution, many of which will be pursued in the spectroscopic and multi-wavelength follow-ups discussed later in Chapter 8. Figure 7.1 closes this Chapter with a visual representation of the analysed field, all systems detected in this Thesis and the locations of some of the most interesting objects detected.

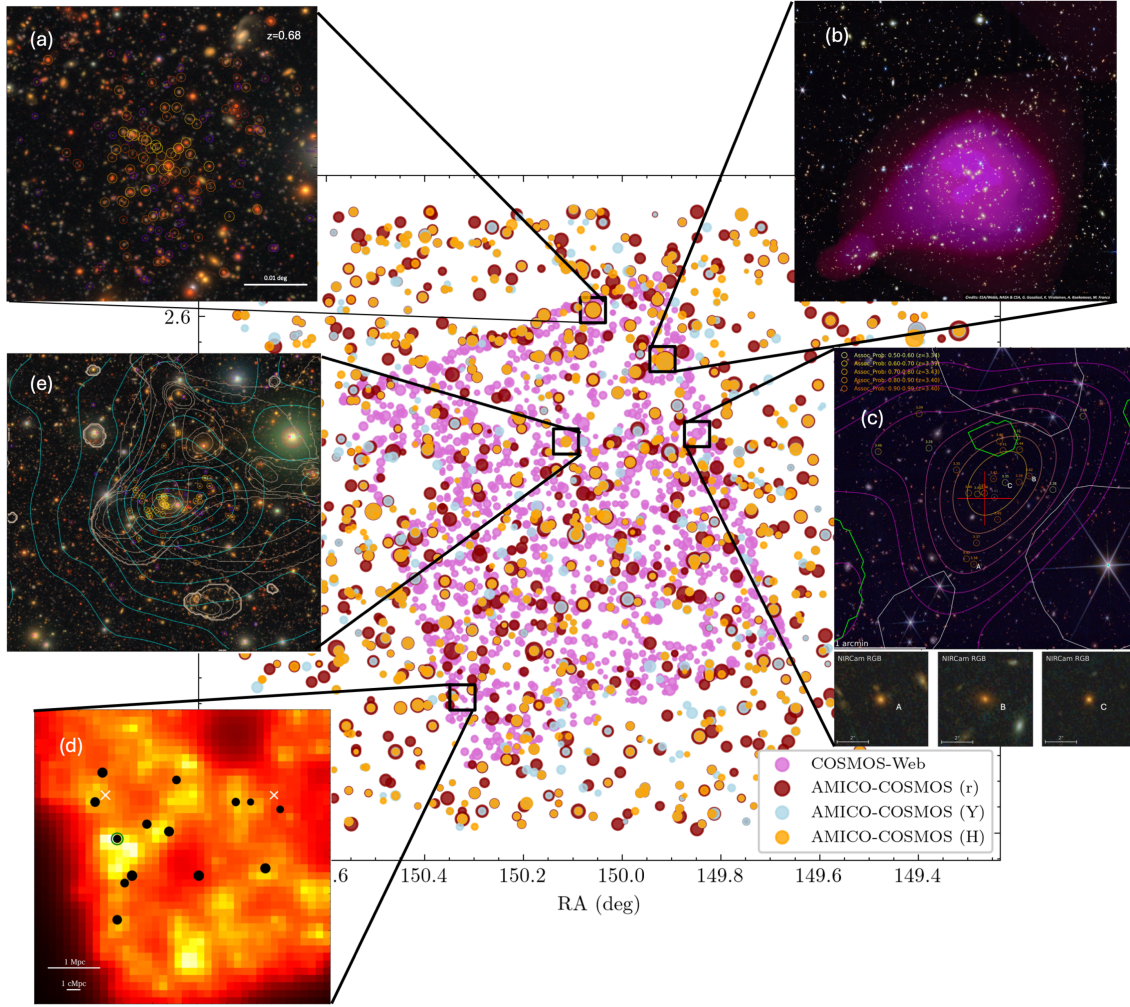


Figure 7.1. A visualisation of all systems detected in the COSMOS and COSMOS-Web field as part of this Thesis work. Different colours refer to different catalogues produced during this work, as indicated in the legend, while the point size scales with the richness, λ_* . Zoom-in panels of five interesting objects detected and characterised in this study are shown: (a) an apparent rich cluster without significant X-ray emission at $z = 0.68$ which was revealed to be a double group aligned along the line of sight, probably in a filament ($\Delta z \sim 0.01$, Figure 3.15); (b) the richest system in the COSMOS-Web field, a galaxy cluster with clear X-ray-emitting ICM at $z \approx 0.7$, the image was selected as [ESA/Webb picture of the month](#) in April 2025 (Figure 1.3); (c) the highest- z red-sequence system detected in our sample and possibly one of the highest ever detected: 3 passive galaxies (3 bottom panels) in a very compact, rich protocluster core at $z = 3.4$ for which we planned a spectroscopic follow-up with JWST; (d) the AmicOne large-scale structure at $z = 2.5 - 3$: the first and largest protocluster candidate detected with AMICO in real data, composed by 14 different protocluster cores (black points) clustered together over 20 cMpc on the sky plane (Figure 4.13); (e) the (super-)group detected by Smolčić et al. (2017) via wide-angle-tail radio-BGG, as seen by AMICO (Figure 3.8).

8

Future Prospects

The work presented in this Thesis has laid the foundations for a new generation of studies on galaxy groups, protoclusters, and their galaxy populations. The catalogues and tools developed in this Thesis open a wealth of possibilities for future exploration, continuing what we have achieved so far—from the spectroscopic confirmation of high-redshift candidates and the refinement of mass calibration relations, to the study of galaxy populations, the intragroup medium, and the large-scale environments in which these systems are embedded. The next stages of this work will focus on consolidating current findings, testing new hypotheses, and exploiting the unprecedented depth and quality of modern and forthcoming datasets. Below, the main directions and projects that will guide this continuation in the near future are outlined.

- **X-ray aperture photometry and scaling relations for COSMOS-Web groups.** In [Toni et al. \(2024; Chapter 3\)](#), we have performed an X-ray analysis of our AMICO-COSMOS detections both via cross-matching and via direct aperture photometry on the newly detected systems. The same exercise can obviously be performed for the most recent COSMOS-Web group sample presented in [Toni et al. \(2025a; Chapter 4\)](#). Preliminary work has already been carried out to assign masses via mass-observable scaling relations for this sample, showing consistent results with the previous analysis. However, further assessment is needed to complete the calibration of the scaling relations and the X-ray measurements targeted on new candidates, for which we will rely on our ongoing collaboration with A. Finoguenov (Univ. Helsinki). Additionally, we will use mass estimates coming from independent proxies/probes, like recent SZ ([Chéroutier et al., 2025](#)) and weak-lensing

maps of the COSMOS field (Scognamiglio et al. 2026, including GT). This analysis will complement the catalogues we produced and released in this Thesis with mass estimates, providing the community with a complete characterisation of COSMOS groups up to high redshift.

- **X-ray bright and X-ray faint AMICO detections in COSMOS.** As shown in Toni et al. (2024; Chapter 3), around 46% of the AMICO detections in the AMICO-COSMOS catalogue have no significant X-ray emission, a result fairly independent of S/N or redshift cuts. Recently, for the GAMA-eROSITA sample comparison, Popesso et al. (2024) found that X-ray-selected and optically-selected systems seem to form two distinct populations with different structural properties. X-ray-faint clusters show less concentrated X-ray emission, dark matter, and galaxy distributions, consistent with being younger systems. They also found differential distribution in the cosmic web, which we have confirmed for COSMOS in Toni et al. (2025b, Chapter 5), finding generally lower quiescent fractions for X-ray faint groups. Once the above-mentioned X-ray analysis is completed, a comprehensive study of the differences between the two populations over a wide mass and redshift range will be achievable. Using richness estimates, one can also explore "silent" detections (low X-ray emission, high richness) and vice versa ("loud" detections). This approach will help identify and characterise projection effects, peculiar, and hard-to-detect systems and shed light on intrinsic differences between the two populations.
- **X-ray bright and X-ray faint AMICO detections in Fornax.** In line with the project mentioned in the previous point, a forthcoming deep, multi-wavelength survey that provides both high-quality optical/NIR imaging and deep X-ray coverage will be the $\sim 10 \text{ deg}^2$ Euclid Deep Field Fornax (EDFF). As one of the three selected Euclid Deep Fields, the EDFF will be observed at an expected depth of 2 magnitudes fainter than the Euclid wide survey mode. Additionally, the EDFF benefits from a dedicated 1000-hour XMM-Newton Heritage program, the Fornax project¹, and includes the Chandra Deep Field South (CDFS)². The synergy between XMM-Newton and Euclid in the Fornax field is expected to enable the combined multi-wavelength study of more than 500 groups over more than 10 Gyrs of the Universe's history, bridging the gap between the early assembly of large-scale structures and the mature clusters observed in the local Universe. The methods and results of this Thesis are going to be crucial for the application of AMICO (as one of the official algorithms of Euclid) to the Euclid Deep Fields and to the interpretation of the results in relation to the cross-matching and comparison with the X-rays.
- **BGG properties in relation to the host halo.** A comprehensive study of the structural

¹<https://fornax.cosmostat.org/>

²<https://cxc.harvard.edu/cdo/cdfs.html>

evolution of BGGs in the group catalogue presented in Toni et al. (2025a; Chapter 4) has been recently carried out by Gozaliasl et al. (2025, including GT). This analysis provides insights into the morphological and star formation history of these central galaxies in group-scale haloes over ~ 12 Gyrs. In a complementary way, interesting insights can be gained by studying the properties of BGGs (such as size, SFR, stellar mass, total mass, age, quiescence and other crucial ingredients not included before like the presence of an AGN), put into relation to the hosting halo properties (such as richness, total halo mass, position in the LSS, dynamical state and multi-wavelength emission). This would help shed light on the still-debated influence that the group environment has on the formation and evolution of central galaxies and its disentanglement with respect to intrinsic galaxy properties (the so-called *nature vs nurture* problem).

- **Evolution of the luminosity function.** The work done on the AMICO-KiDS cluster sample by Puddu et al. (2021) and the development of the new cylinder-based background subtraction method presented in this Thesis, pose the basis for a robust study of the LF and its evolution using our new COSMOS cluster and group catalogues. The COSMOS catalogues map groups of galaxies down to very low masses (less than 10^{13} solar masses) and cover a small sky area, but at an unprecedented depth and up to a relatively unexplored redshift regime. This will bring important insights into the build-up and evolution of the LF and of the most luminous galaxies. The probabilistic classification method developed in this Thesis will also allow for the study of the LF of quiescent and star-forming galaxy populations separately.
- **Protocluster detection in Euclid.** Euclid is expected to provide the first wide-area statistical census of protoclusters up to $z \sim 4$, enabling studies of their abundance, structure, and role in early galaxy evolution, while also revealing rare, massive systems that test cosmological models (Euclid Collaboration: Böhringer et al., 2025). Ongoing efforts within the collaboration are focused on the detection of protoclusters in the first data, the cross-matching with external data and the systematic comparison of different identification algorithms, with the goal of optimising completeness, purity, and robustness of the resulting samples. The ability of AMICO to detect protocluster cores by using a cluster model, demonstrated in Toni et al. (2025a; Chapter 4), has opened new exciting perspectives for the use of the algorithm in the protocluster regime. Promising active work is underway in this direction.
- **Spectroscopic and millimetric follow-up of high- z candidates.** In this Thesis project, we have identified more than 300 new candidate high- z group and protocluster cores at $z \geq 2$, solely based on photometric redshifts (Toni et al. 2025a; Chapter 4). Despite the excellent quality of photo- z provided by COSMOS2025, spectroscopic confirmation is of

crucial importance, especially for high- z systems. We plan to carry out a medium-large spectroscopic campaign to confirm AMICO protocluster core candidates in the COSMOS-Web field. For these candidates, the strategy would include using NIR spectrographs on facilities such as JWST and Keck to target spectral features of star-forming galaxies, which are prevalent at these cosmic times. Complementarily, the emission features of dusty star-forming members, such as CO lines redshifted into the sub-mm, can be targeted with observatories such as ALMA or NOEMA.

- **Spectroscopic confirmation of quiescence.** Recently, we have submitted a JWST Cycle 5 proposal (PI: Toni) to spectroscopically confirm the nature and internal structure of the rare, compact $z = 3.4$ protocluster core we have discovered, with hints of early quenched star formation from photometry. By characterising the stellar populations and ages of its central galaxies, the intention is to address key open questions on the timing of galaxy evolution, testing whether the assembly of dense structures can accelerate quenching and produce among the oldest quiescent galaxies in the Universe—potential progenitors of early brightest cluster galaxies. Besides this, as stressed in [Toni et al. \(2025b, Chapter 5\)](#), a more extensive spectroscopic follow-up would be necessary to confirm the trends of the red-sequence parameters we have presented, especially with increasing redshift.
- **IRAM 30m BGG campaign.** In this Thesis, we have extended previous efforts in characterising the processing of molecular gas in BCGs to the group regime, targeting 3 star-forming BGGs at $z \sim 0.3$, selected from our AMICO-COSMOS groups. The results suggest different gas consumption scenarios, including rapid exhaustion of molecular gas and rapid quenching. Building upon our first IRAM 30m observations and the previous successful IRAM campaign on distant BCGs, we plan to extend our observations further and enlarge the BGG sample. This will be allowed by the growing (public) spectroscopic coverage over the COSMOS field ([Khostovan et al., 2026, including GT](#))—since precise redshifts are needed to target CO lines— and by the deep JWST observations, which have unveiled many previously unknown systems (as seen in [Toni et al. 2025a; Chapter 4](#)).

Beyond the projects envisioned in this list, several independent teams are using our publicly released catalogues. These resources are being employed for a range of studies, from the environmental effects on galaxy morphology, such as the effect on galaxy bars, to the link between group properties and radio-AGN fractions (in collaboration with F. Ubertosi and I. Delvecchio). The public availability of these catalogues continues to foster new collaborations and expand the scientific reach of this work.

Looking ahead, the landscape for the study of structure formation and galaxy evolution is particularly promising. The synergy between deep imaging surveys, expanding spectroscopic and multi-wavelength coverage, and new facilities such as Euclid, JWST, LSST and the Roman

Space Telescope will offer an unparalleled view of galaxies and their environments across cosmic time. The tools, techniques, and findings presented in this Thesis are well-positioned to make the most of the wealth of data that forthcoming surveys will soon deliver.

The list of future projects outlined here is part of an evolving trajectory—reshaping and growing as new data, new hypotheses and new questions arise. The pace of discovery in this field ensures that every answer and every new resource reveals new directions to explore. In this sense, the methods, the results and the public resources developed in this Thesis represent a gateway: a step forward toward a deeper understanding of how galaxies, structures, and ultimately the cosmic web itself came to be.

IV

Appendix



AMICO run based on spectroscopic redshifts

This Appendix is fully based on Toni et al. (2025a), A&A, 697, A197.

In Chapter 4, Section 4.3.2, we performed a simple a-posteriori association of spectroscopic counterparts to the member galaxies retrieved by applying AMICO to the COSMOS-Web data. The association performed by AMICO is entirely based on photometric redshifts. In this Appendix, we describe instead a group search done with AMICO by using not only photometric redshifts but also spectroscopic redshifts when available.

The availability of almost 100 spectroscopic surveys covering the COSMOS field and collected in the DR1.0 of the compilation by Khostovan et al. (2026, including GT) makes this an ideal chance to test cluster and group detection with AMICO, including the information coming from spectroscopic redshifts. First of all, we associated spectroscopic redshifts with the full cleaned input galaxy catalogue. This yielded a galaxy sample with spectroscopic redshift available for almost 20,000 galaxies, which is 5% of the initial input catalogue. Besides these, we kept all galaxies with the photo- z only, creating a hybrid input catalogue. We built this hybrid catalog by selecting galaxies with spec- z and by shifting their redshift to the spectroscopic values. Then, we added a 1σ error, compatible with the size of the chosen AMICO redshift resolution, namely $\Delta z = 0.01(1+z)$. We then run AMICO detection on the hybrid catalogue with the same parameters described in Section 4.2. This run resulted in a catalogue of 1559 candidate clusters and groups

in the range $0.03 \leq z \leq 3.7$ and with $\lambda_* > 2$. Among these, a total of 595 groups are detected with $S/N_{ocl} > 10.0$. We compared the results of this run with the one based only on photo- z by matching the two catalogues within $dz = 0.05(1+z)$ and $dr = 1 \text{ Mpc}/h$. For simplicity, we refer to the standard catalogue described in Section 4.3 as PhotoCat and to the results of this run with hybrid redshifts as SpecCat. The comparison between the two catalogues resulted in 1501 successfully matched candidates in total, of which 1404 had $0.03 \leq z \leq 3.7$ and $\lambda_* > 2$, that is $\sim 90\%$ and $\sim 83\%$ of SpecCat and PhotoCat, respectively. Then, we looked for possible unmatched detections that have significantly changed their redshift by introducing spec- z s but are still detected above the minimum signal-to-noise. To do so, we matched unmatched detections from SpecCat and PhotoCat, without using the redshift, within a sky separation of 1 arcmin. There are 62 detections in the considered redshift range that are compatible with having the same position in the sky but having a different redshift. We then compared the assigned members to clean from random matches. Among the 62 matches, we found only 4 detections sharing more than 5 associations with $P > 20\%$ changing their redshift of a scatter larger than $0.07(1+z)$, so we attribute most of them to random matches. Then, we analysed the possibility of redshift fragmentation, namely the possibility that a given detection with a sufficient number of members with spectroscopic redshifts and a sufficient number of members with only photometric redshifts ends up being split into the spectroscopic and photometric components, resulting in two distinct detections aligned along the line of sight. This can be studied by looking at the possible correspondence on the sky plane (without redshift information) between a couple of matched detections and an unmatched spectroscopic detection, since this would indicate that the photometric component is still successfully identified, but it also gave origin to a new detection at a different redshift due to the spectroscopic galaxies. Among the list of 151 sky matches within 1 arcmin, we found 42 detections that share at least three members with $P > 50\%$. However, the selected pairings have a redshift scatter between the matched couple and the only-spectroscopic detection, smaller than the average photo- z error. Therefore, in our sample, the redshift fragmentation does not affect the group detection. Finding several examples of pairs coupling with a third detection in the SpecCat may instead indicate that introducing spectroscopic redshifts possibly contributes to reducing over-merging, and the algorithm can more easily distinguish between close-by objects.

The next step is to consolidate the current findings by testing a hybrid run on a larger spectroscopic sample and by performing an AMICO run based solely on spec- z s. One of the main challenges is that most galaxies with spectra are low- to intermediate-mass star-forming systems, which are relatively rare in clusters and groups, at least up to $z \sim 1-2$. Moreover, the current statistics are still too limited to conduct a purely spectroscopic detection with AMICO, since the density of galaxies with spectra is not yet representative of the overall physical density. However, spectroscopic coverage is subject to continuous rapid growth. The COSMOS spectroscopic compilation by [Khostovan et al. \(2026, including GT\)](#) has already nearly doubled since the public

release of DR1.0 (March 2025). With the forthcoming large datasets from facilities such as Euclid, Subaru/PFS, VLT/MOONS, Roman and many others, this type of analysis will soon become feasible.

B

BGG selection: a new hybrid method

This Appendix is fully based on Gozaliasl et al. (2025, including GT), accepted for publication on A&A.

Historically, BCGs and BGGs have been identified either through luminosity-based selection—by choosing the most luminous galaxy in a specific band, typically the r-band—or through stellar mass-based selection, which selects the most massive galaxy within the system (De Lucia & Blaizot, 2007; Koester et al., 2007; Gozaliasl et al., 2019).

We identified BGGs in the COSMOS-Web galaxy group catalogue (Chapter 4) using a hybrid selection method that combines both stellar mass and luminosity criteria. Building upon the framework of Gozaliasl et al. (2019); Gozaliasl et al. (2020), our approach addresses two key limitations of traditional methods: (1) the spatial offset of BGGs from group centres in low-mass systems, and (2) contamination from starburst galaxies in purely luminosity-based selections.

Our methodology examines galaxies within fixed apertures of 250 kpc and 500 kpc (with additional tests at 750 kpc) from the group centres, employing a two-stage selection algorithm. First, we follow the approach outlined in Chapter 6, applying a redshift constraint requiring $\Delta z < 0.05(1 + z)$ relative to the group candidate to ensure the selection remains confined to the group’s redshift space. Within this filtered sample, we identify the luminosity-selected BGG as the galaxy with the lowest apparent magnitude m in the F150W band. However, if the magnitude difference between the first and second brightest galaxies is less than 0.5 mag, then we choose the one with the highest membership probability.

In parallel, we repeat the selection using the median stellar mass derived from LePhare instead of magnitude. In this case, we select the most massive galaxy as the mass-selected BGG, unless the logarithmic stellar mass difference between the most massive and the second most massive galaxy is smaller than 0.25 dex; in that case, we again select the one with the higher membership probability. For all detections, we store both the first and second candidates according to these criteria. This approach ensures that each group has both a mass-selected and a luminosity-selected BGG, providing a robust cross-check of central galaxy identification.

For each group, we evaluate two primary candidates: the galaxy with the highest median stellar mass and the brightest galaxy. We then compare the stellar mass and r -band luminosity of both candidates within a given group. The selection incorporates aperture-dependent corrections, preferentially choosing 250 kpc candidates when: (a) for mass-selected BGGs, the mass difference is ≤ 0.1 dex; and (b) for luminosity-selected BGGs, the magnitude difference is ≤ 0.15 mag relative to 500 kpc candidates. Thus, the final hybrid selection employs the following decision workflow:

1. select the mass-dominant galaxy if its stellar mass exceeds that of the luminosity-selected candidate by ≥ 0.1 dex;
2. choose the luminosity-dominant galaxy if it is brighter by an equivalent mass margin;
3. default to the brighter galaxy when stellar mass differences are < 0.1 dex, giving higher weight to recent star formation activity.

A total of 1,294 BGGs were identified based on stellar mass. Additionally, 384 BGGs were identified using a luminosity-based criterion, selecting the brightest galaxies. For a more comprehensive selection that accounts for both stellar mass and luminosity, we included 803 BGGs selected within a 250 kpc radius from the group centre, and 875 BGGs within a 500 kpc radius. This combined approach provides a more accurate representation of the BGG population, minimising selection biases and capturing galaxies with both high stellar mass and luminosity.

Figure B.1 shows that smaller apertures (250 kpc) bias selections toward lower luminosity and stellar mass values, often misidentifying satellites as BGGs. By combining mass and luminosity metrics across multiple apertures, we significantly reduce these biases and more accurately represent the true BGG population across diverse group environments, from compact to loosely bound systems.

To assess the consistency and robustness of BGG identification methods, we compared stellar mass- and luminosity-based selections across our sample of galaxy groups. The two methods agreed on the BGG in 48.7% of the systems. For the remaining groups, the stellar mass difference between the selected BGGs, $\Delta \log_{10}(M_*)$, has a median of 0 and a standard deviation of 0.46 dex, with 65.7% of the groups exhibiting differences smaller than 0.3 dex, and 5.7% showing catastrophic disagreement ($\Delta \log_{10}(M_*) > 1$ dex). The difference in absolute R -band magnitude, ΔM_R , shows a wider spread, with a standard deviation of 0.62 mag.

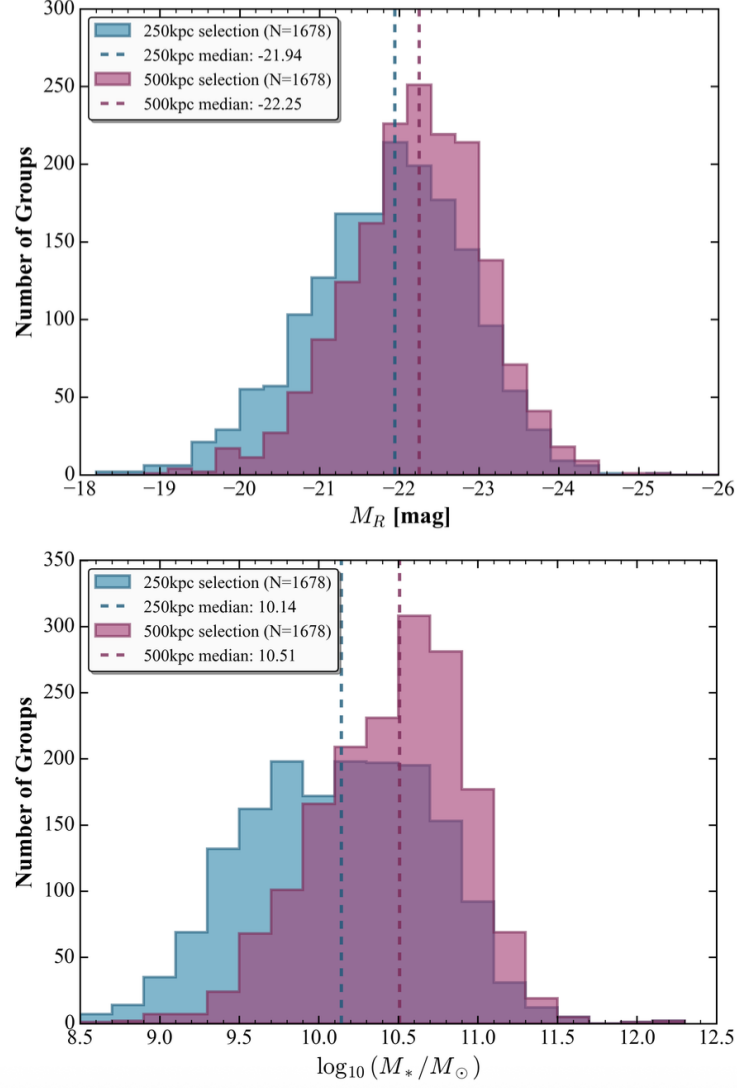


Figure B.1. Selection of BGGs in the COSMOS-Web galaxy group catalogue. The panels show histograms of the number of groups as a function of $\log_{10} M_*/M_\odot$ (bottom panel) and absolute R -band magnitude (M_R , top panel), comparing selections based on two different aperture sizes: 250 kpc (blue) and 500 kpc (magenta). Credits: [Gozaliasl et al. \(2025, including GT\)](#).

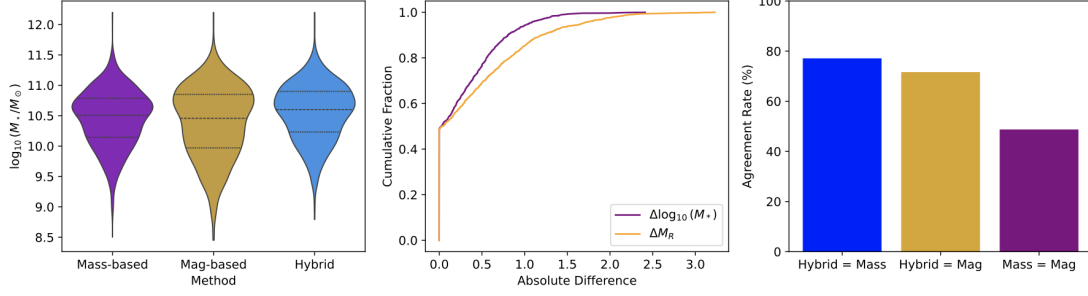


Figure B.2. Comparison of BGG selection methods. *Left:* Violin plots of stellar mass distributions for BGGs selected using mass-based, luminosity-based, and hybrid criteria. *Center:* Cumulative distribution functions (CDFs) of the absolute differences in stellar mass ($\Delta \log_{10}(M_*)$) and R-band magnitude (ΔM_R) between mass- and mag-selected BGGs, showing that $\sim 65.7\%$ of the sample has $\Delta \log_{10}(M_*) < 0.3$ dex, while $\sim 5.7\%$ exceeds 1 dex. *Right:* Agreement rates between selection methods. The hybrid BGG matches the mass-based and mag-based selections in 77.1% and 71.6% of groups, respectively, while the mass and mag selections agree in only 48.7% of cases. Credits: [Gozaliasl et al. \(2025, including GT\)](#).

To mitigate inconsistencies between these two approaches, we implemented the hybrid BGG selection strategy that combines stellar mass and luminosity ranking. The hybrid-selected BGG matches the mass-based BGG in 77.1% of the groups and the magnitude-based BGG in 71.6%, demonstrating improved overlap with both selection criteria. These results confirm that, while mass- and luminosity-based selections individually perform well in many cases, a hybrid method provides a more robust and consistent approach to BGG identification across a diverse range of environments.

A summary of these comparisons is presented in Figure B.2, which displays: (left) violin plots of the stellar mass distributions for each selection method, (center) cumulative distribution functions (CDFs) of absolute differences in stellar mass and *R*-band magnitude between mass- and magnitude-selected BGGs, and (right) a bar chart quantifying the agreement rates among the different selection approaches. The steep rise in the CDF at $\Delta = 0$ reflects the significant number of exact matches, while the bar chart highlights the strong alignment of the hybrid method with both individual strategies.

We construct the BGG catalogue by identifying the most massive and most luminous galaxy within each group with the hybrid method just described. Each group is characterised by its central coordinates (RA_{GR}, DEC_{GR}), redshift (Z_{GR}) from the COSMOS-Web group catalogue (Chapter 4), and an associated group radius defined as the hybrid radius derived from the detection algorithm.

The physical properties of the BGGs are obtained from the COSMOS2025 multi-band catalogue ([Shuntov et al., 2025a](#)). We use the unique galaxy identifier ID_{COSMOS2025_SE++} for matching.

The BGG catalogue we produced includes the following columns:

- **Group_ID**: Unique identifier of the galaxy group;
- **RA_GR, DEC_GR**: Right ascension and declination of the group centre (in degrees);
- **Z_GR**: Redshift of the group;
- **Hybrid_radius**: Estimated physical radius of the group (in arcminutes);
- **ID_COSMOS2025_SE++**: BGG source identifier from the *COSMOS2025 SE++* catalogue;
- **RA_DETEC, DEC_DETEC**: Right ascension and declination of the detected BGG (in degrees);
- **logM**: Stellar mass of the BGG, in units of $\log_{10}(M_{\odot})$;
- **M_R**: Absolute rest-frame magnitude in the *R* band;
- **flag_type**: Quality or classification flag indicating the BGG selection type or ambiguity in group membership.

This catalogue serves as the basis for our subsequent structural and star-forming analyses of BGGs across cosmic time, and has already been used to study the evolution and characteristics of the size–mass relation for BGGs (Gozaliasl et al., 2025, including GT).

Evolution of the BGG size–mass relation since $z = 3.7$

At the highest stellar masses, galaxies are predominantly central systems in group-scale haloes, exhibiting quenching signatures shaped by both internal stellar evolution and environmentally driven hierarchical assembly (Peng et al., 2012; Weinmann et al., 2006; Dekel et al., 2009; Tacchella et al., 2015; Carollo et al., 2016; van der Burg et al., 2014). Their effective radii increase by factors of ~ 2 – 3 from $z = 2$ to $z = 0$, with gas-poor (dry) mergers dominating size growth above $\log(M_{\star}/M_{\odot}) \sim 11$, while below this mass, the observed growth mainly results from the continuous addition of newly quenched, larger galaxies (Williams et al., 2010; Belli et al., 2015; Toft et al., 2007; Carollo et al., 2013; Cassata et al., 2013; Saracco et al., 2014). Stellar age differences between compact (older) and extended (younger) quiescent galaxies at fixed mass further support this dual evolutionary channel (Saracco et al., 2011; Belli et al., 2015; Fagioli et al., 2016). Distinct quenching mechanisms operate across the mass spectrum: gradual gas depletion preserves disk morphology, whereas rapid quenching through gas-rich mergers or AGN feedback induces central compaction and brief starburst phases on ~ 100 – 200 Myr timescales (Barro et al., 2013; Zolotov et al., 2015; Dubois et al., 2015; Choi et al., 2015, 2017). Leveraging deep JWST/NIRCam imaging,

this study quantifies the size–mass relation of brightest group galaxies (BGGs) up to $z = 4$, probing their structural evolution during the critical epoch from peak star formation to quenching (Yang et al., 2025). Consistent with COSMOS-Web results, star-forming galaxies maintain nearly constant structural scaling relations, while quiescent systems show a compactness threshold at $\log \Sigma_* \sim 9.5\text{--}10 M_\odot \text{ kpc}^{-2}$ (Yang et al., 2025).

Complementary analyses indicate that massive quiescent galaxies ($\log M_*/M_\odot > 11.4$) experience rapid quenching and compactification below $z \sim 2$, followed by post-quenching size growth dominated by dry minor mergers and possibly enhanced by AGN-induced “puffing up” or tidal interactions (Faisst et al., 2017).

This analysis has been performed on the COSMOS-Web group catalogue (Chapter 4), which provides the most comprehensive JWST-based view of BGG evolution across ~ 12 Gyr, revealing how structure, star formation efficiency, and environment jointly shape massive galaxy growth. To further ensure the robustness of our group and the selection of BGG, we restricted our analysis to groups with $\lambda_* > 4$ corresponding to groups with enough members to minimise contamination and projection effects, and applied a stellar mass threshold of $\log(M_*/M_\odot) > 9$ to ensure completeness across the redshift range.

Classification of BGGs: Star-forming and quiescent

To investigate the evolutionary pathways of BGGs, these were classified into star-forming (SFGs) and quiescent (QGs) systems using a combination of rest-frame colour–colour diagnostics and redshift-dependent sSFR thresholds. This dual approach provides a robust separation by accounting for both photometric colours and intrinsic star formation activity. Three classification methods were adopted: (i) the rest-frame NUVrJ diagram following Ilbert et al. (2013), (ii) a redshift-dependent sSFR threshold based on Pacifici et al. (2016), and (iii) a hybrid method that combines both criteria to ensure high purity and minimise contamination from dusty or transitional systems. The sSFR criterion defines quiescence as $\log_{10}(\text{sSFR}) < \log_{10}(0.2/t_{\text{obs}}(z))$, where $t_{\text{obs}}(z)$ is the cosmic time at redshift z . The hybrid classification requires galaxies to meet both criteria to be labelled quiescent and to fail both to be labelled star-forming, excluding ambiguous cases that satisfy only one.

Applying this combined scheme to the selected BGG sample across $0.08 < z < 3.7$ yields a clean and conservative division between SFGs and QGs. As shown in Figure B.3, SFGs dominate at high redshift ($z \gtrsim 2$), while the quiescent fraction rises steadily toward lower redshift, marking the build-up of massive, passive systems.

Structural measurements of BGGs

Structural properties of BGGs were quantified via the two-dimensional Sérsic profile fitting on JWST/NIRCam imaging by Yang et al. (2025). Each galaxy was modelled with a single-

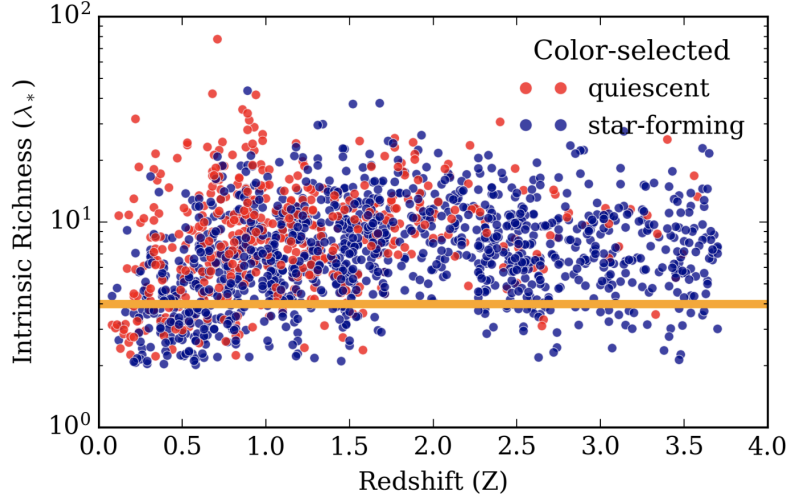


Figure B.3. The intrinsic richness, λ_* , for the sample of detected groups and its trend with redshift, BGGs classified as quiescent and star-forming are shown with red and blue points, respectively. The dashed horizontal orange line represents the level of λ_* limit we apply in this study. Credits: [Gozaliasl et al. \(2025, including GT\)](#).

component Sérsic function (Sersic, 1968) using the `Galight` software (Ding et al., 2020) built on the `Lenstronomy` framework (Birrer et al., 2021), which accounts for neighbouring sources and includes accurate PSF convolution. Empirical PSFs were constructed from COSMOS-Web star stacks (Yang et al., 2025), achieving $\sim 0.03''$ precision across all NIRCcam filters (F115W, F150W, F277W, F444W), and galaxies with poor fits or contamination were excluded. Structural parameters were validated against HST/ACS F814W measurements (Koekemoer et al., 2007), showing typical uncertainties of 0.1–0.2 dex in $\log R_e$ with no significant redshift-dependent bias. To ensure rest-frame optical consistency, a redshift-dependent filter selection was adopted, corresponding to $\sim 7000\text{--}11000 \text{ \AA}$ (F115W for $z < 0.4$, F150W for $0.4 < z < 1$, F277W for $1 < z < 3$, F444W for $3 < z < 4$).

Results

The structural evolution of COSMOS-Web BGGs has been investigated from $z \sim 4$ to the present, focusing on their size–mass relation, redshift evolution, and size distributions as a function of star formation state and stellar mass. Using the combined rest-frame colour and sSFR classification described above, it was found that star-forming BGGs and quiescent BGGs follow distinct but parallel size–mass relations. At fixed stellar mass, SFGs are systematically larger than QGs, with shallower slopes ($\alpha_{\text{SF}} \sim 0.1\text{--}0.2$) compared to QGs ($\alpha_{\text{QG}} \sim 0.4\text{--}0.6$), consistent with disk-versus-

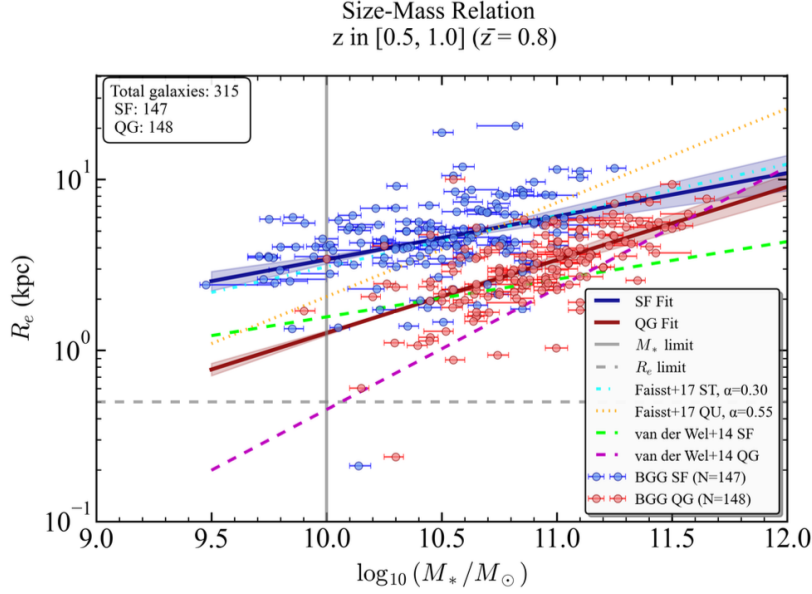


Figure B.4. Size–mass relation of BGGs classified using the consensus (color+sSFR) method, here for the redshift bin $z = 0.5 - 1.0$. Blue circles represent star-forming BGGs, and red circles show quiescent BGGs. The solid lines indicate best-fit power-law relations for SFGs and QGs, respectively, with shaded bands showing 1σ uncertainties. Cyan and orange dotted lines show comparison relations from Faisst et al. (2017), while lime and magenta lines show van der Wel et al. (2014) results. Credits: Gozaliasl et al. (2025, including GT).

spheroid-dominated morphologies (see Figure B.4). The normalisation of the relation decreases with redshift, and BGGs appear ~ 0.1 – 0.2 dex smaller than field galaxies at fixed mass, suggesting accelerated compacting or suppressed late-time accretion in dense group environments (cf. van der Wel et al., 2014; Faisst et al., 2017).

The slope and intrinsic scatter of the size–mass relation evolve systematically with time, as seen in Figure B.5. Quiescent BGGs exhibit steeper slopes and a stronger redshift dependence, implying mass-dependent size growth likely driven by dry mergers, whereas star-forming systems evolve more uniformly, consistent with secular disk growth. The intrinsic scatter $\sigma(\log R_e)$ increases toward high redshift (up to ~ 0.4 dex at $z > 2$), reflecting enhanced structural diversity in early epochs. At fixed stellar mass ($\log M_*/M_\odot = 10.7$), both populations experience significant size growth following $R_e \propto (1+z)^{-\alpha}$, with $\alpha_{\text{SF}} \approx 1.0$ and $\alpha_{\text{QG}} \approx 1.2$, corresponding to size growth factors of ~ 4.3 and ~ 6.6 , respectively, since $z \sim 3.5$.

The one-dimensional size distributions further reveal that QGs are consistently more compact and less scattered ($\sigma \sim 0.25$ dex) than SFGs ($\sigma \sim 0.3$ – 0.35 dex), which show extended, right-skewed tails toward large sizes. When split by stellar mass, low-mass BGGs ($\log M_*/M_\odot < 10.75$) display broader and more asymmetric size distributions than massive ones, while high-mass quiescents

dominate the compact end at $z < 1.5$. These results indicate that the structural evolution of BGGs is strongly shaped by environment: quiescent centrals undergo early compacting followed by merger-driven size growth, whereas star-forming centrals evolve more gradually through extended disk building and gas accretion.

This analysis presented the first comprehensive measurement of BGG structural evolution over $0.08 < z < 3.7$ using COSMOS-Web NIRCam data. BGGs were found to follow global size–mass trends (e.g., [van der Wel et al., 2014](#); [Faisst et al., 2017](#)) but remain systematically more compact than field galaxies at fixed stellar mass, especially among quiescent systems. The steeper size evolution of quiescent BGGs compared to star-forming ones supports a two-phase scenario in which early compaction and quenching are followed by merger-driven size growth in dense environments. This environmental dependence likely reflects earlier assembly, reduced late-time gas accretion, and enhanced minor merger activity in group-scale haloes. Together, these results demonstrate that BGGs are key laboratories for studying environmentally regulated galaxy evolution and provide critical constraints for models linking quenching, morphology, and halo growth across ~ 12 Gyr of cosmic history.

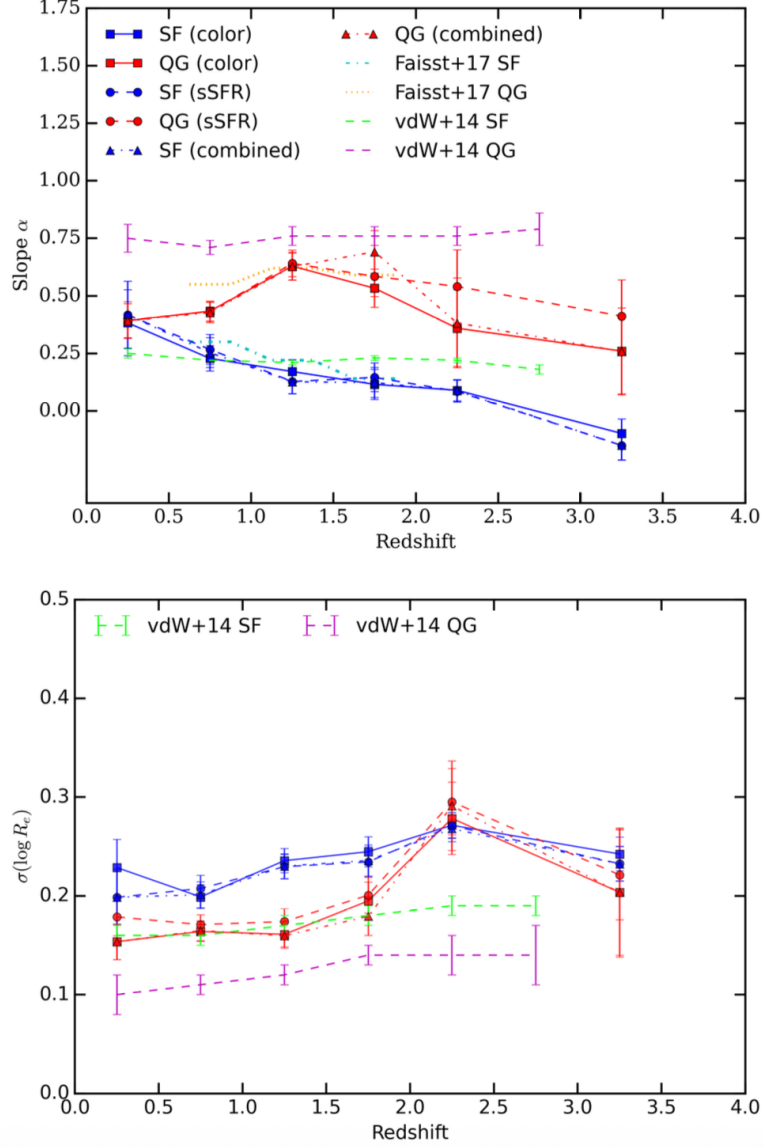
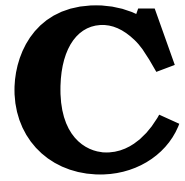


Figure B.5. Redshift evolution of the size–mass relation slope (top) and intrinsic scatter (bottom) for BGGs, classified as star-forming (blue) and quiescent (red) using three methods: color–color (solid lines with squares), specific star formation rate (sSFR; dashed lines with circles), and a combined color+sSFR criterion (dash-dotted lines with triangles). Quiescent BGGs display consistently steeper slopes, especially at $z \sim 1-3$, while star-forming BGGs maintain nearly flat relations across redshift. Credits: [Gozaliasl et al. \(2025, including GT\)](#).



Clusters as first seen by Euclid

This Appendix is fully based on Euclid Collaboration: Bhargava et al. (2025, including GT), accepted as part of the A&A Special Issue "Euclid Quick Data Release (Q1)".

The Euclid mission (Laureijs et al., 2011) is a European Space Agency survey designed to map the geometry of the Universe through optical and NIR imaging and spectroscopy. Euclid is expected to observe over 1.5 billion galaxies up to $z = 2$ over about $14,000 \text{ deg}^2$ (Euclid Collaboration: Mellier et al., 2025) and to detect and characterise around 10^5 clusters and groups (Sartoris et al., 2016; Euclid Collaboration: Adam et al., 2019).

The work outlined in Euclid Collaboration: Bhargava et al. (2025, including GT) presents the validation of Euclid's galaxy cluster detection pipeline, applied to the first real Euclid data from the Quick Data Release 1 (Q1)¹—a non-contiguous 63 deg^2 region in the Euclid Deep Fields observed to the nominal Wide-Survey depth. This dataset is a technical validation dataset, not suitable yet for cosmological inference, but crucial to assess the performance of the detection algorithms (AMICO and PZWav), evaluate photometric-redshift accuracy, input catalogue cleaning and quality cuts, and compare Euclid-detected clusters with known systems from X-ray, SZ, and optical surveys. This analysis aims to provide a training ground in preparation to the forthcoming Data Release 1 (DR1), for the delivery of the first statistically meaningful cluster samples for cosmological analyses.

¹<https://www.euclid-ec.org/science/q1/>

Input data preparation

Euclid’s Q1 provides the first public release with available imaging data obtained with the VIS and NISP instruments. It covers three non-contiguous regions in the Euclid Deep Fields: the Euclid Deep Field North (EDFN), the Euclid Deep Field Fornax (EDFF), and the Euclid Deep Field South (EDFS).

Galaxy selection was based on the outputs of the OU-MER and OU-LE3 processing units of the Euclid Science Ground Segment (Euclid Collaboration: Aussel et al., 2025). Euclid data are complemented by ancillary ground-based optical photometry from overlapping surveys such as UNIONS, DES, and HSC (Gwyn et al., 2025; Abbott et al., 2021; Aihara et al., 2022). A series of quality filters and masks were applied to ensure a uniform input sample for cluster detection. Sources must be detected in the VIS image ($\text{VIS_DET} = 1$), unsaturated, covered in all bands, and classified as non-stellar with $\text{POINT_LIKE_PROB} \leq 0.95$. Regions affected by bright stars, diffraction spikes, or incomplete inter-band coverage were masked, and sources were limited to $H < 24$. The resulting number counts are consistent across the three patches, peaking near the adopted magnitude limit.

Photometric redshifts were computed using the template-fitting code Phosphoros (Euclid Collaboration: Tucci et al., 2025) and validated against spectroscopic redshifts. This yielded a stable scatter and a small bias below 5%. The photo- z distributions tend to underestimate the true redshift uncertainties by $\sim 65\%$ in the EDFN and $\sim 15\%$ in the EDFF. This underestimation can introduce spurious detections, motivating adjustments in the cluster detection algorithms to properly account for photo- z errors². Ongoing work on photometric bias correction and error calibration is expected to improve the robustness of redshift estimates in forthcoming releases (Euclid Collaboration: Tucci et al., 2025).

Cluster catalogue construction

The Euclid cluster workflow relies on two independent detection algorithms, AMICO and PZWav, as we have described in Section 2.5.

Both algorithms were applied independently to the Q1 dataset using identical galaxy selections and masks. The resulting preliminary cluster lists were cross-matched to produce a high-confidence joint catalogue up to $z = 1.5$, which is less than the planned Euclid cluster upper redshift bound due to the limits of this dataset in terms of selection and redshift quality. The following matching criterion was adopted: 1 Mpc and $dz=0.1(1+z)$, with S/N as ranking property in the matching priority. In total, 426 clusters were jointly detected by AMICO and PZWav in the Q1 fields. Figure C.1 shows an example of a candidate cluster within the Euclid VIS image (Euclid

²to this end the experience and insights gained in this Thesis work were essential for the correct handling of anomalous photo- z uncertainties (e.g. very peaked distributions) as it has been discussed for the cleaning of COSMOS data in Chapters 3 and 4.

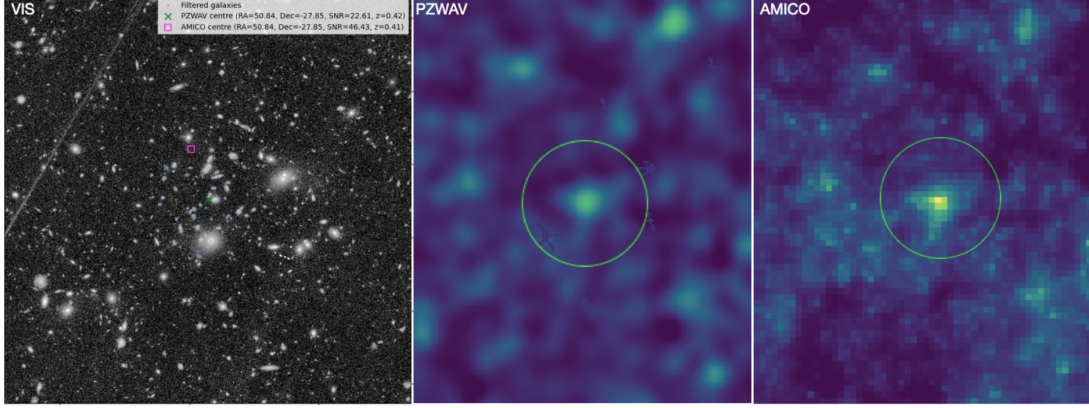


Figure C.1. *Left:* VIS zoom-in of $2' \times 2'$ at the location of a PZWav and AMICO-detected cluster, at redshift $z_p = 0.42$, located in the EDFF. The cluster has a Euclid richness of $\lambda = 56.1$. The middle and right panels show PZWav and AMICO amplitude maps, respectively, centred at the cluster position. The radius of the green circle is 6 arcminutes. Credits: Euclid Collaboration: Bhargava et al. (2025, including GT).

Collaboration: McCracken et al., 2025) that is detected in both the PZWav and AMICO density maps. Figure C.2 shows the comparison in terms of redshift and centre of matched detection as identified by AMICO and by PZWav.

Cluster richness in this study is estimated using RICH-CL, which assigns membership probabilities based on galaxy cluster-centric distance, H -band magnitude, and photometric redshift, following the methods of Castignani & Benoist (2016) and Andreon (2016). The richness is defined as the sum of the probabilities for galaxies brighter than $H_*(z_{\text{cluster}}) + 1.5$, where $H_*(z)$ is derived from the passive evolution of a burst galaxy model (Fioc & Rocca-Volmerange, 1997) and calibrated following Lin et al. (2006).

Six examples of clusters detected in Q1 in Euclid images are shown in Figure C.3, while Figure C.4 shows an example of a strong lens in one of the jointly detected systems.

Cross-matching with external catalogues

To validate the Q1 cluster detections, the joint catalogue was cross-correlated with a wide set of external meta-catalogues and individual surveys selected across multiple wavelengths. These include X-ray catalogues such as MCXC (Sadibekova et al., 2024) and eROSITA (Bulbul et al., 2024; Kluge et al., 2024); Sunyaev–Zeldovich (SZ) catalogues such as MCSZ (Tarrío et al., in preparation) and ComPRASS (Tarrío et al., 2019); and optical catalogues including the DES Y1 redMaPPer (Rykoff et al., 2016; Abbott et al., 2020) and the Abell catalogue (Abell, 1958; Abell et al., 1989). Additional compilations such as LC² (Sereno, 2015) and the MCCD (Euclid Collaboration: Melin et al., 2025) provided high-quality dynamical information. A two-way matching algorithm

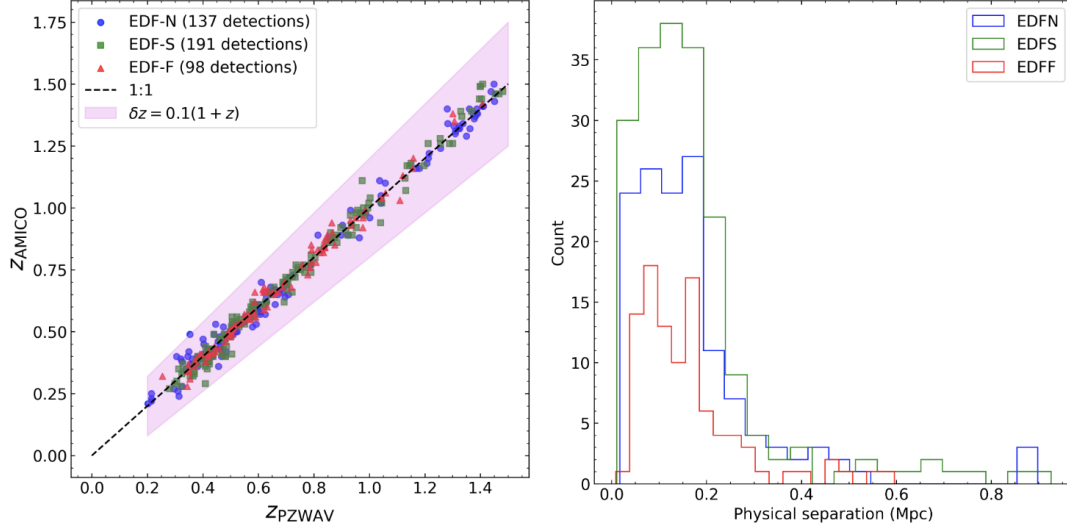


Figure C.2. *Left:* Redshift correlation between the photometric redshift estimate from the cluster finders AMICO and PZWav for all cluster detections in the Q1 region. *Right:* Euclid-based centring offset distribution between AMICO and PZWav centres for the catalogues in the three Q1 fields. Credits: [Euclid Collaboration: Bhargava et al. \(2025, including GT\)](#).

was implemented, identifying pairs of detections mutually recognised as nearest neighbours within thresholds in sky distance and redshift as presented in [Euclid Collaboration: Melin et al. \(2025\)](#).

Across the three fields, a total of 110 Euclid clusters were successfully matched to external catalogues, with the majority of associations arising from redMaPPer and eROSITA. The comparison of photometric redshifts in the EDFS shows a typical scatter of $\sigma_z \simeq 0.05(1+z)$, improving to $\sim 0.03(1+z)$ for $z > 0.3$, although Euclid photo- z uncertainties appear overestimated by factors of ~ 1.7 for PZWav and ~ 2.9 for AMICO. The positional offsets between Euclid and external detections peak below or near 200 kpc, consistent with expected differences in centring definitions and the positional accuracy of SZ and X-ray instruments.

Comparison between Euclid optical richness and external mass proxies (X-ray luminosity, temperature, or SZ flux) reveals a clear correlation (see Figure C.5), confirming that the Euclid richness is a reliable mass tracer. SZ signal from *Planck* legacy maps was also extracted across three richness bins. This comparative analysis has also suggested that unmatched Euclid systems might be gas-poor or not massive enough.

Among unmatched detections, most occupy the lower richness and mass range, reflecting the high S/N cut applied in Q1 and therefore the incomplete sampling of faint or low-mass systems. Nevertheless, cross-matching with NED-listed sources, MaDCoWS2 ([Thongkham et al., 2024](#)), and the catalogue by [Wen & Han \(2024\)](#) yielded 77 potential new Euclid clusters, 48 of

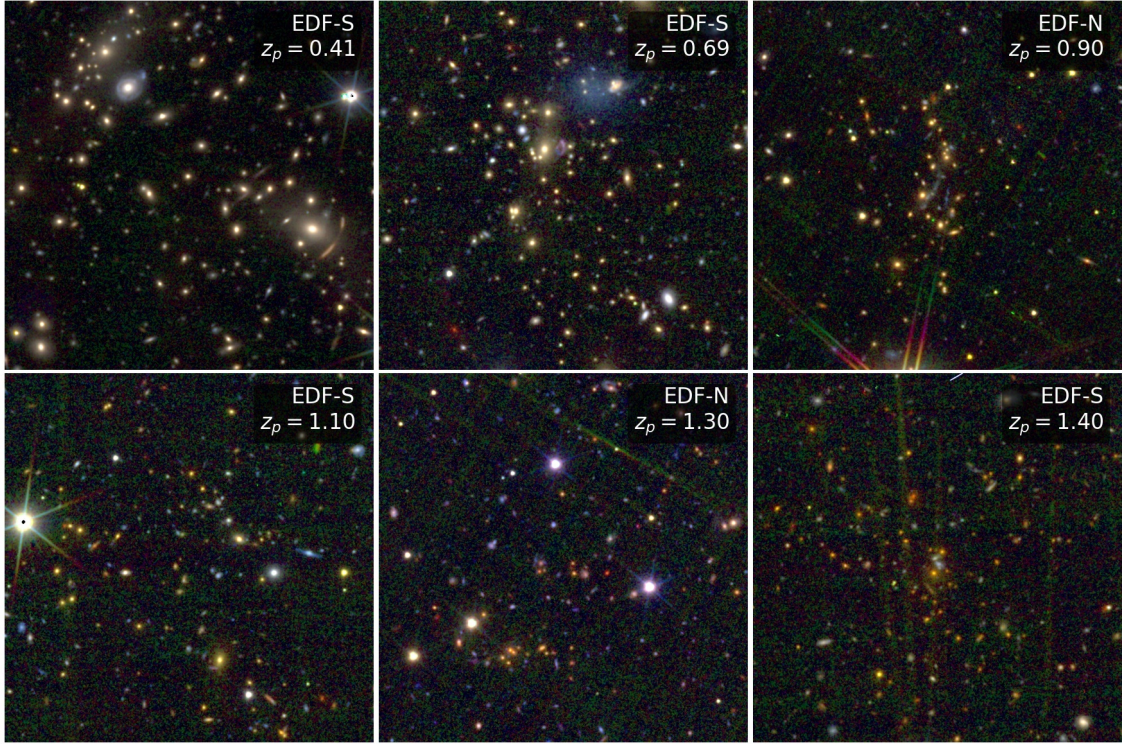


Figure C.3. VIS I and NISP Y- and H-band combined colour postage stamps of $2' \times 2'$ for six of the highest S/N PZWav- and AMICO-detected clusters in six selected redshift ranging from low to high redshift. Credits: [Euclid Collaboration: Bhargava et al. \(2025, including GT\)](#).

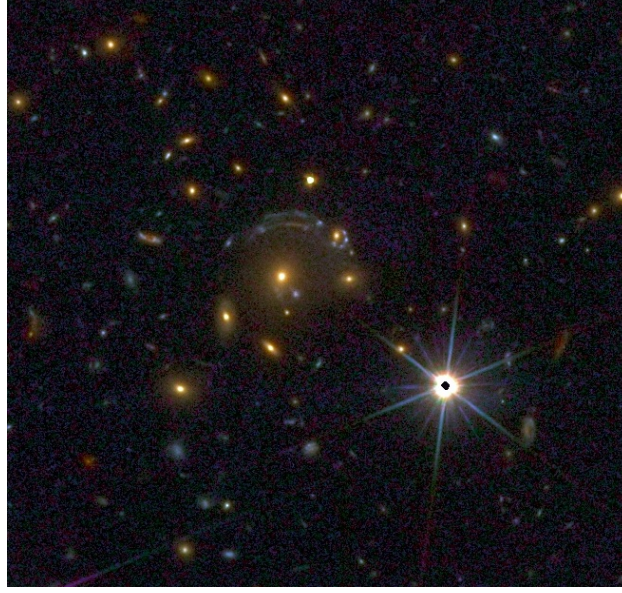


Figure C.4. Zoomed-in image of a strong lens in a PZWav and AMICO-detected cluster with a richness $\lambda = 58.7$ and redshift $z_p = 0.58$ in the EDFs. This cluster is also reported in the catalogue of giant arcs found in the Euclid Q1 data release (Euclid Collaboration: Bergamini et al., 2025). Credits: Euclid Collaboration: Bhargava et al. (2025, including GT).

which lie at $z \geq 1$. Visual inspections confirm that the majority correspond to genuine galaxy overdensities, though about 20 detections are likely spurious. Furthermore, 58 systems were identified at $1.5 \leq z \leq 2$, of which 15 show convincing evidence of real, high-redshift clusters. These candidates highlight Euclid’s capability to probe the distant Universe, where a substantial fraction of its cosmological constraining power is expected to come from (Sartoris et al., 2016).

External spectroscopic validation

The external spectroscopic validation was carried out in the EDFF and EDFN fields using publicly available redshift catalogues, as the EFDS field lacks sufficient spectroscopic coverage. For each Q1 detection, galaxies with secure external spectroscopic redshifts were selected within a cylindrical aperture of radius 1 Mpc and redshift window $0.1(1+z_d)$ centred on the PZWav cluster photometric redshift z_d . The search was then refined by applying an additional photometric-redshift filter of $0.04(1+z_d)$ to isolate the main redshift peak and reduce contamination along the line of sight.

The cluster spectroscopic redshift was computed using the biweight estimator after expanding the search radius to 2 Mpc and excluding galaxies with velocity offsets exceeding 3000 km s^{-1} . Reliable spectroscopic redshifts were obtained for 39 of the 98 clusters in the EDFF and 51 of the 137 clusters in the EDFN, with more than three spectroscopic members per system. Approximately two-thirds of these clusters have at least five confirmed members, yielding accurate redshift

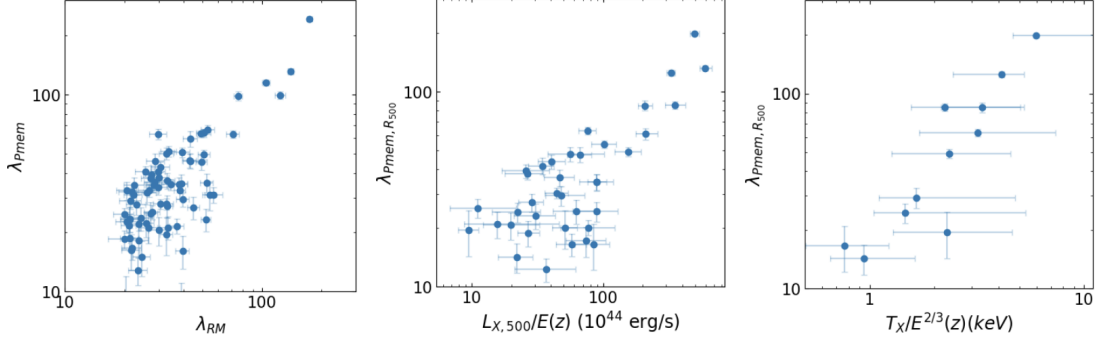


Figure C.5. Correlation of the Euclid richness λ_{pmem} against RedMaPPer richness (left), eROSITA mass proxies, the X-ray luminosity $L_{X,500}$ (middle) and X-ray temperature $T_{X,500}$ (right) for the matched systems within the EDFs. Credits: [Euclid Collaboration: Bhargava et al. \(2025, including GT\)](#).

estimates.

A comparison between the spectroscopic and photometric redshifts shows good overall agreement, with only mild systematic biases observed in the EDF for $0.5 < z < 0.9$ and in the EDFN below $z \approx 0.7$. These trends mirror those seen in the galaxy photometric redshift calibration. Supplementary checks with the internal Euclid OU-SPE spectroscopy confirmed the consistency of these results, although the limited number of available spectra prevents full spectroscopic confirmation.

Conclusions

We have validated the Euclid cluster detection workflow on real Q1 data, making it its first application beyond simulations such as the Flagship mock catalogues ([Euclid Collaboration: Castander et al., 2025](#)). After implementing optimised galaxy selection and rigorous photometric-quality validation, the AMICO and PZWav algorithms were run on the three Euclid Q1 fields, yielding a joint catalogue of 426 high- S/N detections within $0.2 \leq z_p \leq 1.5$. Each detection was characterised in terms of richness, colour-magnitude relation, centring offsets, and underwent visual inspection.

Cross-matching with external multi-wavelength catalogues confirmed 349 of these systems, while 77 remain as plausible new cluster candidates, validated through photometric redshift distributions and visual inspection. Among these, eight systems are spectroscopically confirmed with at least five members within 2 Mpc. The comparison of redshift and centring between the two independent finders shows excellent internal consistency. Moreover, positive correlations between Euclid richness and external mass proxies such as SZ total mass and X-ray luminosity and temperature confirm the physical reliability of the detections.

The current limitations mainly arise from residual issues in the galaxy catalogues—particularly for bright, low-redshift galaxies—and from the large photometric-redshift scatter above $z > 1.5$, both of which are expected to be solved in the upcoming DR1 release. Further improvements are expected in contamination rejection, masking procedures, and photometric calibration. Overall, these results establish the robustness of the Euclid cluster workflow, demonstrating its capability to deliver reliable optical cluster detections and laying the foundation for cosmological analyses with upcoming Euclid data releases.

D

Details on the ML training & on the cylinder method

This Appendix is fully based on Toni et al. (2025b), A&A, in press.

In this Appendix, we provide additional details about the selection and the labelling of the dataset chosen to train the machine learning model used for galaxy classification in Chapter 5, Section 5.2. Additionally, we discuss the choice of the parameters entering the computation of the quiescent fraction with the cylinder method, presented in Chapter 5, Section 5.3.2.

Additional details about the training-set choice

Here we provide additional details on the selection and use of the training set, highlighting its strengths and limitations in the context of the analysis presented in Chapter 5 of this Thesis. Firstly, we chose not to train our model directly on the target dataset because doing so would require extracting data from it and splitting it into training and testing subsets, thereby drastically reducing the statistical completeness of the final catalogue that we aim to fully analyse. Instead, we train the model on an independent labelled dataset (COSMOS2015) while applying it to the full target sample (COSMOS2025). Using real data from the same field ensures that the model is exposed to the true photometric uncertainties, systematics, blending issues, and observational conditions of our survey, which simulations might fail to fully reproduce.

COSMOS2025 lies entirely within the COSMOS2015 footprint, so some galaxies appear in both catalogues. However, the two surveys differ in instrument sets, depths, PSFs, and photometric methods, resulting in different galaxy characterisation, including different rest-frame magnitudes. From the ML perspective, what matters is overlap in feature space, not sky coordinates. We verified that no galaxy has identical or effectively duplicate feature vectors in both datasets, meaning there is no risk of data leakage or overfitting. COSMOS2015 thus provides an independent, realistic training set for our application to COSMOS2025.

The redshift range of the training set extends to $z = 6$, while the target set reaches $z \sim 13$. However, galaxies with $z > 6$ constitute only 0.6% of the target sample and are negligible for our analysis. The main gap in the training set we chose in our approach lies instead in the depth mismatch: the training set is limited to objects brighter than $H = 25$ mag, while the target set extends to 27.3 mag. This forces the model to extrapolate beyond its training domain, where faint objects have larger photometric uncertainties and somewhat broader colour distributions. Consequently, classification performance may degrade at the faint end, potentially increasing systematic misclassifications.

To assess the practical impact, we examined galaxies fainter than $H = 25$ mag in the classified target set, using the widely adopted NUVrJ diagram as a reference. We analysed (1) the distribution of predicted probabilities and (2) classification metrics in classifying quiescent galaxies¹ (precision, recall, F1 score) for faint versus bright objects. The probability distribution for faint galaxies was found to be broadly consistent with that of the brighter sample, with a slight increase in intermediate probabilities ($P_{\text{blue}} \sim 70\%$), reflecting model caution in ambiguous cases. Compared to the bright end, the faint sample shows increased precision ($\sim +10\%$) but substantially reduced recall ($\sim -30\%$), leading to a net modest decrease in the F1 score ($\sim -10\%$). This indicates that the classifier becomes more conservative at faint magnitudes: quiescence predictions are generally correct, but a larger fraction of quiescent galaxies are missed. In other words, the model favours purity over completeness in the unrepresented faint regime. We stress that this comparison measures consistency with NUVrJ rather than ground-truth validation, so it does not provide a direct correction for faint-end behaviour.

Large uncertainties are in general expected in this regime, given the larger photometric errors at the faint end, and our model’s conservative response is appropriate. Importantly, only $\sim 25\%$ of group member galaxies fall in this faint regime (reducing to $\sim 14\%$ when considering confident associations with $P_{\text{mem}} > 75\%$), limiting its impact on our main scientific goals. In conclusion, when deriving quiescent fractions, the faint end may be characterised by a selection that is purer but less complete, consistent with the intrinsic limitations of the data. Overall, the results indicate that the model effectively copes with fainter, noisier data, becoming more conservative, without

¹the performances for star-forming galaxies, which are the majority class, are not relevant here because they do not vary with magnitude, remaining stable above $\sim 96\%$.

compromising the analysis of galaxy population trends.

Additional details about the training-set labelling thresholds

Here we provide further information about the choice of labelling criteria and thresholds and their z -dependence.

The NUVrJ criterion classifies galaxies using rest-frame $NUV - r$ and $r - J$ colours, identifying as star-forming the galaxies with $NUV - r < 3(r - J) + 1$ or $NUV - r < 3.1$. The method was introduced by Ilbert et al. (2013) as an alternative to the very commonly adopted UVJ diagnostic. It was developed as an alternative with extended validity up to high redshift, given its larger dynamical range, lower sensitivity to uncertainties, and better sampling at $z > 2$. The NUVrJ criterion has been widely tested and employed for galaxy classification up to $z \sim 4$ and beyond (e.g. Ilbert et al., 2013; Weaver et al., 2023b; Gould et al., 2023; Shuntov et al., 2025a; Pearson et al., 2023).

Another variation that complements the NUVrJ criterion is the NUVrK threshold (Arnouts et al., 2013; Davidzon et al., 2016), which replaces the $r - J$ with the $r - K$ colour. Galaxies are classified as star-forming when $NUV - r < 1.37(r - K) + 2.6$ or $NUV - r < 3.15$ or $r - K > 1.3$. The NUVrK criterion, while with similar behavior and z -validity to the NUVrJ, combines the sensitivity of the $NUV - r$ color to the recent star formation history of the galaxy with the sensitivity of the $r - K$ color to stellar aging and dust attenuation, and helps also to better separate the influence of AGNs (e.g. Siudek et al., 2023; Moutard et al., 2020; Pearson et al., 2023).

As for what concerns the sSFR cut, this has been widely used in the literature by adopting different thresholds, generally between 10^{-10} yr^{-1} (e.g. Wu et al., 2018) and 10^{-11} yr^{-1} (e.g. Ilbert et al., 2013). We decided to adopt an intermediate fixed cut at $10^{-10.5} \text{ yr}^{-1}$ which is the same adopted in Euclid (Euclid Collaboration: Humphrey et al., 2023), following the work of Bisigello et al. (2020). However, it is true that the sSFR threshold to separate quiescent and star-forming galaxies is not uniform and can vary with cosmic time and with galaxy properties (e.g. Bisigello et al., 2020; Pacifici et al., 2016; Franx et al., 2008; Pearson et al., 2023; Wetzel et al., 2013).

In particular, Franx et al. (2008) have proposed a threshold proportional to the Hubble parameter, $H(z)$, where quiescent galaxies have $\text{sSFR} < 0.3H(z)$. In an analogous way, Pacifici et al. (2016) have more recently proposed a threshold dependent on the inverse of the age of the Universe, $t_U(z)$, with quiescent galaxies having $\text{sSFR} < 0.2/t_U(z)$. Figure D.1 shows the density of the selected COSMOS2015 star-forming (blue) and quiescent (red) galaxies according to the standard NUVrJ diagnostic in the sSFR vs. z diagram. The solid black line indicates the fixed threshold in sSFR we applied in this work, while the dashed lines mark the time-dependent threshold suggested by Pacifici et al. (2016), in green, and the one by Franx et al. (2008), in purple. The fixed and time-dependent thresholds diverge significantly at $z > 2$, when low-sSFR galaxies are, however, expected to be much rarer. The thresholds become more consistent with each other around the

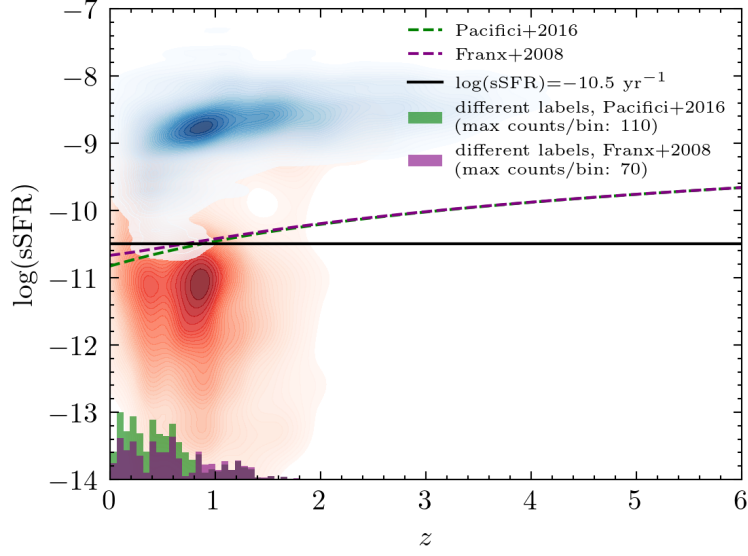


Figure D.1. Density of COSMOS2015 star-forming (blue) and quiescent (red) galaxies (according to NUVrJ classification) in the sSFR– z plane. The solid black line shows the fixed sSFR threshold used in this work; dashed lines indicate the time-dependent thresholds of [Pacifici et al. 2016](#) (green) and [Franx et al. 2008](#) (purple). Thresholds diverge at $z > 2$, but converge near $z \sim 0.7\text{--}0.9$, where most COSMOS galaxies lie. The bottom histograms show the redshift distribution of galaxies that would change classification under the two time-dependent thresholds, as reported in the legend.

crossing-point at $z \sim 0.7\text{--}0.9$ where most of the galaxies in COSMOS are (e.g. [Iovino et al., 2016](#)), as visible in the plot. At the bottom of the panel in [Figure D.1](#), we also show the distribution in redshift of all galaxies that would change their label if a time-dependent threshold was adopted. Counts for specific changes between the three classes of galaxies are also reported in [Table D.1](#). Two main conclusions can be drawn from these results: (1) the impact of a time-dependent sSFR is significant only below the crossing-points ($z \lesssim 0.7$) and affecting a small fraction of galaxies (overall $\sim 0.2\text{--}0.4\%$ of the galaxy sample); (2) this effect is mainly manifested in slightly reducing the number of quiescent galaxies ($\sim 1\text{--}2\%$), and the number of green valley galaxies ($\sim 1\text{--}8\%$) in favor of a few hundreds ($\sim 0.1\text{--}0.3\%$) of new examples of star-forming galaxies (especially when applying [Pacifici et al. \(2016\)](#)’s threshold). Given that the effect is limited in redshift, affecting only a small fraction of the galaxies, and that it reduces examples for the minority class (quiescent galaxies) and increases those for the majority class (star-forming galaxies) close to the decision boundary, it is not expected to drastically impact the learning process and performance. In the redshift range we analysed, the simple choice of $10^{-10.5} \text{ yr}^{-1}$ to separate the two populations is therefore well justified and closer to time-dependent relations with respect to, for instance a threshold at 10^{-11} yr^{-1} , which would largely overestimate the number of star-forming galaxies.

Finally, for the Sa-colour evolution diagram ([Andreon, 2006](#); [Radovich et al., 2020](#)), a few

Table D.1. Number of galaxies which would have been labelled differently if the threshold by [Franx et al. \(2008\)](#) / [Pacifci et al. \(2016\)](#) was considered instead of a fixed threshold at $10^{-10.5} \text{ yr}^{-1}$. Rows indicate labels assigned in this study, and columns indicate new labels that would have been assigned with time-dependent thresholds.

	Quiescent	Star-forming	Green
Quiescent	–	255 / 458	27 / 84
Star-forming	19 / 17	–	204 / 123
Green	42 / 26	229 / 427	–

more aspects have to be taken into account. The classification method is interesting because it uses apparent colours and not rest-frame colours. The use of an evolving Sa-model is what makes it redshift-consistent: each galaxy is classified as quiescent or star-forming, given the typical colour of a Sa-type galaxy (a spiral galaxy with properties close to those of lenticular and elliptical galaxies according to Hubble’s morphological classification scheme) at that specific redshift. However, the main limitation of this method is in the model-dependence itself and the fact that the chosen colour must bracket the $4000\text{\AA}$ break and therefore change accordingly as redshift increases.

In this study, we modelled the evolution of Sa-colours with the GALEV evolutionary synthesis interface ([Kotulla et al., 2009](#)), choosing typical Sa parameters and formation redshift. We decided to use this criterion only over a limited redshift range; that is, $0.38 \leq z \leq 1.4$. The lower bound is due to the absence of g-band photometry in the COSMOS2015 catalogue caused by poor seeing ([Laigle et al., 2016](#)), which would be the blue-bracket of the $4000\text{\AA}$ break at $z < 0.38$. The upper limit is instead given by the $4000\text{\AA}$ break shifting into the near-infrared. Given that the uncertainties of our model increase in this regime, we chose to be conservative and have the ML algorithms rely only on the other three methods at higher redshifts.

Figure D.2 shows the variation with redshift (upper panel), and with galaxy stellar mass (lower panel) of the percentage of classified quiescent galaxies over the total, for the four methods and for the final combined labelling we adopted (purple). As expected, even if quiescent galaxies are only 7% of the overall sample, the percentage over the total is very redshift- and mass-dependent. The panels also clearly illustrate the strong consistency between the colour-colour cuts and the sSFR threshold across all redshifts and stellar masses, while highlighting a noticeable overestimate in the number of quiescent galaxies derived from the Sa-colour method compared to the other approaches. The peak around $z \sim 0.8$ is likely influenced by a mismatch between our model and the observed colours when transitioning from the i to the z band. Apart from this effect, the number of quiescent galaxies remains systematically 15 – 20% higher across nearly all redshifts and masses – an excess over standard colour-colour cuts similar to that recently reported by [Asadi et al. \(2025\)](#). Interesting insights on this discrepancy could be gained by investigating a complete and spectroscopically characterised sample of green valley galaxies, which comprise most of the

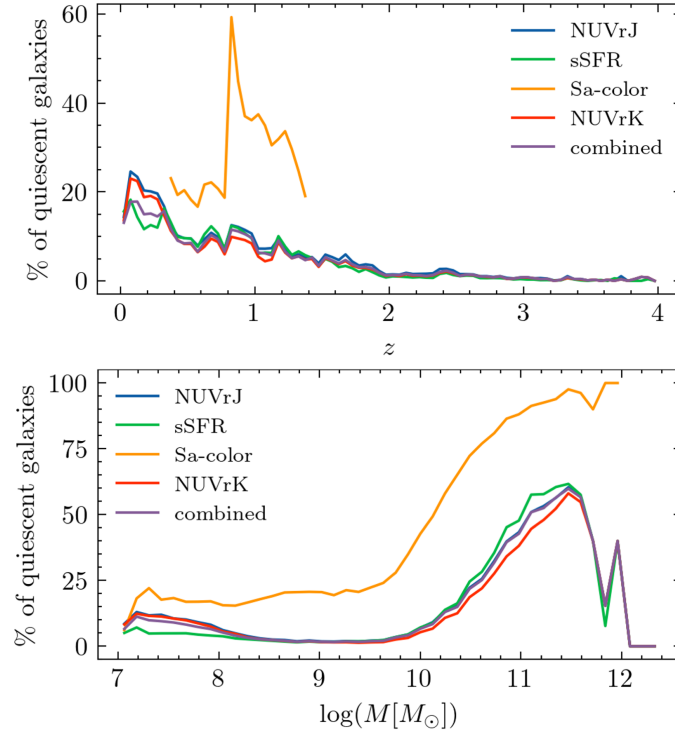


Figure D.2. Variation in bins of redshift (upper panel) and galaxy stellar mass (lower panel) of the percentage of classified quiescent galaxies over the total, for the NUVrJ colour-colour cut (blue), the sSFR threshold (green), the Sa-colour method (orange), and the NUVrK colour-colour cut (red). The purple line shows the final combined labelling we adopted for the training and testing sets in Chapter 5.

ambiguous and transitional cases in the classification.

In our specific application, as shown in Figure D.2, the resulting overall galaxy classification (purple line) is, by construction, driven by the consensus among methods and is therefore effectively constrained by the other three criteria. Nevertheless, the Sa-colour method provides valuable information in ambiguous or edge cases and is particularly helpful when UV photometry is missing. Training the ML models on a labelled dataset that integrates multiple independent and complementary methods, while basing the labelling on their consensus, is one of the key strengths of our approach.

Additional details about the cylinder method parameters

In Section 5.3.2, we have introduced a new method to characterise AMICO galaxy group members that (nearly) eliminates the model-dependence (and therefore the dependence on magnitude and position), while retaining the information of the redshift probability distribution. In short, this method exploits the concept of considering a volume around each detection and subtracting the density of field galaxies at the corresponding redshift to remove field contamination. Two key requirements come with this formalism: the appropriate choice of the cylinder size and the handling of negative densities, since these may arise from the group–field subtraction.

The choice of the depth of the cylinder is motivated by the typical photometric redshift uncertainty in the redshift range of interest. The depth is variable and redshift dependent, set to $\pm 0.01(1+z)$, and allows for a consistent sampling of members along the line of sight. Since galaxies are weighted by their redshift probability distribution, the exact choice of depth is generally less constraining and less impactful than the choice of the sky radius.

For choosing the cylinder sky radius, we started from an initial guess of $r = 0.5 \text{ Mpc}/h$, and tested both r and $2r$ as fixed cylinder radii. The differences in the resulting measurements and their richness dependence suggested that an adaptive radius, varying with the distribution of galaxies in each group – and therefore tracing its richness and spatial extent – would be a more robust and physically motivated choice. To determine this adaptive radius, we computed the RMS of all galaxies associated with a group, weighting each galaxy by its membership probability. The distribution of the resulting RMS radii and their relation to richness (λ_*) are shown in Fig. D.3. Although the top panel indicates that $0.5 \text{ Mpc}/h$ was a reasonable initial guess, the adaptive radius consistently captures the group region across different group sizes, as demonstrated by its trend with richness (bottom panel).

As mentioned above, the subtraction of average field-galaxy densities from the group densities within the cylinder can theoretically produce negative densities (see Eq. 5.4). In other words, in the ratio of Eq. 5.5: (1) the numerator, ρ_{red} , can be negative (the density of red galaxies is higher outside the cylinder than inside, i.e. a particularly blue group); (2) the term ρ_{blue} can be negative (the density of blue galaxies is higher outside the cylinder than inside, i.e. a particularly

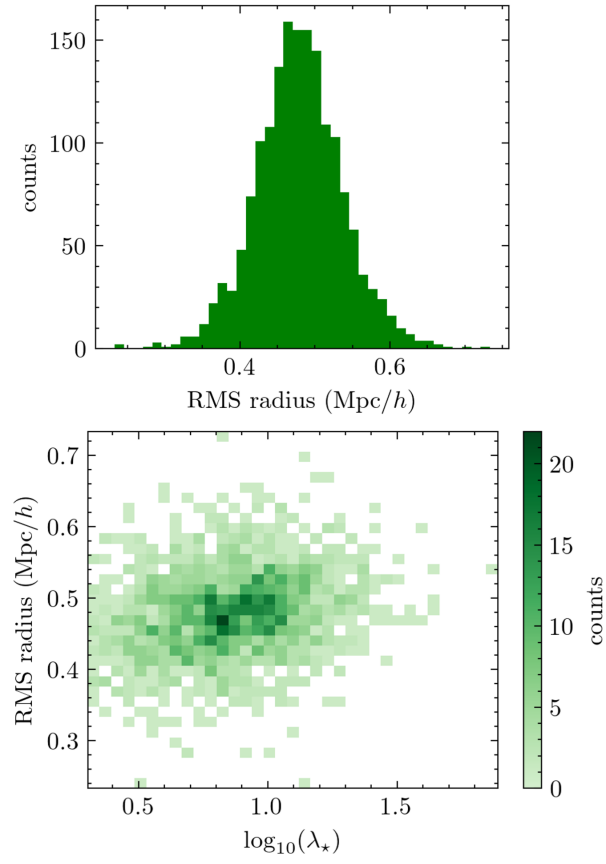


Figure D.3. Top panel: distribution of RMS radii in Mpc/h for the studied group sample. Bottom panel: trend of RMS radii with richness, $\log_{10}(\lambda_*)$, shown as counts per bin.

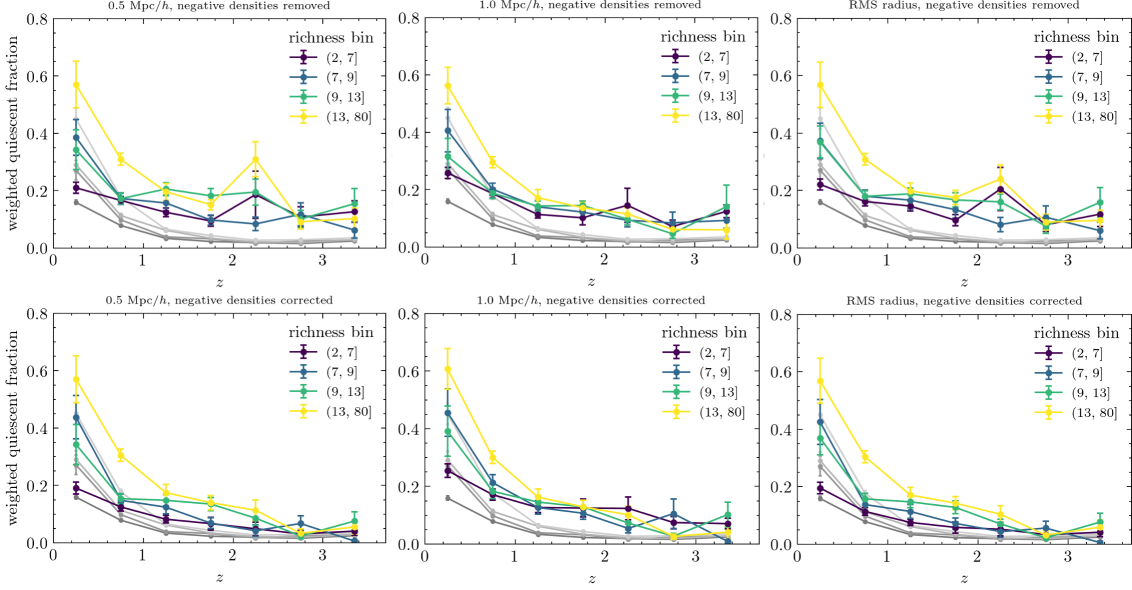


Figure D.4. Purity-weighted quiescent fraction as a function of redshift for different richness bins, using fixed radii of 0.5 Mpc/h (left), 1.0 Mpc/h (middle), and the RMS radius (right). In the top panels, we discard all systems with negative densities, while in the bottom panels, we apply the corrections described in the text. Grayscale trend lines are reported using the pure-membership method in the same bins, for comparison. The RMS-based measurement with corrected negative densities (bottom right panel) is the method adopted in Sect. 5.3.2 and shown in Fig. 5.5.

red group); (3) the denominator can be negative overall (the group appears as an underdensity relative to the global field, but overdense with respect to the local field).

Cases (1) and (2), affecting a total of about 800 groups, are physically meaningful. We therefore flagged them and included them in our analysis by replacing negative densities with close-to-zero positive constants to avoid divergences. Case (3) affects only 10 low-richness groups (when using the adaptive RMS radius). If real, such systems might correspond to overdensities embedded within locally underdense regions, such as filament edges or voids. However, these detections are excluded from the present analysis, as the cylinder-based estimate cannot be determined reliably.

Figure D.4 shows the weighted quiescent fraction as a function of redshift for different richness bins using 0.5 Mpc/h (left column), 1.0 Mpc/h (middle column), and the adaptive RMS radius (right column), comparing the results obtained by discarding all groups with any case of negative densities (top row) to those obtained after correcting negative densities in cases (1) and (2) as described above (bottom row). The RMS-based quiescent fraction with these corrections is the final method adopted for the results presented and discussed in Sect. 5.3.2 and shown in Fig. 5.5.

Bibliography

- Abbott, T. M. C., Agüena, M., Alarcon, A., et al. 2020, Physical Review D, 102, 023509, doi: [10.1103/PhysRevD.102.023509](https://doi.org/10.1103/PhysRevD.102.023509)
- Abbott, T. M. C., Adamów, M., Agüena, M., et al. 2021, Astrophysical Journal Supplement Series, 255, 20, doi: [10.3847/1538-4365/ac00b3](https://doi.org/10.3847/1538-4365/ac00b3)
- Abell, G. O. 1958, The Astrophysical Journal Supplement Series, 3, 211, doi: [10.1086/190036](https://doi.org/10.1086/190036)
- Abell, G. O., Corwin, Harold G., J., & Olowin, R. P. 1989, Astrophysical Journal Supplement Series, 70, 1, doi: [10.1086/191333](https://doi.org/10.1086/191333)
- Abril-Melgarejo, V., Epinat, B., Mercier, W., et al. 2021, Astronomy and Astrophysics, 647, A152, doi: [10.1051/0004-6361/202038818](https://doi.org/10.1051/0004-6361/202038818)
- Adami, C., & Mazure, A. 1999, Astronomy and Astrophysics Supplement Series, 134, 393, doi: [10.1051/aas:1999145](https://doi.org/10.1051/aas:1999145)
- Agüena, M., Benoist, C., da Costa, L. N., et al. 2021, Monthly Notices of the Royal Astronomical Society, 502, 4435, doi: [10.1093/mnras/stab264](https://doi.org/10.1093/mnras/stab264)
- Aguerri, J. A. L., & Zarattini, S. 2021, Universe, 7, 132, doi: [10.3390/universe7050132](https://doi.org/10.3390/universe7050132)
- Aguerri, J. A. L., Girardi, M., Boschin, W., et al. 2011, Astronomy and Astrophysics, 527, A143, doi: [10.1051/0004-6361/201015364](https://doi.org/10.1051/0004-6361/201015364)
- Aihara, H., AlSayyad, Y., Ando, M., et al. 2019, 71, 114, doi: [10.1093/pasj/psz103](https://doi.org/10.1093/pasj/psz103)
- . 2022, Publications of the Astronomical Society of Japan, 74, 247, doi: [10.1093/pasj/psab122](https://doi.org/10.1093/pasj/psab122)
- Alberts, S., Pope, A., Brodwin, M., et al. 2016, The Astrophysical Journal, 825, 72, doi: [10.3847/0004-637X/825/1/72](https://doi.org/10.3847/0004-637X/825/1/72)
- Allen, S. W., Evrard, A. E., & Mantz, A. B. 2011, Annual Review of Astronomy and Astrophysics, 49, 409, doi: [10.1146/annurev-astro-081710-102514](https://doi.org/10.1146/annurev-astro-081710-102514)

- Andreon, S. 2006, *Monthly Notices of the Royal Astronomical Society*, 369, 969, doi: [10.1111/j.1365-2966.2006.10353.x](https://doi.org/10.1111/j.1365-2966.2006.10353.x)
- Andreon, S. 2016, *Astronomy and Astrophysics*, 587, A158, doi: [10.1051/0004-6361/201526852](https://doi.org/10.1051/0004-6361/201526852)
- Andreon, S., & Huertas-Company, M. 2011, *Astronomy and Astrophysics*, 526, A11, doi: [10.1051/0004-6361/201015610](https://doi.org/10.1051/0004-6361/201015610)
- Andreon, S., Newman, A. B., Trinchieri, G., et al. 2014, *Astronomy and Astrophysics*, 565, A120, doi: [10.1051/0004-6361/201323077](https://doi.org/10.1051/0004-6361/201323077)
- Arango-Toro, R. C., Ilbert, O., Ciesla, L., et al. 2025, *Astronomy and Astrophysics*, 696, A159, doi: [10.1051/0004-6361/202452519](https://doi.org/10.1051/0004-6361/202452519)
- Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, *Monthly Notices of the Royal Astronomical Society*, 310, 540, doi: [10.1046/j.1365-8711.1999.02978.x](https://doi.org/10.1046/j.1365-8711.1999.02978.x)
- Arnouts, S., Moscardini, L., Vanzella, E., et al. 2002, *Monthly Notices of the Royal Astronomical Society*, 329, 355, doi: [10.1046/j.1365-8711.2002.04988.x](https://doi.org/10.1046/j.1365-8711.2002.04988.x)
- Arnouts, S., Le Floch, E., Chevallard, J., et al. 2013, *Astronomy and Astrophysics*, 558, A67, doi: [10.1051/0004-6361/201321768](https://doi.org/10.1051/0004-6361/201321768)
- Arribas, S., Perna, M., Rodríguez Del Pino, B., et al. 2024, *Astronomy and Astrophysics*, 688, A146, doi: [10.1051/0004-6361/202348824](https://doi.org/10.1051/0004-6361/202348824)
- Asadi, V., Chartab, N., Zonoozi, A. H., et al. 2025, *The Astrophysical Journal*, 993, 123, doi: [10.3847/1538-4357/ae0a2c](https://doi.org/10.3847/1538-4357/ae0a2c)
- Ascaso, B., Mei, S., & Benítez, N. 2015, *Monthly Notices of the Royal Astronomical Society*, 453, 2515, doi: [10.1093/mnras/stv1597](https://doi.org/10.1093/mnras/stv1597)
- Ata, M., Lee, K.-G., Vecchia, C. D., et al. 2022, *Nature Astronomy*, 6, 857, doi: [10.1038/s41550-022-01693-0](https://doi.org/10.1038/s41550-022-01693-0)
- Aymerich, G., Douspis, M., Pratt, G. W., et al. 2024, *Astronomy and Astrophysics*, 690, A238, doi: [10.1051/0004-6361/202449513](https://doi.org/10.1051/0004-6361/202449513)
- Bacon, R., Accardo, M., Adjali, L., et al. 2010, in *Society of photo-optical instrumentation engineers (SPIE) conference series*, Vol. 7735, *Ground-based and airborne instrumentation for astronomy III*, ed. I. S. McLean, S. K. Ramsay, & H. Takami, 773508, doi: [10.1117/12.856027](https://doi.org/10.1117/12.856027)
- Bahar, Y. E., Bulbul, E., Clerc, N., et al. 2022, *Astronomy and Astrophysics*, 661, A7, doi: [10.1051/0004-6361/202142462](https://doi.org/10.1051/0004-6361/202142462)

- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, Publications of the Astronomical Society of the Pacific, 93, 5, doi: [10.1086/130766](https://doi.org/10.1086/130766)
- Balogh, M. L., McGee, S. L., Mok, A., et al. 2014, 443, 2679, doi: [10.1093/mnras/stu1332](https://doi.org/10.1093/mnras/stu1332)
- Balogh, M. L., van der Burg, R. F. J., Muzzin, A., et al. 2021, Monthly Notices of the Royal Astronomical Society, 500, 358, doi: [10.1093/mnras/staa3008](https://doi.org/10.1093/mnras/staa3008)
- Banerjee, P., Szabo, T., Pierpaoli, E., et al. 2018, New Astronomy, 58, 61, doi: [10.1016/j.newast.2017.07.008](https://doi.org/10.1016/j.newast.2017.07.008)
- Barro, G., Faber, S. M., Pérez-González, P. G., et al. 2013, Astrophysical Journal, 765, 104, doi: [10.1088/0004-637X/765/2/104](https://doi.org/10.1088/0004-637X/765/2/104)
- Bartelmann, M. 1996, Arcs from a universal dark-matter halo profile., arXiv, doi: [10.48550/arXiv.astro-ph/9602053](https://doi.org/10.48550/arXiv.astro-ph/9602053)
- . 2010, Classical and Quantum Gravity, 27, 233001, doi: [10.1088/0264-9381/27/23/233001](https://doi.org/10.1088/0264-9381/27/23/233001)
- Bartelmann, M., & Schneider, P. 2001, Physics Reports, 340, 291, doi: [10.1016/S0370-1573\(00\)00082-X](https://doi.org/10.1016/S0370-1573(00)00082-X)
- Bauer, A. H., Baltay, C., Ellman, N., et al. 2012, The Astrophysical Journal, 749, 56, doi: [10.1088/0004-637X/749/1/56](https://doi.org/10.1088/0004-637X/749/1/56)
- Baxter, D. C., Cooper, M. C., Balogh, M. L., et al. 2023, Monthly Notices of the Royal Astronomical Society, 526, 3716, doi: [10.1093/mnras/stad2995](https://doi.org/10.1093/mnras/stad2995)
- Bellagamba, F., Maturi, M., Hamana, T., et al. 2011, Monthly Notices of the Royal Astronomical Society, 413, 1145, doi: [10.1111/j.1365-2966.2011.18202.x](https://doi.org/10.1111/j.1365-2966.2011.18202.x)
- Bellagamba, F., Roncarelli, M., Maturi, M., & Moscardini, L. 2018, Monthly Notices of the Royal Astronomical Society, 473, 5221, doi: [10.1093/mnras/stx2701](https://doi.org/10.1093/mnras/stx2701)
- Bellagamba, F., Sereno, M., Roncarelli, M., et al. 2019, Monthly Notices of the Royal Astronomical Society, 484, 1598, doi: [10.1093/mnras/stz090](https://doi.org/10.1093/mnras/stz090)
- Belli, S., Newman, A. B., & Ellis, R. S. 2015, Astrophysical Journal, 799, 206, doi: [10.1088/0004-637X/799/2/206](https://doi.org/10.1088/0004-637X/799/2/206)
- Benavides, J. A., Sales, L. V., & Abadi, M. G. 2020, Monthly Notices of the Royal Astronomical Society, 498, 3852, doi: [10.1093/mnras/staa2636](https://doi.org/10.1093/mnras/staa2636)
- Benitez, N., Dupke, R., Moles, M., et al. 2014, J-PAS: The Javalambre-Physics of the Accelerated Universe Astrophysical Survey, arXiv, doi: [10.48550/arXiv.1403.5237](https://doi.org/10.48550/arXiv.1403.5237)

- Benoist, C. 2014, 8. <https://ui.adsabs.harvard.edu/abs/2014beecs.confE...8B>
- Benítez, N. 2000, *The Astrophysical Journal*, 536, 571, doi: [10.1086/308947](https://doi.org/10.1086/308947)
- Bertin, E., & Arnouts, S. 1996, 117, 393, doi: [10.1051/aas:1996164](https://doi.org/10.1051/aas:1996164)
- Bertin, E., Schefer, M., Apostolakis, N., et al. 2022, *Astrophysics Source Code Library*, ascl:2212.018. <https://ui.adsabs.harvard.edu/abs/2022ascl.soft12018B>
- Bianconi, M., Smith, G. P., Haines, C. P., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 473, L79, doi: [10.1093/mnrasl/slx167](https://doi.org/10.1093/mnrasl/slx167)
- Birrer, S., Shajib, A., Gilman, D., et al. 2021, *The Journal of Open Source Software*, 6, 3283, doi: [10.21105/joss.03283](https://doi.org/10.21105/joss.03283)
- Bisigello, L., Kuchner, U., Conselice, C. J., et al. 2020, *Monthly Notices of the Royal Astronomical Society*, 494, 2337, doi: [10.1093/mnras/staa885](https://doi.org/10.1093/mnras/staa885)
- Blakeslee, J. P., Holden, B. P., Franx, M., et al. 2006, *The Astrophysical Journal*, Volume 644, Issue 1, pp. 30-53., 644, 30, doi: [10.1086/503539](https://doi.org/10.1086/503539)
- Bleem, L. E., Stalder, B., de Haan, T., et al. 2015, *The Astrophysical Journal Supplement Series*, 216, 27, doi: [10.1088/0067-0049/216/2/27](https://doi.org/10.1088/0067-0049/216/2/27)
- Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, *Annual Review of Astronomy and Astrophysics*, 51, 207, doi: [10.1146/annurev-astro-082812-140944](https://doi.org/10.1146/annurev-astro-082812-140944)
- Bond, J. R., Kofman, L., & Pogosyan, D. 1996, *Nature*, 380, 603, doi: [10.1038/380603a0](https://doi.org/10.1038/380603a0)
- Bonoli, S., Marín-Franch, A., Varela, J., et al. 2021, *Astronomy and Astrophysics*, 653, A31, doi: [10.1051/0004-6361/202038841](https://doi.org/10.1051/0004-6361/202038841)
- Boselli, A., & Gavazzi, G. 2006, *Publications of the Astronomical Society of the Pacific*, 118, 517, doi: [10.1086/500691](https://doi.org/10.1086/500691)
- Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, 686, 1503, doi: [10.1086/591786](https://doi.org/10.1086/591786)
- Brodwin, M., Stanford, S. A., Gonzalez, A. H., et al. 2013, *The Astrophysical Journal*, 779, 138, doi: [10.1088/0004-637X/779/2/138](https://doi.org/10.1088/0004-637X/779/2/138)
- Brunner, H., Liu, T., Lamer, G., et al. 2022, *Astronomy and Astrophysics*, 661, A1, doi: [10.1051/0004-6361/202141266](https://doi.org/10.1051/0004-6361/202141266)
- Bruzual, G., & Charlot, S. 2003, *Monthly Notices of the Royal Astronomical Society*, 344, 1000, doi: [10.1046/j.1365-8711.2003.06897.x](https://doi.org/10.1046/j.1365-8711.2003.06897.x)

- Budzynski, J. M., Koposov, S. E., McCarthy, I. G., McGee, S. L., & Belokurov, V. 2012, *Monthly Notices of the Royal Astronomical Society*, 423, 104, doi: [10.1111/j.1365-2966.2012.20663.x](https://doi.org/10.1111/j.1365-2966.2012.20663.x)
- Bulbul, E., Liu, A., Kluge, M., et al. 2024, *Astronomy and Astrophysics*, 685, A106, doi: [10.1051/0004-6361/202348264](https://doi.org/10.1051/0004-6361/202348264)
- Busillo, V., Covone, G., Sereno, M., et al. 2023, *Monthly Notices of the Royal Astronomical Society*, 524, 5050, doi: [10.1093/mnras/stad2190](https://doi.org/10.1093/mnras/stad2190)
- Böhringer, H., & Werner, N. 2010, *Astronomy and Astrophysics Review*, 18, 127, doi: [10.1007/s00159-009-0023-3](https://doi.org/10.1007/s00159-009-0023-3)
- Calvi, R., Castignani, G., & Dannerbauer, H. 2023, *Astronomy and Astrophysics*, 678, A15, doi: [10.1051/0004-6361/202346200](https://doi.org/10.1051/0004-6361/202346200)
- Calzetti, D. 2013, in *Secular Evolution of Galaxies*, ed. J. Falcón-Barroso & J. H. Knapen, 419, doi: [10.48550/arXiv.1208.2997](https://doi.org/10.48550/arXiv.1208.2997)
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *The Astrophysical Journal*, 533, 682, doi: [10.1086/308692](https://doi.org/10.1086/308692)
- Calzetti, D., Kennicutt, R. C., Adamo, A., et al. 2025, *Astrophysical Journal*, 991, 198, doi: [10.3847/1538-4357/adfbe0](https://doi.org/10.3847/1538-4357/adfbe0)
- Capak, P., Aussel, H., Ajiki, M., et al. 2007, *The Astrophysical Journal Supplement Series*, 172, 99, doi: [10.1086/519081](https://doi.org/10.1086/519081)
- Capak, P. L., Teplitz, H. I., Brooke, T. Y., Laher, R., & Science Center, S. 2013, in *American Astronomical Society Meeting Abstracts*, Vol. 221, 340.06
- Cappelluti, N., Brusa, M., Hasinger, G., et al. 2009, 497, 635, doi: [10.1051/0004-6361/200810794](https://doi.org/10.1051/0004-6361/200810794)
- Carollo, C. M., Cibinel, A., Lilly, S. J., et al. 2013, *Astrophysical Journal*, 776, 71, doi: [10.1088/0004-637X/776/2/71](https://doi.org/10.1088/0004-637X/776/2/71)
- . 2016, *Astrophysical Journal*, 818, 180, doi: [10.3847/0004-637X/818/2/180](https://doi.org/10.3847/0004-637X/818/2/180)
- Casey, C., Kartaltepe, J., & Cosmos-Web. 2022, 240, 203.02. <https://ui.adsabs.harvard.edu/abs/2022AAS...24020302C>
- Casey, C. M., Scoville, N. Z., Sanders, D. B., et al. 2014, *The Astrophysical Journal*, 796, 95, doi: [10.1088/0004-637X/796/2/95](https://doi.org/10.1088/0004-637X/796/2/95)
- Casey, C. M., Cooray, A., Capak, P., et al. 2015, *The Astrophysical Journal*, 808, L33, doi: [10.1088/2041-8205/808/2/L33](https://doi.org/10.1088/2041-8205/808/2/L33)

- Casey, C. M., Kartaltepe, J. S., Drakos, N. E., et al. 2023, *The Astrophysical Journal*, 954, 31, doi: [10.3847/1538-4357/acc2bc](https://doi.org/10.3847/1538-4357/acc2bc)
- Cassata, P., Giavalisco, M., Williams, C. C., et al. 2013, *Astrophysical Journal*, 775, 106, doi: [10.1088/0004-637X/775/2/106](https://doi.org/10.1088/0004-637X/775/2/106)
- Castignani, G., & Benoist, C. 2016, *Astronomy and Astrophysics*, 595, A111, doi: [10.1051/0004-6361/201528009](https://doi.org/10.1051/0004-6361/201528009)
- Castignani, G., Chiaberge, M., Celotti, A., & Norman, C. 2014a, *The Astrophysical Journal*, 792, 113, doi: [10.1088/0004-637X/792/2/113](https://doi.org/10.1088/0004-637X/792/2/113)
- Castignani, G., Chiaberge, M., Celotti, A., Norman, C., & De Zotti, G. 2014b, *The Astrophysical Journal*, 792, 114, doi: [10.1088/0004-637X/792/2/114](https://doi.org/10.1088/0004-637X/792/2/114)
- Castignani, G., Combes, F., & Salomé, P. 2020a, *Astronomy and Astrophysics*, 635, L10, doi: [10.1051/0004-6361/201937155](https://doi.org/10.1051/0004-6361/201937155)
- Castignani, G., Combes, F., Salomé, P., et al. 2019, *Astronomy and Astrophysics*, 623, A48, doi: [10.1051/0004-6361/201834066](https://doi.org/10.1051/0004-6361/201834066)
- Castignani, G., Combes, F., Salomé, P., Edge, A., & Jablonka, P. 2025, *Astronomy and Astrophysics*, 700, A197, doi: [10.1051/0004-6361/202453289](https://doi.org/10.1051/0004-6361/202453289)
- Castignani, G., Combes, F., Salomé, P., & Freundlich, J. 2020b, *Astronomy and Astrophysics*, 635, A32, doi: [10.1051/0004-6361/201936148](https://doi.org/10.1051/0004-6361/201936148)
- Castignani, G., Pandey-Pommier, M., Hamer, S. L., et al. 2020c, *Astronomy and Astrophysics*, 640, A65, doi: [10.1051/0004-6361/202038081](https://doi.org/10.1051/0004-6361/202038081)
- Castignani, G., Jablonka, P., Combes, F., et al. 2020d, *Astronomy and Astrophysics*, 640, A64, doi: [10.1051/0004-6361/201937190](https://doi.org/10.1051/0004-6361/201937190)
- Castignani, G., Radovich, M., Combes, F., et al. 2022a, *Astronomy and Astrophysics*, 667, A52, doi: [10.1051/0004-6361/202243689](https://doi.org/10.1051/0004-6361/202243689)
- Castignani, G., Combes, F., Jablonka, P., et al. 2022b, *Astronomy and Astrophysics*, 657, A9, doi: [10.1051/0004-6361/202040141](https://doi.org/10.1051/0004-6361/202040141)
- Castignani, G., Radovich, M., Combes, F., et al. 2023, *Astronomy and Astrophysics*, 672, A139, doi: [10.1051/0004-6361/202245380](https://doi.org/10.1051/0004-6361/202245380)
- Cautun, M., van de Weygaert, R., Jones, B. J. T., & Frenk, C. S. 2014, *Monthly Notices of the Royal Astronomical Society*, 441, 2923, doi: [10.1093/mnras/stu768](https://doi.org/10.1093/mnras/stu768)

- Cautun, M. C., & van de Weygaert, R. 2011, The DTFE public software: The Delaunay Tessellation Field Estimator code, arXiv, doi: [10.48550/arXiv.1105.0370](https://doi.org/10.48550/arXiv.1105.0370)
- Cerulo, P., Couch, W. J., Lidman, C., et al. 2016, Monthly Notices of the Royal Astronomical Society, Volume 457, Issue 2, p.2209-2235, 457, 2209, doi: [10.1093/mnras/stw080](https://doi.org/10.1093/mnras/stw080)
- Chabrier, G. 2003, Publications of the Astronomical Society of the Pacific, 115, 763, doi: [10.1086/376392](https://doi.org/10.1086/376392)
- Chapman, S. C., Hill, R., Aravena, M., et al. 2024, The Astrophysical Journal, 961, 120, doi: [10.3847/1538-4357/ad0b77](https://doi.org/10.3847/1538-4357/ad0b77)
- Chartab, N., Mobasher, B., Darvish, B., et al. 2020, The Astrophysical Journal, 890, 7, doi: [10.3847/1538-4357/ab61fd](https://doi.org/10.3847/1538-4357/ab61fd)
- Chary, R., & Elbaz, D. 2001, Astrophysical Journal, 556, 562, doi: [10.1086/321609](https://doi.org/10.1086/321609)
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. 2011, SMOTE: Synthetic Minority Over-sampling Technique, arXiv, doi: [10.48550/arXiv.1106.1813](https://doi.org/10.48550/arXiv.1106.1813)
- Chen, T., & Guestrin, C. 2016, XGBoost: A Scalable Tree Boosting System, doi: [10.48550/arXiv.1603.02754](https://doi.org/10.48550/arXiv.1603.02754)
- Chiang, Y.-K., Overzier, R., & Gebhardt, K. 2014, The Astrophysical Journal, 782, L3, doi: [10.1088/2041-8205/782/1/L3](https://doi.org/10.1088/2041-8205/782/1/L3)
- Chiang, Y.-K., Overzier, R. A., Gebhardt, K., et al. 2015, The Astrophysical Journal, 808, 37, doi: [10.1088/0004-637X/808/1/37](https://doi.org/10.1088/0004-637X/808/1/37)
- Choi, E., et al. 2015, ApJ, 813, 35, doi: [10.1088/0004-637X/813/1/35](https://doi.org/10.1088/0004-637X/813/1/35)
- . 2017, ApJ, 844, 31, doi: [10.3847/1538-4357/aa7a10](https://doi.org/10.3847/1538-4357/aa7a10)
- Chu, A., Sarron, F., Durret, F., & Márquez, I. 2022, Astronomy and Astrophysics, 666, A54, doi: [10.1051/0004-6361/202243504](https://doi.org/10.1051/0004-6361/202243504)
- Chéroutrier, D., Macías-Pérez, J. F., Désert, F. X., et al. 2025, Astronomy and Astrophysics, 700, A30, doi: [10.1051/0004-6361/202554450](https://doi.org/10.1051/0004-6361/202554450)
- Civano, F., Marchesi, S., Comastri, A., et al. 2016, The Astrophysical Journal, 819, 62, doi: [10.3847/0004-637X/819/1/62](https://doi.org/10.3847/0004-637X/819/1/62)
- Clerc, N., & Finoguenov, A. 2023, in Handbook of X-ray and Gamma-ray Astrophysics, 123, doi: [10.1007/978-981-16-4544-0_117-1](https://doi.org/10.1007/978-981-16-4544-0_117-1)

- Coenda, V., Muriel, H., & Martínez, H. J. 2012, *Astronomy and Astrophysics*, 543, A119, doi: [10.1051/0004-6361/201118318](https://doi.org/10.1051/0004-6361/201118318)
- Coil, A. L., Blanton, M. R., Burles, S. M., et al. 2011, 741, 8, doi: [10.1088/0004-637X/741/1/8](https://doi.org/10.1088/0004-637X/741/1/8)
- Comparat, J., Richard, J., Kneib, J.-P., et al. 2015, 575, A40, doi: [10.1051/0004-6361/201424767](https://doi.org/10.1051/0004-6361/201424767)
- Conroy, C. 2013, , 51, 393, doi: [10.1146/annurev-astro-082812-141017](https://doi.org/10.1146/annurev-astro-082812-141017)
- Cool, R. J., Moustakas, J., Blanton, M. R., et al. 2013, 767, 118, doi: [10.1088/0004-637X/767/2/118](https://doi.org/10.1088/0004-637X/767/2/118)
- Cossairt, A., Buehlmann, M., Kovacs, E., et al. 2022, eprint arXiv:2203.08768, arXiv:2203.08768, doi: [10.48550/arXiv.2203.08768](https://doi.org/10.48550/arXiv.2203.08768)
- Coupon, J. 2018, *Astrophysics Source Code Library*, ascl:1802.002. <https://ui.adsabs.harvard.edu/abs/2018ascl.soft02002C>
- Coupon, J., Czakon, N., Bosch, J., et al. 2018, *Publications of the Astronomical Society of Japan*, 70, S7, doi: [10.1093/pasj/psx047](https://doi.org/10.1093/pasj/psx047)
- Cucciati, O., Zamorani, G., Lemaux, B. C., et al. 2014, *Astronomy and Astrophysics*, 570, A16, doi: [10.1051/0004-6361/201423811](https://doi.org/10.1051/0004-6361/201423811)
- Cucciati, O., Lemaux, B. C., Zamorani, G., et al. 2018, *Astronomy and Astrophysics*, 619, A49, doi: [10.1051/0004-6361/201833655](https://doi.org/10.1051/0004-6361/201833655)
- Cutri, R. M., Wright, E. L., Conrow, T., et al. 2021, *VizieR Online Data Catalog: AllWISE Data Release*
- Daddi, E., Dannerbauer, H., Stern, D., et al. 2009, *The Astrophysical Journal*, 694, 1517, doi: [10.1088/0004-637X/694/2/1517](https://doi.org/10.1088/0004-637X/694/2/1517)
- Daddi, E., Valentino, F., Rich, R. M., et al. 2021, *Astronomy and Astrophysics*, 649, A78, doi: [10.1051/0004-6361/202038700](https://doi.org/10.1051/0004-6361/202038700)
- Daddi, E., Rich, R. M., Valentino, F., et al. 2022, *The Astrophysical Journal*, 926, L21, doi: [10.3847/2041-8213/ac531f](https://doi.org/10.3847/2041-8213/ac531f)
- Damjanov, I., Zahid, H. J., Geller, M. J., Fabricant, D. G., & Hwang, H. S. 2018, 234, 21, doi: [10.3847/1538-4365/aaa01c](https://doi.org/10.3847/1538-4365/aaa01c)
- Dannerbauer, H., Kurk, J. D., De Breuck, C., et al. 2014, *Astronomy and Astrophysics*, 570, A55, doi: [10.1051/0004-6361/201423771](https://doi.org/10.1051/0004-6361/201423771)

- Darvish, B., Sobral, D., Mobasher, B., et al. 2014, *The Astrophysical Journal*, 796, 51, doi: [10.1088/0004-637X/796/1/51](https://doi.org/10.1088/0004-637X/796/1/51)
- Darvish, B., Scoville, N. Z., Martin, C., et al. 2020, *The Astrophysical Journal*, 892, 8, doi: [10.3847/1538-4357/ab75c3](https://doi.org/10.3847/1538-4357/ab75c3)
- Davidzon, I., Cucciati, O., Bolzonella, M., et al. 2016, *Astronomy and Astrophysics*, 586, A23, doi: [10.1051/0004-6361/201527129](https://doi.org/10.1051/0004-6361/201527129)
- de Jong, J. T. A., Verdoes Kleijn, G. A., Erben, T., et al. 2017, *Astronomy and Astrophysics*, 604, A134, doi: [10.1051/0004-6361/201730747](https://doi.org/10.1051/0004-6361/201730747)
- De Lucia, G., & Blaizot, J. 2007, *Monthly Notices of the Royal Astronomical Society*, 375, 2, doi: [10.1111/j.1365-2966.2006.11287.x](https://doi.org/10.1111/j.1365-2966.2006.11287.x)
- De Lucia, G., Springel, V., White, S. D. M., Croton, D., & Kauffmann, G. 2006, 366, 499, doi: [10.1111/j.1365-2966.2005.09879.x](https://doi.org/10.1111/j.1365-2966.2005.09879.x)
- De Lucia, G., Weinmann, S., Poggianti, B. M., Aragón-Salamanca, A., & Zaritsky, D. 2012, *Monthly Notices of the Royal Astronomical Society*, 423, 1277, doi: [10.1111/j.1365-2966.2012.20983.x](https://doi.org/10.1111/j.1365-2966.2012.20983.x)
- De Lucia, G., Poggianti, B. M., Aragón-Salamanca, A., et al. 2007, *Monthly Notices of the Royal Astronomical Society*, 374, 809, doi: [10.1111/j.1365-2966.2006.11199.x](https://doi.org/10.1111/j.1365-2966.2006.11199.x)
- De Propris, R., Colless, M., Driver, S. P., et al. 2003, *Monthly Notices of the Royal Astronomical Society*, 342, 725, doi: [10.1046/j.1365-8711.2003.06510.x](https://doi.org/10.1046/j.1365-8711.2003.06510.x)
- Dekel, A., et al. 2009, *Nature*, 457, 451, doi: [10.1038/nature07648](https://doi.org/10.1038/nature07648)
- DESI Collaboration, Adame, A. G., Aguilar, J., et al. 2024, *The Astronomical Journal*, 168, 58, doi: [10.3847/1538-3881/ad3217](https://doi.org/10.3847/1538-3881/ad3217)
- DESI Collaboration, Abdul-Karim, M., Adame, A. G., et al. 2025, *Data Release 1 of the Dark Energy Spectroscopic Instrument*, arXiv, doi: [10.48550/arXiv.2503.14745](https://doi.org/10.48550/arXiv.2503.14745)
- Despali, G., Giocoli, C., Bonamigo, M., Limousin, M., & Tormen, G. 2017, *Monthly Notices of the Royal Astronomical Society*, 466, 181, doi: [10.1093/mnras/stw3129](https://doi.org/10.1093/mnras/stw3129)
- Despali, G., Giocoli, C., & Tormen, G. 2014, *Monthly Notices of the Royal Astronomical Society*, 443, 3208, doi: [10.1093/mnras/stu1393](https://doi.org/10.1093/mnras/stu1393)
- Dey, A., Schlegel, D. J., Lang, D., et al. 2019, *The Astronomical Journal*, 157, 168, doi: [10.3847/1538-3881/ab089d](https://doi.org/10.3847/1538-3881/ab089d)

- Diener, C., Lilly, S. J., Knobel, C., et al. 2013, *The Astrophysical Journal*, 765, 109, doi: [10.1088/0004-637X/765/2/109](https://doi.org/10.1088/0004-637X/765/2/109)
- Diener, C., Lilly, S. J., Ledoux, C., et al. 2015, *The Astrophysical Journal*, 802, 31, doi: [10.1088/0004-637X/802/1/31](https://doi.org/10.1088/0004-637X/802/1/31)
- Ding, X., Silverman, J., Treu, T., et al. 2020, *Astrophysical Journal*, 888, 37, doi: [10.3847/1538-4357/ab5b90](https://doi.org/10.3847/1538-4357/ab5b90)
- Dong, C., Lee, K.-G., Ata, M., Horowitz, B., & Momose, R. 2023, *The Astrophysical Journal*, 945, L28, doi: [10.3847/2041-8213/acba89](https://doi.org/10.3847/2041-8213/acba89)
- Dressler, A. 1980, *The Astrophysical Journal*, 236, 351, doi: [10.1086/157753](https://doi.org/10.1086/157753)
- Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, *Monthly Notices of the Royal Astronomical Society*, 413, 971, doi: [10.1111/j.1365-2966.2010.18188.x](https://doi.org/10.1111/j.1365-2966.2010.18188.x)
- Driver, S. P., Robotham, A. S. G., Obreschkow, D., et al. 2022, *Monthly Notices of the Royal Astronomical Society*, 515, 2138, doi: [10.1093/mnras/stac581](https://doi.org/10.1093/mnras/stac581)
- Dubois, Y., et al. 2015, *MNRAS*, 452, 1502, doi: [10.1093/mnras/stv1426](https://doi.org/10.1093/mnras/stv1426)
- Dunne, D. A., Webb, T. M. A., Noble, A., et al. 2021, *The Astrophysical Journal*, 909, L29, doi: [10.3847/2041-8213/abeb6f](https://doi.org/10.3847/2041-8213/abeb6f)
- Durret, F., Laganá, T. F., & Haider, M. 2011, *Astronomy and Astrophysics*, 529, A38, doi: [10.1051/0004-6361/201015978](https://doi.org/10.1051/0004-6361/201015978)
- Eckert, D., Gaspari, M., Gastaldello, F., Le Brun, A. M. C., & O’Sullivan, E. 2021, *Universe*, 7, 142, doi: [10.3390/universe7050142](https://doi.org/10.3390/universe7050142)
- Eckmiller, H. J., Hudson, D. S., & Reiprich, T. H. 2011, *Astronomy and Astrophysics*, 535, A105, doi: [10.1051/0004-6361/201116734](https://doi.org/10.1051/0004-6361/201116734)
- Edge, A., Sutherland, W., Kuijken, K., et al. 2013, *The Messenger*, 154, 32. <https://ui.adsabs.harvard.edu/abs/2013Msngr.154...32E>
- Edge, A. C. 2001, *Monthly Notices of the Royal Astronomical Society*, 328, 762, doi: [10.1046/j.1365-8711.2001.04802.x](https://doi.org/10.1046/j.1365-8711.2001.04802.x)
- Edward, A. H., Balogh, M. L., Bahé, Y. M., et al. 2024, *Monthly Notices of the Royal Astronomical Society*, 527, 8598, doi: [10.1093/mnras/stad3751](https://doi.org/10.1093/mnras/stad3751)
- Eigenthaler, P., & Zeilinger, W. W. 2009, *Astronomische Nachrichten*, 330, 978, doi: [10.1002/asna.200911274](https://doi.org/10.1002/asna.200911274)

- Einasto, J. 1965, Trudy Astrofizicheskogo Instituta Alma-Ata, 5, 87. <https://ui.adsabs.harvard.edu/abs/1965TrAlm...5...87E>
- Einasto, M., Einasto, J., Müller, V., Heinämäki, P., & Tucker, D. L. 2003, Astronomy and Astrophysics, 401, 851, doi: [10.1051/0004-6361:20021727](https://doi.org/10.1051/0004-6361:20021727)
- Einasto, M., Einasto, J., Tenjes, P., et al. 2024, Astronomy and Astrophysics, 681, A91, doi: [10.1051/0004-6361/202347504](https://doi.org/10.1051/0004-6361/202347504)
- Eisenhardt, P. R. M., Brodwin, M., Gonzalez, A. H., et al. 2008, The Astrophysical Journal, 684, 905, doi: [10.1086/590105](https://doi.org/10.1086/590105)
- Eke, V. R., Baugh, C. M., Cole, S., et al. 2004, Monthly Notices of the Royal Astronomical Society, 348, 866, doi: [10.1111/j.1365-2966.2004.07408.x](https://doi.org/10.1111/j.1365-2966.2004.07408.x)
- Elvis, M., Civano, F., Vignali, C., et al. 2009, 184, 158, doi: [10.1088/0067-0049/184/1/158](https://doi.org/10.1088/0067-0049/184/1/158)
- Epinat, B. 2021, in Extragalactic spectroscopic surveys: Past, present and future of galaxy evolution (GALSPEC2021), 77, doi: [10.5281/zenodo.4721886](https://doi.org/10.5281/zenodo.4721886)
- Epinat, B., Contini, T., Mercier, W., et al. 2024, Astronomy and Astrophysics, 683, A205, doi: [10.1051/0004-6361/202348038](https://doi.org/10.1051/0004-6361/202348038)
- Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. 1996, A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. <https://ui.adsabs.harvard.edu/abs/1996kddm.conf..226E>
- Euclid Collaboration: Adam, R., Vannier, M., et al. 2019, Astronomy and Astrophysics, 627, A23, doi: [10.1051/0004-6361/201935088](https://doi.org/10.1051/0004-6361/201935088)
- Euclid Collaboration: Aussel, H., Tereno, I., Schirmer, M., et al. 2025, A&A, submitted, Euclid Q1 SI, arXiv:2503.15302. <https://arxiv.org/abs/2503.15302>
- Euclid Collaboration: Bergamini, P., Meneghetti, M., Acebron, A., et al. 2025, A&A, submitted, Euclid Q1 SI, arXiv:2503.15330. <https://arxiv.org/abs/2503.15330>
- Euclid Collaboration: Bhargava, S., Benoist, C., et al. 2025, Euclid Quick Data Release (Q1). First detections from the galaxy cluster workflow, arXiv, doi: [10.48550/arXiv.2503.19196](https://doi.org/10.48550/arXiv.2503.19196)
- Euclid Collaboration: Böhringer, H., Chon, G., et al. 2025, Astronomy and Astrophysics, 693, A59, doi: [10.1051/0004-6361/202451683](https://doi.org/10.1051/0004-6361/202451683)
- Euclid Collaboration: Castander, F. J., Fosalba, P., et al. 2025, Astronomy and Astrophysics, 697, A5, doi: [10.1051/0004-6361/202450853](https://doi.org/10.1051/0004-6361/202450853)

- Euclid Collaboration: Humphrey, , A., Bisigello, L., et al. 2023, *Astronomy and Astrophysics*, 671, A99, doi: [10.1051/0004-6361/202244307](https://doi.org/10.1051/0004-6361/202244307)
- Euclid Collaboration: Ingoglia, , L., Sereno, M., et al. 2025, *Astronomy and Astrophysics*, 695, A280, doi: [10.1051/0004-6361/202452122](https://doi.org/10.1051/0004-6361/202452122)
- Euclid Collaboration: McCracken, H. J., Benson, K., Dolding, C., et al. 2025, *A&A*, submitted, Euclid Q1 SI, arXiv:2503.15303. <https://arxiv.org/abs/2503.15303>
- Euclid Collaboration: Melin, , J. B., Stanford, S. A., et al. 2025, Euclid preparation. Methodology for validating the Euclid Catalogue of Galaxy Clusters using external data, arXiv, doi: [10.48550/arXiv.2509.06805](https://doi.org/10.48550/arXiv.2509.06805)
- Euclid Collaboration: Mellier, , Y., Abdurro'uf, et al. 2025, *Astronomy and Astrophysics*, 697, A1, doi: [10.1051/0004-6361/202450810](https://doi.org/10.1051/0004-6361/202450810)
- Euclid Collaboration: Sereno, , M., Farrens, S., et al. 2024, *Astronomy and Astrophysics*, 689, A252, doi: [10.1051/0004-6361/202348680](https://doi.org/10.1051/0004-6361/202348680)
- Euclid Collaboration: Tucci, M., Paltani, S., Hartley, W. G., et al. 2025, *A&A*, submitted, Euclid Q1 SI, arXiv:2503.15306. <https://arxiv.org/abs/2503.15306>
- Fabian, A. C., Sanders, J. S., Ettori, S., et al. 2000, *Monthly Notices of the Royal Astronomical Society*, 318, L65, doi: [10.1046/j.1365-8711.2000.03904.x](https://doi.org/10.1046/j.1365-8711.2000.03904.x)
- Fagioli, M., Carollo, C. M., Renzini, A., et al. 2016, *Astrophysical Journal*, 831, 173, doi: [10.3847/0004-637X/831/2/173](https://doi.org/10.3847/0004-637X/831/2/173)
- Faisst, A. L., Carollo, C. M., Capak, P. L., et al. 2017, *Astrophysical Journal*, 839, 71, doi: [10.3847/1538-4357/aa697a](https://doi.org/10.3847/1538-4357/aa697a)
- Farrens, S., Abdalla, F. B., Cypriano, E. S., Sabiu, C., & Blake, C. 2011, *Monthly Notices of the Royal Astronomical Society*, 417, 1402, doi: [10.1111/j.1365-2966.2011.19356.x](https://doi.org/10.1111/j.1365-2966.2011.19356.x)
- Fassbender, R., Nastasi, A., Böhringer, H., et al. 2011, *Astronomy and Astrophysics*, 527, L10, doi: [10.1051/0004-6361/201016169](https://doi.org/10.1051/0004-6361/201016169)
- Figueira, M., Pollo, A., Malek, K., et al. 2022, *Astronomy and Astrophysics*, 667, A29, doi: [10.1051/0004-6361/202141701](https://doi.org/10.1051/0004-6361/202141701)
- Finoguenov, A., Guzzo, L., Hasinger, G., et al. 2007, *The Astrophysical Journal Supplement Series*, 172, 182, doi: [10.1086/516577](https://doi.org/10.1086/516577)
- Fioc, M., & Rocca-Volmerange, B. 1997, *Astronomy and Astrophysics*, 326, 950

- Fisher, R. A. 1936, *Annals of Eugenics*, 7, 179, doi: [10.1111/j.1469-1809.1936.tb02137.x](https://doi.org/10.1111/j.1469-1809.1936.tb02137.x)
- Fixsen, D. J. 2009, *The Astrophysical Journal*, 707, 916, doi: [10.1088/0004-637X/707/2/916](https://doi.org/10.1088/0004-637X/707/2/916)
- Fogarty, K., Postman, M., Li, Y., et al. 2019, *The Astrophysical Journal*, 879, 103, doi: [10.3847/1538-4357/ab22a4](https://doi.org/10.3847/1538-4357/ab22a4)
- Forrest, B., Lemaux, B. C., Shah, E., et al. 2023, *Monthly Notices of the Royal Astronomical Society*, 526, L56, doi: [10.1093/mnras/slad114](https://doi.org/10.1093/mnras/slad114)
- Franck, J. R., & McGaugh, S. S. 2016, *The Astrophysical Journal*, 833, 15, doi: [10.3847/0004-637X/833/1/15](https://doi.org/10.3847/0004-637X/833/1/15)
- Franco, M., Casey, C. M., Koekemoer, A. M., et al. 2025, COSMOS-Web: Comprehensive Data Reduction for Wide-Area JWST NIRCам Imaging, arXiv, doi: [10.48550/arXiv.2506.03256](https://doi.org/10.48550/arXiv.2506.03256)
- Franx, M., van Dokkum, P. G., Förster Schreiber, N. M., et al. 2008, *The Astrophysical Journal*, 688, 770, doi: [10.1086/592431](https://doi.org/10.1086/592431)
- Fraser-McKelvie, A., Brown, M. J. I., & Pimbblet, K. A. 2014, *Monthly Notices of the Royal Astronomical Society*, 444, L63, doi: [10.1093/mnras/flu117](https://doi.org/10.1093/mnras/flu117)
- Freyer, D. T., Sanders, D. B., Surace, J. A., et al. 2009, *Astronomical Journal*, 138, 1261, doi: [10.1088/0004-6256/138/5/1261](https://doi.org/10.1088/0004-6256/138/5/1261)
- Freundlich, J., Combes, F., Tacconi, L. J., et al. 2019, *Astronomy and Astrophysics*, 622, A105, doi: [10.1051/0004-6361/201732223](https://doi.org/10.1051/0004-6361/201732223)
- Friedman, J. H. 2000, ResearchGate, doi: [10.1214/aos/1013203451](https://doi.org/10.1214/aos/1013203451)
- Fritz, A., Scodreggio, M., Ilbert, O., et al. 2014, *Astronomy and Astrophysics*, 563, A92, doi: [10.1051/0004-6361/201322379](https://doi.org/10.1051/0004-6361/201322379)
- Fu, S., Zhang, Y., Avestruz, C., & Coronel, R. 2024, *Monthly Notices of the Royal Astronomical Society*, 530, 2251, doi: [10.1093/mnras/stae950](https://doi.org/10.1093/mnras/stae950)
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *Astronomy and Astrophysics*, 616, A1, doi: [10.1051/0004-6361/201833051](https://doi.org/10.1051/0004-6361/201833051)
- Gallazzi, A., Bell, E. F., Zibetti, S., Brinchmann, J., & Kelson, D. D. 2014, *The Astrophysical Journal*, 788, 72, doi: [10.1088/0004-637X/788/1/72](https://doi.org/10.1088/0004-637X/788/1/72)
- García-Vergara, C., Rybak, M., Hodge, J., et al. 2022, *The Astrophysical Journal*, 927, 65, doi: [10.3847/1538-4357/ac469d](https://doi.org/10.3847/1538-4357/ac469d)

- Gaspari, M., Brighenti, F., D’Ercole, A., & Melioli, C. 2011, *Monthly Notices of the Royal Astronomical Society*, 415, 1549, doi: [10.1111/j.1365-2966.2011.18806.x](https://doi.org/10.1111/j.1365-2966.2011.18806.x)
- Geach, J. E., Smail, I., Moran, S. M., et al. 2011, *The Astrophysical Journal*, 730, L19, doi: [10.1088/2041-8205/730/2/L19](https://doi.org/10.1088/2041-8205/730/2/L19)
- Geach, J. E., Sobral, D., Hickox, R. C., et al. 2012, *Monthly Notices of the Royal Astronomical Society*, 426, 679, doi: [10.1111/j.1365-2966.2012.21725.x](https://doi.org/10.1111/j.1365-2966.2012.21725.x)
- George, M. R., Leauthaud, A., Bundy, K., et al. 2011, *The Astrophysical Journal*, 742, 125, doi: [10.1088/0004-637X/742/2/125](https://doi.org/10.1088/0004-637X/742/2/125)
- . 2012, *The Astrophysical Journal*, 757, 2, doi: [10.1088/0004-637X/757/1/2](https://doi.org/10.1088/0004-637X/757/1/2)
- Giacconi, R., Bechtold, J., Branduardi, G., et al. 1979, *The Astrophysical Journal*, 234, L1, doi: [10.1086/183099](https://doi.org/10.1086/183099)
- Gilbank, D. G., Baldry, I. K., Balogh, M. L., Glazebrook, K., & Bower, R. G. 2010, *Monthly Notices of the Royal Astronomical Society*, 405, 2594, doi: [10.1111/j.1365-2966.2010.16640.x](https://doi.org/10.1111/j.1365-2966.2010.16640.x)
- Gilbank, D. G., Gladders, M. D., Yee, H. K. C., & Hsieh, B. C. 2011, *The Astronomical Journal*, 141, 94, doi: [10.1088/0004-6256/141/3/94](https://doi.org/10.1088/0004-6256/141/3/94)
- Giocoli, C., Marulli, F., Moscardini, L., et al. 2021, *Astronomy and Astrophysics*, 653, A19, doi: [10.1051/0004-6361/202140795](https://doi.org/10.1051/0004-6361/202140795)
- Giocoli, C., Palmucci, L., Lesci, G. F., et al. 2024, *Astronomy and Astrophysics*, 687, A79, doi: [10.1051/0004-6361/202449561](https://doi.org/10.1051/0004-6361/202449561)
- Giodini, S., Pierini, D., Finoguenov, A., et al. 2009, *The Astrophysical Journal*, 703, 982, doi: [10.1088/0004-637X/703/1/982](https://doi.org/10.1088/0004-637X/703/1/982)
- Girardi, M., & Giuricin, G. 2000, *The Astrophysical Journal*, 540, 45, doi: [10.1086/309314](https://doi.org/10.1086/309314)
- Girardi, M., Aguerri, J. A. L., De Grandi, S., et al. 2014, *Astronomy and Astrophysics*, 565, A115, doi: [10.1051/0004-6361/201323311](https://doi.org/10.1051/0004-6361/201323311)
- Gladders, M. D., & Yee, H. K. C. 2000, *The Astronomical Journal*, 120, 2148, doi: [10.1086/301557](https://doi.org/10.1086/301557)
- Gladders, M. D., Yee, H. K. C., Majumdar, S., et al. 2007, *The Astrophysical Journal*, 655, 128, doi: [10.1086/509909](https://doi.org/10.1086/509909)
- Gonzalez, A. 2014, 7. <https://ui.adsabs.harvard.edu/abs/2014becs.confE...7G>

- Gonzalez, A. H., Gettings, D. P., Brodwin, M., et al. 2019, *The Astrophysical Journal Supplement Series*, 240, 33, doi: [10.3847/1538-4365/aafad2](https://doi.org/10.3847/1538-4365/aafad2)
- Gouin, C., Aghanim, N., Bonjean, V., & Douspis, M. 2020, *Astronomy and Astrophysics*, 635, A195, doi: [10.1051/0004-6361/201937218](https://doi.org/10.1051/0004-6361/201937218)
- Gould, K. M. L., Brammer, G., Valentino, F., et al. 2023, *The Astronomical Journal*, 165, 248, doi: [10.3847/1538-3881/accadc](https://doi.org/10.3847/1538-3881/accadc)
- Gozaliasl, G., Finoguenov, A., Khosroshahi, H. G., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 475, 2787, doi: [10.1093/mnras/sty003](https://doi.org/10.1093/mnras/sty003)
- Gozaliasl, G., Finoguenov, A., Tanaka, M., et al. 2019, *Monthly Notices of the Royal Astronomical Society*, 483, 3545, doi: [10.1093/mnras/sty3203](https://doi.org/10.1093/mnras/sty3203)
- Gozaliasl, G., Finoguenov, A., Khosroshahi, H. G., et al. 2020, *Astronomy and Astrophysics*, 635, A36, doi: [10.1051/0004-6361/201936745](https://doi.org/10.1051/0004-6361/201936745)
- Gozaliasl, G., Finoguenov, A., Babul, A., et al. 2024, *Astronomy and Astrophysics*, 690, A315, doi: [10.1051/0004-6361/202449543](https://doi.org/10.1051/0004-6361/202449543)
- Gozaliasl, G., Yang, L., Kartaltepe, J. S., et al. 2025, *Astronomy and Astrophysics*, 703, A129, doi: [10.1051/0004-6361/202556085](https://doi.org/10.1051/0004-6361/202556085)
- Grebel, E. K. 2005, in (eprint: arXiv:astro-ph/0508147: AIP), 161–174, doi: [10.1063/1.1893349](https://doi.org/10.1063/1.1893349)
- Greco, J. P., Hill, J. C., Spergel, D. N., & Battaglia, N. 2015, *The Astrophysical Journal*, 808, 151, doi: [10.1088/0004-637X/808/2/151](https://doi.org/10.1088/0004-637X/808/2/151)
- Grishin, K., Mei, S., & Ilić, S. 2023, *Astronomy and Astrophysics*, 677, A101, doi: [10.1051/0004-6361/202345976](https://doi.org/10.1051/0004-6361/202345976)
- Gursky, H., Kellogg, E., Murray, S., et al. 1971, *The Astrophysical Journal*, 167, L81, doi: [10.1086/180765](https://doi.org/10.1086/180765)
- Gwyn, S., McConnachie, A. W., Cuillandre, J.-C., et al. 2025, *Astronomical Journal*, 170, 324, doi: [10.3847/1538-3881/ae03ab](https://doi.org/10.3847/1538-3881/ae03ab)
- Hamer, S. L., Edge, A. C., Swinbank, A. M., et al. 2012, *Monthly Notices of the Royal Astronomical Society*, 421, 3409, doi: [10.1111/j.1365-2966.2012.20566.x](https://doi.org/10.1111/j.1365-2966.2012.20566.x)
- Hao, J., Koester, B. P., McKay, T. A., et al. 2009, *The Astrophysical Journal*, Volume 702, Issue 1, pp. 745-758 (2009), 702, 745, doi: [10.1088/0004-637X/702/1/745](https://doi.org/10.1088/0004-637X/702/1/745)

- Harish, S., Kartaltepe, J. S., Liu, D., et al. 2025, *The Astrophysical Journal*, 992, 45, doi: [10.3847/1538-4357/adfa1e](https://doi.org/10.3847/1538-4357/adfa1e)
- Hasinger, G., Cappelluti, N., Brunner, H., et al. 2007, *The Astrophysical Journal Supplement Series*, 172, 29, doi: [10.1086/516576](https://doi.org/10.1086/516576)
- Hasinger, G., Capak, P., Salvato, M., et al. 2018, *The Astrophysical Journal*, 858, 77, doi: [10.3847/1538-4357/aabacf](https://doi.org/10.3847/1538-4357/aabacf)
- Hatch, N. A., De Breuck, C., Galametz, A., et al. 2011, *Monthly Notices of the Royal Astronomical Society*, 410, 1537, doi: [10.1111/j.1365-2966.2010.17538.x](https://doi.org/10.1111/j.1365-2966.2010.17538.x)
- Hennawi, J. F., Prochaska, J. X., Cantalupo, S., & Arrigoni-Battaia, F. 2015, *Science*, 348, 779, doi: [10.1126/science.aaa5397](https://doi.org/10.1126/science.aaa5397)
- Hennig, C., Mohr, J. J., Zenteno, A., et al. 2017, *Monthly Notices of the Royal Astronomical Society*, 467, 4015, doi: [10.1093/mnras/stx175](https://doi.org/10.1093/mnras/stx175)
- Herrero-Illana, R., Pérez-Torres, M. Á., Randriamanakoto, Z., et al. 2017, *Monthly Notices of the Royal Astronomical Society*, 471, 1634, doi: [10.1093/mnras/stx1672](https://doi.org/10.1093/mnras/stx1672)
- Heymans, C., Van Waerbeke, L., Miller, L., et al. 2012, *Monthly Notices of the Royal Astronomical Society*, 427, 146, doi: [10.1111/j.1365-2966.2012.21952.x](https://doi.org/10.1111/j.1365-2966.2012.21952.x)
- Hickson, P. 1982, *The Astrophysical Journal*, 255, 382, doi: [10.1086/159838](https://doi.org/10.1086/159838)
- Hill, J. C., Baxter, E. J., Lidz, A., Greco, J. P., & Jain, B. 2018, *Physical Review D*, 97, 083501, doi: [10.1103/PhysRevD.97.083501](https://doi.org/10.1103/PhysRevD.97.083501)
- Hilton, M., Sifón, C., Naess, S., et al. 2021, *The Astrophysical Journal Supplement Series*, 253, 3, doi: [10.3847/1538-4365/abd023](https://doi.org/10.3847/1538-4365/abd023)
- Hitomi Collaboration, Aharonian, F., Akamatsu, H., et al. 2016, *Nature*, 535, 117, doi: [10.1038/nature18627](https://doi.org/10.1038/nature18627)
- Hoekstra, H., Bartelmann, M., Dahle, H., et al. 2013, *Space Science Reviews*, 177, 75, doi: [10.1007/s11214-013-9978-5](https://doi.org/10.1007/s11214-013-9978-5)
- Holden, B. P., Stanford, S. A., Eisenhardt, P., & Dickinson, M. 2004, 127, 2484, doi: [10.1086/382722](https://doi.org/10.1086/382722)
- Hung, C.-L., Casey, C. M., Chiang, Y.-K., et al. 2016, *The Astrophysical Journal*, 826, 130, doi: [10.3847/0004-637X/826/2/130](https://doi.org/10.3847/0004-637X/826/2/130)
- Hung, D., Lemaux, B. C., Cucciati, O., et al. 2025, *The Astrophysical Journal*, 980, 155, doi: [10.3847/1538-4357/ada616](https://doi.org/10.3847/1538-4357/ada616)

- Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, *Astronomy and Astrophysics*, 457, 841, doi: [10.1051/0004-6361:20065138](https://doi.org/10.1051/0004-6361:20065138)
- Ilbert, O., Capak, P., Salvato, M., et al. 2009, *The Astrophysical Journal*, 690, 1236, doi: [10.1088/0004-637X/690/2/1236](https://doi.org/10.1088/0004-637X/690/2/1236)
- Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, *Astronomy and Astrophysics*, 556, A55, doi: [10.1051/0004-6361/201321100](https://doi.org/10.1051/0004-6361/201321100)
- Ilbert, O., Arnouts, S., Le Floch, E., et al. 2015, *Astronomy and Astrophysics*, 579, A2, doi: [10.1051/0004-6361/201425176](https://doi.org/10.1051/0004-6361/201425176)
- Ingoglia, L., Covone, G., Sereno, M., et al. 2022, *Monthly Notices of the Royal Astronomical Society*, 511, 1484, doi: [10.1093/mnras/stac046](https://doi.org/10.1093/mnras/stac046)
- Iovino, A., Petropoulou, V., Scodreggio, M., et al. 2016, *Astronomy and Astrophysics*, 592, A78, doi: [10.1051/0004-6361/201527673](https://doi.org/10.1051/0004-6361/201527673)
- Ito, K., Kashikawa, N., Toshikawa, J., et al. 2019, *The Astrophysical Journal*, 878, 68, doi: [10.3847/1538-4357/ab1f0c](https://doi.org/10.3847/1538-4357/ab1f0c)
- Ito, K., Tanaka, M., Valentino, F., et al. 2023, *The Astrophysical Journal*, 945, L9, doi: [10.3847/2041-8213/acb49b](https://doi.org/10.3847/2041-8213/acb49b)
- Ivezić, , Kahn, S. M., Tyson, J. A., et al. 2019, *The Astrophysical Journal*, 873, 111, doi: [10.3847/1538-4357/ab042c](https://doi.org/10.3847/1538-4357/ab042c)
- Jackson, T. M., Pasquali, A., Pacifici, C., et al. 2020, *Monthly Notices of the Royal Astronomical Society*, 497, 4262, doi: [10.1093/mnras/staa2306](https://doi.org/10.1093/mnras/staa2306)
- Jansen, F., Lumb, D., Altieri, B., et al. 2001, *Astronomy and Astrophysics*, 365, L1, doi: [10.1051/0004-6361:20000036](https://doi.org/10.1051/0004-6361:20000036)
- Jarrett, T. H., Cluver, M. E., Magoulas, C., et al. 2017, *Astrophysical Journal*, 836, 182, doi: [10.3847/1538-4357/836/2/182](https://doi.org/10.3847/1538-4357/836/2/182)
- Jin, S., Sillassen, N. B., Magdis, G. E., et al. 2024, *Astronomy and Astrophysics*, 683, L4, doi: [10.1051/0004-6361/202348540](https://doi.org/10.1051/0004-6361/202348540)
- Jones, L. R., Ponman, T. J., Horton, A., et al. 2003, *Monthly Notices of the Royal Astronomical Society*, 343, 627, doi: [10.1046/j.1365-8711.2003.06702.x](https://doi.org/10.1046/j.1365-8711.2003.06702.x)
- Jung, S. L., Rennehan, D., Saeedzadeh, V., et al. 2022, *Monthly Notices of the Royal Astronomical Society*, 515, 22, doi: [10.1093/mnras/stac1622](https://doi.org/10.1093/mnras/stac1622)

- Kaasra, J. S., Ferrigno, C., Tamura, T., et al. 2001, *Astronomy and Astrophysics*, 365, L99, doi: [10.1051/0004-6361:20000041](https://doi.org/10.1051/0004-6361:20000041)
- Kartalpe, J. S., Sanders, D. B., Le Flo'h, E., et al. 2010, *The Astrophysical Journal*, 709, 572, doi: [10.1088/0004-637X/709/2/572](https://doi.org/10.1088/0004-637X/709/2/572)
- Kartalpe, J. S., Sanders, D. B., Silverman, J. D., et al. 2015, *The Astrophysical Journal*, 806, L35, doi: [10.1088/2041-8205/806/2/L35](https://doi.org/10.1088/2041-8205/806/2/L35)
- Kashino, D., Silverman, J. D., Sanders, D., et al. 2019, *The Astrophysical Journal Supplement Series*, 241, 10, doi: [10.3847/1538-4365/ab06c4](https://doi.org/10.3847/1538-4365/ab06c4)
- Kennicutt, R. C., & Evans, N. J. 2012, *Annual Review of Astronomy and Astrophysics*, 50, 531, doi: [10.1146/annurev-astro-081811-125610](https://doi.org/10.1146/annurev-astro-081811-125610)
- Kennicutt, Jr., R. C. 1998, , 36, 189, doi: [10.1146/annurev.astro.36.1.189](https://doi.org/10.1146/annurev.astro.36.1.189)
- Kepner, J., Fan, X., Bahcall, N., et al. 1999, *The Astrophysical Journal*, 517, 78, doi: [10.1086/307160](https://doi.org/10.1086/307160)
- Khostovan, A. A., Kartalpe, J. S., Salvato, M., et al. 2026, *The Astrophysical Journal Supplement Series*, 282, 6, doi: [10.3847/1538-4365/ae1cb9](https://doi.org/10.3847/1538-4365/ae1cb9)
- King, I. R. 1966, *The Astronomical Journal*, 71, 64, doi: [10.1086/109857](https://doi.org/10.1086/109857)
- Kitayama, T., Ueda, S., Okabe, N., et al. 2023, *Publications of the Astronomical Society of Japan*, 75, 311, doi: [10.1093/pasj/psac110](https://doi.org/10.1093/pasj/psac110)
- Kiyota, T., Ando, M., Tanaka, M., et al. 2025, *The Astrophysical Journal*, 980, 104, doi: [10.3847/1538-4357/ada5f4](https://doi.org/10.3847/1538-4357/ada5f4)
- Kluge, M., Neureiter, B., Riffeser, A., et al. 2020, *The Astrophysical Journal Supplement Series*, 247, 43, doi: [10.3847/1538-4365/ab733b](https://doi.org/10.3847/1538-4365/ab733b)
- Kluge, M., Comparat, J., Liu, A., et al. 2024, *Astronomy and Astrophysics*, 688, A210, doi: [10.1051/0004-6361/202349031](https://doi.org/10.1051/0004-6361/202349031)
- Kneib, J. P., Ellis, R. S., Smail, I., Couch, W. J., & Sharples, R. M. 1996, *The Astrophysical Journal*, 471, 643, doi: [10.1086/177995](https://doi.org/10.1086/177995)
- Knobel, C., Lilly, S. J., Iovino, A., et al. 2009, 697, 1842, doi: [10.1088/0004-637X/697/2/1842](https://doi.org/10.1088/0004-637X/697/2/1842)
- . 2012, *The Astrophysical Journal*, 753, 121, doi: [10.1088/0004-637X/753/2/121](https://doi.org/10.1088/0004-637X/753/2/121)
- Kodama, T., & Arimoto, N. 1997, *Astronomy and Astrophysics*, 320, 41, doi: [10.48550/arXiv.astro-ph/9609160](https://doi.org/10.48550/arXiv.astro-ph/9609160)

- Koekemoer, A. M., Aussel, H., Calzetti, D., et al. 2007, *The Astrophysical Journal Supplement Series*, 172, 196, doi: [10.1086/520086](https://doi.org/10.1086/520086)
- Koester, B. P., McKay, T. A., Annis, J., et al. 2007, *The Astrophysical Journal*, 660, 221, doi: [10.1086/512092](https://doi.org/10.1086/512092)
- Kotulla, R., Fritze, U., Weilbacher, P., & Anders, P. 2009, *Monthly Notices of the Royal Astronomical Society*, 396, 462, doi: [10.1111/j.1365-2966.2009.14717.x](https://doi.org/10.1111/j.1365-2966.2009.14717.x)
- Koulouridis, E., Clerc, N., Sadibekova, T., et al. 2021, *Astronomy and Astrophysics*, 652, A12, doi: [10.1051/0004-6361/202140566](https://doi.org/10.1051/0004-6361/202140566)
- Koyama, Y., Polletta, M. d. C., Tanaka, I., et al. 2021, *Monthly Notices of the Royal Astronomical Society*, 503, L1, doi: [10.1093/mnras/503/1/1](https://doi.org/10.1093/mnras/503/1/1)
- Kramer, C., Penalver, J., & Greve, A. 2013. <https://www.semanticscholar.org/paper/Improvement-of-the-IRAM-30-m-telescope-beam-pattern-Kramer-Penalver/5a7a715e355c61930c2080053493d32b450f91dd>
- Kravtsov, A. V., & Borgani, S. 2012, *Annual Review of Astronomy and Astrophysics*, 50, 353, doi: [10.1146/annurev-astro-081811-125502](https://doi.org/10.1146/annurev-astro-081811-125502)
- Kravtsov, A. V., Vikhlinin, A. A., & Meshcheryakov, A. V. 2018, *Astronomy Letters*, 44, 8, doi: [10.1134/S1063773717120015](https://doi.org/10.1134/S1063773717120015)
- Kroupa, P. 2002, *Science*, 295, 82, doi: [10.1126/science.1067524](https://doi.org/10.1126/science.1067524)
- Kugel, R., Schaye, J., Schaller, M., et al. 2024, *Monthly Notices of the Royal Astronomical Society*, 534, 2378, doi: [10.1093/mnras/stae2218](https://doi.org/10.1093/mnras/stae2218)
- Kuijken, K., Heymans, C., Dvornik, A., et al. 2019, *Astronomy and Astrophysics*, 625, A2, doi: [10.1051/0004-6361/201834918](https://doi.org/10.1051/0004-6361/201834918)
- Kukstas, E., Balogh, M. L., McCarthy, I. G., et al. 2023, *Monthly Notices of the Royal Astronomical Society*, 518, 4782, doi: [10.1093/mnras/stac3438](https://doi.org/10.1093/mnras/stac3438)
- Käfer, F., Finoguenov, A., Eckert, D., et al. 2019, *Astronomy and Astrophysics*, 628, A43, doi: [10.1051/0004-6361/201935124](https://doi.org/10.1051/0004-6361/201935124)
- Kéruzoré, F., Mayet, F., Pratt, G. W., et al. 2020, *Astronomy and Astrophysics*, 644, A93, doi: [10.1051/0004-6361/202038933](https://doi.org/10.1051/0004-6361/202038933)
- Köhlinger, F., Hoekstra, H., & Eriksen, M. 2015, *Monthly Notices of the Royal Astronomical Society*, 453, 3107, doi: [10.1093/mnras/stv1852](https://doi.org/10.1093/mnras/stv1852)

- Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, *The Astrophysical Journal Supplement Series*, 224, 24, doi: [10.3847/0067-0049/224/2/24](https://doi.org/10.3847/0067-0049/224/2/24)
- Laigle, C., Pichon, C., Arnouts, S., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 474, 5437, doi: [10.1093/mnras/stx3055](https://doi.org/10.1093/mnras/stx3055)
- Lan, T.-W., Ménard, B., & Mo, H. 2016, *Monthly Notices of the Royal Astronomical Society*, 459, 3998, doi: [10.1093/mnras/stw898](https://doi.org/10.1093/mnras/stw898)
- Lang, D., Hogg, D. W., & Mykytyn, D. 2016, *The Tractor: Probabilistic astronomical source detection and measurement*
- Laporte, N., Zitrin, A., Dole, H., et al. 2022, *Astronomy and Astrophysics*, 667, L3, doi: [10.1051/0004-6361/202244719](https://doi.org/10.1051/0004-6361/202244719)
- Lauer, T. R., Faber, S. M., Richstone, D., et al. 2007, *Astrophysical Journal*, 662, 808, doi: [10.1086/518223](https://doi.org/10.1086/518223)
- Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, *Euclid Definition Study Report*, doi: [10.48550/arXiv.1110.3193](https://doi.org/10.48550/arXiv.1110.3193)
- Leauthaud, A., Finoguenov, A., Kneib, J.-P., et al. 2010, *The Astrophysical Journal*, 709, 97, doi: [10.1088/0004-637X/709/1/97](https://doi.org/10.1088/0004-637X/709/1/97)
- Lee, K.-G., Hennawi, J. F., Stark, C., et al. 2014, *The Astrophysical Journal*, 795, L12, doi: [10.1088/2041-8205/795/1/L12](https://doi.org/10.1088/2041-8205/795/1/L12)
- Lee, K.-G., Hennawi, J. F., White, M., et al. 2016, *The Astrophysical Journal*, 817, 160, doi: [10.3847/0004-637X/817/2/160](https://doi.org/10.3847/0004-637X/817/2/160)
- Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G., & Byler, N. 2017, *Astrophysical Journal*, 837, 170, doi: [10.3847/1538-4357/aa5ffe](https://doi.org/10.3847/1538-4357/aa5ffe)
- Lemaux, B. C., Cucciati, O., Tasca, L. A. M., et al. 2014, *Astronomy and Astrophysics*, 572, A41, doi: [10.1051/0004-6361/201423828](https://doi.org/10.1051/0004-6361/201423828)
- Lemaux, B. C., Le Fèvre, O., Cucciati, O., et al. 2018, *Astronomy and Astrophysics*, 615, A77, doi: [10.1051/0004-6361/201730870](https://doi.org/10.1051/0004-6361/201730870)
- Lesci, G. F., Marulli, F., Moscardini, L., et al. 2022a, *Astronomy and Astrophysics*, 659, A88, doi: [10.1051/0004-6361/202040194](https://doi.org/10.1051/0004-6361/202040194)
- Lesci, G. F., Nanni, L., Marulli, F., et al. 2022b, *Astronomy and Astrophysics*, 665, A100, doi: [10.1051/0004-6361/202243538](https://doi.org/10.1051/0004-6361/202243538)

- Lesci, G. F., Marulli, F., Moscardini, L., et al. 2025, *Astronomy and Astrophysics*, 703, A25, doi: [10.1051/0004-6361/202555285](https://doi.org/10.1051/0004-6361/202555285)
- Libeskind, N. I., van de Weygaert, R., Cautun, M., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 473, 1195, doi: [10.1093/mnras/stx1976](https://doi.org/10.1093/mnras/stx1976)
- Lidman, C., Suherli, J., Muzzin, A., et al. 2012, *Monthly Notices of the Royal Astronomical Society*, 427, 550, doi: [10.1111/j.1365-2966.2012.21984.x](https://doi.org/10.1111/j.1365-2966.2012.21984.x)
- Lietzen, H., Tempel, E., Heinämäki, P., et al. 2012, *Astronomy and Astrophysics*, 545, A104, doi: [10.1051/0004-6361/201219353](https://doi.org/10.1051/0004-6361/201219353)
- Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, *The Astrophysical Journal Supplement Series*, 172, 70, doi: [10.1086/516589](https://doi.org/10.1086/516589)
- Lilly, S. J., Le Brun, V., Maier, C., et al. 2009, *The Astrophysical Journal Supplement Series*, 184, 218, doi: [10.1088/0067-0049/184/2/218](https://doi.org/10.1088/0067-0049/184/2/218)
- Lin, Y.-T., & Mohr, J. J. 2004, *The Astrophysical Journal*, 617, 879, doi: [10.1086/425412](https://doi.org/10.1086/425412)
- Lin, Y.-T., Mohr, J. J., Gonzalez, A. H., & Stanford, S. A. 2006, *Astrophysical Journal Letters*, 650, L99, doi: [10.1086/508940](https://doi.org/10.1086/508940)
- Lisenfeld, U., Xu, C. K., Gao, Y., et al. 2019, *Astronomy and Astrophysics*, 627, A107, doi: [10.1051/0004-6361/201935536](https://doi.org/10.1051/0004-6361/201935536)
- Liu, A., Bulbul, E., Ghirardini, V., et al. 2022, *Astronomy and Astrophysics*, 661, A2, doi: [10.1051/0004-6361/202141120](https://doi.org/10.1051/0004-6361/202141120)
- Liu, H.-Y., Liu, W.-J., Dong, X.-B., et al. 2019, *Astrophysical Journal Supplement Series*, 243, 21, doi: [10.3847/1538-4365/ab298b](https://doi.org/10.3847/1538-4365/ab298b)
- Lloyd-Davies, E. J., Romer, A. K., Mehrrens, N., et al. 2011, *Monthly Notices of the Royal Astronomical Society*, 418, 14, doi: [10.1111/j.1365-2966.2011.19117.x](https://doi.org/10.1111/j.1365-2966.2011.19117.x)
- Lopes, P. A. A., Ribeiro, A. L. B., & Brambila, D. 2024, *Monthly Notices of the Royal Astronomical Society*, 527, L19, doi: [10.1093/mnras/slad134](https://doi.org/10.1093/mnras/slad134)
- Lovisari, L., Ettori, S., Gaspari, M., & Giles, P. A. 2021, *Universe*, 7, 139, doi: [10.3390/universe7050139](https://doi.org/10.3390/universe7050139)
- Lovisari, L., Reiprich, T. H., & Schellenberger, G. 2015, *Astronomy and Astrophysics*, 573, A118, doi: [10.1051/0004-6361/201423954](https://doi.org/10.1051/0004-6361/201423954)
- Lu, X., Ye, X., & Cheng, Y. 2024, doi: [10.1016/j.engappai.2024.108107](https://doi.org/10.1016/j.engappai.2024.108107)

- Lundberg, S., & Lee, S.-I. 2017, A Unified Approach to Interpreting Model Predictions, arXiv, doi: [10.48550/arXiv.1705.07874](https://doi.org/10.48550/arXiv.1705.07874)
- Ma, Z., & Yan, H. 2015, *Astrophysical Journal*, 811, 58, doi: [10.1088/0004-637X/811/1/58](https://doi.org/10.1088/0004-637X/811/1/58)
- Malek, K., Buat, V., Roehlly, Y., et al. 2018, *Astronomy and Astrophysics*, 620, A50, doi: [10.1051/0004-6361/201833131](https://doi.org/10.1051/0004-6361/201833131)
- Mamon, G. A., Biviano, A., & Boué, G. 2013, 429, 3079, doi: [10.1093/mnras/sts565](https://doi.org/10.1093/mnras/sts565)
- Mandelbaum, R., Lanusse, F., Leauthaud, A., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 481, 3170, doi: [10.1093/mnras/sty2420](https://doi.org/10.1093/mnras/sty2420)
- Mantz, A., Allen, S. W., Ebeling, H., Rapetti, D., & Drlica-Wagner, A. 2010, *Monthly Notices of the Royal Astronomical Society*, 406, 1773, doi: [10.1111/j.1365-2966.2010.16993.x](https://doi.org/10.1111/j.1365-2966.2010.16993.x)
- Marini, I., Borgani, S., Saro, A., et al. 2021, *Monthly Notices of the Royal Astronomical Society*, 507, 5780, doi: [10.1093/mnras/stab2518](https://doi.org/10.1093/mnras/stab2518)
- Marini, I., Popesso, P., Dolag, K., et al. 2025, *Astronomy and Astrophysics*, 698, A191, doi: [10.1051/0004-6361/202554677](https://doi.org/10.1051/0004-6361/202554677)
- Marinoni, C., Davis, M., Newman, J. A., & Coil, A. L. 2002, *The Astrophysical Journal*, 580, 122, doi: [10.1086/343092](https://doi.org/10.1086/343092)
- Martinet, N., Durret, F., Guennou, L., et al. 2015, *Astronomy and Astrophysics*, 575, A116, doi: [10.1051/0004-6361/201423796](https://doi.org/10.1051/0004-6361/201423796)
- Martizzi, D., Teyssier, R., & Moore, B. 2012, *Monthly Notices of the Royal Astronomical Society*, 420, 2859, doi: [10.1111/j.1365-2966.2011.19950.x](https://doi.org/10.1111/j.1365-2966.2011.19950.x)
- Martínez, H. J., Coenda, V., & Muriel, H. 2013, *Astronomy and Astrophysics*, 557, A61, doi: [10.1051/0004-6361/201321931](https://doi.org/10.1051/0004-6361/201321931)
- Masters, D. C., Stern, D. K., Cohen, J. G., et al. 2019, 877, 81, doi: [10.3847/1538-4357/ab184d](https://doi.org/10.3847/1538-4357/ab184d)
- Maturi, M., Bellagamba, F., Radovich, M., et al. 2019, *Monthly Notices of the Royal Astronomical Society*, 485, 498, doi: [10.1093/mnras/stz294](https://doi.org/10.1093/mnras/stz294)
- Maturi, M., Meneghetti, M., Bartelmann, M., Dolag, K., & Moscardini, L. 2005, *Astronomy and Astrophysics*, 442, 851, doi: [10.1051/0004-6361:20042600](https://doi.org/10.1051/0004-6361:20042600)
- Maturi, M., Finoguenov, A., Lopes, P. A. A., et al. 2023, *Astronomy and Astrophysics*, 678, A145, doi: [10.1051/0004-6361/202245323](https://doi.org/10.1051/0004-6361/202245323)

- Maturi, M., Radovich, M., Moscardini, L., et al. 2025, *Astronomy and Astrophysics*, 701, A201, doi: [10.1051/0004-6361/202554340](https://doi.org/10.1051/0004-6361/202554340)
- McCarthy, I. G., Schaye, J., Ponman, T. J., et al. 2010, *Monthly Notices of the Royal Astronomical Society*, 406, 822, doi: [10.1111/j.1365-2966.2010.16750.x](https://doi.org/10.1111/j.1365-2966.2010.16750.x)
- McClintock, T., Varga, T. N., Gruen, D., et al. 2019, *Monthly Notices of the Royal Astronomical Society*, 482, 1352, doi: [10.1093/mnras/sty2711](https://doi.org/10.1093/mnras/sty2711)
- McConachie, I., Wilson, G., Forrest, B., et al. 2022, *The Astrophysical Journal*, 926, 37, doi: [10.3847/1538-4357/ac2b9f](https://doi.org/10.3847/1538-4357/ac2b9f)
- . 2025, *The Astrophysical Journal*, 978, 17, doi: [10.3847/1538-4357/ad8f36](https://doi.org/10.3847/1538-4357/ad8f36)
- McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, *Astronomy and Astrophysics*, 544, A156, doi: [10.1051/0004-6361/201219507](https://doi.org/10.1051/0004-6361/201219507)
- McDonald, M., Benson, B., Veilleux, S., Bautz, M. W., & Reichardt, C. L. 2013, *The Astrophysical Journal*, 765, L37, doi: [10.1088/2041-8205/765/2/L37](https://doi.org/10.1088/2041-8205/765/2/L37)
- McDonald, M., Swinbank, M., Edge, A. C., et al. 2014, *The Astrophysical Journal*, 784, 18, doi: [10.1088/0004-637X/784/1/18](https://doi.org/10.1088/0004-637X/784/1/18)
- McDonald, M., Stalder, B., Bayliss, M., et al. 2016, *The Astrophysical Journal*, 817, 86, doi: [10.3847/0004-637X/817/2/86](https://doi.org/10.3847/0004-637X/817/2/86)
- McGee, S. L., Balogh, M. L., Bower, R. G., Font, A. S., & McCarthy, I. G. 2009, *Monthly Notices of the Royal Astronomical Society*, 400, 937, doi: [10.1111/j.1365-2966.2009.15507.x](https://doi.org/10.1111/j.1365-2966.2009.15507.x)
- McLure, R. J., Dunlop, J. S., Cullen, F., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 476, 3991, doi: [10.1093/mnras/sty522](https://doi.org/10.1093/mnras/sty522)
- McNab, K., Balogh, M. L., van der Burg, R. F. J., et al. 2021, *Monthly Notices of the Royal Astronomical Society*, 508, 157, doi: [10.1093/mnras/stab2558](https://doi.org/10.1093/mnras/stab2558)
- McNamara, B. R., & Nulsen, P. E. J. 2007, *Annual Review of Astronomy and Astrophysics*, 45, 117, doi: [10.1146/annurev.astro.45.051806.110625](https://doi.org/10.1146/annurev.astro.45.051806.110625)
- McNamara, B. R., Russell, H. R., Nulsen, P. E. J., et al. 2014, *The Astrophysical Journal*, 785, 44, doi: [10.1088/0004-637X/785/1/44](https://doi.org/10.1088/0004-637X/785/1/44)
- Mei, S., Holden, B. P., Blakeslee, J. P., et al. 2009, *The Astrophysical Journal*, 690, 42, doi: [10.1088/0004-637X/690/1/42](https://doi.org/10.1088/0004-637X/690/1/42)

- Menci, N., Rosati, P., Gobat, R., et al. 2008, *The Astrophysical Journal*, Volume 685, Issue 2, pp. 863-874 (2008), 685, 863, doi: [10.1086/591016](https://doi.org/10.1086/591016)
- Mendes de Oliveira, C., & Hickson, P. 1991, *The Astrophysical Journal*, 380, 30, doi: [10.1086/170559](https://doi.org/10.1086/170559)
- Meneghetti, M., Bartelmann, M., & Moscardini, L. 2003, *Monthly Notices of the Royal Astronomical Society*, 340, 105, doi: [10.1046/j.1365-8711.2003.06276.x](https://doi.org/10.1046/j.1365-8711.2003.06276.x)
- Merson, A. I., Baugh, C. M., Helly, J. C., et al. 2013, *Monthly Notices of the Royal Astronomical Society*, 429, 556, doi: [10.1093/mnras/sts355](https://doi.org/10.1093/mnras/sts355)
- Milkeraitis, M., van Waerbeke, L., Heymans, C., et al. 2010, *Monthly Notices of the Royal Astronomical Society*, 406, 673, doi: [10.1111/j.1365-2966.2010.16720.x](https://doi.org/10.1111/j.1365-2966.2010.16720.x)
- Miyazaki, S., Oguri, M., Hamana, T., et al. 2018, *Publications of the Astronomical Society of Japan*, 70, S27, doi: [10.1093/pasj/psx120](https://doi.org/10.1093/pasj/psx120)
- Moneti, A., McCracken, H. J., Hudelot, W., et al. 2023, *VizieR Online Data Catalog*, 2373, II/373. <https://ui.adsabs.harvard.edu/abs/2023yCat.2373....0M>
- Morishita, T., Roberts-Borsani, G., Treu, T., et al. 2023, *The Astrophysical Journal*, 947, L24, doi: [10.3847/2041-8213/acb99e](https://doi.org/10.3847/2041-8213/acb99e)
- Moustakas, J., Buhler, J., Scholte, D., Dey, B., & Khederlarian, A. 2023, *FastSpecFit: Fast spectral synthesis and emission-line fitting of DESI spectra*. <http://ascl.net/2308.005>
- Moutard, T., Malavasi, N., Sawicki, M., Arnouts, S., & Tripathi, S. 2020, *Monthly Notices of the Royal Astronomical Society*, 495, 4237, doi: [10.1093/mnras/staa1434](https://doi.org/10.1093/mnras/staa1434)
- Muldrew, S. I., Hatch, N. A., & Cooke, E. A. 2015, *Monthly Notices of the Royal Astronomical Society*, 452, 2528, doi: [10.1093/mnras/stv1449](https://doi.org/10.1093/mnras/stv1449)
- Mushotzky, R., Loewenstein, M., Arnaud, K. A., et al. 1996, *The Astrophysical Journal*, 466, 686, doi: [10.1086/177541](https://doi.org/10.1086/177541)
- Myers, A. D., Moustakas, J., Bailey, S., et al. 2023, *The Astronomical Journal*, 165, 50, doi: [10.3847/1538-3881/aca5f9](https://doi.org/10.3847/1538-3881/aca5f9)
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *The Astrophysical Journal*, Volume 490, Issue 2, pp. 493-508., 490, 493, doi: [10.1086/304888](https://doi.org/10.1086/304888)
- Negri, A., Dalla Vecchia, C., Aguerri, J. A. L., & Bahé, Y. 2022, *Monthly Notices of the Royal Astronomical Society*, 515, 2121, doi: [10.1093/mnras/stac1481](https://doi.org/10.1093/mnras/stac1481)

- Newman, A. B., Rudie, G. C., Blanc, G. A., et al. 2020, *The Astrophysical Journal*, 891, 147, doi: [10.3847/1538-4357/ab75ee](https://doi.org/10.3847/1538-4357/ab75ee)
- Oguri, M., & Marshall, P. J. 2010, *Monthly Notices of the Royal Astronomical Society*, 405, 2579, doi: [10.1111/j.1365-2966.2010.16639.x](https://doi.org/10.1111/j.1365-2966.2010.16639.x)
- Oke, J. B., & Gunn, J. E. 1983, *The Astrophysical Journal*, 266, 713, doi: [10.1086/160817](https://doi.org/10.1086/160817)
- Old, L. J., Balogh, M. L., van der Burg, R. F. J., et al. 2020, *Monthly Notices of the Royal Astronomical Society*, 493, 5987, doi: [10.1093/mnras/staa579](https://doi.org/10.1093/mnras/staa579)
- Olivares, V., Salome, P., Combes, F., et al. 2019, *Astronomy and Astrophysics*, 631, A22, doi: [10.1051/0004-6361/201935350](https://doi.org/10.1051/0004-6361/201935350)
- Olivares, V., Salomé, P., Hamer, S. L., et al. 2022, *Astronomy and Astrophysics*, 666, A94, doi: [10.1051/0004-6361/202142475](https://doi.org/10.1051/0004-6361/202142475)
- Oppenheimer, B. D., Babul, A., Bahé, Y., Butsky, I. S., & McCarthy, I. G. 2021, *Universe*, 7, 209, doi: [10.3390/universe7070209](https://doi.org/10.3390/universe7070209)
- Osmond, J. P. F., & Ponman, T. J. 2004, *Monthly Notices of the Royal Astronomical Society*, 350, 1511, doi: [10.1111/j.1365-2966.2004.07742.x](https://doi.org/10.1111/j.1365-2966.2004.07742.x)
- Overzier, R. A. 2016, *Astronomy and Astrophysics Review*, 24, 14, doi: [10.1007/s00159-016-0100-3](https://doi.org/10.1007/s00159-016-0100-3)
- Pacifici, C., Kassin, S. A., Weiner, B. J., et al. 2016, *The Astrophysical Journal*, 832, 79, doi: [10.3847/0004-637X/832/1/79](https://doi.org/10.3847/0004-637X/832/1/79)
- Pannella, M., Elbaz, D., Daddi, E., et al. 2015, *The Astrophysical Journal*, 807, 141, doi: [10.1088/0004-637X/807/2/141](https://doi.org/10.1088/0004-637X/807/2/141)
- Paul, S., John, R. S., Gupta, P., & Kumar, H. 2017, *Monthly Notices of the Royal Astronomical Society*, 471, 2, doi: [10.1093/mnras/stx1488](https://doi.org/10.1093/mnras/stx1488)
- Paulino-Afonso, A., Sobral, D., Darvish, B., et al. 2018, 620, A186, doi: [10.1051/0004-6361/201832688](https://doi.org/10.1051/0004-6361/201832688)
- Payerne, C., Zhang, Z., Agüena, M., et al. 2025, *Astronomy and Astrophysics*, 700, A34, doi: [10.1051/0004-6361/202554107](https://doi.org/10.1051/0004-6361/202554107)
- Pearson, W. J., Pistis, F., Figueira, M., et al. 2023, *Astronomy and Astrophysics*, 679, A35, doi: [10.1051/0004-6361/202346396](https://doi.org/10.1051/0004-6361/202346396)
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, *Journal of Machine Learning Research*, 12, 2825, doi: [10.48550/arXiv.1201.0490](https://doi.org/10.48550/arXiv.1201.0490)

- Peng, Y., Maiolino, R., & Cochrane, R. 2015, *Nature*, 521, 192, doi: [10.1038/nature14439](https://doi.org/10.1038/nature14439)
- Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010, *The Astrophysical Journal*, 721, 193, doi: [10.1088/0004-637X/721/1/193](https://doi.org/10.1088/0004-637X/721/1/193)
- Peng, Y.-j., et al. 2012, *Astrophysical Journal*, 757, 4, doi: [10.1088/0004-637X/757/1/4](https://doi.org/10.1088/0004-637X/757/1/4)
- Pereira-Santaella, M., Alonso-Herrero, A., Colina, L., et al. 2015, *Astronomy and Astrophysics*, 577, A78, doi: [10.1051/0004-6361/201425359](https://doi.org/10.1051/0004-6361/201425359)
- Peterson, J. R., Kahn, S. M., Paerels, F. B. S., et al. 2003, *The Astrophysical Journal*, 590, 207, doi: [10.1086/374830](https://doi.org/10.1086/374830)
- Phipps, F., Bogdán, Á., Lovisari, L., et al. 2019, *Astrophysical Journal*, 875, 141, doi: [10.3847/1538-4357/ab107c](https://doi.org/10.3847/1538-4357/ab107c)
- Pierre, M., Pacaud, F., Adami, C., et al. 2016, *Astronomy and Astrophysics*, 592, A1, doi: [10.1051/0004-6361/201526766](https://doi.org/10.1051/0004-6361/201526766)
- Piffaretti, R., Arnaud, M., Pratt, G. W., Pointecouteau, E., & Melin, J. B. 2011, *Astronomy and Astrophysics*, 534, A109, doi: [10.1051/0004-6361/201015377](https://doi.org/10.1051/0004-6361/201015377)
- Pillepich, A., Nelson, D., Hernquist, L., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 475, 648, doi: [10.1093/mnras/stx3112](https://doi.org/10.1093/mnras/stx3112)
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2013, *Astronomy and Astrophysics*, 557, A52, doi: [10.1051/0004-6361/201220941](https://doi.org/10.1051/0004-6361/201220941)
- . 2016, *Astronomy and Astrophysics*, 594, A27, doi: [10.1051/0004-6361/201525823](https://doi.org/10.1051/0004-6361/201525823)
- Pointecouteau, E., Arnaud, M., & Pratt, G. W. 2005, *Astronomy and Astrophysics*, 435, 1, doi: [10.1051/0004-6361:20042569](https://doi.org/10.1051/0004-6361:20042569)
- Poitevineau, R., Castignani, G., & Combes, F. 2023, *Astronomy and Astrophysics*, 672, A164, doi: [10.1051/0004-6361/202244560](https://doi.org/10.1051/0004-6361/202244560)
- Polletta, M., Soucail, G., Dole, H., et al. 2021, *Astronomy and Astrophysics*, 654, A121, doi: [10.1051/0004-6361/202140612](https://doi.org/10.1051/0004-6361/202140612)
- Popesso, P., Biviano, A., Böhringer, H., & Romaniello, M. 2006, *Astronomy and Astrophysics*, 445, 29, doi: [10.1051/0004-6361:20052954](https://doi.org/10.1051/0004-6361:20052954)
- Popesso, P., Biviano, A., Bulbul, E., et al. 2024, *Monthly Notices of the Royal Astronomical Society*, 527, 895, doi: [10.1093/mnras/stad3253](https://doi.org/10.1093/mnras/stad3253)

- Postman, M., Lubin, L. M., Gunn, J. E., et al. 1996, *The Astronomical Journal*, 111, 615, doi: [10.1086/117811](https://doi.org/10.1086/117811)
- Postman, M., Coe, D., Benítez, N., et al. 2012, *Astrophysical Journal Supplement Series*, 199, 25, doi: [10.1088/0067-0049/199/2/25](https://doi.org/10.1088/0067-0049/199/2/25)
- Pratt, G. W., Böhringer, H., Croston, J. H., et al. 2007, *Astronomy and Astrophysics*, 461, 71, doi: [10.1051/0004-6361:20065676](https://doi.org/10.1051/0004-6361:20065676)
- Privon, G. C., Baum, S. A., O’Dea, C. P., et al. 2012, *The Astrophysical Journal*, 747, 46, doi: [10.1088/0004-637X/747/1/46](https://doi.org/10.1088/0004-637X/747/1/46)
- Puddu, E., Radovich, M., Sereno, M., et al. 2021, *Astronomy and Astrophysics*, 645, A9, doi: [10.1051/0004-6361/202039259](https://doi.org/10.1051/0004-6361/202039259)
- Radovich, M., Tortora, C., Bellagamba, F., et al. 2020, *Monthly Notices of the Royal Astronomical Society*, 498, 4303, doi: [10.1093/mnras/staa2705](https://doi.org/10.1093/mnras/staa2705)
- Ragagnin, A., Saro, A., Singh, P., & Dolag, K. 2021, *Monthly Notices of the Royal Astronomical Society*, Volume 500, Issue 4, pp.5056-5071, 500, 5056, doi: [10.1093/mnras/staa3523](https://doi.org/10.1093/mnras/staa3523)
- Raj, M. A., Iodice, E., Napolitano, N. R., et al. 2020, *Astronomy and Astrophysics*, 640, A137, doi: [10.1051/0004-6361/202038043](https://doi.org/10.1051/0004-6361/202038043)
- Ramella, M., Boschini, W., Fadda, D., & Nonino, M. 2001, *Astronomy and Astrophysics*, 368, 776, doi: [10.1051/0004-6361:20010071](https://doi.org/10.1051/0004-6361:20010071)
- Raouf, M., Smith, R., Khosroshahi, H. G., et al. 2021, *The Astrophysical Journal*, 908, 123, doi: [10.3847/1538-4357/abd47d](https://doi.org/10.3847/1538-4357/abd47d)
- Reeves, A. M. M., Balogh, M. L., van der Burg, R. F. J., et al. 2021, *Monthly Notices of the Royal Astronomical Society*, 506, 3364, doi: [10.1093/mnras/stab1955](https://doi.org/10.1093/mnras/stab1955)
- Renzini, A. 1997, *The Astrophysical Journal*, 488, 35, doi: [10.1086/304696](https://doi.org/10.1086/304696)
- Rieke, M. J., Kelly, D. M., Misselt, K., et al. 2023, *Publications of the Astronomical Society of the Pacific*, 135, 028001, doi: [10.1088/1538-3873/acac53](https://doi.org/10.1088/1538-3873/acac53)
- Romanello, M., Marulli, F., Moscardini, L., et al. 2024, *Astronomy and Astrophysics*, 682, A72, doi: [10.1051/0004-6361/202348305](https://doi.org/10.1051/0004-6361/202348305)
- Romeo, A. D., Kang, X., Contini, E., et al. 2015, *Astronomy and Astrophysics*, 581, A50, doi: [10.1051/0004-6361/201526226](https://doi.org/10.1051/0004-6361/201526226)

- Rosani, G., Caminha, G. B., Caputi, K. I., & Deshmukh, S. 2020, 633, A159, doi: [10.1051/0004-6361/201935782](https://doi.org/10.1051/0004-6361/201935782)
- Rosati, P., Borgani, S., & Norman, C. 2002, Annual Review of Astronomy and Astrophysics, 40, 539, doi: [10.1146/annurev.astro.40.120401.150547](https://doi.org/10.1146/annurev.astro.40.120401.150547)
- Rosati, P., Tozzi, P., Gobat, R., et al. 2009, Astronomy and Astrophysics, 508, 583, doi: [10.1051/0004-6361/200913099](https://doi.org/10.1051/0004-6361/200913099)
- Rozo, E., & Rykoff, E. S. 2014, The Astrophysical Journal, 783, 80, doi: [10.1088/0004-637X/783/2/80](https://doi.org/10.1088/0004-637X/783/2/80)
- Rudnick, G., von der Linden, A., Pelló, R., et al. 2009, The Astrophysical Journal, Volume 700, Issue 2, pp. 1559-1588 (2009)., 700, 1559, doi: [10.1088/0004-637X/700/2/1559](https://doi.org/10.1088/0004-637X/700/2/1559)
- Russell, H. R., Fabian, A. C., McNamara, B. R., et al. 2014, Monthly Notices of the Royal Astronomical Society, 444, 629, doi: [10.1093/mnras/stu1469](https://doi.org/10.1093/mnras/stu1469)
- Russell, H. R., McNamara, B. R., Fabian, A. C., et al. 2019, Monthly Notices of the Royal Astronomical Society, 490, 3025, doi: [10.1093/mnras/stz2719](https://doi.org/10.1093/mnras/stz2719)
- Rybicki, G. B., & Lightman, A. P. 1979, Radiative processes in astrophysics. <https://ui.adsabs.harvard.edu/abs/1979rpa..book.....R>
- Rykoff, E. S., Rozo, E., Busha, M. T., et al. 2014, The Astrophysical Journal, 785, 104, doi: [10.1088/0004-637X/785/2/104](https://doi.org/10.1088/0004-637X/785/2/104)
- Rykoff, E. S., Rozo, E., Hollowood, D., et al. 2016, The Astrophysical Journal Supplement Series, 224, 1, doi: [10.3847/0067-0049/224/1/1](https://doi.org/10.3847/0067-0049/224/1/1)
- Sadibekova, T., Arnaud, M., Pratt, G. W., Tarrío, P., & Melin, J. B. 2024, Astronomy and Astrophysics, 688, A187, doi: [10.1051/0004-6361/202449427](https://doi.org/10.1051/0004-6361/202449427)
- Saintonge, A., Tacconi, L. J., Fabello, S., et al. 2012, Astrophysical Journal, 758, 73, doi: [10.1088/0004-637X/758/2/73](https://doi.org/10.1088/0004-637X/758/2/73)
- Salerno, J. M., Martínez, H. J., & Muriel, H. 2019, Monthly Notices of the Royal Astronomical Society, 484, 2, doi: [10.1093/mnras/sty3456](https://doi.org/10.1093/mnras/sty3456)
- Salim, S., Boquien, M., & Lee, J. C. 2018, The Astrophysical Journal, 859, 11, doi: [10.3847/1538-4357/aabf3c](https://doi.org/10.3847/1538-4357/aabf3c)
- Salomé, P., & Combes, F. 2003, Astronomy and Astrophysics, 412, 657, doi: [10.1051/0004-6361:20031438](https://doi.org/10.1051/0004-6361:20031438)

- Salomé, P., Combes, F., Revaz, Y., et al. 2011, *Astronomy and Astrophysics*, 531, A85, doi: [10.1051/0004-6361/200811333](https://doi.org/10.1051/0004-6361/200811333)
- Salvati, L., Saro, A., Bocquet, S., et al. 2022, *The Astrophysical Journal*, 934, 129, doi: [10.3847/1538-4357/ac7ab4](https://doi.org/10.3847/1538-4357/ac7ab4)
- Sanders, D. B., & Mirabel, I. F. 1996, , 34, 749, doi: [10.1146/annurev.astro.34.1.749](https://doi.org/10.1146/annurev.astro.34.1.749)
- Saracco, P., Longhetti, M., & Gargiulo, A. 2011, *Monthly Notices of the Royal Astronomical Society*, 412, 2707, doi: [10.1111/j.1365-2966.2010.18098.x](https://doi.org/10.1111/j.1365-2966.2010.18098.x)
- Saracco, P., Casati, A., Gargiulo, A., et al. 2014, *Astronomy and Astrophysics*, 567, A94, doi: [10.1051/0004-6361/201423495](https://doi.org/10.1051/0004-6361/201423495)
- Sarazin, C. L. 1986, *Reviews of Modern Physics*, 58, 1, doi: [10.1103/RevModPhys.58.1](https://doi.org/10.1103/RevModPhys.58.1)
- Sarron, F., & Conselice, C. J. 2021, *Monthly Notices of the Royal Astronomical Society*, 506, 2136, doi: [10.1093/mnras/stab1844](https://doi.org/10.1093/mnras/stab1844)
- Sarron, F., Martinet, N., Durret, F., & Adami, C. 2018, *Astronomy and Astrophysics*, 613, A67, doi: [10.1051/0004-6361/201731981](https://doi.org/10.1051/0004-6361/201731981)
- Sartoris, B., Biviano, A., Fedeli, C., et al. 2016, *Monthly Notices of the Royal Astronomical Society*, 459, 1764, doi: [10.1093/mnras/stw630](https://doi.org/10.1093/mnras/stw630)
- Sawicki, M., Arnouts, S., Huang, J., et al. 2019, *Monthly Notices of the Royal Astronomical Society*, 489, 5202, doi: [10.1093/mnras/stz2522](https://doi.org/10.1093/mnras/stz2522)
- Schechter, P. 1976, *Astrophysical Journal*, Vol. 203, p. 297-306 (1976), 203, 297, doi: [10.1086/154079](https://doi.org/10.1086/154079)
- Schellenberger, G., & Reiprich, T. H. 2017, *Monthly Notices of the Royal Astronomical Society*, 471, 1370, doi: [10.1093/mnras/stx1583](https://doi.org/10.1093/mnras/stx1583)
- Schneider, P., van Waerbeke, L., Jain, B., & Kruse, G. 1998, *Monthly Notices of the Royal Astronomical Society*, 296, 873, doi: [10.1046/j.1365-8711.1998.01422.x](https://doi.org/10.1046/j.1365-8711.1998.01422.x)
- Scognamiglio, D., Leroy, G., Harvey, D., et al. 2026, *Nature Astronomy*, doi: [10.1038/s41550-025-02763-9](https://doi.org/10.1038/s41550-025-02763-9)
- Scoville, N., Aussel, H., Brusa, M., et al. 2007, *The Astrophysical Journal Supplement Series*, 172, 1, doi: [10.1086/516585](https://doi.org/10.1086/516585)
- Seppi, R., Comparat, J., Nandra, K., et al. 2023, *Astronomy and Astrophysics*, 671, A57, doi: [10.1051/0004-6361/202245138](https://doi.org/10.1051/0004-6361/202245138)

- Sereno, M. 2015, *Monthly Notices of the Royal Astronomical Society*, 450, 3665, doi: [10.1093/mnras/stu2505](https://doi.org/10.1093/mnras/stu2505)
- Sereno, M. 2016, *Monthly Notices of the Royal Astronomical Society*, 455, 2149, doi: [10.1093/mnras/stv2374](https://doi.org/10.1093/mnras/stv2374)
- Sereno, M., Ettori, S., Lesci, G. F., et al. 2020, *Monthly Notices of the Royal Astronomical Society*, 497, 894, doi: [10.1093/mnras/staa1902](https://doi.org/10.1093/mnras/staa1902)
- Sersic, J. L. 1968, *Atlas de Galaxias Australes*
- Shan, H., Kneib, J.-P., Tao, C., et al. 2012, *The Astrophysical Journal*, 748, 56, doi: [10.1088/0004-637X/748/1/56](https://doi.org/10.1088/0004-637X/748/1/56)
- Shen, Y., Richards, G. T., Strauss, M. A., et al. 2011, *Astrophysical Journal Supplement Series*, 194, 45, doi: [10.1088/0067-0049/194/2/45](https://doi.org/10.1088/0067-0049/194/2/45)
- Shimakawa, R., Koyama, Y., Röttgering, H. J. A., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 481, 5630, doi: [10.1093/mnras/sty2618](https://doi.org/10.1093/mnras/sty2618)
- Shimasaku, K., Ouchi, M., Okamura, S., et al. 2003, *The Astrophysical Journal*, 586, L111, doi: [10.1086/374880](https://doi.org/10.1086/374880)
- Shuntov, M., Akins, H. B., Paquereau, L., et al. 2025a, *Astronomy and Astrophysics*, 704, A339, doi: [10.1051/0004-6361/202555799](https://doi.org/10.1051/0004-6361/202555799)
- Shuntov, M., Ilbert, O., Toft, S., et al. 2025b, *Astronomy and Astrophysics*, 695, A20, doi: [10.1051/0004-6361/202452570](https://doi.org/10.1051/0004-6361/202452570)
- Sillassen, N. B., Jin, S., Magdis, G. E., et al. 2024, *Astronomy and Astrophysics*, 690, A55, doi: [10.1051/0004-6361/202450760](https://doi.org/10.1051/0004-6361/202450760)
- Silverman, J. D., Kashino, D., Sanders, D., et al. 2015, *The Astrophysical Journal Supplement Series*, 220, 12, doi: [10.1088/0067-0049/220/1/12](https://doi.org/10.1088/0067-0049/220/1/12)
- Sinha, M., Berlind, A. A., McBride, C. K., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 478, 1042, doi: [10.1093/mnras/sty967](https://doi.org/10.1093/mnras/sty967)
- Siudek, M., Mezcua, M., & Krywult, J. 2023, *Monthly Notices of the Royal Astronomical Society*, 518, 724, doi: [10.1093/mnras/stac3092](https://doi.org/10.1093/mnras/stac3092)
- Skelton, R. E., Bell, E. F., & Somerville, R. S. 2012, 753, 44, doi: [10.1088/0004-637X/753/1/44](https://doi.org/10.1088/0004-637X/753/1/44)
- Smolčić, V., Schinnerer, E., Finoguenov, A., et al. 2007, *The Astrophysical Journal Supplement Series*, 172, 295, doi: [10.1086/516583](https://doi.org/10.1086/516583)

- Smolčić, V., Novak, M., Bondi, M., et al. 2017, *Astronomy and Astrophysics*, 602, A1, doi: [10.1051/0004-6361/201628704](https://doi.org/10.1051/0004-6361/201628704)
- Solomon, P. M., Downes, D., Radford, S. J. E., & Barrett, J. W. 1997, *The Astrophysical Journal*, 478, 144, doi: [10.1086/303765](https://doi.org/10.1086/303765)
- Solomon, P. M., & Vanden Bout, P. A. 2005, *Annual Review of Astronomy and Astrophysics*, 43, 677, doi: [10.1146/annurev.astro.43.051804.102221](https://doi.org/10.1146/annurev.astro.43.051804.102221)
- Soo, J. Y. H., Moraes, B., Joachimi, B., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 475, 3613, doi: [10.1093/mnras/stx3201](https://doi.org/10.1093/mnras/stx3201)
- Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, *The Astrophysical Journal Supplement Series*, 214, 15, doi: [10.1088/0067-0049/214/2/15](https://doi.org/10.1088/0067-0049/214/2/15)
- Spitler, L. R., Labbé, I., Glazebrook, K., et al. 2012, *The Astrophysical Journal*, 748, L21, doi: [10.1088/2041-8205/748/2/L21](https://doi.org/10.1088/2041-8205/748/2/L21)
- Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Nature*, 435, 629, doi: [10.1038/nature03597](https://doi.org/10.1038/nature03597)
- Stapelberg, S., Carrasco, M., & Maturi, M. 2019, *Monthly Notices of the Royal Astronomical Society*, 482, 1824, doi: [10.1093/mnras/sty2784](https://doi.org/10.1093/mnras/sty2784)
- Steidel, C. C., Adelberger, K. L., Dickinson, M., et al. 1998, *The Astrophysical Journal*, 492, 428, doi: [10.1086/305073](https://doi.org/10.1086/305073)
- Stern, D., Assef, R. J., Benford, D. J., et al. 2012, *Astrophysical Journal*, 753, 30, doi: [10.1088/0004-637X/753/1/30](https://doi.org/10.1088/0004-637X/753/1/30)
- Stern, J., & Laor, A. 2012, *Monthly Notices of the Royal Astronomical Society*, 423, 600, doi: [10.1111/j.1365-2966.2012.20901.x](https://doi.org/10.1111/j.1365-2966.2012.20901.x)
- Story, K. T., Reichardt, C. L., Hou, Z., et al. 2013, 779, 86, doi: [10.1088/0004-637X/779/1/86](https://doi.org/10.1088/0004-637X/779/1/86)
- Stoher, L., Norberg, P., & Baugh, C. M. 2019, *Monthly Notices of the Royal Astronomical Society*, 485, L126, doi: [10.1093/mnrasl/slz045](https://doi.org/10.1093/mnrasl/slz045)
- Stott, J. P., Collins, C. A., Burke, C., Hamilton-Morris, V., & Smith, G. P. 2011, *Monthly Notices of the Royal Astronomical Society*, 414, 445, doi: [10.1111/j.1365-2966.2011.18404.x](https://doi.org/10.1111/j.1365-2966.2011.18404.x)
- Stott, J. P., Pimblett, K. A., Edge, A. C., Smith, G. P., & Wardlow, J. L. 2009, *Monthly Notices of the Royal Astronomical Society*, Volume 394, Issue 4, pp. 2098-2108., 394, 2098, doi: [10.1111/j.1365-2966.2009.14477.x](https://doi.org/10.1111/j.1365-2966.2009.14477.x)
- Stott, J. P., Smail, I., Edge, A. C., et al. 2007, *The Astrophysical Journal*, 661, 95, doi: [10.1086/514329](https://doi.org/10.1086/514329)

- Strazzullo, V., Gobat, R., Daddi, E., et al. 2013, *The Astrophysical Journal*, 772, 118, doi: [10.1088/0004-637X/772/2/118](https://doi.org/10.1088/0004-637X/772/2/118)
- Strazzullo, V., Daddi, E., Gobat, R., et al. 2016, *The Astrophysical Journal*, 833, L20, doi: [10.3847/2041-8213/833/2/L20](https://doi.org/10.3847/2041-8213/833/2/L20)
- Sun, M. 2012, *New Journal of Physics*, 14, 045004, doi: [10.1088/1367-2630/14/4/045004](https://doi.org/10.1088/1367-2630/14/4/045004)
- Sunyaev, R. A., & Zeldovich, Y. B. 1970, *Comments on Astrophysics and Space Physics*, 2, 66.
<https://ui.adsabs.harvard.edu/abs/1970CoASP...2...66S>
- Sérsic, J. L. 1963, *Boletín de la Asociación Argentina de Astronomía La Plata Argentina*, 6, 41.
<https://ui.adsabs.harvard.edu/abs/1963BAAA....6...41S>
- Taamoli, S., Nezhad, N., Mobasher, B., et al. 2024a, *The Astrophysical Journal*, 977, 263, doi: [10.3847/1538-4357/ad94f3](https://doi.org/10.3847/1538-4357/ad94f3)
- Taamoli, S., Mobasher, B., Chartab, N., et al. 2024b, *The Astrophysical Journal*, 966, 18, doi: [10.3847/1538-4357/ad32c5](https://doi.org/10.3847/1538-4357/ad32c5)
- Tacchella, S., Carollo, C. M., Renzini, A., et al. 2015, *Science*, 348, 314, doi: [10.1126/science.1261094](https://doi.org/10.1126/science.1261094)
- Tacconi, L. J., Genzel, R., Saintonge, A., et al. 2018, *The Astrophysical Journal*, 853, 179, doi: [10.3847/1538-4357/aaa4b4](https://doi.org/10.3847/1538-4357/aaa4b4)
- Takey, A., Durret, F., Márquez, I., et al. 2019, *Monthly Notices of the Royal Astronomical Society*, 486, 4863, doi: [10.1093/mnras/stz1097](https://doi.org/10.1093/mnras/stz1097)
- Tanaka, M., Finoguenov, A., Lilly, S. J., et al. 2012, *Publications of the Astronomical Society of Japan*, 64, 22, doi: [10.1093/pasj/64.2.22](https://doi.org/10.1093/pasj/64.2.22)
- Tanaka, M., Onodera, M., Shimakawa, R., et al. 2024, *The Astrophysical Journal*, 970, 59, doi: [10.3847/1538-4357/ad5316](https://doi.org/10.3847/1538-4357/ad5316)
- Taniguchi, Y., Kajisawa, M., Kobayashi, M. A. R., et al. 2015, *Publications of the Astronomical Society of Japan*, 67, 104, doi: [10.1093/pasj/psv106](https://doi.org/10.1093/pasj/psv106)
- Tarrió, P., Melin, J. B., & Arnaud, M. 2019, *Astronomy and Astrophysics*, 626, A7, doi: [10.1051/0004-6361/201834979](https://doi.org/10.1051/0004-6361/201834979)
- Tarrió, P., Melin, J. B., & Arnaud, M. 2018, *Astronomy and Astrophysics*, 614, A82, doi: [10.1051/0004-6361/201731984](https://doi.org/10.1051/0004-6361/201731984)
- Temi, P., Amblard, A., Gitti, M., et al. 2018, *The Astrophysical Journal*, 858, 17, doi: [10.3847/1538-4357/aab9b0](https://doi.org/10.3847/1538-4357/aab9b0)

- Tempel, E., Tamm, A., Gramann, M., et al. 2014, *Astronomy and Astrophysics*, 566, A1, doi: [10.1051/0004-6361/201423585](https://doi.org/10.1051/0004-6361/201423585)
- The Dark Energy Survey Collaboration. 2005, arXiv e-prints, astro, doi: [10.48550/arXiv.astro-ph/0510346](https://doi.org/10.48550/arXiv.astro-ph/0510346)
- Thongkham, K., Gonzalez, A. H., Brodwin, M., et al. 2024, *The Astrophysical Journal*, 967, 123, doi: [10.3847/1538-4357/ad3c44](https://doi.org/10.3847/1538-4357/ad3c44)
- Tian, D.-C., Yang, Y., Wen, Z.-L., & Xia, J.-Q. 2025, *The Astrophysical Journal Supplement Series*, 276, 21, doi: [10.3847/1538-4365/ad8bbd](https://doi.org/10.3847/1538-4365/ad8bbd)
- Toft, S., van Dokkum, P., Franx, M., et al. 2007, *Astrophysical Journal*, 671, 285, doi: [10.1086/521810](https://doi.org/10.1086/521810)
- Toni, G., Maturi, M., Finoguenov, A., Moscardini, L., & Castignani, G. 2024, *Astronomy and Astrophysics*, 687, A56, doi: [10.1051/0004-6361/202348832](https://doi.org/10.1051/0004-6361/202348832)
- Toni, G., Gozaliasl, G., Maturi, M., et al. 2025a, *Astronomy and Astrophysics*, 697, A197, doi: [10.1051/0004-6361/202553759](https://doi.org/10.1051/0004-6361/202553759)
- Toni, G., Maturi, M., Castignani, G., et al. 2025b, COSMOS-Web galaxy groups: Evolution of red sequence and quiescent galaxy fraction, arXiv, doi: [10.48550/arXiv.2509.08040](https://doi.org/10.48550/arXiv.2509.08040)
- Tricottet, M., Mamon, G. A., & Díaz-Giménez, E. 2025, *Astronomy and Astrophysics*, 699, A329, doi: [10.1051/0004-6361/202451713](https://doi.org/10.1051/0004-6361/202451713)
- Trobbiani, L., Maturi, M., Giocoli, C., Moscardini, L., & Panebianco, G. 2025, *Astronomy and Astrophysics*, 699, A275, doi: [10.1051/0004-6361/202553889](https://doi.org/10.1051/0004-6361/202553889)
- Tully, R. B. 1987, *The Astrophysical Journal*, 321, 280, doi: [10.1086/165629](https://doi.org/10.1086/165629)
- Ubertosi, F., Gitti, M., Temi, P., et al. 2025, *The Astrophysical Journal*, 989, 128, doi: [10.3847/1538-4357/adf19c](https://doi.org/10.3847/1538-4357/adf19c)
- van der Burg, R. F. J., et al. 2014, *A&A*, 561, A79, doi: [10.1051/0004-6361/201322313](https://doi.org/10.1051/0004-6361/201322313)
- van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, *Astrophysical Journal*, 788, 28, doi: [10.1088/0004-637X/788/1/28](https://doi.org/10.1088/0004-637X/788/1/28)
- van der Wel, A., Bezanson, R., D'Eugenio, F., et al. 2021, 256, 44, doi: [10.3847/1538-4365/ac1356](https://doi.org/10.3847/1538-4365/ac1356)
- van Weeren, R. J., de Gasperin, F., Akamatsu, H., et al. 2019, *Space Science Reviews*, 215, 16, doi: [10.1007/s11214-019-0584-z](https://doi.org/10.1007/s11214-019-0584-z)
- Vikhlinin, A., McNamara, B. R., Forman, W., et al. 1998, 502, 558, doi: [10.1086/305951](https://doi.org/10.1086/305951)

- Viola, M., Cacciato, M., Brouwer, M., et al. 2015, *Monthly Notices of the Royal Astronomical Society*, 452, 3529, doi: [10.1093/mnras/stv1447](https://doi.org/10.1093/mnras/stv1447)
- Violino, G., Ellison, S. L., Sargent, M., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 476, 2591, doi: [10.1093/mnras/sty345](https://doi.org/10.1093/mnras/sty345)
- Voit, G. M. 2005, *Reviews of Modern Physics*, 77, 207, doi: [10.1103/RevModPhys.77.207](https://doi.org/10.1103/RevModPhys.77.207)
- Von Der Linden, A., Best, P. N., Kauffmann, G., & White, S. D. 2006, in , 183.05. <https://ui.adsabs.harvard.edu/abs/2006AAS...20918305V>
- Wang, G. C. P., Hill, R., Chapman, S. C., et al. 2021, *Monthly Notices of the Royal Astronomical Society*, 508, 3754, doi: [10.1093/mnras/stab2800](https://doi.org/10.1093/mnras/stab2800)
- Wang, T., Elbaz, D., Daddi, E., et al. 2016, *The Astrophysical Journal*, 828, 56, doi: [10.3847/0004-637X/828/1/56](https://doi.org/10.3847/0004-637X/828/1/56)
- Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022, *The Astrophysical Journal Supplement Series*, 258, 11, doi: [10.3847/1538-4365/ac3078](https://doi.org/10.3847/1538-4365/ac3078)
- Weaver, J. R., Zalesky, L., Kokorev, V., et al. 2023a, 269, 20, doi: [10.3847/1538-4365/acf850](https://doi.org/10.3847/1538-4365/acf850)
- Weaver, J. R., Davidzon, I., Toft, S., et al. 2023b, *Astronomy and Astrophysics*, 677, A184, doi: [10.1051/0004-6361/202245581](https://doi.org/10.1051/0004-6361/202245581)
- Webb, T. M. A., Lowenthal, J., Yun, M., et al. 2017, *The Astrophysical Journal*, 844, L17, doi: [10.3847/2041-8213/aa7749](https://doi.org/10.3847/2041-8213/aa7749)
- Weinmann, S. M., et al. 2006, *MNRAS*, 372, 1161, doi: [10.1111/j.1365-2966.2006.10922.x](https://doi.org/10.1111/j.1365-2966.2006.10922.x)
- Weisskopf, M. C., Tananbaum, H. D., Van Speybroeck, L. P., & O'Dell, S. L. 2000, in , eprint: [arXiv:astro-ph/0004127](https://arxiv.org/abs/astro-ph/0004127), 2–16, doi: [10.1117/12.391545](https://doi.org/10.1117/12.391545)
- Wen, Z. L., & Han, J. L. 2018, *Monthly Notices of the Royal Astronomical Society*, 481, 4158, doi: [10.1093/mnras/sty2533](https://doi.org/10.1093/mnras/sty2533)
- Wen, Z. L., & Han, J. L. 2024, *Astrophysical Journal Supplement Series*, 272, 39, doi: [10.3847/1538-4365/ad409d](https://doi.org/10.3847/1538-4365/ad409d)
- Wen, Z. L., Han, J. L., & Liu, F. S. 2012, *The Astrophysical Journal Supplement Series*, 199, 34, doi: [10.1088/0067-0049/199/2/34](https://doi.org/10.1088/0067-0049/199/2/34)
- Wen, Z. L., Han, J. L., & Yang, F. 2018, *Monthly Notices of the Royal Astronomical Society*, 475, 343, doi: [10.1093/mnras/stx3189](https://doi.org/10.1093/mnras/stx3189)

- Werner, S. V., Hatch, N. A., Matharu, J., et al. 2023a, *Monthly Notices of the Royal Astronomical Society*, 523, 91, doi: [10.1093/mnras/stad1410](https://doi.org/10.1093/mnras/stad1410)
- Werner, S. V., Cypriano, E. S., Gonzalez, A. H., et al. 2023b, *Monthly Notices of the Royal Astronomical Society*, 519, 2630, doi: [10.1093/mnras/stac3273](https://doi.org/10.1093/mnras/stac3273)
- Wetzel, A. R., Tinker, J. L., Conroy, C., & van den Bosch, F. C. 2013, *Monthly Notices of the Royal Astronomical Society*, 432, 336, doi: [10.1093/mnras/stt469](https://doi.org/10.1093/mnras/stt469)
- Whitaker, K. E., Labbé, I., van Dokkum, P. G., et al. 2011, *The Astrophysical Journal*, 735, 86, doi: [10.1088/0004-637X/735/2/86](https://doi.org/10.1088/0004-637X/735/2/86)
- Williams, R. J., Quadri, R. F., Franx, M., et al. 2010, *Astrophysical Journal*, 713, 738, doi: [10.1088/0004-637X/713/2/738](https://doi.org/10.1088/0004-637X/713/2/738)
- Willis, J. P., Clerc, N., Bremer, M. N., et al. 2013, *Monthly Notices of the Royal Astronomical Society*, 430, 134, doi: [10.1093/mnras/sts540](https://doi.org/10.1093/mnras/sts540)
- Wilman, D. J., Balogh, M. L., Bower, R. G., et al. 2005, *Monthly Notices of the Royal Astronomical Society*, 358, 71, doi: [10.1111/j.1365-2966.2005.08744.x](https://doi.org/10.1111/j.1365-2966.2005.08744.x)
- Wilson, C. D., Scoville, N., Madden, S. C., & Charmandaris, V. 2000, *Astrophysical Journal*, 542, 120, doi: [10.1086/309504](https://doi.org/10.1086/309504)
- Wilson, D., Nayyeri, H., Cooray, A., & Häußler, B. 2020, *The Astrophysical Journal*, 888, 83, doi: [10.3847/1538-4357/ab5a79](https://doi.org/10.3847/1538-4357/ab5a79)
- Wright, R. H., Sabatke, D., & Telfer, R. 2022, 12180, 121803P, doi: [10.1117/12.2632087](https://doi.org/10.1117/12.2632087)
- Wu, P.-F., van der Wel, A., Gallazzi, A., et al. 2018, *The Astrophysical Journal*, 855, 85, doi: [10.3847/1538-4357/aab0a6](https://doi.org/10.3847/1538-4357/aab0a6)
- Wylezalek, D., Vernet, J., De Breuck, C., et al. 2013, *Monthly Notices of the Royal Astronomical Society*, 428, 3206, doi: [10.1093/mnras/sts264](https://doi.org/10.1093/mnras/sts264)
- Yang, L., Kartaltepe, J. S., Franco, M., et al. 2025, *The Astrophysical Journal Supplement Series*, 281, 68, doi: [10.3847/1538-4365/ae0e1b](https://doi.org/10.3847/1538-4365/ae0e1b)
- Young, J. S., Xie, S., Tacconi, L., et al. 1995, *The Astrophysical Journal Supplement Series*, 98, 219, doi: [10.1086/192159](https://doi.org/10.1086/192159)
- Yuan, T., Nanayakkara, T., Kacprzak, G. G., et al. 2014, *The Astrophysical Journal*, 795, L20, doi: [10.1088/2041-8205/795/1/L20](https://doi.org/10.1088/2041-8205/795/1/L20)
- Zamojski, M. A., Schiminovich, D., Rich, R. M., et al. 2007, 172, 468, doi: [10.1086/516593](https://doi.org/10.1086/516593)

- Zandivarez, A., Díaz-Giménez, E., Taverna, A., Rodriguez, F., & Merchán, M. 2024, *Astronomy and Astrophysics*, 691, A6, doi: [10.1051/0004-6361/202451471](https://doi.org/10.1051/0004-6361/202451471)
- Zarattini, S., Aguerri, J. A. L., Tarrío, P., & Corsini, E. M. 2023, *Astronomy and Astrophysics*, 676, A133, doi: [10.1051/0004-6361/202346238](https://doi.org/10.1051/0004-6361/202346238)
- Zatloukal, M., Röser, H. J., Wolf, C., Hippelein, H., & Falter, S. 2007, 474, L5, doi: [10.1051/0004-6361:20078063](https://doi.org/10.1051/0004-6361:20078063)
- Zavala, J. A., Casey, C. M., Scoville, N., et al. 2019, *The Astrophysical Journal*, 887, 183, doi: [10.3847/1538-4357/ab5302](https://doi.org/10.3847/1538-4357/ab5302)
- Zeimann, G. R., Stanford, S. A., Brodwin, M., et al. 2013, *Astrophysical Journal*, 779, 137, doi: [10.1088/0004-637X/779/2/137](https://doi.org/10.1088/0004-637X/779/2/137)
- Zenteno, A., Mohr, J. J., Desai, S., et al. 2016, 462, 830, doi: [10.1093/mnras/stw1649](https://doi.org/10.1093/mnras/stw1649)
- Zenteno, A., Kluge, M., Kharkrang, R., et al. 2025, *Astronomy and Astrophysics*, 698, A171, doi: [10.1051/0004-6361/202452440](https://doi.org/10.1051/0004-6361/202452440)
- Zhang, C., Ramos-Ceja, M. E., Pacaud, F., & Reiprich, T. H. 2020, *Astronomy and Astrophysics*, 642, A17, doi: [10.1051/0004-6361/201937329](https://doi.org/10.1051/0004-6361/201937329)
- Zhao, D., Conselice, C. J., Aragón-Salamanca, A., et al. 2017, *Monthly Notices of the Royal Astronomical Society*, 464, 1393, doi: [10.1093/mnras/stw2406](https://doi.org/10.1093/mnras/stw2406)
- Zheng, Y., Xu, K., Zhao, D., et al. 2025, *The Astrophysical Journal*, 984, 193, doi: [10.3847/1538-4357/adc39a](https://doi.org/10.3847/1538-4357/adc39a)
- Zolotov, A., Dekel, A., Mandelker, N., et al. 2015, *Monthly Notices of the Royal Astronomical Society*, 450, 2327, doi: [10.1093/mnras/stv740](https://doi.org/10.1093/mnras/stv740)
- Zu, Y., Weinberg, D. H., Rozo, E., et al. 2014, *Monthly Notices of the Royal Astronomical Society*, 439, 1628, doi: [10.1093/mnras/stu033](https://doi.org/10.1093/mnras/stu033)
- Álvarez Márquez, J., Burgarella, D., Heinis, S., et al. 2016, *Astronomy and Astrophysics*, 587, A122, doi: [10.1051/0004-6361/201527190](https://doi.org/10.1051/0004-6361/201527190)

