

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

SCUOLA DI SCIENZE
Dipartimento di Fisica e Astronomia
Corso di Laurea Magistrale in Astrofisica e Cosmologia

**Detection and characterization
of galaxy clusters in the
COSMOS field with the
AMICO algorithm**

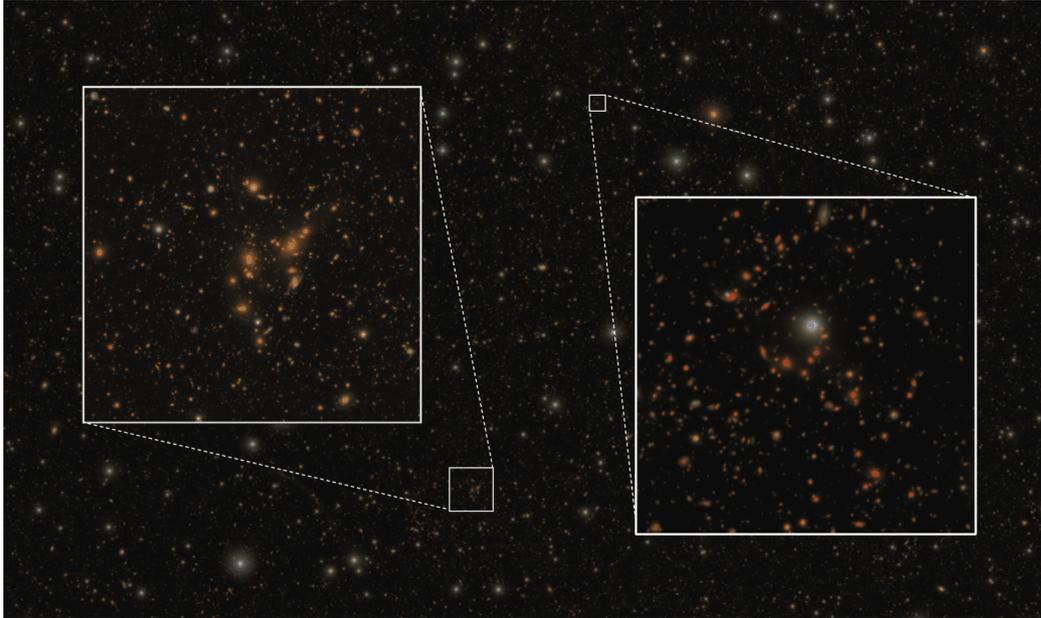
Tesi di Laurea Magistrale

Relatori:
Prof. Dr.
Lauro Moscardini

Priv.-Doz. Dr.
Matteo Maturi

Presentata da:
Greta Toni

Anno Accademico 2020/2021



Two clusters of galaxies in the COSMOS field as detected by the AMICO algorithm within the study described in this Thesis work. On the left a small nearby cluster and on the right a bigger and more distant one. The colours allow to appreciate by eye the effect of cosmological redshift. Colour composite images (g,r,i) from Subaru Hyper-Supreme-Cam DR3 (Aihara et al., 2021).

Abstract

Galaxy clusters are objects whose importance in modern cosmology and in astrophysics has been well established. In this work we made use of the AMICO algorithm to detect clusters in COSMOS2015, a photometric galaxy catalogue of the COSMOS field (Laigle et al., 2016). We divided our study in two different analyses. First, we based the cluster search on the r -band photometry and produced a catalogue of 301 galaxy clusters in the range $0 < z < 1.25$, of which 175 with $SNR > 3$. Then, we extended the cluster search to higher redshift, by basing the analysis on the H -band photometry and we obtained a total of 351 detections in the range $0 < z \lesssim 1.8$ with $3 < SNR \lesssim 8$, giving a total number of 481 clusters detected over the COSMOS field area. In the process, we also identified the cluster member galaxies by assigning them a probabilistic membership. During the cluster search, from a technical point of view, we tested the usability of AMICO, for instance, for the case of the high redshifts available in COSMOS and of its small covered area. Thus, we introduced some new methods within AMICO, to allow its application on new kinds of data-sets. We introduced a new way to generate from the data-set a complete visibility mask, we tested a regularization method for the noise and we introduced a new way to attenuate cluster contribution to the noise model.

Thanks to the rich multi-wavelength covering of the COSMOS field, we identified the X-ray counterparts of our optical detections, by making use of the publicly available catalogues produced by Gozaliasl et al. (2019) and George et al. (2011) and of the X-ray 0.5-2 keV emission map (Gozaliasl et al., 2019). This comparison has been carried out with the main goal of testing the consistency between the retrieved AMICO catalogues and the X-ray catalogues. The X-ray catalogues were also used to compare the reliability of our two different catalogues, the r -band one and the H -band one, with respect to a common reference. Moreover, such a comparison allowed us to calibrate the AMICO mass proxies scaling relations (for amplitude and intrinsic richness), using the available X-rays mass estimates.

Sommario

Gli ammassi di galassie sono oggetti la cui importanza in cosmologia e astrofisica è da tempo ben nota. Il principale scopo dell'analisi descritta in questo lavoro di Tesi è l'identificazione di ammassi di galassie grazie all'applicazione dell'algoritmo "AMICO" a un catalogo fotometrico di galassie del campo COSMOS, il catalogo COSMOS2015 (Laigle et al., 2016). Lo studio è stato diviso in due diverse analisi. In principio, la ricerca di ammassi è stata basata sulla fotometria nella banda r , con la quale è stato prodotto un catalogo di 301 ammassi di galassie nell'intervallo $0 < z < 1.25$, dei quali 175 con $SNR > 3$. Successivamente, la ricerca è stata estesa a più alto *redshift* basando l'analisi sulla fotometria nella banda H , per la quale abbiamo ottenuto un totale di 351 identificazioni nel range $0 < z \lesssim 1.8$ con $3 < SNR \lesssim 8$, per un numero complessivo di 481 ammassi identificati nel campo COSMOS. Durante il processo, sono state identificate anche le galassie membre degli ammassi assegnando loro un'appartenenza probabilistica al relativo ammasso. Durante la nostra analisi, da un punto di vista tecnico, abbiamo testato l'usabilità di AMICO per esempio nel caso degli alti *redshift* disponibili per il campo COSMOS e nel caso della sua ristretta area. Dunque, alcuni nuovi metodi sono stati introdotti in AMICO per consentirne l'applicazione a nuovi tipi di dati. Abbiamo introdotto un nuovo modo di generare una maschera ottica completa, abbiamo testato un metodo di regolarizzazione del rumore e introdotto un nuovo modo di attenuare il contributo degli ammassi al modello del rumore.

Grazie alla ricca copertura multi-banda del campo COSMOS, abbiamo identificato anche la controparte negli *X-rays* delle nostre identificazioni ottiche, usando i cataloghi disponibili per il campo COSMOS, prodotti da Gozaliasl et al. (2019) e da George et al. (2011) e usando la mappa di emissione a 0.5-2 keV (Gozaliasl et al., 2019). Il confronto è stato eseguito con lo scopo principale di testare la consistenza tra i cataloghi ricavati con AMICO e i cataloghi X. Questi ultimi sono stati anche usati per confrontare l'attendibilità dei cataloghi ottenuti durante le nostre due diverse analisi, considerando un comune catalogo di riferimento. Inoltre, questa analisi di confronto ci ha permesso di calibrare le relazioni degli indicatori di massa forniti da AMICO

(*amplitude* e ricchezza intrinseca) usando le stime di massa disponibili nei cataloghi X.

Contents

| | |
|--|-----------|
| Introduction | 1 |
| 1 Introduction to cosmology and galaxy clusters | 5 |
| 1.1 Cosmological framework | 6 |
| 1.2 Clusters and groups of galaxies | 11 |
| 1.2.1 Cosmology with galaxy clusters | 12 |
| 1.2.2 Cluster content and detection | 16 |
| 2 The AMICO algorithm | 21 |
| 2.1 Basic principles | 22 |
| 2.2 Application of the filter | 23 |
| 2.2.1 Amplitude variance | 25 |
| 2.2.2 Local background | 25 |
| 2.3 Input parameters | 26 |
| 2.4 Detection procedure | 30 |
| 2.4.1 Membership probability assignment | 32 |
| 2.4.2 Cleaning process | 33 |
| 2.5 Output | 33 |
| 3 The galaxy catalogue: COSMOS2015 | 35 |
| 3.1 The COSMOS field | 36 |
| 3.2 The COSMOS2015 catalogue | 37 |
| 3.2.1 Included data | 38 |
| 3.2.2 Object selection and photometric redshifts | 40 |
| 3.3 Selection of the input sample | 42 |
| 3.3.1 Flag selection | 42 |
| 3.3.2 FLAG_PETER analysis | 45 |
| 3.4 Final galaxy catalogue | 46 |

| | | |
|----------|--|------------|
| 4 | Applying AMICO to the COSMOS2015 galaxy catalogue | 51 |
| 4.1 | Generation of the visibility mask | 52 |
| 4.1.1 | Correction for UltraVISTA data | 56 |
| 4.2 | The cluster model | 58 |
| 4.2.1 | Radial distribution | 58 |
| 4.2.2 | Luminosity function | 60 |
| 4.3 | The noise assessment | 65 |
| 4.3.1 | Cluster galaxy imprint removal | 66 |
| 4.3.2 | Regularization of the noise | 70 |
| 5 | The catalogue of cluster candidates | 75 |
| 5.1 | Cluster candidates of Catalogue r | 76 |
| 5.1.1 | AMICO richness and mass proxies | 77 |
| 5.1.2 | Amplitude map and optical image inspection | 80 |
| 5.2 | Catalogue H: extending the cluster search up to $z \sim 1.8$ | 83 |
| 5.3 | Matching Catalogue r with Catalogue H | 87 |
| 6 | Matching with X-ray catalogues | 93 |
| 6.1 | COSMOS X-ray cluster catalogues | 94 |
| 6.1.1 | Redshift and membership assignment | 95 |
| 6.2 | Matching with AMICO catalogues | 96 |
| 6.2.1 | Catalogue H <i>vs</i> Gozaliasl+2019 | 98 |
| 6.2.2 | Catalogue r <i>vs</i> Gozaliasl+2019 | 99 |
| 6.2.3 | AMICO catalogues <i>vs</i> George+2011 | 100 |
| 6.3 | X-ray counterparts in the 0.5-2 keV emission map | 104 |
| 6.4 | Mass-observable scaling relations | 106 |
| 6.5 | Catalogue r <i>vs</i> Catalogue H in the X-rays | 109 |
| | Conclusions and future outlook | 113 |
| | A Further analyses on the galaxy catalogue | 117 |
| | B Minimum noise value | 121 |
| | C Examples of detected clusters | 123 |
| | Bibliography | 129 |

List of Figures

| | | |
|-----|--|----|
| 1.1 | The complementarity of cluster and CMB constraints on the cosmological parameters Ω_m and σ_8 for a Λ CDM cosmology and the sensitivity of the cluster cumulative mass function to the cosmological models. <i>Credits in caption.</i> | 15 |
| 3.1 | The COSMOS field with the limits of the areas covered by different surveys in the optical bands and in the near-IR. <i>Credits in caption.</i> | 39 |
| 3.2 | COSMOS2015 photometric redshift 1σ positive and negative errors as a function of photometric redshift for different intervals of i^+ magnitudes. <i>Credits in caption.</i> | 41 |
| 3.3 | The full COSMOS2015 catalogue, with 1182089 objects included. . | 43 |
| 3.4 | The positions of the flagged galaxies colour-coded according to FLAG_PETER values. | 45 |
| 3.5 | The magnitude distributions of all galaxies and the magnitude error as a function of magnitude for r -band and H -band galaxy catalogues. | 47 |
| 3.6 | The number density map of the final galaxy catalogues. Each pixel value indicates the number of galaxies falling into the area of the pixel. | 48 |
| 3.7 | The redshift distribution of the two catalogues used for the r -band analysis and for the H -band analysis in the range $0 < z < 2$ | 49 |
| 4.1 | The preliminary mask, as extracted by AMICO from the catalogue of masked objects. | 53 |
| 4.2 | The two masks retrieved from the masked objects and from the unmasked objects and the composite mask created adding the masked pixels of the former to the latter. | 55 |
| 4.3 | The final composite mask used for the H -band analysis. | 56 |

| | | |
|------|---|----|
| 4.4 | The redshift evolution of the characteristic magnitude m_* extracted from the Bruzual and Charlot (2003) CSP model used by Hennig et al. (2017). <i>Credits in caption.</i> | 62 |
| 4.5 | The redshift evolution of the parameter α for the red sequence galaxies of a cluster sample. <i>Credits in caption.</i> | 64 |
| 4.6 | The evolution in redshift of the absolute characteristic magnitude of red sequence galaxies of a cluster sample. <i>Credits in caption.</i> | 65 |
| 4.7 | The preliminary noise model in the 2D-space (z, r magnitude). | 66 |
| 4.8 | The composite mask generated with the aim of masking cluster areas and exclude member galaxies from the computation of the noise. | 67 |
| 4.9 | The COSMOS field divided in the 4 tiles used for the application of the <i>cleanNoise</i> method. | 68 |
| 4.10 | The final noise model computed by extracting the median of each pixel from the 4 maps of the 4 tiles | 69 |
| 4.11 | The final noise model computed without smoothing and with a smoothing of 2 pixels in z | 70 |
| 4.12 | Sketch of the regularized noise model. | 72 |
| 4.13 | The redshift evolution of the four filter constants: normalization, average background, cluster variance and local background variance correction factor. | 74 |
| 5.1 | Properties of the full Catalogue r. | 78 |
| 5.2 | Slice of the amplitude map and the cleaned amplitude map computed during the r -band run at $z = 0.465$ | 81 |
| 5.3 | Two examples of clusters detected in the r -band analysis. | 82 |
| 5.4 | Properties of the full Catalogue H. | 85 |
| 5.5 | Two examples of high-redshift clusters detected in the H -band analysis. | 86 |
| 5.6 | Separation in sky position and redshift of the successfully matched detections between Catalogue r and H. | 87 |
| 5.7 | The redshift and SNR distributions of clusters detected in the Catalogue r and the Catalogue H split in matched and non-matched. | 88 |
| 5.8 | Four examples of non-matched detections, present just in one of the two catalogues. | 91 |
| 5.9 | Three examples of matched detections, present in both catalogues. | 92 |

| | | |
|------|--|-----|
| 6.1 | Schematic of the different redshift ranges covered by the two X-ray cluster catalogues used for the comparison and the two catalogues produced within this study. | 97 |
| 6.2 | Angular positions in the sky of the detections of the 4 catalogues analyzed in the Chapter. | 98 |
| 6.3 | The separation distributions of the obtained successful matches between Catalogue H and Gozaliasl+2019. | 99 |
| 6.4 | The separation distributions of the obtained successful matches between Catalogue r and Gozaliasl+2019. | 100 |
| 6.5 | Examples of clusters found both in Catalogue r and in George+2011, helping in the comparison between membership assignments. | 101 |
| 6.6 | The distributions of matches and non-matches with the X-ray catalogue depending on X-ray flag values. | 103 |
| 6.7 | The combined XMM-Newton+Chandra 0.5-2 keV wavelet-filtered emission map overlapped with the full Catalogue r. | 105 |
| 6.8 | The combined XMM-Newton+Chandra 0.5-2 keV wavelet-filtered emission map overlapped with the full Catalogue H. | 106 |
| 6.9 | The X-ray virial mass $M_{vir,X}$ relation with amplitude and intrinsic richness for the Catalogue r and the Catalogue H. | 108 |
| 6.10 | The 0.1-2.5 keV rest frame X-ray luminosity (L_X) relation with amplitude and intrinsic richness for the Catalogue r and the Catalogue H. | 109 |
| 6.11 | The detection of a massive structure at the edge of the COSMOS field, with centres as identified during the r magnitude run and during the H magnitude run, compared with the X-ray emission peak. | 110 |
| 6.12 | The redshift distribution of the non-matched clusters from the comparison performed in Section 5.3 in comparison with X-ray detections. | 111 |
| A.1 | The presence of potential cluster galaxies among the masked objects according to the flag FLAG_PETER. | 118 |
| B.1 | Effect of the values chosen for the minimum threshold of the noise on the 4 filter constants at $z < 0.8$ | 121 |
| C.1 | A sample of 12 detections extracted from the resulting catalogue of the analysis performed with r -band photometry. | 124 |

| | | |
|-----|---|-----|
| C.2 | A sample of 12 detections extracted from the resulting catalogue of the analysis performed with <i>H</i> -band photometry. | 125 |
|-----|---|-----|

List of Tables

| | | |
|-----|--|-----|
| 3.1 | Summary of the final selected samples of objects from COSMOS2015, their properties and the selection constraints. | 50 |
| 5.1 | Summary of the samples of clusters detected within our analyses of the COSMOS field. | 90 |
| 6.1 | Summary of the successfully matched structures within this comparison with the X-ray available catalogues of clusters. | 103 |

Introduction

Galaxy clusters are peculiar objects whose importance in modern cosmology and in astrophysics has been well established.

As the most massive structures that had the time to form in the present-day Universe, clusters are tracers of the density peaks in the large-scale matter distribution and therefore they are strongly coupled with the expansion history of the Universe, as well as with the formation and evolution of structures.

Galaxy cluster detection is a field that has a long history and developed over the years. The most successful methods include, for instance observations of the hot and highly ionized gas clusters are embedded in, which makes them very bright X-ray sources (e.g. [Rosati, Borgani, and Norman, 2002](#); [Sarazin, 1986](#)) or the detection of distortions in the cosmic microwave background spectrum ([Sunyaev and Zeldovich \(1970\)](#) effect; e.g. [Bleem et al., 2015](#)).

Modern optical and near-infrared surveys also aim at pushing the limit of our knowledge of what the Universe is composed of and how it evolves, by exploiting the information that clusters encapsulate. Open questions of cosmology and modern physics, like the nature of dark matter and dark energy, can be explored with the use of clusters allowing to achieve a better understanding of our Universe. This requires robust methods to detect clusters and infer mass estimates.

In addition to matched filters (e.g. [Postman et al., 1996](#)), a variety of algorithms for the detection of galaxy clusters in optical and near-infrared surveys have been developed to contribute to the advances of cluster cosmology ([Allen, Evrard, & Mantz, 2011](#)), exploiting many different kinds of methods, using different galaxy properties and different assumptions (e.g. [Euclid Collaboration et al., 2019](#)).

In the study described in this Thesis work, galaxy clusters have been

detected using the AMICO algorithm (Adaptive Matched Identifier of Clustered Objects) (Bellagamba et al., 2011; Bellagamba et al., 2018). The extremely good performances of AMICO have recently placed it as one of the two algorithms for cluster selection officially adopted by the ESA *Euclid* mission¹ (Amendola et al., 2013; Euclid Collaboration et al., 2019; Laureijs et al., 2011). The AMICO algorithm has distinguished itself in the context of the *Euclid* Cluster Finder Challenge (CFC; Euclid Collaboration et al., 2019) in terms of *completeness* and *purity*, often used to define the quality of a cluster sample² (Allen, Evrard, & Mantz, 2011; Euclid Collaboration et al., 2019).

In the framework of this Thesis work, we performed a cluster detection with the AMICO algorithm in the 2deg² COSMOS field (Scoville et al., 2007). The COSMOS survey offers the possibility to access high-quality multi-wavelength data and this allowed us to extend our analysis and include a comparison between our detections and the ones contained in the two publicly available X-ray cluster catalogues for the COSMOS field (George et al., 2011; Gozaliasl et al., 2019).

The analysis presented in this Thesis work has a twofold purpose.

Our prime goal has been to produce a new and deep catalogue of galaxy clusters for the COSMOS 2deg² field up to $z \sim 2$ by running the AMICO code on the photometric galaxy catalogue COSMOS2015 (Laigle et al., 2016). We have made use of the *r*-band photometry to search for clusters up to $z \sim 1.25$ and extended the study with the use of the *H*-band photometry that provided results up to $z \sim 1.8$. Once we obtained, analyzed and compared the two catalogues resulting from this cluster detection, we performed a comparison with two X-ray catalogues. In the context of this comparison we tested the reliability of our catalogues and we retrieved a calibration of the AMICO mass proxy scaling relations with the available X-rays mass estimates.

Nevertheless, as additional purpose of this work, we aimed at making the AMICO code flexible to adapt also to such a depth and a reduced covered area. Applying the code to peculiar sets of data gives indeed the chance to test and introduce new features in the algorithm which allow its use in different contexts and improve its flexibility, in view of its future applications.

This Thesis work is structured according to the following outline:

¹<http://sci.esa.int/euclid/>

²they are indicators, with respect to a reference catalogue, of the amount of detections over the total true clusters and the true detections over the total detections, respectively.

- in [Chapter 1](#) an introductory background will be given about foundations of cosmology, galaxy clusters and their role as cosmological probes. The treatment is not intended to be complete but just aims at providing the reader with the preparatory concepts to follow the discussion of the following Chapters;
- in [Chapter 2](#) the AMICO algorithm will be presented and described, starting from its basic working principles, the methods, the relevant quantities, the required inputs and the provided outputs;
- in [Chapter 3](#) we will present the galaxy catalogue used in this study, the COSMOS2015 catalogue [Laigle et al. \(2016\)](#). After introducing the COSMOS field and the catalogue, the discussion will describe the selection and the cleaning of the data-set performed for the specific purpose of this analysis. We will eventually be able to present the selected galaxy samples used as input for our cluster detection;
- in [Chapter 4](#) the steps that have led to the application of AMICO to the COSMOS2015 catalogue will be discussed. The Chapter is divided in 3 Sections, each of them dedicated to an input needed by the AMICO algorithm to compute the filter: the visibility mask ([Section 4.1](#)), the cluster model ([Section 4.2](#)) and finally the noise model and its assessment ([Section 4.3](#)).
- in [Chapter 5](#) the results of this cluster detection will be presented. The catalogue of cluster candidates will be described in terms of statistical properties, provided mass proxies and membership assignment.
- in [Chapter 6](#) the multi-wavelength analysis will be presented. After briefly introducing the used X-ray catalogues ([George et al., 2011](#); [Gozaliasl et al., 2019](#)) we will discuss the statistical properties of the three-dimensional matching, the calibration of the scaling relations for the mass proxies and the test of the reliability of our analyses.

Eventually, we will conclude with some considerations about methods and results of this work and mention its possible future perspectives.

Chapter 1

Introduction to cosmology and galaxy clusters

Clusters of galaxies are the largest gravitationally bound structures in the present-time Universe. They are rare and non-linear structures, linked with very high sensitivity to the expansion history of the Universe and to the formation of structures. This makes clusters a well-established powerful cosmological probe, namely a test for the cosmological model that best describes our Universe, its properties and its evolution.

The mass budget of clusters has been shown to be represented by just a few percent of stars ($\sim 1 - 5\%$). The most massive luminous component is indeed the hot and rarefied gas clusters are embedded in. Nevertheless, it is long-time known that the dominant component in terms of mass in clusters is the dark matter component ($\sim 85 - 90\%$), which is detectable through its gravitational effects but emits no electromagnetic radiation ([Zwicky, 1933; 1937](#)).

In this Chapter, after an introduction to the essential concepts needed to construct the cosmological framework, we will move to the main character of this Thesis work: galaxy clusters. The focus of the discussion will be the detection and the measurement of the mass of these structures in view of their use in the context of cosmological studies.

It should be stressed that this Chapter is by no means a complete description neither of cosmology nor of galaxy clusters. This brief overview simply aims at introducing the essential information about the object of study of this Thesis work in view of its importance in modern cosmology. Nevertheless, clusters of galaxies are very interesting and rare objects, constituting

not only a powerful cosmological tool but also an important laboratory for the study of astrophysical processes and matter physics¹.

1.1 Cosmological framework

Cosmology originates from fundamental questions about the Universe as an object of study and aims at understanding its global behaviour as well as the formation and evolution of cosmic structures. The currently favoured cosmological model to describe the geometry, the composition and the evolution of the Universe is the so-called Λ CDM (Lambda Cold Dark Matter) model. This model predicts the presence, besides ordinary matter and radiation, of cold dark matter (CDM) and dark energy. The former is thought to be composed of non-relativistic particles which only interact gravitationally. The latter is a form of energy linked to the observed accelerated expansion of the Universe. Cosmological models are described by a set of quantities called cosmological parameters, which can be observationally constrained.

The assumption on which cosmological models in modern cosmology are based on is the so-called *cosmological principle*, namely that on scales typically larger than hundreds of Mpc ² the Universe is homogeneous and isotropic. Thus, there is no privileged position or direction: matter distribution and properties of the Universe look the same everywhere and in every direction. A fundamental evidence supporting the assumption of this principle is the observation of the nearly isotropic cosmic microwave background (CMB), the relic radiation of the origin of the Universe, measured for the first time by [Penzias and Wilson \(1965\)](#).

This is the principle, the foundations of the cosmological description of the Universe are based on: the theory that describes Gravity, the force field that dominates on large scales, the Theory of General Relativity (GR) by [Einstein \(1916\)](#), in addition to the mathematical background of models, based on the Friedmann-Lemaître-Robertson-Walker metric (e.g. [Friedmann, 1922](#); [1924](#); [Robertson, 1935](#)). The functional form of the metric, expressing an element

¹this Chapter is mainly based on the review work by [Allen, Evrard, and Mantz \(2011\)](#), the overview given by Chapters 1, 2, 6 and 7 of the book by [Cimatti, Fraternali, and Nipoti \(2019\)](#) and the treatment provided in Chapters 1, 2 and 12 of the book by [Dodelson and Schmidt \(2020\)](#). Additional references are provided within the text. For further details we refer the reader to the aforementioned review and books and the references therein.

²1 Mpc $\approx 3.086 \times 10^{24}$ cm

of space-time in spherical coordinates (r, θ, ϕ) is

$$ds^2 = c^2 dt^2 - a(t)^2 \left[\frac{dr^2}{1 - kr^2} + r^2(\sin^2 \theta d\phi^2 + d\theta^2) \right], \quad (1.1)$$

where c is the finite speed of light ³, t is the proper cosmic time, $a(t)$ is the scale factor that will be later on addressed and k is the curvature parameter defining the geometry. Observational evidences support the Euclidean geometry scenario ($k = 0$, flat Universe) (see e.g. [Planck Collaboration et al., 2020](#)).

Hubble-Lemaître law

An observational cornerstone of modern cosmology is the Hubble-Lemaître law first derived in the studies of [Lemaître \(1927\)](#) and [Hubble \(1929\)](#) on distances and velocities of a galaxy sample. These two quantities were found to be proportional with positive proportionality coefficient, meaning that increasing separation between galaxies implies increasing velocity in departing from each other. This has been interpreted as a sign of the expansion of the Universe. The proportionality coefficient between velocity and distance is the so-called Hubble parameter $H(t)$, whose value at the present time is referred to as H_0 . This recession, often called *Hubble flow*, is not to be intended as a movement of intrinsic galaxy positions in space but like a movement of objects fixed in space due to the expansion of space itself. It becomes therefore useful to make use of the co-moving observer concept and the associated co-moving distance x that does not depend on time, as the physical distance $r(t)$ does. The two are linked by a dimensionless time function called *scale factor* $a(t)$, as $x = r(t)/a(t)$. Therefore, the Hubble-Lemaître law reads:

$$v(t) = \frac{dr(t)}{dt} = x \frac{da}{dt} = \frac{da/dt}{a} r(t), \quad (1.2)$$

where the Hubble parameter is

$$H(t) \stackrel{\text{def}}{=} \frac{\dot{a}}{a}, \quad (1.3)$$

i.e. the ratio between the time derivative⁴ of the scale factor and the scale factor itself.

³ $c \approx 2.9979 \times 10^{10} \text{ cm s}^{-1}$

⁴by convention, we make use of the overdot to indicate derivative with respect to time.

The Hubble constant is expressed as a velocity over a distance, generally indicated with $H_0 = 100h \text{ km s}^{-1} \text{ Mpc}^{-1}$, being h a dimensionless constant also used to express observables in a parameter-independent way, since the precise value of H_0 remains object of an open debate (see e.g. [Freedman and Madore, 2010](#) for a review).

Cosmological redshift

Now that the context of an expanding, homogeneous and isotropic Universe has been set, a fundamental quantity can be introduced: the cosmological redshift.

Due to the expansion of the Universe, every extragalactic object can be identified by a redshift, z , given by of the rest-frame wavelength shift of any measured spectral feature of that object. This shift, for receding objects is towards longer wavelengths and is given by

$$z \stackrel{\text{def}}{=} \frac{\lambda_{obs} - \lambda_{em}}{\lambda_{em}} = \frac{1}{a(t_{em})} - 1, \quad (1.4)$$

where the subscripts refer to observation and emission. Redshift and scale factor mark distances and therefore specific cosmic times. High-redshift objects are those whose photons travelled the longest distances and were emitted the longest time ago from today. Redshift is the most commonly used observable related to distance for extragalactic distant objects.

All the considerations made above hold true only in case of sufficiently distant extragalactic objects, since on small scales the expansion of the Universe may not dominate over intrinsic proper motion. A clear example with respect to an observer in our Galaxy is Andromeda, a spiral galaxy that is actually approaching us and is therefore characterized by a slight blueshift ($z < 0$) being its proper motion with respect to us dominant over the expansion of the Universe on such a small scale.

A distance expression can be defined by considering the so-called *luminosity distance* d_L starting from the relation between measured flux F and intrinsic luminosity L of a cosmological source: $F = \frac{L}{4\pi d_L^2}$. Thus, the luminosity distance can be written as

$$d_L \stackrel{\text{def}}{=} (1 + z)x, \quad (1.5)$$

where x is the comoving distance. Given the intrinsic and measured size of a source, l and θ , respectively, the *angular diameter distance*, d_A , can be

defined as

$$d_A = \frac{l}{\theta}. \quad (1.6)$$

The two distances are linked by the relation

$$d_A = \frac{d_L}{(1+z)^2}, \quad (1.7)$$

in a way that

$$d_A = \frac{x}{1+z}. \quad (1.8)$$

Friedmann equations and Λ

Once it has been observed, the expansion of the Universe requires the introduction of evolutionary equations for the scale factor that quantify how the space scales with time given its geometry. This description is expressed by the Friedmann equations, based on General Relativity equations, on the cosmological principle and on the modelling of matter as a fluid. The two equations describing $a(t)$ are⁵:

$$\dot{a}^2 = \frac{8\pi G}{3}\rho a^2 - kc^2, \quad (1.9)$$

$$\ddot{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right)a, \quad (1.10)$$

where G is the gravitational constant⁶ and ρ and p are the density and the pressure of the fluid, respectively. Combining the Friedmann equations (1.9, 1.10), the *fluid equation* can be obtained as

$$\dot{\rho} + \frac{3\dot{a}}{a}\left(\rho + \frac{p}{c^2}\right) = 0. \quad (1.11)$$

If we choose one of the three equations, this is not independent of the others, so two of them are sufficient for the description of the dynamics of $a(t)$.

At the time of the formulation of the Friedmann equations, the Universe was thought to be static. Nevertheless, the hypothesis of static Universe

⁵for simplicity we make the time dependence of $a(t)$ implicit.

⁶ $G \approx 6.67 \times 10^{-8} \text{ cm}^3 \text{ g}^{-1} \text{ s}^{-2}$

together with Equation 1.9, leads to the *unphysical* implication that density and pressure must have different signs. To make up for this inconsistency Einstein introduced already in 1917 a term Λ , to modify his gravity theory. We have evidence today of the non-static nature of the Universe, but the Λ constant is nevertheless still used to put into play a new component of the Universe that would explain its accelerated expansion, as first proved by distant supernova observations (Perlmutter et al., 1999; Riess et al., 1998). It is nowadays common in the most used and accepted cosmological frameworks to identify the cosmological constant Λ with a form of *dark energy*, of still unknown physical nature.

In case we assume flat geometry as suggested before and thanks to the definition of $H(t)$ in Equation 1.3, the first Friedman equation (1.9) can be simply rewritten as

$$H^2(t) = \frac{8\pi G}{3}\rho, \quad (1.12)$$

which by isolating the mass density gives the expression for the so-called *critical density*

$$\rho_{crit} \stackrel{\text{def}}{=} \frac{3H^2(t)}{8\pi G}, \quad (1.13)$$

which has a present-time value of $\sim 10^{-29}h^2 \text{ g cm}^{-3}$. This is the characteristic density that defines a Universe with Euclidean geometry.

The critical density constitutes the reference to build the *density parameter*. This parameter is a useful quantity to describe the amount of density, and therefore energy, that is contributed by each component the Universe is made of, today and with time dependence:

$$\Omega_s \stackrel{\text{def}}{=} \frac{\rho_s(t)}{\rho_{crit}(t)}. \quad (1.14)$$

The s subscript indicates the different components of the Universe: baryonic and dark matter (m), photons (γ), cosmological constant (Λ). The parameter of this latter reads:

$$\Omega_\Lambda = \frac{\Lambda c^2}{3H(t)}, \quad (1.15)$$

and dominates the cosmological landscape in today Universe.

The Friedmann equations for a flat Universe, rewritten by making use of the density parameters of the different components and their scaling with redshift z reads:

$$E^2(z) = \frac{H^2(z)}{H_0^2} = \left[\Omega_{m,0}(1+z)^3 + \Omega_{\gamma,0}(1+z)^4 + \Omega_{\Lambda,0} \right], \quad (1.16)$$

where the *expansion function* has been introduced as $E(z) \equiv H(z)/H_0$.

Concordance cosmological model

The currently favoured adoption for the cosmological framework is the co-called *concordance* or *standard* cosmological model, generally referred to as Λ CDM cosmology, to indicate the role of the cosmological constant and of cold dark matter (CDM). At the present time, a recent determination of present-day values of cosmological parameters in a standard cosmology Universe has been presented by [Planck Collaboration et al. \(2020\)](#) and obtained in the context of the *Planck* space mission⁷. The results are consistent with a flat Universe with $h_0 = 0.674 \pm 0.005$, $\Omega_{m,0} = 0.315 \pm 0.007$ and $\Omega_{\Lambda,0} = 0.685 \pm 0.007$ completing the budget.

In the context of this Thesis, we have set the full analysis in a standard Λ CDM cosmological background with $\Omega_{m,0} = 0.3$, $\Omega_{\Lambda,0} = 0.7$ and $h = 0.7$.

1.2 Clusters and groups of galaxies

The majority of the galaxies in the Universe tend to aggregate and be found in gravitationally bound structures rather than being isolated. These structures can include different amounts of galaxies spanning from a few of them to a few thousands. In literature, aggregations of galaxies are commonly called *clusters* or *groups*. The distinction between the two is not univocally defined by a clear threshold, but it usually consists in a difference in mass or richness, i.e. number of galaxies, of the structure. Assemblies of galaxies that have characteristic masses lower than 10^{14} solar masses⁸ and host less than ~ 50 member galaxies have characteristic size $\lesssim 1 \text{ Mpc}$ and are usually referred to as *groups*. It is instead common to name as *clusters* all the

⁷<https://www.esa.int/Planck>

⁸1 solar mass = $1 \text{ M}_{\odot} \approx 2 \times 10^{33} \text{ g}$

gravitationally bound structures that are more extended, more massive and richer in galaxy content with respect to groups.

In the framework of this Thesis work, we will have to deal with galaxy aggregations with the characteristics of both classes of objects, being most of the times at the edge of the commonly adopted thresholds to discriminate between the two.

Thus, for the sake of simplicity we will make use of the two terms interchangeably, typically favouring the more general term *cluster*.

1.2.1 Cosmology with galaxy clusters

As it is by now well known, galaxies formed and aggregate, demonstrating the Universe is well beyond being devoid of inhomogeneities. This has been jointly shown by observations, e.g. the SDSS galaxy survey⁹ (Blanton et al., 2003; Blanton et al., 2017; Zehavi et al., 2011) and N-body simulations like the *Millennium* simulation (Springel, Frenk, & White, 2006), displaying the large-scale structure of the Universe (LSS), i.e. how galaxies are distributed on large scales, which is fundamental to understand the history and the evolution of our Universe. In the current standard scenario, it is thought that the formation of structures has been driven by the growth of initial density perturbations of the dark matter component, under the effect of gravity (Bond, Kofman, & Pogosyan, 1996). This growth can be initially treated as a linear evolution but if the initial overdensity becomes sufficiently large the evolution cannot be rigorously analytically described (one has to resort to numerical simulations or analytical approximations). These small-scale perturbations have grown therefore non-linear and successively clumped together in a hierarchical way to form larger structures.

The largest structures that had the chance to form up to now, in the present-day Universe, are clusters of galaxies.

By holding this peculiar position in the large-scale structure, galaxy clusters are very powerful tools for cosmological studies, only weakly dependent on baryon physical processes but closely coupled with extreme sensitivity to the expansion history of the Universe and to structure formation (Allen, Evrard, & Mantz, 2011).

The gravitational amplification of the small fluctuations of the primordial Universe can be characterized by the evolution of a primordial power-law

⁹for further details see <https://www.sdss.org> and e.g. Stoughton et al. (2002).

power spectrum of fluctuations:

$$P_{prim}(k) \sim k^{n_s}, \quad (1.17)$$

where k is wavenumber and n_s the spectral index, proved to be consistent with a Harrison-Zeldovich spectrum ($n_s = 1$) (Harrison, 1970; Planck Collaboration et al., 2020; Zeldovich, 1972) and the expectations of simple inflationary models¹⁰ (e.g. Baumann and Peiris, 2008). The power spectrum of cosmological perturbations evolves with the cosmic expansion as

$$P(k, a) = D^2(a)T^2(k)P_{prim}(k) \quad (1.18)$$

where the expansion comes into play via the squared dependence on D , the *growth factor* of density fluctuations and where $T(k)$ is a transfer function carrying information about the evolution elapsed between the end of inflation and the equivalence.

By smoothing the power spectrum in Equation 1.18 in the k -space one can obtain $\sigma^2(M, a)$, the mass variance of the dark matter fluctuations that evolved linearly. If evaluated on scales of 8 Mpc/ h and in the present-time Universe, this quantity yields a well-known cosmological parameter, the power spectrum normalization σ_8 .

Nevertheless, it is nowadays well established that most of the matter composing the Universe is not visible and therefore the spacial clustering of observable matter is not directly described by the matter power spectrum but by a modified or *biased* version of it (Kaiser, 1984):

$$P_{obs}(k, a) = b^2(M, a)P_m(k, a), \quad (1.19)$$

where $b(M, a)$ accounts for this modification, commonly referred to as *bias factor* and depends on mass and cosmic epoch. Due to their position and role in the LSS, galaxy clusters are highly biased, namely they have large bias factors, in such a way that the cluster power spectrum results strongly enhanced (up to several tens amplification) over the matter spectrum (Allen, Evrard, & Mantz, 2011; Desjacques, Jeong, & Schmidt, 2018; Tinker et al., 2010) and therefore strongly coupled with the cosmic growth history.

A fundamental element of the modeling given by the Press-Schechter formalism (Press & Schechter, 1974), the established mathematical background

¹⁰suggesting the existence of *inflation*, an epoch in the very early Universe ($\sim 10^{-36} - 10^{-34}$ s) during which the Universe expanded exponentially fast

of the statistical study of the distribution of structures, together with its developments (like Bond et al., 1991; Bond, Kofman, and Pogosyan, 1996; Sheth and Tormen, 1999 among the others) is the so-called *mass function*. This function, expressing the mean number density of a population of structures per co-moving volume is given by

$$\frac{dn}{d \ln M} = \frac{\bar{\rho}}{M} f(\sigma) \left| \frac{d \ln \sigma}{d \ln M} \right|, \quad (1.20)$$

where $\bar{\rho}$ is the mean density of matter, σ the aforementioned mass variance and $f(\sigma)$ a model-dependent function of it, also known as *multiplicity function*. Calibrations and dependencies are generally tuned relying on numerical simulations (e.g. Despali et al., 2016; Shirasaki, Ishiyama, and Ando, 2021).

In order to connect theory and observations, we can now introduce the actual counts of clusters, main goal of most of modern wide surveys, expressed as N , the number of clusters expected for a bin in mass and redshift, in a solid angle Ω :

$$N(M, z) = \frac{\Omega}{4\pi} \int dz \frac{dV}{dz} \int d \ln M \frac{dn}{d \ln M}, \quad (1.21)$$

where the cosmological background comes into play both in the $\frac{dV}{dz}$ term, the co-moving volume element that has a dependence on the scale factor and of course with the just introduced (1.20) *mass function* $\frac{dn}{d \ln M}$. This latter is sensitive to cosmology thanks to the dependence on the mean density, i.e. on the Ω_m parameter and through the variance σ which encapsulates the linear power spectrum, and its defining parameters such as index and normalization (and therefore σ_8). Since the mass M is not a direct observable, in Equation 1.21 it is more appropriate to see N as a function of an actual observable and consequently take into account in the right-hand-side a relation that links this observable to the mass. The determination of this relation (the *mass-observable relation*) is a crucial point for cluster studies and involves the astrophysical processes taking place within the system (Pratt et al., 2019; Singh et al., 2020).

Galaxy clusters, as the most massive bound structures in the LSS, define the high-mass end of the mass function and are therefore extremely sensitive to the choice of cosmological model.

This makes them a unique tool to constrain cosmological parameters, in a complementary way with respect to other cosmological probes such as CMB

observations, galaxy clustering and weak lensing measurements. The best constrained parameters from cluster studies (like as we discussed Ω_m and σ_8) (e.g. [Abdullah, Klypin, and Wilson, 2020](#); [Lesci et al., 2020](#); [Sartoris et al., 2016](#)) have degeneracies that are different with respect to other methods and therefore they help shedding light on parameter measurements and uncertainties (e.g. [Rosati, Borgani, and Norman, 2002](#); [Rozo et al., 2010](#); [Vikhlinin et al., 2009](#)). This can be appreciated thanks to the examples in [Figure 1.1](#). On the left panel the complementarity between CMB and cluster constraints on the cosmological parameters Ω_m and σ_8 for a Λ CDM cosmology is shown: in this case the combination of the two yields an improvement of nearly a factor of two with respect to the CMB results only ([Rozo et al., 2010](#)). On the right panel the sensitivity of massive cluster mass function to the cosmological choice of density parameters is shown, especially evident with increasing redshift: cluster abundance constraints are particularly sensitive to the density parameter Ω_m , in a more significant way with respect to the Ω_Λ parameter ([Rosati, Borgani, & Norman, 2002](#)).

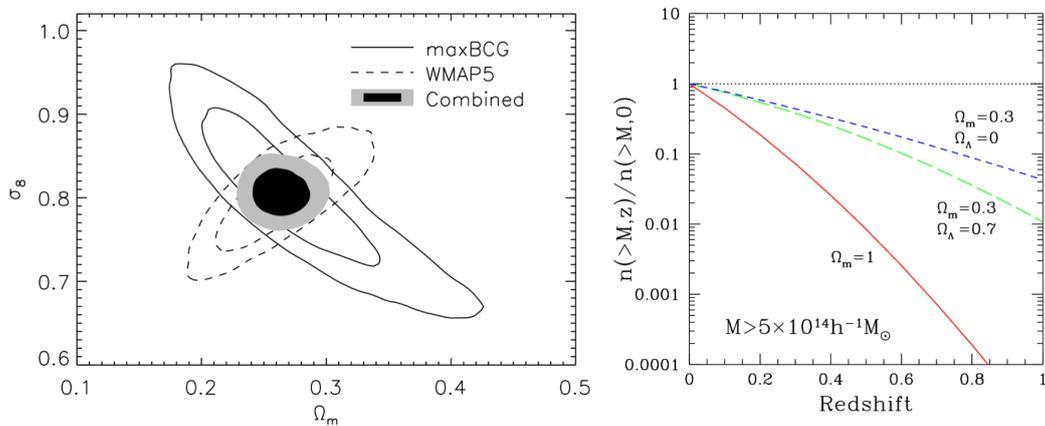


Figure 1.1: *Left:* the complementarity of cluster and CMB constraints on the cosmological parameters Ω_m and σ_8 for a Λ CDM cosmology. Solid lines indicate the 68% and 95% confidence regions constrained in the space $\sigma_8 - \Omega_m$ by [Rozo et al. \(2010\)](#) with the maxBCG cluster catalogue ([Koester et al., 2007](#)). Dashed lines indicate the same but for CMB constraints, data from the WMAP 5-year results ([Dunkley et al., 2009](#)). The filled black and grey regions mark the combined constraints. *Credits:* [Rozo et al. \(2010\)](#).

Right: The sensitivity of the cluster cumulative mass function to the cosmological models. Redshift evolution of the mass function for massive clusters and for three different sets of parameters: $\Omega_m = 1$ (solid red line), $\Omega_m = 0.3$ with $\Omega_\Lambda = 0.7$ (dashed green line) and $\Omega_m = 0.3$ with $\Omega_\Lambda = 0$ (blue dashed line); $\sigma_8 = (0.5, 0.8, 0.8)$ respectively. *Credits:* [Rosati, Borgani, and Norman \(2002\)](#)

1.2.2 Cluster content and detection

In light of the importance of clusters in cosmological studies briefly discussed above, this Section will move the attention on the practical key points of exploiting galaxy clusters: how it is possible to detect and study them given their main characteristics.

Galactic component

One of the first and most relevant catalogues of galaxy clusters dates back to 1958, when Abell listed more than 2 thousands galaxy clusters by examining the observation results of the National Geographic Society - Palomar Observatory Sky Survey (1959) and by identifying clusters as overdensities of galaxies. The main source of optical emission from clusters is indeed the stellar emission of the galactic component.

Abell (1958) described the identified structures as "rich clusters" classifying them according to specific criteria, for instance by considering galaxies brighter than $m_3 + 2$, being m_3 the third brightest galaxy magnitude of the cluster and within a characteristic radius known as *Abell radius*. Since then the optical identification and classification of clusters have followed the same approach: overdensities of galaxies in the optical surveys are identified by making use of the sky position but also of the redshift, to eliminate projection effects of galaxies that appear part of the structure but are instead distant, and by classifying them depending on properties such as the galaxy content, generally used as indicator of mass.

Cluster galaxies exhibit specific properties regarding colours and star formation rate. Elliptical and S0, i.e. the so-called early-type galaxies usually dominate the cluster population, especially in the central regions and in structures with regular shape. On the contrary star-forming galaxies, i.e. spirals and irregulars, tend to dominate the field galaxy population, namely the one located in non-overdense regions, unlike clusters or groups. This bi-modality in the distribution of galaxy morphology (and therefore colours) has been found indeed to correlate with the density of the environment in which galaxies are located (Dressler, 1980). Small-scale environmental high density, influences the evolution of galaxies that are more likely to undergo galaxy-galaxy interactions or galaxy-gas interactions.

The central region of a galaxy cluster is normally dominated by a giant elliptical, the most luminous galaxy in the cluster (among the most luminous

ones in the Universe), also called *BCG* (brightest cluster galaxy), possibly the aftermath of subsequent accretions of smaller galaxies. The term *cD* or *central dominant* is also used to refer to the central object dominating the centre of the cluster, whenever it is characterized by a diffuse and extended envelope of stars.

Environmental effects can show their influence not only in optical colours, but also in radio wavelengths. Radio emitting galaxies connected with an active nucleus are often harboured by clusters, exhibiting peculiar structures like bent and deformed radio jets due to the interaction with the surrounding medium (e.g. [Smolčić et al., 2007](#)). The presence of emission which is not attributable to active nuclei, i.e. to point-like sources, is generally referred to as radio diffuse emission (e.g. radio halos and relics) and represents a very interesting discovery. This emission, attributable to the synchrotron radiation mechanism¹¹, requires the presence of magnetic fields and sufficiently energetic particles in the emission locations and therefore represents an important laboratory for the investigation of underlying physical processes ([van Weeren et al., 2019](#)).

As already mentioned, detection of galaxy clusters in the optical (and near-infrared) range of the electromagnetic spectrum has a long history and has matured over the years. Ongoing and future wide-field optical and near-IR surveys are going to aim at the achievement of a new and deep understanding of many astrophysical processes and cosmological open questions, exploiting detected galaxy clusters ([Euclid Collaboration et al., 2019](#)). A variety of detection algorithms are presents nowadays in literature. Starting from the development of matched filter methods (e.g. [Postman et al., 1996](#)), many different techniques have been developed over the years to contribute to the modern era of optical cluster cosmology ([Allen, Evrard, & Mantz, 2011](#); [Euclid Collaboration et al., 2019](#)). A few examples are the *BCG* methods (e.g. [Koester et al., 2007](#)) the *Voronoi tessellation* methods (e.g. [Ramella et al., 2001](#)), *friends-of-friends* algorithms (e.g. [Wen, Han, and Liu, 2012](#)), *wavelet filtering* techniques (e.g. [Gonzalez, 2014](#)), together with a series of successful applications of the *cluster red-sequence method* (e.g. [Rykoff et al., 2014](#)). All of these methods and many others have obviously advantages and drawbacks, different properties needed and different ranges of reliability ([Euclid Collaboration et al., 2019](#)).

¹¹due to the motion of relativistic particles around magnetic field lines.

X-ray emission from ICM

Despite galaxy clusters are optically localized via the emission of the galactic component, in terms of mass the dominant baryonic component is the hot and highly ionized plasma they are embedded in. This abundant gas is referred to as intracluster medium (ICM) and it has typical densities spanning between a few 10^{-2} and 10^{-4} cm^{-3} with temperatures of $10^7 - 10^8$ K (Molendi, 2004).

Galaxy clusters are extremely bright X-ray sources and this has allowed their observation in this region of the electromagnetic spectrum ever since the first X-ray extragalactic surveys were permitted (Giacconi et al., 1972; Rosati, Borgani, & Norman, 2002; Sarazin, 1986). This diffuse X-ray emission is attributable to the gaseous component, the ICM, that is confined into the gravitational potential well of the galaxy cluster halo. Given its high temperatures, the ICM is responsible of a thermal free-free bremsstrahlung emission¹², the main origin of typical luminosities of the order of $10^{43} \text{ erg s}^{-1} \leq L_X \leq 10^{45} \text{ erg s}^{-1}$ (Rosati, Borgani, & Norman, 2002; Sarazin, 1986).

Observationally speaking, the X-ray emission of galaxy cluster is a powerful tool to detect them, to constrain their position and size, to associate central members by searching for optical counterparts and to retrieve information about their morphology. Moreover, X-ray observations also give the opportunity to carry out spectral analyses that shed light on the density and temperature distributions of clusters and their metal content (thanks to the presence of emission lines in the X-ray spectra).

The main advantages of X-ray identification of clusters are the availability of sufficiently precise and simple relations between observables (e.g. luminosity, temperature) and mass (e.g. Mantz et al., 2010) and the fact that it is a technique nearly insensitive to projection effect, for the nature of the emission itself (Allen, Evrard, & Mantz, 2011).

Sunyaev-Zeldovich effect

Galaxy cluster detection in the context of X-ray and optical surveys has a long history. Nevertheless, these are not the only possibilities, and further techniques have been developed over the years.

Galaxy clusters are also detectable in the microwaves (e.g. Bleem et al., 2015), thanks to the so called Sunyaev-Zeldovich Effect, SZ for short (Sun-

¹²due to the deceleration of electrons interacting with ions in an ionized plasma.

yaev & Zeldovich, 1970). This is how the Inverse Compton scattering taking place between the CMB photons and the free electrons of the ICM is known in literature. The CMB radiation is well described by a black-body emission with temperature $T_{CMB} \simeq 2.726$ K (Fixsen, 2009). The CMB photons are much less energetic than the cluster electrons, so that the interaction between the two and the consequent energy gain performed by the CMB photon become visible as a distortion in the black-body CMB spectrum at the respective cluster location. In particular this distortion consists in a shift with a consequent lower intensity for frequencies $\nu < 218$ GHz and higher intensity for $\nu > 218$ GHz.

The big advantage of SZ detection method is that it is very weakly dependent on redshift, being the source of photons the CMB, so it is particularly suited for the detection of massive high- z clusters, but is less sensitive to low masses than X-ray and optical detections (Allen, Evrard, & Mantz, 2011). The main observable is the Compton parameter y , which is proportional to gas density and temperature.

Gravitational lensing

Modern survey data and the development of suited algorithms opened the possibility to search for galaxy clusters and infer their mass thanks to the distortion of background sources whose emitted photons have bent according to GR, because of the interposition of a massive object (e.g. Shan et al., 2012; Stapelberg, Carrasco, and Maturi, 2019).

The big advantage of gravitational lensing is that it provides a way to infer the cluster mass that does not depend on the physical state of the system, as for instance for X-ray measurements (Bartelmann, 2010). Strong and weak lensing effects and combinations of the two have provided over the years, with the use of different techniques, precise measurements of cluster masses and good constraints on mass-observable scaling relations (e.g. Bellagamba et al., 2019; Leauthaud et al., 2010).

High redshift clusters

As already mentioned, galaxy clusters are also very important laboratories for the study of astrophysical processes including those underlying the formation and evolution of galaxies. This is the reason why, once a cluster has been detected it is crucial to have a reliable method to associate mem-

ber galaxies to the cluster to then study their properties. This is how the correlation between galaxy morphology, colour and star formation activity can be related to the environment as previously discussed. The dominance of red, early-type galaxies in clusters has been shown to characterize clusters at least up to $z \sim 1$ (e.g. [George et al., 2011](#)). Nevertheless, this might not be strictly true for more distant structures, that have been shown to harbour star-forming rates more consistent to that of the field as well as exhibiting more irregular shapes: according to several studies, e.g. [Alberts et al. \(2016\)](#) and [Brodwin et al. \(2013\)](#), $z \sim 1.4$ could be drawn as a threshold.

High- z cluster detection (z up to $\sim 1.5 - 2$) has been possible by making use of different techniques and observables. In optical and near-IR surveys aiming at detecting high- z clusters, it is particularly important to take into account the evolution of the galactic population and the effect of redshift to make a consistent choice of techniques and filters to be used (e.g. [Strazzullo et al., 2016](#); see [Overzier, 2016](#) for a review).

Even if there is no univocal definition of it in the literature, the term *protocluster* is used to refer to structures thought to be the progenitors of the less distant structures that we observe today, often still undergoing aggregation processes ([Muldrew, Hatch, & Cooke, 2015](#)). The study of protoclusters represents an interesting benchmark to investigate the processes underlying cluster evolution.

Chapter 2

The AMICO algorithm

AMICO (Adaptive Matched Identifier of Clustered Objects) (presented and described in [Bellagamba et al., 2011; 2018](#)) is an algorithm developed for the detection of clustered objects among a set of data, such as galaxy clusters or groups in a photometric catalogue of galaxies. The algorithm is based on a technique that is able to extract a specific signal from a set of data affected by noisy background, allowing to maximise the signal-to-noise ratio of the clustered object. This technique is known as Optimal Filtering.

Further descriptions and applications of this kind of technique can be found in literature, e.g. in [Maturi et al. \(2005\)](#).

The working principle of AMICO corresponds to that of a Matched Filter i.e., in case of photometric detections, an Optimal Filter when the background is assumed to be homogeneous (as it was shown in [Bellagamba et al., 2011](#)).

In Matched Filters, the signal can be identified as a resemblance to an a-priori model given for the expected signal to detect. In the framework of galaxy clusters or groups detection, this model can be for instance defined by a radial profile and a luminosity function and can vary with the redshift.

An important advantage of Optimal Matched Filtering formalism with respect to other kinds of methods is its extreme flexibility. Instead of being tailored for a specific data-set type, it can fit for different surveys in a wide range of redshifts and it allows to compare and consistently combine optical data with gravitational lensing, S-Z and X-ray observables. The algorithm itself chooses the set of properties from those which are available in the data-set and weights the elements of the catalogue according to these properties' relevance in the cluster detection ([Bellagamba et al., 2018](#)).

Another important feature of the current version of the AMICO Optimal Filter is the ability of the algorithm to extract and remove in an iterative way the imprint of a detected cluster in order to allow the detection of lower SNR (signal-to-noise ratio) objects that may be located in the surrounding area. This deblending technique is of course a point of crucial importance especially in deep surveys, characterized by a high number of detected clusters per square degree.

In this Chapter, the AMICO algorithm will be introduced starting from the general principles on which it is based. This will give the opportunity to see which quantities are playing a role in the construction of the filter and which input parameters are needed by the algorithm. Eventually, the actual detection procedure and the provided outputs will be shortly presented.

2.1 Basic principles

The basic working principle of Optimal Filtering resides in the assumption that a data set can be modelled as the sum of a signal component and a noise component. In the context of photometric catalogues of galaxies the data component is the galaxy density $D(\boldsymbol{\theta}, \mathbf{m}, z)$, a function of the angular position $\boldsymbol{\theta}$, the redshift z and an array of properties \mathbf{m} . This array consists most of the times in a single photometric magnitude or more than one, but can potentially contain any other available property of catalogue galaxies, such as shape or morphological type. The galaxy density can be therefore written as follows:

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

Thus, the signal component given by the galaxies belonging to the cluster is the model $M_c(\boldsymbol{\theta} - \boldsymbol{\theta}_c, \mathbf{m}, z)$, i.e. the expected distribution of galaxies in the cluster centered in $(\boldsymbol{\theta}_c, z_c)$, properly normalized with the normalization factor $A(\boldsymbol{\theta}_c, z_c)$, namely the *amplitude* of the cluster. Instead, $N(\mathbf{m}, z)$ is the noise component associated with the contribution given by field galaxies.

The signal amplitude A can be seen as the normalization factor that accounts for the number of galaxies belonging to the cluster and it can be estimated as a convolution between the data D and a kernel extracted via constrained minimization approach. The convolution with this *filter* function Ψ generates an unbiased and minimum variance estimate of the number of galaxies belonging to the cluster:

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

where B is the background quantified as the average contribution of field galaxies to the signal amplitude (it has to be subtracted), n is simply the dimension of the given array of galaxy properties \mathbf{m} and α is a normalization constant defined as follows:

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

This normalization constant is introduced 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 of assumption for the noise to be homogeneous and given by random Poissonian counts, the optimal filter can be expressed as the ratio between the behaviour of the cluster properties, i.e. the model, and the noise itself. In this way, in case the model is dominant with respect to the noise, the field galaxies are automatically suppressed by the filter:

$$\Psi_c(\boldsymbol{\theta} - \boldsymbol{\theta}_c, \mathbf{m}, z) = \frac{M_c(\boldsymbol{\theta} - \boldsymbol{\theta}_c, \mathbf{m}, z)}{N(\mathbf{m}, z)}. \quad (2.4)$$

If dealing with catalogues of galaxies, of course, the amplitude signal can be discretized by running over all the i -th elements of the catalogue and equation 2.2 can be rewritten as follows:

$$A(\boldsymbol{\theta}_c, z_c) = \alpha^{-1}(z_c) \sum_{i=1}^{N_{gal}} \frac{M_c(\boldsymbol{\theta}_i - \boldsymbol{\theta}_c, \mathbf{m}_i) p_i(z_c)}{N(\mathbf{m}_i, z_c)} - B(z_c), \quad (2.5)$$

where the redshift dependence of the model has dropped and is now contained in the photometric-redshift distribution for each i -th galaxy $p_i(z)$, being negligible with respect to the galactic redshift uncertainty.

2.2 Application of the filter

In order to compute some of the important quantities that play a role in this first phase of an AMICO identification process, a fundamental function has to be introduced: the typical redshift probability distribution for a galaxy at a given redshift (e.g. the one of the cluster candidate, z_c) (2018).

This distribution $q(z_c, z)$, can be either known a-priori or retrieved directly from the data-set as:

$$q(z_c, z) = \frac{\sum_{i=1}^{N_{gal}} p_i(z - z_c + z_{peak,i}) p_i(z_c)}{\sum_{i=1}^{N_{gal}} p_i(z_c)}, \quad (2.6)$$

with $z_{peak,i}$ being the redshift with the highest probability for the i -th galaxy.

The expression in Equation 2.6 appears in the form of a weighted mean in which the probability distribution of each galaxy is weighted by its value in correspondence of z_c and then moved in such a way that it peaks there.

This function can be computed as an overall property for every object of the data-set or as a function of a property, dividing the data in subsamples, for instance according to different magnitudes to capture differences in the accuracy due to photometric quality (Maturi et al., 2019).

Once the typical redshift probability distribution has been introduced, other fundamental quantities in the application of the filter can be defined.

The first one is the normalization constant $\alpha(z_c)$, that was already introduced and defined in Equation 2.7, and can now be written as

$$\alpha(z_c) = \int \frac{M_c^2(\boldsymbol{\theta} - \boldsymbol{\theta}_c, \mathbf{m}, z_c) q^2(z_c, z)}{N(\mathbf{m}, z_c)} d^2\theta d^m m dz. \quad (2.7)$$

The second one is, for instance, $\beta(z_c)$, the average background contribution to the filter, namely the expectation value of the summation in Equation 2.5 in case the galaxy distribution is entirely generated by field galaxies. This quantity corresponds to:

$$\beta(z_c) = \int \Psi_c(\boldsymbol{\theta} - \boldsymbol{\theta}_c, \mathbf{m}, z_c) q(z_c, z) N(\mathbf{m}, z_c) d^2\theta d^m m dz = \quad (2.8)$$

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

This shows that, for how it was constructed, $\beta(z_c)$ is also the number of galaxies belonging to the cluster model at that specific redshift, z_c .

The background filter constant $\beta(z_c)$, is used to compute the estimate of the background that has to be subtracted in Equations 2.2 and 2.5, and it is done as follows:

$$B(z_c) = \alpha^{-1}(z_c) \beta(z_c), \quad (2.10)$$

where α is the normalization filter constant, introduced before.

Subtracting this quantity guarantees that whenever in a certain position the density of galaxies corresponds to that of the field, the amplitude is zero.

2.2.1 Amplitude variance

A third constant that contributes to the construction of the filter is the cluster variance constant $\gamma(z_c)$ given by

$$\gamma(z_c) = \int \Psi_c^2(\boldsymbol{\theta} - \boldsymbol{\theta}_c, \mathbf{m}, z_c) q^3(z_c, z) M_c(\boldsymbol{\theta} - \boldsymbol{\theta}_c, \mathbf{m}, z_c) d^2\theta d^m m dz = \quad (2.11)$$

$$= \int \frac{M_c^3(\boldsymbol{\theta} - \boldsymbol{\theta}_c, \mathbf{m}, z_c) q^3(z_c, z)}{N^2(\mathbf{m}, z_c)} d^2\theta d^m m dz. \quad (2.12)$$

This quantity contributes to analytically estimate the uncertainty, the expected r.m.s. of the amplitude, which is

$$\sigma_A^2(\boldsymbol{\theta}_c, z_c) = \alpha^{-1}(z_c) + A(\boldsymbol{\theta}_c, z_c) \frac{\gamma(z_c)}{\alpha^2(z_c)}, \quad (2.13)$$

where the two terms summed together account for the stochastic background fluctuations and for Poissonian fluctuations produced by galaxy members of a cluster with amplitude A as retrieved by the filter.

2.2.2 Local background

According to the principle the AMICO algorithm is based on, the definition of the filter has worked so far under the assumption of a uniform noise $N(\mathbf{m}, z_c)$ over the surveyed area. As it has already been mentioned, the noise, i.e. the population of field galaxies, in this way contributes to the amplitude with random Poissonian counts and mean $B(z_c)$.

Nevertheless, the large-scale structure of the Universe produces density correlations on scales larger than that of the clusters, whose contribution has to be subtracted from the final amplitude map.

In order to make up for the presence of these structures, AMICO can correct the noise by applying a local background correction $f(\boldsymbol{\theta}, z_c)$, obtained as follows: for each redshift slice, a density map is computed, with density values for each angular position; then, the peaks likely due to overdense structures, i.e. galaxy clusters, are removed with a $\kappa - \sigma$ clipping applied to the map; eventually a smoothing is computed slice by slice with scale sufficiently larger than the filter scale and a circular top-hat kernel.

The function $f(\boldsymbol{\theta}, z_c)$ is obtained by dividing each slice of the map by its mean value and is computed slice by slice for each angular position.

Whenever the local background correction is considered, the relevant quantities can be corrected a-posteriori, with the exact same outcome as if they were modified at the beginning:

$$A_{new}(\boldsymbol{\theta}_{\mathbf{c}}, z_c) = \frac{S(\boldsymbol{\theta}_{\mathbf{c}}, z_c) - f(\boldsymbol{\theta}, z_c)\beta(z_c)}{\alpha(z_c)}, \quad (2.14)$$

$$\sigma_{A, new}^2(\boldsymbol{\theta}_{\mathbf{c}}, z_c) = f(\boldsymbol{\theta}, z_c)\alpha^{-1}(z_c) + A_{new}(\boldsymbol{\theta}_{\mathbf{c}}, z_c)\frac{\gamma(z_c)}{\alpha^2(z_c)}, \quad (2.15)$$

where the $S(\boldsymbol{\theta}_{\mathbf{c}}, z_c)$ term is the summation term in Equation 2.5, namely

$$S(\boldsymbol{\theta}_{\mathbf{c}}, z_c) = \sum_{i=1}^{N_{gal}} \frac{M_c(\boldsymbol{\theta}_i - \boldsymbol{\theta}_{\mathbf{c}}, \mathbf{m}_i)p_i(z_c)}{N(\mathbf{m}_i, z_c)}. \quad (2.16)$$

2.3 Input parameters

Once the basic working principles of AMICO with the relevant quantities have been introduced it is worth spending a few more words on some of the quantities that, in this formalism, AMICO needs as inputs.

This Section aims at briefly introducing them, but just after the introduction of the data-set they will be further presented and addressed in the specific context of this work.

The photometric catalogue of galaxies. The first and most trivial input, for instance in case of galaxy cluster detection, is the photometric catalogue of observed galaxies with their locations and characterizing properties.

In this context galaxies are represented as data points $\mathbf{x}_i = (\boldsymbol{\theta}_i, \mathbf{m}_i, p_i(z))$ identified by sky coordinates $\boldsymbol{\theta}_i$, an array of properties \mathbf{m}_i and the photometric redshift probability distribution $p_i(z)$ of the i -th galaxy. 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 such as size, morphology or different magnitudes.

The redshift probability distribution can be made available in the photometric catalogue or can be analytically retrieved in case the redshift information is given in the form: $z_{peak}, z_{min}, z_{max}$. The first one is the most probable

estimate of z and the other two are the values of z at a certain confidence level.

In this latter case, a redshift probability distribution can be associated to each galaxy by computing a double Gaussian centered in z_{peak} and defined as follows (Bellagamba et al., 2011):

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma_m} \exp\left(-\frac{(z - z_{peak})^2}{2\sigma^2}\right), \quad (2.17)$$

where the two different values for σ and σ_m are

$$\sigma = \begin{cases} z_{peak} - z_{min}, & \text{if } z < z_{peak} \\ z_{max} - z_{peak}, & \text{if } z > z_{peak} \end{cases} \quad \text{and} \quad (2.18)$$

$$\sigma_m = \frac{z_{max} - z_{min}}{2}. \quad (2.19)$$

The galaxy redshift distribution $p(z)$ plays a crucial role in some of the main quantities introduced in Section 2.1 and 2.2. The definition of an extended distribution according to the uncertainty of the redshift itself links, in the formalism of Matched Filters, the analysis of the cluster made on redshift slices to the 3-dimensional one. For instance, in the computation of the memberships, treated in Section 2.4.1 it will be possible to see how galaxies are assigned to clustered structures being weighted with their probability distribution at the redshift of the candidate structure.

Despite AMICO has been developed in a very general and flexible way, so that it can be applied to a variety of different survey data types, it is fundamental to know the chosen input data-set in order to set and adjust all the other input parameters accordingly. For instance, the chosen survey has a depth and an angular extension, specific sources of the data, specific bands of observation, methods used to retrieve the observables, quality of the photometry and so forth.

Understanding and processing the input data-set is the first crucial step of every data-based analysis, including of course cluster identification from a catalogue of galaxies. This is the reason why a full chapter will be dedicated to the description and the processing of the chosen survey data (Chapter 3).

The mask. Another input parameter needed by AMICO is the mask, namely the footprint of the survey, which contains information about how to

identify areas of the field that were contaminated in the phase of observation. These areas have to be treated carefully during the scientific analysis in order to avoid the introduction of spurious objects in the data-set. This is what happens, for instance, in proximity of bright foreground stars, where pixels saturate and can show artifact patterns and then bias the retrieved information.

Usually, all sources are left as part of the catalogue, but are flagged as affected so that one can filter the data-set depending on the needed sample. Each survey and each analysis carried out on that survey has indeed its specifically suited mask (Coupon et al., 2018).

Within AMICO, the problem of having masked areas can be handled by simply restricting the integrals appearing in the main quantities, such as the filter constants α , β and γ over the only 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, with respect to a completely safe one (Bellagamba et al., 2018).

The cluster model. It has already been mentioned that the model, in the formalism of Optimal Filtering, is the template of the filter, namely it shapes the expected signal to be detected. Thus, in the context of galaxy catalogues, it is the expected galaxy distribution within a cluster and it is one of the main inputs of AMICO.

The analytic expression that describes a generic model is made of two different components, i.e. the radial and the magnitude distribution of galaxies:

$$M_c(r, \mathbf{m}) = \sum_i \Theta_i(r) \Phi_i(\mathbf{m}), \quad (2.20)$$

where M_c is a function of the distance from the center of the cluster, $r = |\boldsymbol{\theta}_i - \boldsymbol{\theta}_c|$ and of the array of galactic properties (most of the times a magnitude). The sum runs over the i index, i.e. the considered populations/components of the cluster (e.g. central massive galaxies and satellites, red and blue populations etc.) and Θ and Φ are the distributions of galaxies in radius and magnitude, respectively. The expected distribution of galaxies in a cluster has also a dependence on redshift z_c , that was here made implicit. The

redshift dependence is here treated through the use of the specific redshift probability of the galaxies.

An AMICO model template is a cube in which each redshift slice contains information about the galaxy density depending on radius and galaxy properties, such as the magnitude. It is computed by running over the galaxy catalogue, assigning the contribution of the elements to a specific bin, normalizing the value and then splitting in radial and magnitude distribution components. The template can be also used to set the binning of the quantities in all the following analysis. The units of the model are $\text{mag}^{-n} \text{deg}^{-2}$, i.e. it is a hypercube of dimension $n + 2$ (2018).

In order to make a practical choice of functions to be used as radial and magnitude distribution one may generally rely on known analysis of cluster samples in which the properties of the galactic populations were studied, according to the chosen magnitude band and the redshift range that must be consistent. In Section 4.2, the chosen parameters and relations will be described for the specific choices of magnitude bands and redshift made within this work.

The noise. The working principle of Optimal Filtering resides in the assumption that a data-set can be shaped as in Equation 2.1. In this context, $N(\mathbf{m}, z)$ was introduced as the noise component that accounts for the distribution of galaxies belonging to the field.

Undoubtedly, one can expect that, given the relations introduced in the previous Section, the noise is a parameter of fundamental importance both in the basic assumptions of this formalism and in the computation of the filter and its constants.

Within AMICO, the noise can be directly extracted from the data-set and it is generally approximated to the overall galaxy distribution of the sample. This assumption holds in case the field is large enough for the cluster galaxies to give a negligible contribution to the noise. Each galaxy contribution is added to the n properties-redshift bin it belongs to, the value is then normalized over the bin volume and according to the unmasked (or effective) area.

The AMICO noise model is a $(n+1)$ -dimensional hypercube, that contains information about the galaxy density binned in redshift and magnitude (or galaxy properties in general). According to the basic assumption behind Optimal Filtering techniques, the noise is in this way given by random counts

and therefore it has constant mean over the area of a given redshift slice.

As it possible to appreciate in Equation 2.4, and as a consequence in all the quantities deriving from the filter, the position of the noise at the denominator requires the introduction of a minimum threshold for the noise value, attributed to the empty pixels, i.e. pixels in which none of the galaxies' contribution falls. A minimum value of the noise model is required to prevent numerical divergencies of the filter, defined with the field galaxy distribution at the denominator. This concept, and the role of the noise in the computation of filter constants will become crucial in the adjustment of the initialization parameters (see Section 4.3).

2.4 Detection procedure

As soon as the amplitude map is computed, i.e. an estimate for the amplitude is given for each point in the 3D-grid made of angular positions and redshift, the second phase of AMICO detection can start: the selection of the candidate clusters.

According to the modeling of the data as a sum of cluster signal contribution and noise contribution (Eq. 2.1) and according to the estimate given for the amplitude (Eq. 2.2), the Gaussian likelihood of being a cluster can be computed as

$$\mathcal{L}(\boldsymbol{\theta}_c, z_c) = - \int \frac{D_{obs}(\boldsymbol{\theta}, \mathbf{m}, z) - D(\boldsymbol{\theta}, \mathbf{m}, z)}{N(\mathbf{m}, z_c)} = \quad (2.21)$$

$$= \mathcal{L}_0 + A^2(\boldsymbol{\theta}_c, z_c)\alpha(z_c) \quad (2.22)$$

with \mathcal{L}_0 being a negative constant over the position grid. The data-sets D are here referring to the observed galaxies and to the galaxy data as modelled in Equation 2.1, respectively.

It is intuitive that the biggest clusters are associated with a biggest improvement of the likelihood if the model is centered in their locations, so that maxima of amplitude correspond to maxima of the likelihood at fixed redshift. Nevertheless, to execute the detection also the information about the signal-to-noise ratio has to be used, since the quadratic dependence in Equation 2.22 cancels out the information about the sign of the amplitude. This approach returns high values of likelihood both for galaxy overdensities and underdensities, making the likelihood not able to distinguish between the two, i.e. positive amplitude and negative amplitude cases respectively.

Leaning on the amplitude information directly as object of the analysis, makes the filter linear with respect to the data, i.e. sensitive to the distinction between overdensities and underdensities (Bellagamba et al., 2011).

The signal-to-noise ratio, SNR for short, is in this formalism simply computed by dividing the amplitude by its variance:

$$SNR = \frac{A}{\sigma_A}. \quad (2.23)$$

In case the local background correction is applied, the likelihood follows the process of correction described before, analogously to the other quantities, so that Equation 2.22 becomes:

$$\mathcal{L}_{new}(\boldsymbol{\theta}_c, z_c) = \mathcal{L}_0 + \frac{A^2(\boldsymbol{\theta}_c, z_c)\alpha(z_c)}{f(\boldsymbol{\theta}, z_c)}. \quad (2.24)$$

Cluster selection is therefore performed with the identification of the position with the largest value of likelihood and the largest value of signal-to-noise ratio, with positive amplitude.

In the 3-dimensional amplitude map, each pixel has an amplitude estimate assigned, according to Equation 2.2, that is the amplitude of the signal in case a clustered structure is actually centered in that location. Since the maxima correspond, talking about peaks of likelihood or peaks of amplitude does not make any difference in the context of cluster selection from the map. The center of the cluster candidate is chosen on a regular position grid and is not restricted to galaxy positions.

In order to convert a 3D amplitude map into a catalogue of detections, AMICO operates in an iterative way, through a procedure of selection, membership assignment and cleaning. The number of detected clusters N_{det} depends, of course, on the minimum threshold chosen for the signal-to-noise ratio, which is an input parameter of AMICO: SNR_{min} .

The iteration is run over the j -th detections from 1 to N_{det} , following this procedure: first of all, every pixel with SNR below the threshold is rejected; the first cluster is then identified as the largest likelihood pixel among the remaining ones and some of its properties are saved; a membership probability (see Section 2.4.1) is computed and assigned to each galaxy located close to the candidate center; the map is cleaned by removing the imprint of the detection (see Section 2.4.2); the values of the output quantities (i.e. likelihood, amplitude and variance) are updated and the iteration can continue up to the last pixel.

2.4.1 Membership probability assignment

During the iterative procedure of detection, the assignment of a membership probability for galaxies surrounding the cluster candidate is performed by AMICO. Given the j -th cluster detection with amplitude A_j and located in $(\boldsymbol{\theta}_j, z_j)$, the probability $P_{i,j}$ of the i -th galaxy of belonging to the j -th detected clustered object is computed as the ratio between the galaxy density of the cluster component alone and the total one:

$$P_{i,j} \equiv \frac{A_j M_j(\boldsymbol{\theta}_i - \boldsymbol{\theta}_j, \mathbf{m}_i) p_i(z_j)}{A_j M_j(\boldsymbol{\theta}_i - \boldsymbol{\theta}_j, \mathbf{m}_i) p_i(z_j) + N(\mathbf{m}_i, z_j)} \quad (2.25)$$

It is obvious that it would be limiting to consider just one cluster at a time because actually it is possible for a galaxy to have a positive probability of belonging to more structures at the same time.

Theoretically speaking, one should take into account not only the signal contribution of the j -th cluster at the denominator of Equation 2.25, but take the total signal summing over all detected clusters.

On the other hand, it is most of the times not possible to have this information a-priori since the cluster candidates are sequentially identified in the iterative process. In the moment in which the j -th cluster and its possible memberships are taken into consideration, there is no information about possible clusters that still have to be identified in the surrounding area.

In order to take into account the probabilities assigned to a specific galaxy in previous detections, the field probability P_{F_i} is introduced in such a way that is equal to 1 for every galaxy and then progressively decreases whenever the galaxy is attributed to a structure:

$$P_{F_i} = 1 - \sum_{k=1}^{j-1} P_{i,k}. \quad (2.26)$$

Thus, Equation 2.25 can be rewritten as:

$$P_{i,j} = P_{F_i} \frac{A_j M_j(\boldsymbol{\theta}_i - \boldsymbol{\theta}_j, \mathbf{m}_i) p_i(z_j)}{A_j M_j(\boldsymbol{\theta}_i - \boldsymbol{\theta}_j, \mathbf{m}_i) p_i(z_j) + N(\mathbf{m}_i, z_j)}, \quad (2.27)$$

where the field probability can be seen as a scaling factor for the membership probability that accounts for how much the galaxy is still available for associations and always guarantees that for each galaxy $P_{i,j} \leq 1$.

Along with the output catalogue of candidate clusters, AMICO provides the list of associated galaxies with their membership probability or probabilities if associated to more than one structure, down to $P_{i,j} = 0$. The maximum number of clusters hosting the same galaxy is limited by a chosen threshold. Whenever this maximum is reached, an additional association is stored if better than the least safe of the list and replaces it.

2.4.2 Cleaning process

The last step of the detection procedure iterated until the last pixel of the map is the cleaning process, a key point in the AMICO algorithm. The aim of this step is to remove the imprint of sequentially detected objects from the map, in order to allow the subsequent detection of smaller and blended structures.

The signal subtraction is performed by exploiting the information about the probability of association in order to weight the contribution of the cluster members. Thus, the re-computation of the amplitude map is done as follows:

$$A_{clean}(\boldsymbol{\theta}_j, z_k) = A(\boldsymbol{\theta}_j, z_k) - \sum_{i=1}^{N_{gal}} P_{i,j} \frac{M_c(\boldsymbol{\theta}_i - \boldsymbol{\theta}_j, \mathbf{m}_i) p_i(z_k)}{N(\mathbf{m}_i, z_k)} \quad (2.28)$$

Computationally speaking this is the most expensive step in the detection procedure, since at every new detection the code has to run over every galaxy and update the map quantities in the surroundings of the detection (Belagamba et al., 2018).

2.5 Output

A typical AMICO output consists of the amplitude map, i.e. the 3D response of the filter depending on the location, the cleaned amplitude map, i.e. the latter after the removal of all the detected structure imprints, the variance map and the cleaned variance map. In addition, AMICO gives the possibility to inspect the values of the filter constants and the average redshift probability distribution. Finally, the prime output of the detection algorithm: the candidate cluster catalogue. The retrieved catalogues are, the catalogue of detections together with some properties of each detection, the refined catalogue of detections with revised positions and additional properties such

as estimates of the richness and the catalogue of members assigned to the detections, with the respective probabilities, as described in Section [2.4.1](#).

Chapter 3

The galaxy catalogue: COSMOS2015

In the context of this project, we made use of the available photometric data that are part of the Cosmic Evolution Survey, known as COSMOS ([Scoville et al., 2007](#)). The COSMOS project has offered over the years the unique chance to access very deep data, ranging from X-ray to radio wavelengths with high resolution and sensitivity. This has made it a popular and powerful tool for many studies concerning, among the others, galaxy and AGN populations, star formation, dark matter and large-scale structure of the Universe.

The used photometric catalogue is the recent COSMOS2015 release ([Laigle et al., 2016](#)), which contains a large amount of objects with very precise photometric redshifts, and near-infrared images and observables from the Ultra-VISTA project (Data Release 2; [McCracken et al., 2012](#)) with depth up to 24.7 AB magnitude¹.

This Chapter will be completely dedicated to the data-set that has been exploited for this study, namely the catalogue of galaxies on which the cluster detection has been based.

The COSMOS field and its main properties will be first summarized, in order to be able to introduce then the chosen catalogue and its role of resource for precise photometric high-redshift analyses. Finally, the steps through which the full catalogue was investigated in order to choose the most suitable sample for this kind of study, will be briefly described.

¹Within this work, all magnitudes are expressed in the AB system ([Oke, 1974](#)).

3.1 The COSMOS field

The Cosmic Evolution Survey (COSMOS) was born with the aim of shedding light on the coupled evolution of galaxies and large-scale structure of the Universe. Its depth, its area of coverage and the multi-wavelength approach to the study of this area of the sky, have made it undoubtedly a pillar of extragalactic astrophysics and observational cosmology. A full description of the goals and the properties of the survey can be found in [Scoville et al. \(2007\)](#).

The COSMOS survey covers an area of approximately 2 equatorial square degrees, boasting a wavelength coverage that spans the entire electromagnetic spectrum, with high sensitivity imaging and spectroscopy retrieved by great part of the main space telescopes and some of the largest and most powerful ground-based observatories. More than 200 scientists involved all over the world for more than 2 million of galaxies so far detected. Starting as a Hubble Space Telescope program, imaged with the Advanced Camera for Surveys (HST-ACS; [Koekemoer et al., 2007](#)), it includes now data from VLA (e.g. [Delvecchio et al., 2017](#); [Smolčić et al., 2017](#)), Spitzer (e.g. [Sanders et al., 2007](#)), Subaru (e.g. [Taniguchi et al., 2015](#)), ESO-VLT (e.g. [Lilly et al., 2007](#)), GALEX (e.g. [Zamojski et al., 2007](#)), XMM-Newton (e.g. [Cappelluti et al., 2009](#); [Hasinger et al., 2007](#)), Chandra (e.g. [Civano et al., 2016](#); [Marchesi et al., 2016](#)), Herschel (e.g. [Lutz et al., 2011](#)) and many others.²

What makes it unique is that, considering the large volume covered in depth, with data that reach redshift of ~ 10 ([Weaver et al., 2022](#)), the COSMOS field is a relatively extended field on the sky plane. This creates the conditions to investigate the overall galaxy evolution as a function of different large-scale-structure environments that are amply mapped within the field ([Taniguchi et al., 2007](#)).

Another key property is the location in the sky of the COSMOS field. It has been selected to cover an area that is close to the celestial equator, a strategic location that makes it easily accessible by a large number of ground-based facilities, both from the northern and the southern terrestrial hemisphere. Moreover, it should be mentioned that the COSMOS field is characterized by good conditions for what concerns galactic extinction, generally lower and more uniform with respect to other equatorial fields ([Scoville](#)

²For further information we refer the reader to the COSMOS survey web page: <https://cosmos.astro.caltech.edu>.

et al., 2007).

Accessible and deep, COSMOS was born as a pioneering survey and with very rich statistics still constitutes today a milestone of multi-wavelength astrophysics.

3.2 The COSMOS2015 catalogue

As already mentioned in the previous Chapter, the prime input of the AMICO algorithm for the detection of galaxy clusters is of course the photometric catalogue of galaxies.

As the object of a high number of campaigns covering the full spectrum, the COSMOS field offers a wide choice of photometric catalogues with deep imaging and extremely precise photometric redshifts. Notable examples are works like the *i*-band selected photometric catalogue presented in Ilbert et al. (2009) and Capak et al. (2007) and the more recent works by McCracken et al. (2012) and Ilbert et al. (2013).

A newly released catalogue has made available to the public a very rich data-set for over a million sources. It is the COSMOS2015 catalogue, described and presented in Laigle et al. (2016). The release contains a large amount of precise data like photometric redshifts and stellar masses for more than half of this million of objects, within the COSMOS survey area of 2 deg^2 . The standard tangent point of COSMOS has been used as center of the field: RA, DEC = (150.1163213, 2.20973097) ³.

The selection of the objects has been performed in the near-IR, making use of the $YJHK_s$ data from the UltraVISTA survey (McCracken et al., 2012) with the addition of the *Y*-band from the Subaru Hyper-Suprime-Cam (HSC) (Miyazaki et al., 2012), boasting the availability of a total of 32 different bands. This makes the catalogue a highly optimized data-set for high-redshift galaxy evolution and large-scale structure studies and therefore also particularly suited for the purpose of this analysis.

The COSMOS2015 catalogue has been chosen as input object catalogue for the study presented in this Thesis work.

³Within this study, all angular coordinates are expressed as J2000 right ascension (RA or R.A.) and declination (DEC), units are degrees.

3.2.1 Included data

The data were retrieved from different sources depending on the observed wavelength. This Section aims at summarizing the sources and the filters used to select the objects and retrieve the sample quantities. Nevertheless, a complete and exhaustive discussion of the observations and the data reduction can be found in [Laigle et al. \(2016\)](#) where, for instance, Table 1 well summarizes used instruments and filter properties.

Optical-ultraviolet. The catalogue includes the near-UV ($0.23\mu\text{m}$) GALEX observations ([Zamojski et al., 2007](#)), the u^* -band from the Mega-Cam of the Canada France Hawaii Telescope (CFHT)⁴, and data from the COSMOS-20 survey that consists of 12 medium bands, 6 broad bands (B , V , g , r , i and z^+) and two narrow bands, obtained with the Subaru Suprime-Cam ([Taniguchi et al., 2007; 2015](#)).

The z -band was replaced by the deeper z^{++} band and the g -band data were rejected due to poor seeing conditions.

Near-infrared. The $YJHK_s$ -band data were mostly obtained within the UltraVISTA survey program (Data Release 2⁵) carried out with the VIR-CAM instrument on the VISTA telescope ([McCracken et al., 2012](#)). The area of coverage of this data-set does not correspond to the full COSMOS field but concerns just the central 1.5 deg^2 . This UltraVISTA release contains both *deep* and *ultra-deep* stripes. The former has a depth of $K_s = 24.0$ (at 3σ in a $3''$ aperture) and covers an effective area over the COSMOS surveyed field of 0.92 deg^2 , the latter has a depth of $K_s = 24.7$ (at 3σ in a $3''$ aperture) for 0.46 deg^2 . This inhomogeneity creates differences in the completeness and depth of the final catalogue depending on the position considered within the COSMOS field. In Figure 3.1, the schematic of COSMOS has been reported from Figure 1 in [Laigle et al. \(2016\)](#), to show the limits of the areas covered by different surveys in the optical bands and in the near-IR. The blue delimited region refers to the COSMOS HST-ACS survey ([Koekemoer et al., 2007](#)). A^{COSMOS} indicates the COSMOS square. The UltraVISTA areas are highlighted by orange (*deep* region) and green contours (*ultra-deep* stripes). It is clearly evident here how the near-IR sample occupies a restricted area of the field with respect to the optical one, and therefore influences the depth

⁴<https://www.cfht.hawaii.edu/Instruments/Imaging/Megacam/>

⁵http://www.eso.org/sci/observing/phase3/data_releases/uvista_dr2.pdf

of the overall catalogue. The background image is the detection image that will be briefly described in Section 3.2.2.

The fraction of the field that is not covered by the UltraVISTA stripes is nevertheless not lacking near-IR photometry, since the catalogue is integrated in this wavelength range with the H and K bands from WIRCAM (McCracken et al., 2010).

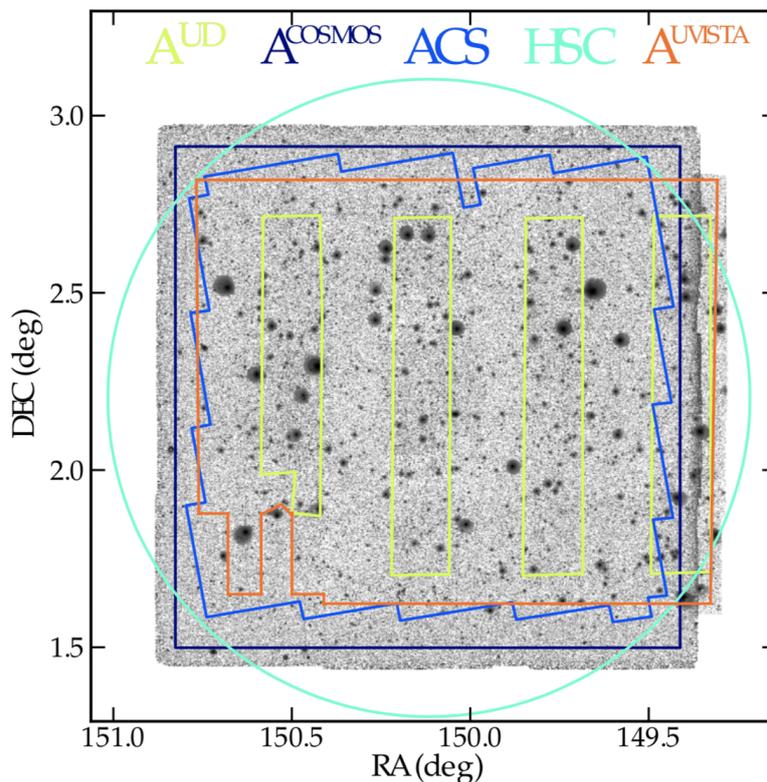


Figure 3.1: The COSMOS field with the limits of the areas covered by different surveys in the optical bands and in the near-IR. The blue delimited region refers to the COSMOS HST-ACS survey (Koekemoer et al., 2007). A^{COSMOS} indicates the COSMOS square. The UltraVISTA areas are highlighted by orange (*deep* region) and green contours (*ultra-deep* stripes). It is clearly evident here how the near-IR sample occupies a restricted area of the field with respect to the optical one, and therefore influences the depth of the overall catalogue. The background image is the detection image briefly described in Section 3.2.2. Credits: Laigle et al., 2016, Figure 1.

In this release also the Y-band data from Subaru Hyper-Supreme-Cam (HSC) (Miyazaki et al., 2012) were added. This was done both to improve the redshift and mass estimates at $1 < z < 1.5$ and to make it more consistent with following HSC images-based catalogues (Laigle et al., 2016).

Mid-Infrared. IRAC data are also included in the catalogue and were mostly retrieved from the SPLASH COSMOS project, S-COSMOS (Sanders et al., 2007), from the Spitzer S-CANDELS, the Extended Mission Deep Survey projects (Ashby et al., 2013; 2015) and other COSMOS programs.

For details about the image PSF homogenisation we refer the reader to the main paper of the catalogue, Laigle et al. (2016).

3.2.2 Object selection and photometric redshifts

In the process of object extraction, optical and near-IR photometry was computed through the use of the SExtractor program (Bertin & Arnouts, 1996) in the *dual-image* mode. The detection image was created combining the near-IR images from UltraVISTA with the z^{++} data from Subaru, in order to account for UV-luminous high-redshift objects as well. This choice was made to reduce the impact that selecting in the near-IR instead of in the i^+ -band has on the number of blue objects detected.

Nevertheless, still around one sixth of the sources are missing in the COSMOS2015 catalogue with respect to previous publicly released catalogues (e.g. Capak et al., 2007). These sources are likely to mostly belong to a class of blue faint star-forming galaxies. These missing objects are expected, since different selection bands obviously sample different galaxy populations.

Available magnitudes are 3" aperture (`MAG_APER`) and total magnitudes (`MAG_AUTO`) with universal offset to transform the former into the latter.

Photometric redshift estimation is based of 3" aperture fluxes and not on the magnitudes, over the all redshift range. It was carried out thanks to the photometric analysis software LEPHARE (Arnouts et al., 2002; Ilbert et al., 2006).

The code performs a χ^2 analysis between fluxes expected by a wide set of galaxy templates, and the observed ones for each galaxy. The obtained χ^2 is then converted into a probability and the probabilities are summed together at each z . The final result is the Probability Distribution Function (PDF), whose median is the reported photometric redshift value and the 1σ uncertainties are the values enclosing 68% of the PDF itself.

It should be noted that this does not apply for X-ray selected sources included in this catalogue that needed a specific tuning and a different method for the redshift determination, described in Salvato et al. (2011).

As already mentioned, COSMOS2015 is an unprecedented catalogue in

terms of photometric redshift accuracy. A comparison with available spectroscopic catalogues has indeed shown very competitive results. For instance, in a test performed by Laigle et al. (2016) using a reliable sample extracted by the zCOSMOS-bright program (Lilly et al., 2007), the redshift accuracy was found to be $\sigma = 0.007$ with catastrophic failures $\eta = 0.51\%$ (sample reached $z = 1.2$) and remained precise even for comparisons at $z > 3$, with $\sigma = 0.021$ and $\eta = 13.2\%$.⁶ Nevertheless, between photometric redshift PDF

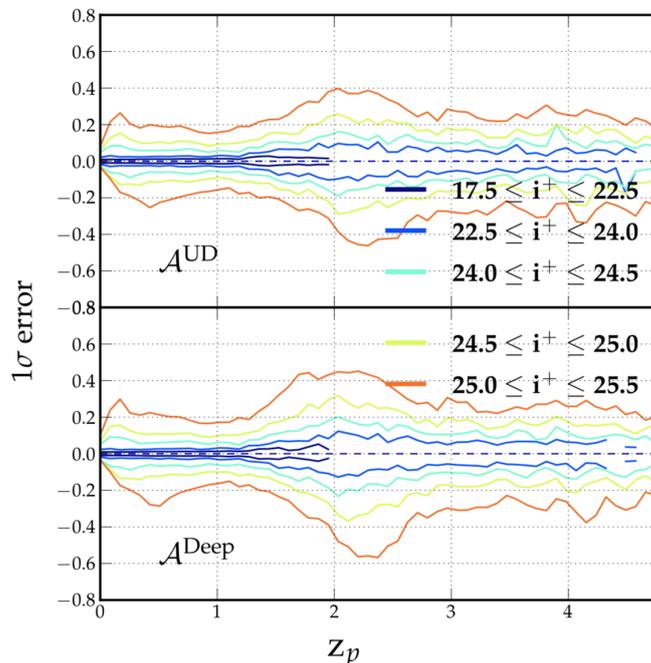


Figure 3.2: Photometric redshift 1σ positive and negative errors as a function of photometric redshift for different intervals of i^+ magnitudes. Top panel shows the UltraVISTA *ultra-deep* stripes and bottom panel the area covered by UltraVISTA data without *ultra-deep* stripes, namely the difference between the orange delimited area and the green stripes in Figure 3.1. The redshift uncertainty increases in the interval $1.4 < z < 2.5$ mainly due to the shift of the Balmer-break to the near-IR region. In this range, the *ultra-deep* stripes are better in terms of redshift errors. The accuracy improves again when the Lyman-break enters the optical range, at redshift higher than ~ 2.5 . Other fluctuations in the photometric uncertainties are due to other intrinsic differences between the filters or are attributable to the phase of calibration or fitting. Credits: Laigle et al., 2016, Figure 14.

uncertainties and spectroscopic comparison, an inconsistency was found so that the photometric redshift errors were underestimated. This issue was

⁶In Laigle et al. (2016) the redshift accuracy measured with spectroscopic samples is defined as $1.48 \cdot \text{median}(|z_p - z_s|/(1 + z_s))$ and η is the percentage of objects with $|z_p - z_s|/(1 + z_s) > 0.15$, where z_p and z_s are the photometric and the spectroscopic redshifts, respectively.

already present in COSMOS catalogues retrieved with analogous methods and was managed by introducing a magnitude-dependent correction factor that does not influence the other physical parameters (Laigle et al., 2016).

Figure 3.2, (2016) shows the negative and positive 1σ uncertainty as a function of redshift and in intervals of magnitudes, for the deep and the ultra-deep regions. It is possible to appreciate different features, including the increase of uncertainty in the interval $1.4 < z < 2.5$ mainly due to the shift of the Balmer-break to the near-IR region. In this range, the *ultra-deep* stripes are also visibly better in terms of redshift uncertainties. The accuracy improves again when the Lyman-break enters the optical range, at redshift higher than ~ 2.5 .

Other fluctuations in the photometric uncertainties are due to other intrinsic differences between the filters or are attributable to the phase of calibration or fitting.

Further details about the additional calibrations, corrections and tests along with the estimation of other physical properties of the sample can be found in Laigle et al. (2016).

3.3 Selection of the input sample

As already mentioned, the process of understanding and selecting the suitable sample that will be the input of the algorithm is of prime importance.

Despite AMICO can work in an n -dimensional space of n properties of each galaxy, in this phase of the study we restricted the analysis to a single magnitude at a time. The chosen magnitudes are two, used for different redshift ranges: the r -band magnitude and the H -band magnitude, extracted, as mentioned in Section 3.2.1, from Subaru Supreme-Cam and UltraVISTA, respectively. All magnitudes used in the context of this work are automatic aperture magnitudes, i.e. total magnitudes (MAG_AUTO in COSMOS2015).

3.3.1 Flag selection

Flagging objects in order to label them according to their quality or position is the common way to include information about sources in catalogues and to make the selection interactively adjustable.

Hereafter the flags of the COSMOS2015 catalogue that are relevant for

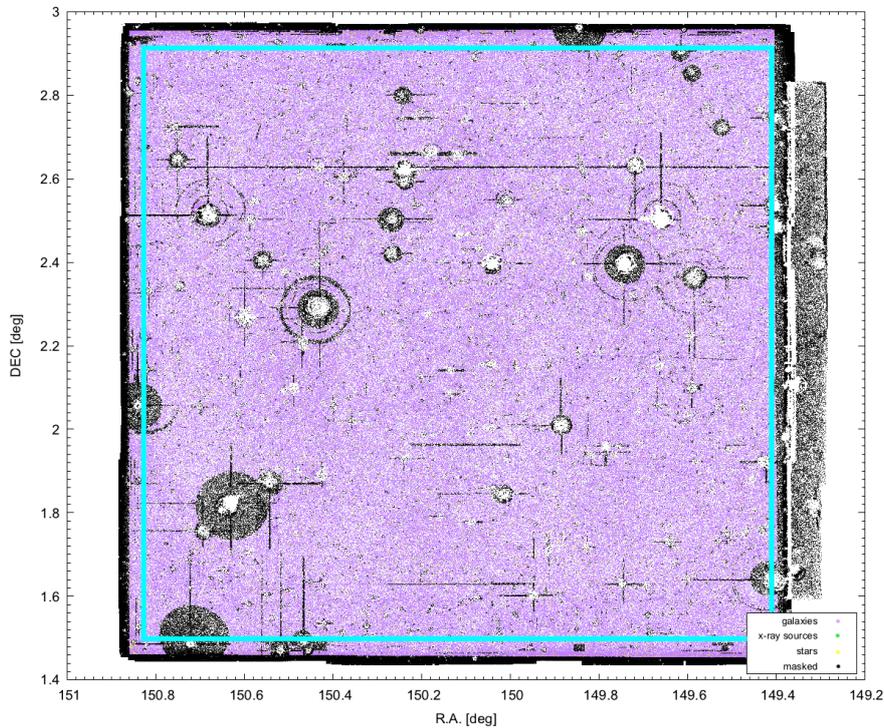


Figure 3.3: The full COSMOS2015 catalogue, with 1182089 objects included. Galaxies are shown in purple, stars in yellow, X-ray selected galaxies in green and objects masked in the optical broad-bands are shown in black ($\text{FLAG_PETER}=0$). The COSMOS field at these wavelengths is clearly affected by the presence of bright foreground stars. The light blue contour delimits the 2 square degree COSMOS area ($\text{FLAG_COSMOS}=1$).

our analysis are listed with their correspondent meaning ⁷:

- **FLAG_HJMCC**: if 0 it selects only the data that are included in the UltraVISTA area, namely the objects with near-IR observables retrieved by this specific survey (in *deep* or *ultra-deep* mode) and not by other instruments;
- **FLAG_COSMOS**: if 1 it simply restricts the catalogue to the standard COSMOS 2 deg^2 area;
- **[MAG]_FLAGS**: it is the internal flag of each magnitude, namely it contains information about the quality of the photometry. It is a SEXTRACTOR extraction flag, i.e. it is a scale that indicates with increasing

⁷For a complete list and description of flags and included quantities, see the official documentation of the catalogue: https://irsa.ipac.caltech.edu/data/COSMOS/gator_docs/uvista_dr2.1.pdf.

degree of warning, the possible issues faced during the source extraction process⁸;

- **FLAG_PETER**: it is the optical masking flag. Objects with flag value > 0 are masked in the optical broad-bands. If 0 the objects selected are safe and unmasked. Value between 1 and 15 are assigned to saturated or bad-regions objects.

Figure 3.3 presents the entire COSMOS2015 catalogue, with selection of masked objects and COSMOS objects made with appropriate flags. Sources masked in the optical broad-bands are shown in black (**FLAG_PETER=0**). The light blue contour delimits the 2 square degree COSMOS area (**FLAG_COSMOS=1**). The COSMOS field at these wavelengths is clearly affected by the presence of bright foreground stars.

The first flag listed above, **FLAG_HJMCC**, was not necessary in this phase of the study. For the *H*-band selected catalogue, only UltraVISTA extracted *H* magnitude has been considered and so the area of coverage has been automatically restricted to the deep region and the ultra-deep stripes. This flag will be nevertheless used to build the visibility mask for the *H*-band.

The second flag was exploited in every sub-sample chosen to always restrict the study to the 2 square degree COSMOS covered area.

The SExtractor extraction flag is a label that can assume the values of 8 different flag bits, expressed in powers of 2 and written in decimal. If the source is affected by more than one problematic issue, the bits value can be summed up. In deep field surveys and in case of clustered structures of galaxies exactly as the ones that are object of this study, it is common that a certain number of sources have to undergo a deblending procedure during the source extraction from the image, since objects can easily be found in positions where emissions overlap each other.

Thus, for the specific purpose of this work, the galaxies selected in our final sample will be restricted to: (0) galaxies with optimal extraction procedure, (1) galaxies with aperture photometry likely to be biased by close sources (or any other aperture with more than 10% of bad pixels) and (2) galaxies that underwent a deblending process. According to the summing-up possibility for multiple flags, we selected for both magnitude analyses: **r_FLAG \leq 3** and **H_FLAG \leq 3**.

⁸For further details see the SExtractor documentation about flagging: <https://sextractor.readthedocs.io/en/latest/Flagging.html>.

For what concerns the masking flag, `FLAG_PETER`, a more detailed selection was needed. This will be addressed in the following Section.

3.3.2 `FLAG_PETER` analysis

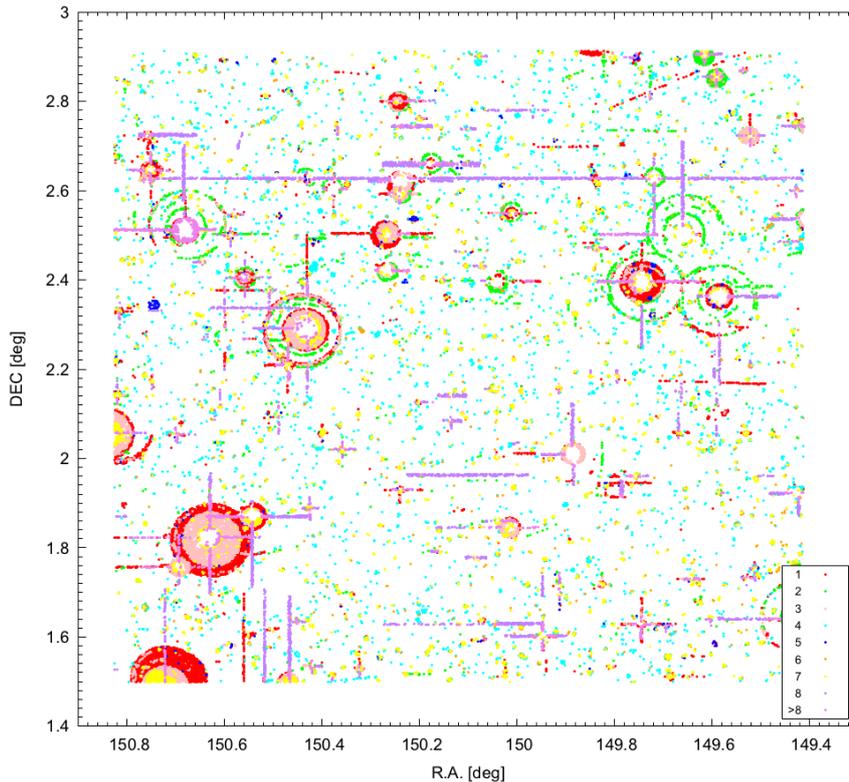


Figure 3.4: The positions of the flagged galaxies colour-coded according to `FLAG_PETER` values. The position of the flagged galaxies gives information about how the problematic issue in the quality of the detection was produced. For instance, some flagged galaxies are directly connected to the presence of bright foreground stars, others are located in positions resembling geometric patterns, and some galaxies are flagged because close to other bright galaxies. The latter case is exactly what can occur in galaxy clusters, so it is fundamental to include these objects in the galaxy catalogue. The most complete and safe sample for our studies was found to be the sample including galaxies with `FLAG_PETER` = 0 (all the unmasked galaxies not shown here) or `FLAG_PETER` = 4 (cyan) or `FLAG_PETER` = 6 (orange).

During the phase of revision of one of the first runs performed on a chosen preliminary sub-sample of the catalogue, it became necessary to modify the selected sample in order to include some galaxies that remained excluded if the choice was just to keep unmasked safe objects (`FLAG_PETER` = 0).

The runs performed on the preliminary catalogue, which was the catalogue of safe objects only, generated unrealistic and visibly incomplete re-

sults. This was attributable to an incomplete mapping of the sky, namely a non-appropriate selection of the catalogue objects for our specific goals.

This is a point of crucial importance, because, on one hand, as already mentioned, masked objects have to be treated carefully in order not to introduce in the sample spurious objects that will lead to misleading detections. On the other hand, the flagging of masked objects can be done for a variety of reasons that may not be relevant for a certain data-set or that may even bias it.

In the process of studying the meaning of the 15 different values of this quality flag associated with the galaxies it was clear that most of the cluster members were rejected when adopting standard selection criteria.

The position of the flagged galaxies, as shown in Figure 3.4, gives indeed information about how the problematic issue in the quality of the detection was produced. For instance, some flagged galaxies are directly connected to a masked area for the presence of bright foreground stars, others are located in positions resembling geometric patterns due to optical or electronic effects, and finally some galaxies are flagged because close to another bright galaxy or close to other galaxies.

The latter case is exactly what can occur in galaxy clusters, so it is fundamental to include these objects in the galaxy catalogue. This became evident when the different runs, with different combinations of included sub-sample, chosen by picking different values of `FLAG_PETER`, were compared (for additional details, see Appendix A).

3.4 Final galaxy catalogue

It has been found as a result of the masked-area analysis that the most complete and safe sample for our purpose is the sample that includes, not only galaxies with `FLAG_PETER` = 0, but also with `FLAG_PETER` = 4 or `FLAG_PETER` = 6, with a total recovery of more than 6 thousand galaxies (mostly cluster galaxies) that were excluded by using standard selection criteria.

After having selected the effectively unmasked objects for our study and having adjusted in a complementary way the catalogues of masked objects, the data-sets were cleaned according to the used magnitude.

As already mentioned, within this study, AMICO has dealt with just one magnitude at a time on two different redshift ranges, first with *r*-band and

then with H -band magnitude.

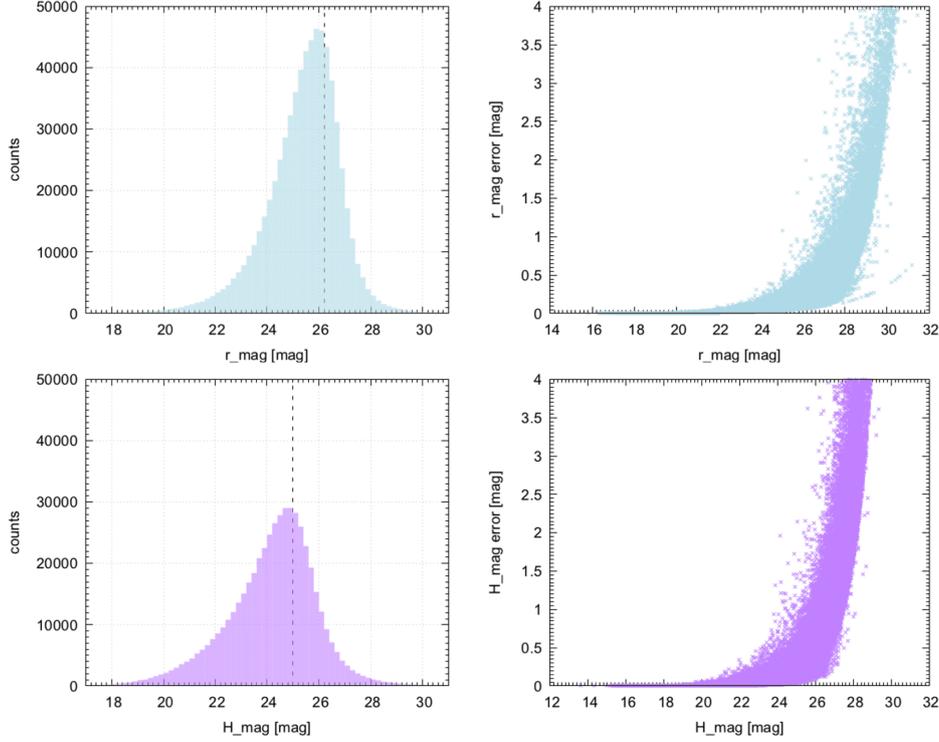


Figure 3.5: The magnitude distributions of all galaxies (left) for r -band magnitude (top) and H -band magnitude (bottom). The shown data are restricted in both cases to the range $17 < \text{mag} < 31$. The dashed vertical lines indicates the magnitude cut, performed to reject the faint-end tail of the distribution, likely to be more affected by large uncertainties. It is possible to appreciate indeed the trend of the magnitude error as a function of magnitude in the right panels. For faint objects the errors dramatically increase. Outliers with error larger than 4 are here not shown.

Therefore, we performed magnitude-cleaning separately for the two magnitude by keeping, first only galaxies with available magnitude and then by restricting the sample to galaxies brighter than the peak magnitude of the overall distribution. Galaxies belonging to the faint-end tail of the magnitude distribution were rejected so as not to contaminate the galaxy sample with noisy or spurious detections. Figure 3.5 shows the magnitude distributions (left panels) for r -band magnitude (top) and H -band magnitude (bottom) with a dashed vertical line indicating the magnitude cut (m_{cut}), performed to reject the galaxies in the faint-end tail of the distribution. These objects are likely to be affected by large uncertainties as visible in the error trends in the right panels. For objects fainter than the peak magnitude, errors dramatically increase, up to unrealistic values.

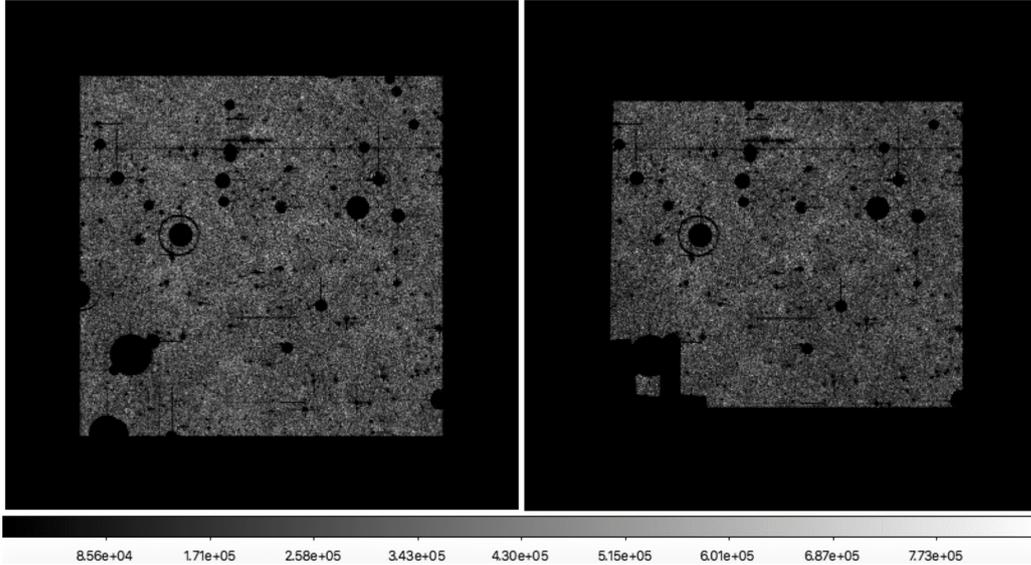


Figure 3.6: The number density map of the final galaxy catalogues. Each pixel value indicates the number of galaxies falling into the area of the pixel. The final selected samples contains 479124 galaxies with a final effective area of 1.759 deg^2 , and 340247 galaxies with effective area of 1.351 deg^2 for the r -band run (left) and H -band run (right) catalogues, respectively.

Obviously, since photometric redshift plays a fundamental role in cluster detection with AMICO, only galaxies with available redshift have been used.

A summary of the final selected catalogues of objects can be found in Table 3.1.

The final selected sample for the r -band magnitude analysis contains 479124 galaxies. The mask has been adjusted accordingly, leading to a final effective area of 1.759 deg^2 , i.e. the 87.95% of the total surveyed area. Each pixel value indicates the number of galaxies falling into the area of the pixel. The average density is $1.192 \times 10^5 \text{ galaxies/deg}^2$.

The final selected catalogue of galaxies for the H -band magnitude analysis contains instead 340247 galaxies, with average galaxy density $8.464 \times 10^4 \text{ galaxies/deg}^2$, on a total effective area of 1.351 deg^2 , i.e. the 67.55% of the COSMOS field. Figure 3.6 shows the galaxy density of the two catalogues.

The redshift distribution of the galaxy number counts for the two final catalogues is shown in Figure 3.7. The two redshift distributions are consistent and marked by roughly the same trends.

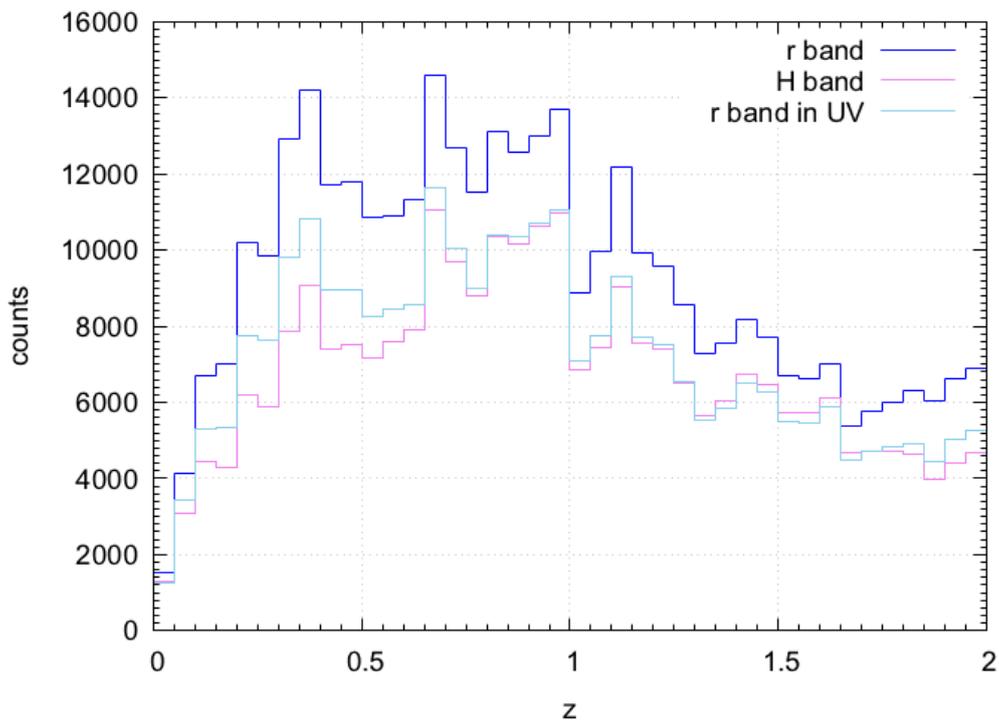


Figure 3.7: The redshift distribution of the two catalogues used for the r -band analysis (blue histogram) and for the H -band analysis (pink histogram) in the range $0 < z < 2$. Despite the poorer statistics of the latter due to the different effective area covered, the two distributions follow roughly the same trend and have the same drop and peak features. The light blue histogram shows the redshift distribution of the r -band galaxy catalogue restricted to the UltraVISTA area for consistency with the H -band one.

| ID | description | FLAG_PETER and constraints | [MAG]_FLAGS | # objects |
|------------------------|---|---|--------------|-----------|
| all_masked_objects | all the masked objects according to chosen criteria that are useful to produce the visibility mask | FLAG_PETER != 0 and (stars or FLAG_PETER != 4 and 6) | - | 69655 |
| all_unmasked_objects | all the unmasked objects according to chosen criteria that are useful to produce the visibility mask | FLAG_PETER = 0 or (galaxies and FLAG_PETER = 4 or 6) | - | 703463 |
| galaxy_catalogue_rband | galaxies selected according to chosen criteria and with cleaned r magnitude will be the input catalogue for AMICO | (r_MAG not null and <=26.2) and z not null and (FLAG_PETER = 0 or 4 or 6) | r_FLAGS <= 3 | 479124 |
| galaxy_catalogue_Hband | galaxies selected according to chosen criteria and with cleaned H magnitude will be the input catalogue for AMICO | (H_MAG not null and <=25.0) and z not null and (FLAG_PETER = 0 or 4 or 6) | H_FLAGS <= 3 | 340247 |

Table 3.1: Summary of the final selected samples of objects from COSMOS2015, their properties and the selection constraints. The galaxy catalogue for the r -band analysis contains 479124 galaxies, with effective area of 1.759 deg^2 . For the H -band analysis we selected instead 340247 galaxies, for a total effective area of 1.351 deg^2 . All selected objects shown here in the # *objects* column have to be intended as objects in the full redshift range, not limited to the analyzed interval.

Chapter 4

Applying AMICO to the COSMOS2015 galaxy catalogue

The AMICO algorithm is a very flexible identifier of clustered objects relying on a very general base structure and it can be applied to a variety of different survey types over a wide redshift range. The purpose of the analysis described in this Thesis is to detect clusters with AMICO on the deep and restricted COSMOS field with the final goal of extending the cluster search up to redshift ~ 2 . This study is scientifically important in order to retrieve a new and deep catalogue of clusters with a unique amount of cluster properties and galaxy membership information for which different benchmarks are available thanks to the rich multi-wavelength covering of the field.

From a technical point of view, it is additionally important to test the flexibility and usability of AMICO on a peculiar survey configuration. Enriching the data-set types on which AMICO is being used just enlarges the testing-field and increases the chances to improve the reliability and the efficiency of the algorithm.

This application of AMICO to the COSMOS field, with its deep and precise photometric data, opened therefore the opportunity to adapt, test and integrate some new methods within AMICO, to allow the use of new kinds of data-sets.

In contrast to recent applications of AMICO, this work is based on a different kind of survey, in terms of depth and covering area for instance. An example of recent application is the catalogue of galaxy clusters obtained by applying AMICO to the KiDS Data Release 3 ([de Jong et al., 2017](#)).

In this study, the input catalogue of galaxies contained a total of about 32 million objects down to magnitude $r = 24$, with an overall area of coverage of 438 deg^2 that led to a catalogue of 7988 candidate clusters in the range $0.1 < z < 0.8$ (Maturi et al., 2019). It is easy to grasp the difference between this application of AMICO and the survey configuration we aim at analyzing in the study described in this Thesis.

In Chapter 3, the prime input of an AMICO detection run has been introduced. Now that the data-set and the selected samples have been presented, it is possible to move to the other key input parameters of AMICO. Presenting the main steps of the process through which the inputs were generated will also allow to point out the improvements and the adaptations performed on the code in order to make it able to face new issues that may be present in the photometric data-set.

In this Chapter, we will discuss the masked objects, the visibility masks and the template, i.e. the cluster model. Finally, the noise model, its process of regularization and assessment will be presented.

4.1 Generation of the visibility mask

Artifacts produced by the instruments, contaminating sources and bright objects saturating the CCD pixels are the most common sources of spurious or biased detections. Several calibration and cleaning procedures are applied in order to reduce the influence of intrinsic instrumental defects, and images are usually inspected to remove other contributions like cosmic-rays, satellite tracks and so forth. A widely adopted technique to deal with the artifacts produced by bright stars is to use star catalogues to identify possibly affected regions (as done in Capak et al., 2007). Bright stars indeed easily saturate and generate both electronic and optical patterns, like leaking to neighbour pixels, creation of the so-called *bleed trails*, diffraction patterns or extended reflection "ghosts" (Coupon et al., 2018).

Normally every object whose properties are not reliably determinable or that are in a significant way affected by one of the image defects, are flagged as masked, namely labelled as poorer quality detections.

Objects that are not flagged as masked have therefore generally the most reliable and safe photometry in the catalogue. Nevertheless, the complete catalogue is made available to make the choice of sources adjustable through flags used to select the galaxies needed for a specific application (Capak et

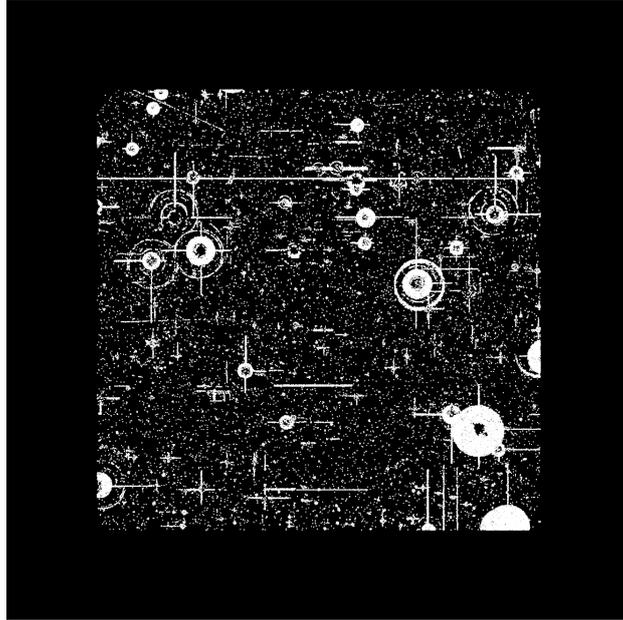


Figure 4.1: The preliminary mask, as extracted by AMICO from the catalogue of masked objects. It has ones (white) assigned to all pixels with more than one masked object. Zeros (black) are assigned to fully unmasked pixels. This mask gives a first idea of the footprint of the COSMOS field in the optical bands, but it is not complete yet. Central areas in correspondence of bright foreground stars are lacking also masked objects and therefore are marked as unmasked. Additional information is required.

al., 2007).

Masking is of fundamental importance not only to prevent the image defect to impact on the used sample, but also to mark the fraction of surveyed area that is effectively available for the study and for the computation of all the relevant quantities.

This is the reason why using an accurate input mask is another important step in the initialization of a galaxy cluster detection with photometric data-sets. In the framework of this study, the optical masks to analyse the COSMOS field have been directly generated through the AMICO package. The AMICO algorithm contains a routine that computes a smoothed number density maps starting from a given catalogue of galaxies. An example of the output of this program is shown in Figure 3.6. The same program can be used to generate binary masks: zeros will be assigned to fully unmasked pixels and ones to fully masked ones. Given the catalogue of masked objects (first row of Table 3.1), we extracted a preliminary mask (Figure 4.1) by assigning ones to all pixels with more than one masked object falling inside the area of the pixel. It should be noted that *objects* are here meant not only

as galaxies but also as stars, and include all kinds of objects and all kinds of photometry in order to increase the statistics and map more precisely the masked area.

This gives a first footprint of the unsafe regions of the COSMOS field in the optical bands. Nevertheless, the information retrieved by the masked object catalogue is not complete yet. As visible in Figure 4.1, central areas in correspondence of bright foreground stars are lacking objects in general and they are in this way interpreted as unmasked and available pixels because lacking masked objects as well.

Additional information is therefore required, and was retrieved by making use of a catalogue of unmasked objects (as anticipated in the second row of Table 3.1). This can be achieved by adding together in a consistent way the presence of masked objects and the absence of unmasked objects to create a more complete footprint of the survey. This is why, in the context of this study, a new possibility has been introduced in AMICO: *the generator of composite masks, **mask2combined***.

This program takes as inputs a mask generated as just described starting from masked objects and an inverted¹ mask generated with unmasked objects and combines them together in a single mask file:

```
masked ← mask from masked objects
unmasked ← inverted mask from unmasked objects
pixel ← 0
while pixel < length(masked) do
  if masked[pixel] == 0 then
    pixel ← pixel + 1
  else
    unmasked[pixel] ← 1
    pixel ← pixel + 1
  end if
end while.
```

An additional advantage of choosing the unmasked catalogue mask as master file on which adding ones is that the area outside the field that obviously is devoid of any detection is automatically marked as masked, delimiting the field. In Figure 4.2 we show the three masks: on the top left panel the preliminary mask obtained by masked objects (the same shown in Figure

¹with zeros instead of ones where the catalogue objects fall, and vice versa.

4.1), on the top right panel the mask used as master file, created with the unmasked objects catalogue and on the bottom panel the composite mask resulted by combining the two. The composite mask gives a more complete footprint of the accessible and not affected regions.

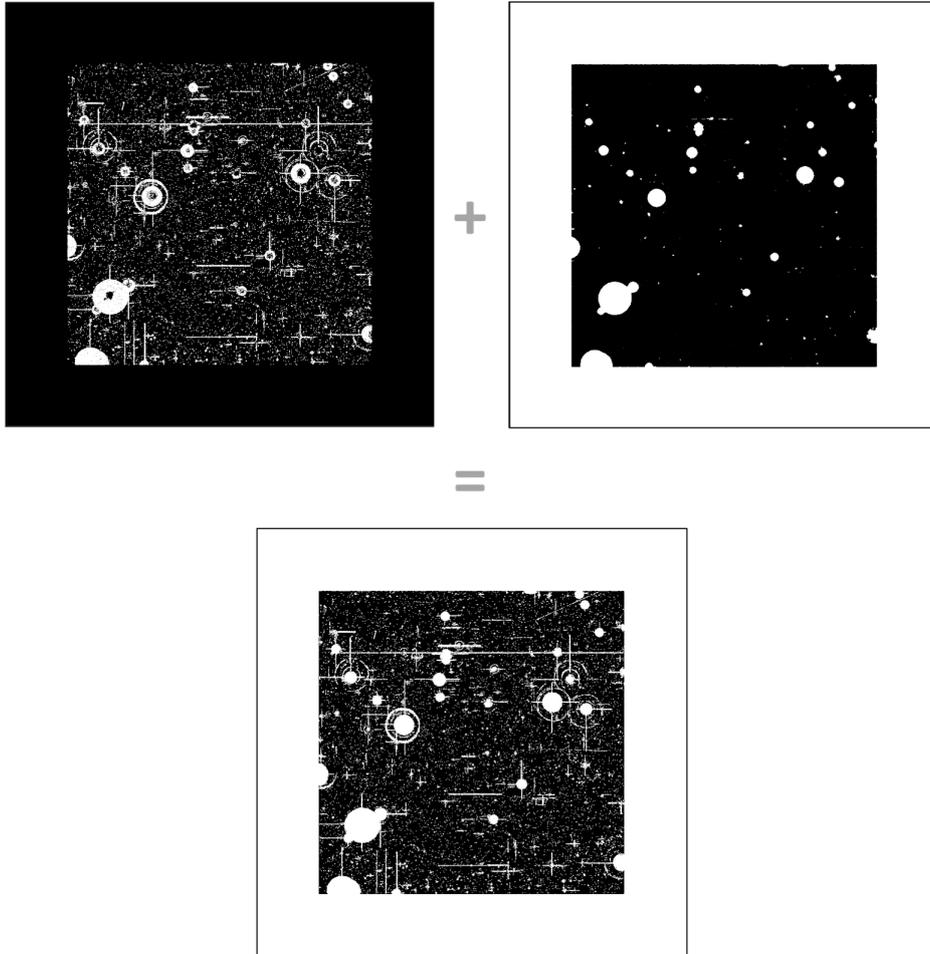


Figure 4.2: The two masks retrieved from the masked objects (top left) and from the unmasked objects (top right) and the composite mask created adding the masked pixels of the former to the latter (bottom). The white pixels are the masked pixels (1 in the binary mask), the black ones are the unmasked pixels (0). The two masks combined together are generated using a different radius for the smoothing kernel. In the stage of creation of the files, the masks are binary, subsequently converted in AMICO non-binary masks, with different resolution, in which each pixel has a value indicating the percentage of the pixel area that is masked.

To avoid the problem of labelling as masked pixels in which objects are lacking simply due to intrinsic fluctuations in the number density of galaxies and not due to photometry problems, the maps have been first smoothed.

For the mask based of unmasked objects we have used a smoothing kernel of 4 pixels in size, for the mask based on masked objects we have used a kernel size of 1 pixel. These values have been derived empirically. An appropriate smoothing allows to reduce the possibilities of losing information where only a few galaxies are available.

The final binary mask has to be converted into a lower resolution, i.e. the resolution used by AMICO during the cluster detection phase, in such a way that each pixel value expresses the fraction of its area that is masked. For instance, a value of 0.56 stays for a pixel masked for the 56% of its area.

4.1.1 Correction for UltraVISTA data

The composite mask has been used as input visibility mask during both the analyses, for r -band and H -band magnitude but a small correction was necessary for the latter, because the source of data is the UltraVISTA program, that displays a similar mask pattern but covers a smaller area.

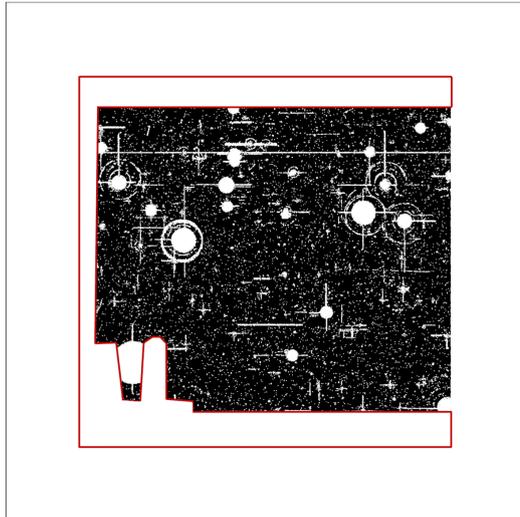


Figure 4.3: The final composite mask used for the H -band analysis. The area delimited by red contours is the portion of the COSMOS field falling out of the UltraVISTA coverage and has been marked as masked by adding a series of ones. To guarantee homogeneity in the covering, a sufficiently large smoothing radius has been used.

In order to account for this difference, a new catalogue has been created by selecting more than 168 thousand objects detected in the COSMOS field but out of the UltraVISTA coverage. This selection has been possible thanks to the FLAG_HJMCC, mentioned in Section 3.3.1. The introduction of masked

pixels accounting for this non-covered area has been computed in the same way as described above, by using *the generator of composite masks*. The obtained mask was used to delimit the survey perimeter and to compute the effective area. The COSMOS field area out of the UltraVISTA coverage is marked with red contours in Figure 4.3, where the mask used for the *H*-band analysis is shown.

4.2 The cluster model

In this Section we introduce the template describing the cluster model used by AMICO for the detection of galaxy clusters.

The analytic expression that describes the generic model is made of two different components (see Equation 2.20), namely the distributions of galaxies in radius and in magnitude as a function of redshift.

As already mentioned, to make a practical choice of the two functions shaping the distributions, it is possible to rely on observational analysis of cluster samples which are consistent with the chosen magnitude band and redshift range.

In the framework of this study, two different models were computed for the two different bands and redshift ranges. The choices and the corresponding references for the radial distribution and for the luminosity distribution of galaxies will be briefly described.

4.2.1 Radial distribution

In this context, the radial distribution in the cluster model has been chosen to be modelled by a Navarro-Frenk-White profile (NFW; Navarro, Frenk, and White, 1997) with parameters tuned according to the scaling relation described in Hennig et al. (2017).

Hennig et al. (2017) based their analysis on the galaxy populations of 74 galaxy clusters selected via Sunyaev-Zeldovich effect. This selection has been performed in the context of the overlap between the millimetric 2500 deg² South Pole Telescope survey (SPT) (see e.g. Story et al., 2013) and the deep optical data from the science verification phase of the Dark Energy Survey (DES)² (The Dark Energy Survey Collaboration, 2005). The resulting sample of massive clusters is homogeneous and unbiased, and boasts of good quality SZ mass estimates. The sample redshift extends up to $z \sim 1.1$ with masses in the interval $4.3 \times 10^{14} M_{\odot} \leq M_{200} \leq 2.9 \times 10^{15} M_{\odot}$.

The subscript 200 refers in literature to the virial sphere in which the mean density is 200 times the critical density of the Universe at the redshift of that specific structure, defined as $\rho_{crit} \equiv 3H^2(z)/8\pi G$. The cosmological context of this SZ cluster population analysis is a Λ CDM cosmology with parameters from Bocquet et al. (2015) ($H_0 = 68.6 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.292$).

²<https://www.darkenergysurvey.org>

Cluster redshifts were measured by using the red-sequence method, i.e. by searching for red galaxy overdensities in the colour-magnitude space in correspondence with the selected clusters. The selected photometric bands are the one containing the 4000 Å break and the adjacent redder band in different redshift intervals. The adopted bands are r , i and z . The redshift evolution of colours was managed by using a composite stellar population model.

One of the properties we need for the galaxy population is the radial distribution of galaxies which can be well described by a projected NFW profile up to $\sim 4R_{200}$.

A 3-dimensional NFW profile (Navarro, Frenk, & White, 1997) models the radial density of galaxies $\rho(r)$ as follows:

$$\frac{\rho(r)}{\rho_{crit}} = \frac{\delta_c}{\frac{r}{r_s} \left(1 + \frac{r}{r_s}\right)^2} \quad (4.1)$$

where the scale radius is the ratio between the virial radius and the concentration parameter, $r_s \equiv R_{200}/c_{200}$ and δ_c is the characteristic density.

For the model used within our cluster detection we used this profile and the parameters estimated by Hennig et al. (2017) by fitting both stacked and individual clusters (both showed no significant evidence for redshift or mass trend).

The projected 2-dimensional NFW profile can be expressed as follows (Navarro, Frenk, and White (1997); and according to Bartelmann, 1996 and Meneghetti, Bartelmann, and Moscardini, 2003):

$$\Theta(x) = \frac{N_0}{x^2 - 1} f(x) \quad (4.2)$$

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

$$f(x) = \begin{cases} 1 - \ln \frac{1+a}{\sqrt{1-x^2}} & x < 1 \\ 1 - 2 \frac{\arctan a}{\sqrt{x^2-1}} & x > 1 \\ 0 & x = 0 \end{cases} \quad (4.3)$$

where

$$a = \begin{cases} \sqrt{\frac{1-x}{1+x}} & x < 1 \\ \sqrt{\frac{x-1}{x+1}} & x > 1 \end{cases}. \quad (4.4)$$

The radial distribution normalization N_0 is the number of galaxies inside R_{200} , also known as N_{200} normalized itself over the virial area, consistently with what was done in Hennig et al. (2017). All the physical quantities, including the number of galaxies N_{200} are obtained by following the scaling relation suggested by Hennig et al. (2017), in their Equation 13 and Table 4.

The scaling relation contains mass and redshift dependence for different observables, indicated as a generic observable $O(M_{200}, z)$:

$$O(M_{200}, z) = A \left(\frac{M_{200}}{M_{piv}} \right)^B \left(\frac{1+z}{1+z_{piv}} \right)^C \quad (4.5)$$

where A is the normalization constant of the relation, B is the index of the mass dependence and C the one of the redshift dependence. The mass pivot point of the relation is $M_{piv} = 6 \times 10^{14} M_{\odot}$ and the pivot redshift is $z_{piv} = 0.46$.

As fully presented in Table 4 of Hennig et al. (2017), the best-fitting parameters, for instance for the overall population of galaxies, were found to be:

$$\begin{cases} A = 71.1 \pm 3.9 \\ B = 0.79 \pm 0.10 \\ C = -0.42 \pm 0.31 \end{cases}$$

for what concerns the number of galaxies, useful for the normalization, N_{200} .

Instead for the concentration c_{200} , the best parameter obtained were:

$$\begin{cases} A = 3.89 \pm 0.52 \\ B = -0.32 \pm 0.18 \\ C = -0.31 \pm 0.45 \end{cases}$$

with an overall mean estimate of the concentration parameter that is, for the full population $c = 3.59_{-0.18}^{+0.20}$, a higher concentration for the red population, $c_{red} = 5.73_{-0.24}^{+0.27}$ and a lower concentration for the blue population $c_{blue} = 1.38_{-0.19}^{+0.21}$.

A set of methods within AMICO build the radial distribution of galaxies inside the cluster model consistently with the assumptions, the models and the parameters that have just been described.

4.2.2 Luminosity function

In line with other applications of AMICO in the context of cluster detection, the luminosity function component of the cluster model has been com-

puted. The luminosity function is assumed to follow the so-called Schechter function (Schechter, 1976),

$$\Phi(m) = 0.4 \ln(10) \Phi_{\star} 10^{-0.4(m-m_{\star})(\alpha+1)} \exp[-10^{-0.4(m-m_{\star})}], \quad (4.6)$$

where the three parameters are the characteristic magnitude m_{\star} , the normalization Φ_{\star} and the faint-end slope of the distribution α .

Model for r -band magnitude

The first stage of our cluster detection covers the redshift interval $0 < z < 1.25$ and is performed by making use of the automatic r -band magnitude (MAG_AUTO) provided in the COSMOS2015 catalogue, obtained with the Subaru Supreme-Cam. The needed parameters for the luminosity function have to be consistent with the redshift range and the magnitude band that is being analyzed.

In a study carried out by Zenteno et al. (2016), the three parameters of the Schechter function in Equation 4.6 were tested and studied in the context of galaxy populations within clusters. The selected cluster sample they used consists of the 26 most massive structures detected within the South-Pole-Telescope Sunyaev-Zeldovich survey. The sample extends up to $z \sim 1.13$ with masses of $1.2 \times 10^{15} M_{\odot} \leq M_{200} \leq 2.7 \times 10^{15} M_{\odot}$. The computations have been done according to a Λ CDM cosmology, with estimates of the parameters given by Komatsu et al. (2011) ($H_0 = 70.2 \text{ kms}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.272$).

In order to constrain the three Schechter parameters without large uncertainties, they fixed one of the three parameters and extract the other two. The parameter m_{\star} has been fixed by using red sequence composite stellar population (CSP) models for some selected bands. The adopted SSP (Simple Stellar Population) models are the ones presented in Bruzual and Charlot (2003) along with the Python interface EZGAL by Mancone and Gonzalez (2012). The chosen evolutionary models are built with a Salpeter Initial Mass Function (IMF) (Salpeter, 1955) and a star formation decay time of 0.4 Gyr, at $z = 3$. This approach is justified by the fact that according to several studies the evolution of the parameter m_{\star} is well described by passively evolving stellar population already formed at high redshift (see e.g. de Propris et al., 1999, Mancone et al., 2010).

This approach was also tested and confirmed, always within the work by Zenteno et al. (2016), by letting m_{\star} as a free parameter.

The final result for the other Schechter parameters for the full population was found to be: $\Phi_{\star} = 2.24_{-0.20}^{+0.23}$ and $\alpha = -1.06_{-0.03}^{+0.04}$.

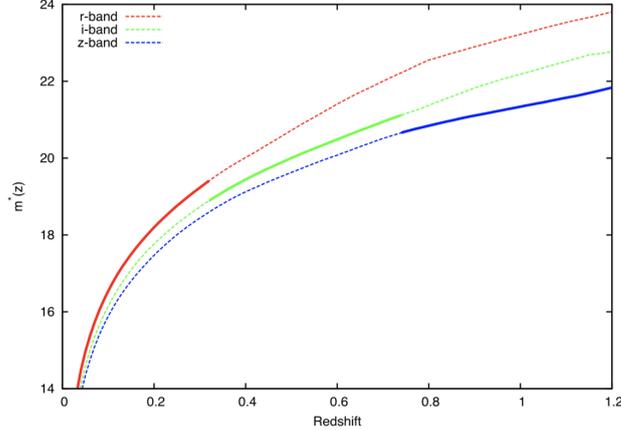


Figure 4.4: The redshift evolution of the characteristic magnitude m_{\star} extracted from the Bruzual and Charlot (2003) CSP model used by Hennig et al. (2017). The three colours represent the three bands used in this study. The values for the r -band magnitude (red) have been exploited to build the luminosity function within AMICO. Credits: Hennig et al. (2017), Figure 3.

The same procedure to study the evolution of m_{\star} has been exploited by Hennig et al. (2017), this time with a Chabrier IMF (Chabrier, 2003). The redshift evolution of the m_{\star} parameter is shown in Figure 4.4, from Hennig et al. (2017). The values that in this way were retrieved for the characteristic magnitude m_{\star} as well as the faint-end slope from Zenteno et al. (2016), are the main parameters describing the Schechter luminosity function, with cut at $m_{\star} + 2$, chosen to build the cluster model.

The choice is consistent with the use of the r -band magnitude in a the redshift range chosen here to be $0 < z < 1.25$, i.e. with an allowance of ~ 0.15 with respect to the context of the study, that we included by simple linear extrapolation of the curves.

The model for the r -band magnitude has been created directly within the AMICO algorithm, as described in Section 2.3. This model has been adopted for the chosen redshift range with redshift resolution step $\Delta z = 0.01$. The radial extremes of the model are set to $R_{min} = 0.0 \text{ deg}$ and $R_{max} = 2.0 \text{ deg}$ with step $\Delta R = 0.002$. The r -band magnitude range has been limited to m_{cut} and extended to include the brightest galaxies: $11.2 < r < 26.2$ with step $\Delta r = 0.5$. The cluster typical mass has been set to $M_{200} = 10^{14} M_{\odot}/h$, to which all the relevant quantities such as R_{200} and N_{200} correspond.

Model for H -band magnitude

The second phase of our cluster detection covers the full redshift interval ($0 < z < 2.0$) and is performed by making use of the automatic H -band magnitude extracted by the COSMOS2015 catalogue and belonging to the Data Release 2 of the UltraVISTA survey.

Higher redshift implies the use of a redder band, and therefore an adjustment of the model that could respond to the redshift range shift and to the filter change. The luminosity function for galaxies within a cluster is here therefore computed by relying on different parameters, coming from a study concerning higher-redshift structures.

The study that have been chosen to be suitable for this second analysis is the work of [Andreon et al. \(2014\)](#).

In this work, a detailed study of the $z = 1.803$ JKCS041 cluster has been presented. The used data are the near-IR HST data in the Y and H band with the addition of grism spectroscopy, exploited to investigate in details several physical properties and relations of such a high-redshift structure. Within the same work a comparison of JKCS041 with a sample of 41 lower-redshift clusters known in literature was performed. This comparison led to the formulation of a relation for the evolutionary trend of the faint-end slope (α) of the luminosity function, assuming once again that the population within clusters is successfully fitted by a Schechter function.

The selected sample of 42 clusters up to $z = 1.803$ is well fitted in the $\alpha - z$ space by

$$\alpha(z) = b(z - 0.5) + a, \quad (4.7)$$

with b being the evolution of α per unit redshift and a being the mean value of α at $z = 0.5$ ([Andreon et al., 2014](#)). The results found by [Andreon et al. \(2014\)](#) for these parameters are

$$\begin{cases} a = -0.98 \pm 0.03 \\ b = 0.08 \pm 0.09 \end{cases}$$

which, according to the value of the term b , is consistent with a statistical lack of evolution for the faint-end slope over the entire redshift interval.

Given the faint-end slope evolution in Equation 4.7, the cluster model has been constructed by using the mean value of α . The fitting analysis performed by [Andreon et al. \(2014\)](#) is displayed in Figure 4.5.

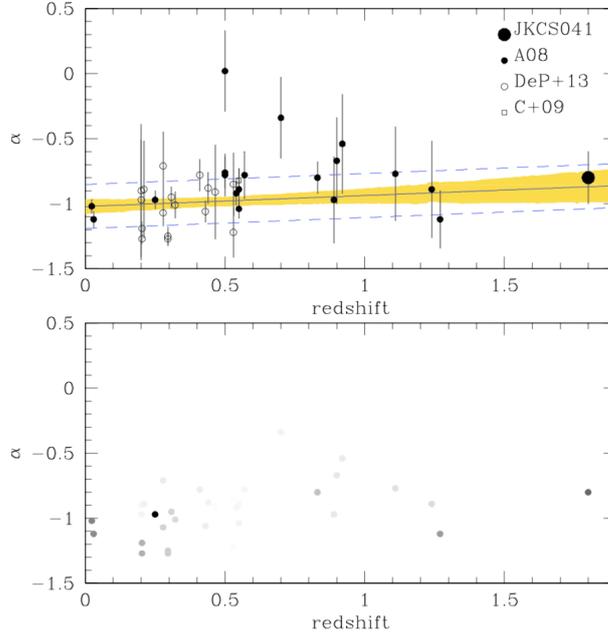


Figure 4.5: The redshift evolution of the parameter α for the red sequence galaxies of the cluster sample selected by Andreon et al. (2014). On top the cluster sample is shown with the mean model and 68% error range. The highest redshift cluster is JKCS041, the others are from data extracted from the literature (see Andreon et al. (2014), Section 5.1 for the full references; Andreon, 2008; Crawford, Bershady, and Hoessel, 2009; De Propris, Phillipps, and Bremer, 2013). On bottom the same data points are shown, with transparency depending on the attributed weight in the fit. It is the case that darker point are also stacked clusters. Credits: Andreon et al. (2014), Figure 11.

The same holds true for the value of the characteristic magnitude, the other Schechter parameter and its redshift evolution that was tuned according to a BC03 (Bruzual & Charlot, 2003) model with formation redshift $z = 3$. The analysis of this parameter evolution for the sample analyzed by Andreon et al. (2014) is shown in Figure 4.6.

The model for the H -band magnitude has been created in analogy with the r -band one: redshift step $\Delta z = 0.01$ over the interval $0 < z < 2$, radial extremes of the model $R_{min} = 0.0 \text{ deg}$ and $R_{max} = 2.0 \text{ deg}$ with step $\Delta R = 0.002$. In contrast to the r -band analysis, the model for this band has been generated with a new version of AMICO that offers the possibility to cut the radial profile internally and externally. The cut has been here chosen to be $(R_{in}, R_{out}) = (0.05, 1.5) \text{ deg}$. The magnitude range has been limited to m_{cut} : $11 < H < 25$ with same step as before $\Delta H = 0.5$. The cluster typical mass has been again set to $M_{200} = 10^{14} M_{\odot} / h$.

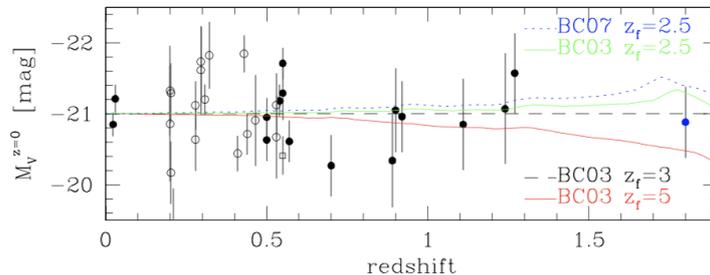


Figure 4.6: The evolution in redshift of the absolute characteristic magnitude of red sequence galaxies of the cluster sampled by Andreon et al. (2014). The lines show different Bruzual and Charlot (2003) models, in the 2003 and 2007 versions. The model with formation redshift $z = 3$ showed good consistency with the absolute magnitude evolution for JKCS041 and for the rest of the cluster sample. *Credits:* Andreon et al. (2014), Figure 12.

4.3 The noise assessment

As already mentioned in Section 2.3, the noise model is a parameter of fundamental importance in this formalism for the role it plays in the computation of the filter. Within AMICO, the noise model can be extracted directly from the overall galaxy distribution of the sample, under the assumption that the contribution of galaxy clusters is negligible.

The small area of the COSMOS data-set poses difficulties due to the limited statistics and the contamination of cluster galaxies that may not give negligible contribution to the noise. A wrong computation of the noise may have a strong impact on the process of construction of the filter because the main constants building up the filter are inversely proportional to the noise model N (e.g. the normalization constant in Equation 2.7) or even to the second power of it (e.g. the cluster variance in Equation 2.11), being the filter itself $\Psi_c = \frac{M_c}{N}$.

In this Section we will address these two aspects of the noise computation in the AMICO formalism in the specific context of the problematic issues encountered during this study.

First, the investigation of the influence of cluster galaxies on the noise will be briefly presented, then the procedure of optimization of the free parameters and their influence on the filter constants will be discussed in detail.

4.3.1 Cluster galaxy imprint removal

In the AMICO formalism, the noise model contains the statistical properties of the overall field galaxy population that are generally approximated to the mean properties of the full population in the data-set. This holds true if one assumes that the field is large enough to neglect cluster galaxy contribution. This turned out to be the case for the 2-deg² COSMOS field.

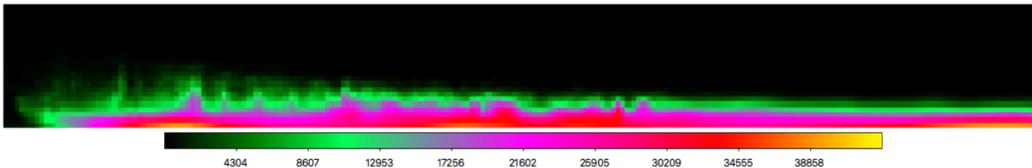


Figure 4.7: The preliminary noise model in the 2D-space (z , r magnitude). By moving to the right one shifts to higher redshift and by moving upwards one moves to brighter objects. The noise model shows stripes of higher noise values localized in redshift. The magnitude range is here set to $[14.2, 26.2]$ with step $\Delta m = 0.5$. This noise model is generated already by considering only safe galaxies according to the flagging of the catalogue, deblended and close-to-bright sources are here excluded.

Figure 4.7 shows the noise model based on the entire galaxy population. Redshift runs along the x -axis and magnitude along the y -axis: by moving to the right one shifts to higher redshift and by moving upwards one moves to brighter objects. The shown models are based on the r -band according to the constraints and the cleaning described in Section 3.4, unless otherwise indicated. The redshift range extends here up to 2 just for displaying, but it should be noted that the r -band analysis is performed up to $z = 1.25$.

The noise model displayed in Figure 4.7 shows the presence of peculiar vertical stripes in the $z-r$ -mag space, marking pixels with noise values larger than the surroundings, localized in redshift. This may be interpreted as the signature of clustered structures, in addition to the typical accumulation points produced by the photo- z s.

In order to investigate the influence of single clusters on the noise, a few different techniques have been tested in the context of this work. The main solutions are here presented:

- masking members from available cluster membership catalogues of the COSMOS field;
- cleaning with *cleanNoise*: a new code within AMICO to remove imprints of clusters directly from noise maps;

- performing a smoothing in redshift of the noise model.

Masking members. The COSMOS field offers the unique possibility to access a large amount of multi-wavelength good quality data to perform different kinds of studies. This includes samples of galaxy clusters produced, for instance, in the X-rays. A publicly available catalogue of cluster with $0 < z < 1$ is the one presented in [George et al. \(2011\)](#), that contains not only a cluster list but also a probabilistic member assignment catalogue. This catalogue will be more extensively presented and used in Chapter 6.

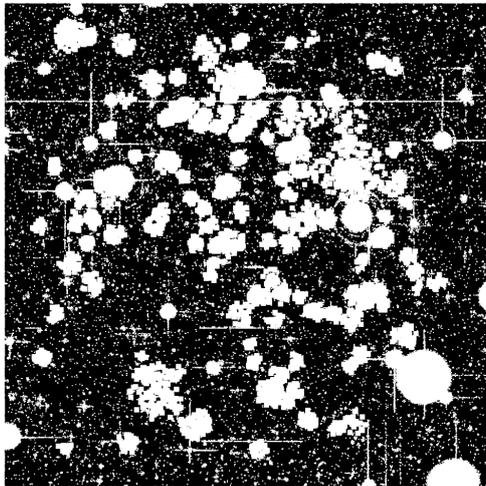


Figure 4.8: The composite mask generated with the aim of masking cluster areas and exclude member galaxies from the computation of the noise. The members are extracted from the catalogue presented in [George et al. \(2011\)](#). Smoothing and resolution have been tuned to cover cluster areas completely. We also suspected that the ring structure on top of the star circle in the center right could resemble the latter, and tried to shift the field and make them match. Nevertheless, the structure is not an artifact but a luminous area in the X-rays, probably due to the presence of two overlapping clusters. This will be expanded in Chapter 6.

We are here particularly interested in the catalogue of cluster members that we have used to generate a mask to exclude from the computation of the noise model all galaxies that have been assigned to a structure with probability $> 50\%$.

To generate the new mask used to exclude these objects we have used once more the *mask2combined* program to include the newly masked pixels in correspondence of the members. The resulting mask is shown in Figure 4.8. Clearly the member galaxies are rejected only during the computation of the noise and not during the detection procedure.

The result of this test has displayed no significant improvement in the removal of the vertical stripes, suggesting they might be not generated completely by cluster galaxies.

All tests performed in the computation of the noise with masking of members did not show sufficient improvement neither in the attenuation of the noise features nor in the $SNRs$ of the resulting candidate output detections.

cleanNoise. This new program introduced within AMICO gives the possibility, by exploiting the statistical analysis routines within AMICO, to extract the median or the average (with $\kappa\sigma$ clipping) of each correspondent pixel out of several noise estimates computed on non-overlapping portions of the survey.

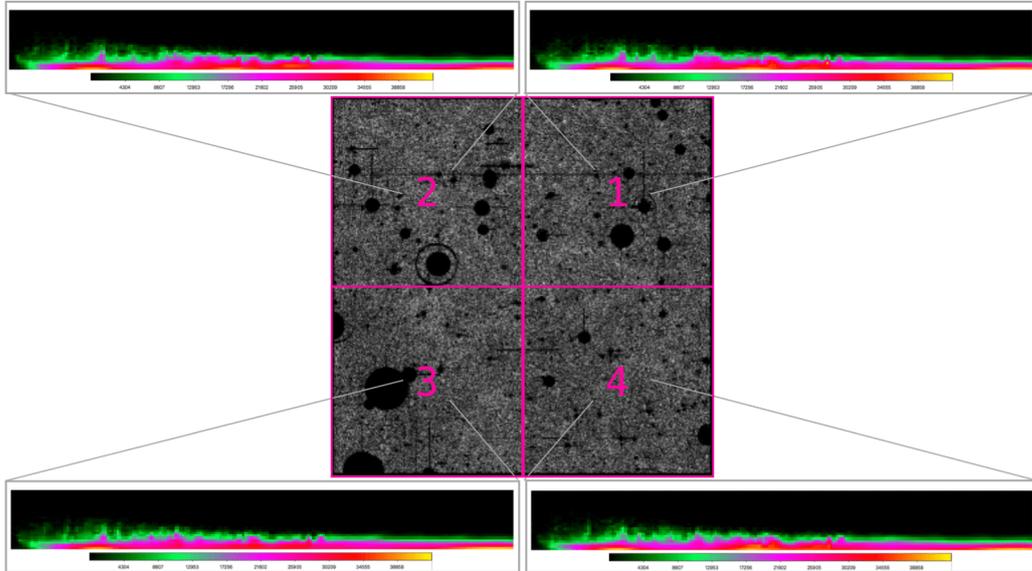


Figure 4.9: The COSMOS field divided in the 4 tiles used for the application of the *cleanNoise* method. For each tile the computed noise is shown. The scales are made consistent. The magnitude range is here set to $[14.2, 26.2]$ with step $\Delta m = 0.5$. In the background the density map shown in Figure 3.6 is displayed. These maps are generated already by considering only safe galaxies according to the flagging of the catalogue, deblended and close-to-bright sources are here excluded.

The idea behind this method is the following: if one of the noise maps shows the imprint of a clustered structure contained in that field, computing the median or the average with the same pixels but extracted by a different area that does not contain that structure would attenuate the imprint of the cluster in the final noise map.

A median can be performed between the pixels or alternatively a $\kappa\sigma$

clipping³ averaging.

The *cleanNoise* method has been tested on the COSMOS field with two different configuration: first by splitting it into 2 stripes and then into 4 adjacent tiles, then combined together with the described procedure. The 2 stripes have been created by dividing the field in upper and lower parts, in two intervals of DEC.

The partition in 4 tiles (2×2) is shown in Figure 4.9, where each noise map displayed is based on the galaxy population and the mask of the corresponding tile. The colour scale is the same in all panels to facilitate the comparison. Differences between the maps are visible, with some of the stripes likely to be generated by local overdensities of galaxies. When looking at the final noise models computed by integrating the information from the tiles (e.g. the result when using the median is shown in Figure 4.10), the imprints result in some cases suppressed but not completely removed.

An attempt to divide the field in 16 tiles (4×4) has also been performed, but discarded because the statistics was in this case too limited.

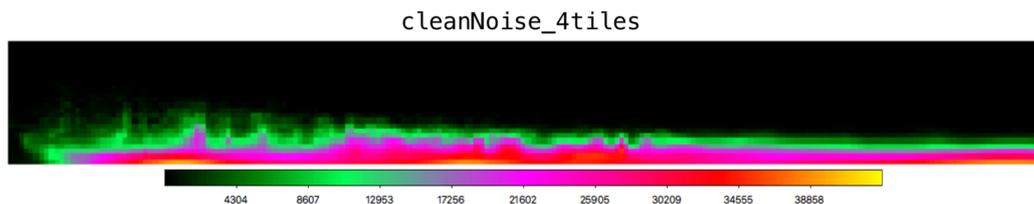


Figure 4.10: The final noise model computed by extracting the median of each pixel from the 4 maps of the 4 tiles in Figure 4.9. The scale is the same. The magnitude range is always set to $[14.2, 26.2]$ with step $\Delta m = 0.5$. Peaks and inhomogeneities have been removed but most of the stripes are still present. This noise model is generated already by considering only safe galaxies according to the flagging of the catalogue, deblended and close-to-bright sources are here excluded.

Smoothing. A more straightforward technique to attenuate the signatures localized in redshift in the noise model is to perform a smoothing of the final map along the redshift direction. On the bottom panel of Figure 4.11 an example of noise smoothing with a top-hat kernel of radius 2 pixel is shown.

The noise smoothing is in general a good technique to adopt, especially in case of poor statistics, to attenuate the impact on the noise of the signatures due to intrinsic peculiarities of the data-set.

³a method for computing statistical properties by rejecting values that are in absolute value larger than a factor κ which multiplies the value of the r.m.s. at each iteration.

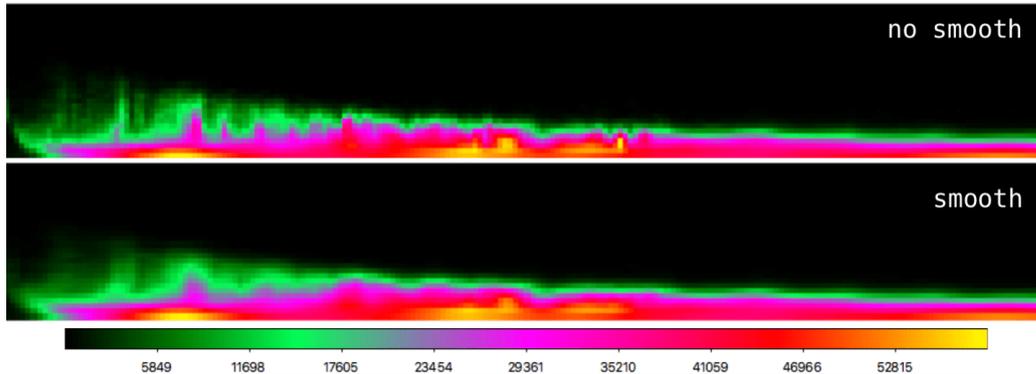


Figure 4.11: The final noise model computed without smoothing (top) and with a smoothing of 2 pixels in z (bottom). The scale is the same. The magnitude range is here set to $[11.2, 26.2]$ with step $\Delta m = 0.5$. Peaks, inhomogeneities and stripes have been attenuated in a significant way via redshift smoothing.

Despite a general substantial improvement in the quality of the results, the smoothing of the noise did not completely solve the problems related to the preliminary results, mainly consisting in few detections, low signal-to-noise ratios, unrealistic z -distribution of the detections and more importantly anomalies in the filter constant trends.

It should be added that it might be very useful to exploit a combination of the presented techniques, for instance by performing a cleaning of the noise with median and then smoothing the result along the redshift direction.

Apparently, the attempted removal or attenuation of the noise stripes proved that the impact of the cluster galaxies on the noise in the chosen data-set was not as significant as expected and could not be a main source of detection problems. The regions with higher noise values, localized in redshift were therefore identified as possible intrinsic accumulation points of photometric redshifts in this type of data-set. Thus, we investigated a new way to regularize the noise by tuning the initialization parameters.

4.3.2 Regularization of the noise

The filter implemented in the AMICO algorithm contains three main constants, introduced in Equations 2.7, 2.11 and 2.8, Chapter 2.

As already mentioned, AMICO offers the possibility of forcing to a chosen finite value the minimum noise in order to avoid numerical divergencies of the filter. This feature implemented in the C++ method `noise::set_minimum(min_val)`, assigns the chosen value (`min_val`) to every

pixel with value below it.

```
noise::set_minimum(min_val)
  for every pixel of noise file do
    val ← value of noise in the pixel
    if val < min_val then
      val ← min_val
    end if
  end for .
```

To improve the quality of the filter constant estimates and to make the filter more stable in case of small samples such as COSMOS, we introduced a new regularization. This method has been used within an application of AMICO to star clusters (described in Thomas Plewa, Master Thesis).

The previously introduced `min_val` has been replaced by two free parameters that are now regulating the noise values:

- `dummy_value_min` which is the actual minimum value that will appear in the regularized noise model;
- `dummy_value_zeros` which is the "dummy value" the program is going to assign to empty (*zeros*) pixels.

With this new version, in addition to introducing a "safety net" for numerical divergencies, the program is also taking care of making the extremes of integration in the filter constants finite.

This can be made by attributing to the second free parameter an arbitrarily large value, and has a significant impact on the computation of the filter quantities.

The new version of the method is therefore the following:

```
noise::set_minimum(dummy_value_min, dummy_value_zeros)
  for every pixel of noise file do
    val ← value of noise in the pixel
    if val < 10-6 then
      val ← dummy_value_zeros
    else if val < dummy_value_min then
      val ← dummy_value_min
    end if
  end for ,
```

where the double condition actually introduces three intervals of values. For noise values larger than `dummy_value_min`, the original noise model is kept, for values between `dummy_value_min` and minimum threshold (e.g. 10^{-6} here) the noise is forced to be equal to the `dummy_value_min` and below `dummy_value_min` it is forced to be equal to `dummy_value_zeros`. The intermediate interval is added to avoid loss of information under the `dummy_value_min` threshold, but it has in the context of this study, no significant influence on the results. A sketch summarizing the different regions of the regularized noise is shown in Figure 4.12.

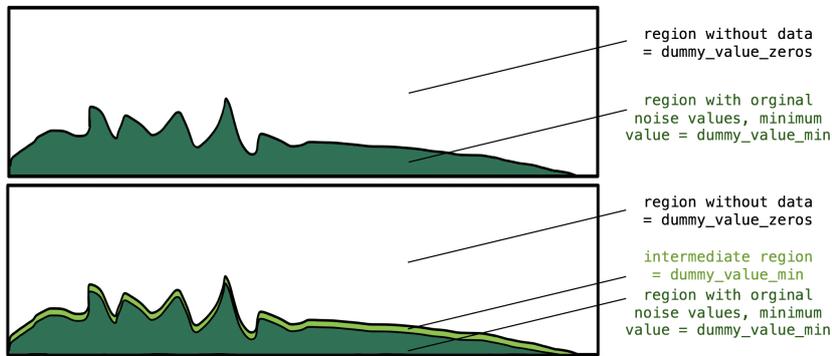


Figure 4.12: Sketch of the regularized noise model. The dark green region represents the noise part that is left as in its original form. The white region indicates the region containing empty pixels, without contribution of galaxies, and assigned therefore a high value to limit the integration of the filter constants. The intermediate area in light green is a stripe where small values of noise are forced to be equal to the chosen minimum threshold. The noise-model 2D-space has always z on the x-axis and r magnitude on the y-axis in this representation. The difference between top and bottom panels is just the presence of the intermediate stripe that helps not to lose information.

The two free parameters have been fine-tuned for the specific case of this application. The choice of the arbitrarily large value has not a big influence on the results, as long as it is sufficiently large to make the integration space finite in the unbound integrals of the filter constants (e.g. in Equation 2.7, 2.11 and derived or related quantities). The `dummy_value_min` has been set in such a way that the noise model pixels with excessively low values are assigned the contribution given by one fictitious galaxy. Since every pixel value is normalized over the pixel volume, the contribution of one galaxy is given by $\frac{1}{\Delta z \Delta m}$, where $\Delta z = 1$ by construction and Δm is the chosen magnitude resolution.

The choice of the minimum value may not be straightforward and it has been therefore tuned by studying its effects on the filter constants (see Appendix B). The problematic issue could be to lose information in low noise

regions of the $z-m$ space, when forcing them to be equal to a large value. Despite that in the selected range of possible value for the minimum, the results have shown good stability when changing the chosen value. The final couple of values chosen for our purpose is `dummy_value_min`, `dummy_value_zeros` = (2, 10^{10}). We used this numerical regularization for both analyses with the addition of a smoothing for the H -band noise model.

The major improvement due to the introduction of a numerical regularization procedure in the filter constants is visible in Figure 4.13. Here we show the constants associated to two runs, one with and one without the use of the new regularization. Clear differences are visible in the redshift evolution trend as well as in the absolute values, especially in the cluster variance γ , which depends on $\sim N^{-2}$, and in the local variance background correction factor, which is computed starting from the value of γ and therefore reproduces its trend.

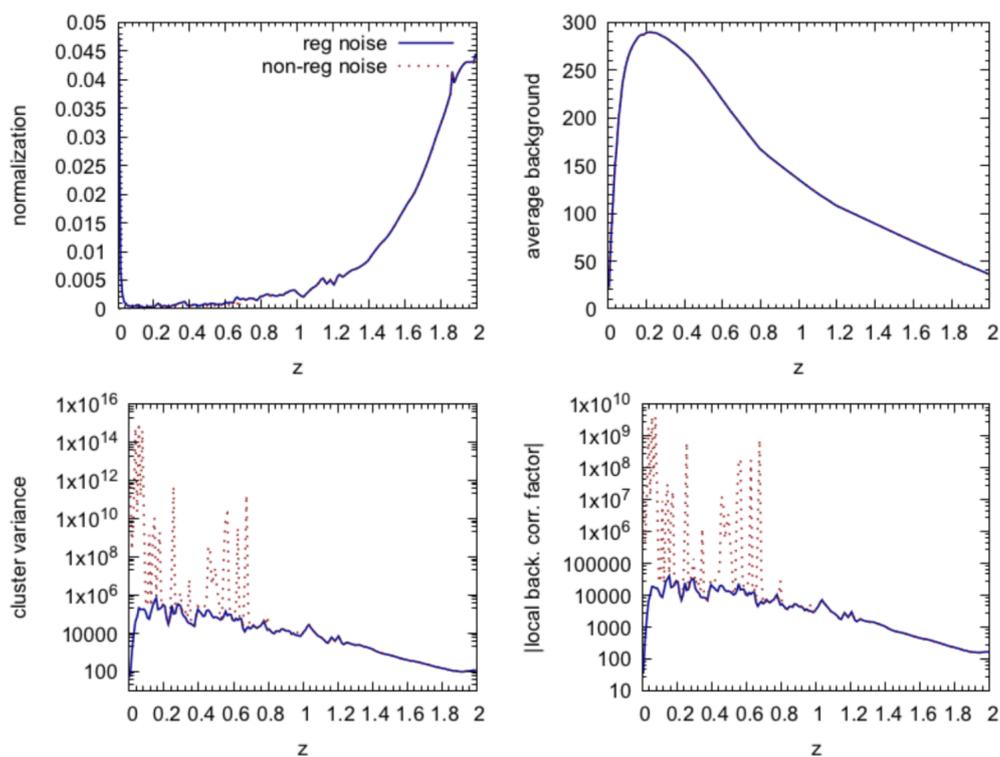


Figure 4.13: The redshift evolution of the four filter constants: normalization, average background, cluster variance and local background variance correction factor. Red dotted line stays for the filter constants without noise regularization. Blue line for regularized noise as described in the text. The introduced parameters, to limit the unbound integrals in the computation of quantities, have a consistent impact on the constants. The numerical regularization makes the constants smoother, particularly at low-intermediate redshift. Showing them together allows to appreciate the difference especially in the case of cluster variance and local background correction (bottom panels).

Chapter 5

The catalogue of cluster candidates

The main goal of this Thesis work has been the application of the AMICO algorithm to the COSMOS field photometric data to detect galaxy clusters.

After having obtained a catalogue of cluster candidates in the range $0 < z < 1.25$ using the r -band photometry, we tried to extend the study to higher redshifts, which is challenging both because of the quality of the data themselves and for the reliability of the model. Cluster modeling is based on observational studies, as discussed in Chapter ???. With increasing redshift, it is very difficult to study with precision galaxy clusters, given the increasing uncertainties and the poor statistics of known structures.

In this Chapter we will present the sample of galaxy cluster candidates obtained during the two distinct analyses, one based on the r -band photometry and one on the H -band photometry. The cluster candidates are provided with relevant cluster properties and galaxy membership information, up to $z \sim 1.8$. We detected a total number of 301 galaxy clusters in the r -band in the range $0 < z < 1.25$ and 351 clusters in the range $0 < z \lesssim 1.8$ detected in the H -band.

As mentioned in Section 2.5, AMICO galaxy cluster catalogues usually consist of a list of candidate clusters and an associated list of members. The catalogue of members contains all the objects of the original galaxy catalogue given as input. All galaxies in this list have the ID of the cluster they have been associated with if there is one (or more than one), with the corresponding association probability and the complementary probability of belonging to the field.

The cluster catalogue properties include the sky position, the redshift, the likelihood of being a cluster, the masked fraction and three different proxies of mass: the amplitude, the apparent richness and the intrinsic richness. First of all, the two cluster catalogues and their main statistical properties will be presented in Sections 5.1 and 5.2. A few examples of optical images of detected clusters will be shown together with the discussion about the membership catalogue. This will be done both for the catalogue extracted in the r analysis and in the H analysis: the resulting catalogues will be from now on referred to as **Catalogue r** and **Catalogue H** , for simplicity.

We will then make a comparison between these two lists of candidates, by performing a three-dimensional matching and by studying in this way the catalogues, their matched and non-matched detections and their reliability.

5.1 Cluster candidates of Catalogue r

The selected galaxy sample for the r -band magnitude analysis, described in Section 3.4 has been used as input catalogue to obtain a list of galaxy cluster candidates. The mask and the model for this band have been constructed as described in Section 4.2. Same holds true for the noise model, for which the numerical regularization has been exploited, as described in Section 4.3. We searched for clustered structures in the entire redshift window ($0 < z < 2$) but only the detections below $z = 1.25$ have been kept, due to the redshift range over which the adopted cluster model holds. The model derives from studies that have a coverage up to $z \sim 1.1$, so this limits the reliability of the modelled signal. Nevertheless, we have kept a margin of ~ 0.15 by limiting the results to $z = 1.25$ in which the extrapolation seemed to lead to realistic results as well. The final galaxy catalogue used in this run contained 479124 selected galaxies in total (of which 264915 with $z < 1.25$).

A total of 301 cluster candidates have been detected with $0 < z < 1.25$ and signal-to-noise ratio up to almost 7. The detected cluster sample has been analyzed, in terms of statistic properties and in terms of associated memberships.

The AMICO version used in this analysis bases the detection procedure order on likelihood and records also detections under the SNR threshold if they have sufficiently large likelihood. The minimum threshold is therefore not strictly defining the minimum SNR of the catalogue in this case. The SNR threshold has been chosen to be $SNR_{min} = 3.0$.

The actual sub-sample with SNR strictly larger than 3.0 contains 175 clusters in the selected redshift range.

Some of the main properties of the detections from the resulting Catalogue r are shown in Figure 5.1. The left panels show the values for each galaxy cluster of three mass proxies: the amplitude A , the intrinsic richness λ_* and the apparent richness λ , from top to bottom respectively. The right panels show the distribution of the full sample in redshift, signal-to-noise ratio and likelihood, from top to bottom respectively. The darker colour here indicates the full sample, while the light-blue colour the detections with the highest SNRs. In the redshift distribution of detections (top right panel), the main drops in detections may be attributable to accumulation points of the photometric redshifts and fluctuations in the number density. These latter are expected for such an area of coverage, tracing the intrinsic density fluctuations of structures along the line of sight in that portion of the sky. In this case some of the features in the cluster redshift distribution may be correlated to the distribution of galaxies in the used catalogue, shown in Figure 3.7. The drop at $z \sim 0.35$ may be also in part attributable to the covering of the 4000 Å break moving between the g and r bands. The distributions in SNR in the central right panel shows that, as expected, the SNR distribution peaks around the chosen threshold (SNR=3.0) and that the detections are not limited to the SNR minimum, since the detection criteria is based on likelihood. In the bottom right panel, we show how the highest-SNR detections are as expected among the largest likelihood structures.

5.1.1 AMICO richness and mass proxies

AMICO provides as output three different mass proxies:

- the amplitude, A , natural output of the filtering process given by Equation 2.2. It gives a mass proxy in units of mass as expressed by the adopted cluster model;
- the apparent richness, λ , i.e. the visible number of galaxies;
- the intrinsic richness, λ_* , i.e. the visible number of galaxies inside the virial region and with a cut in magnitude.

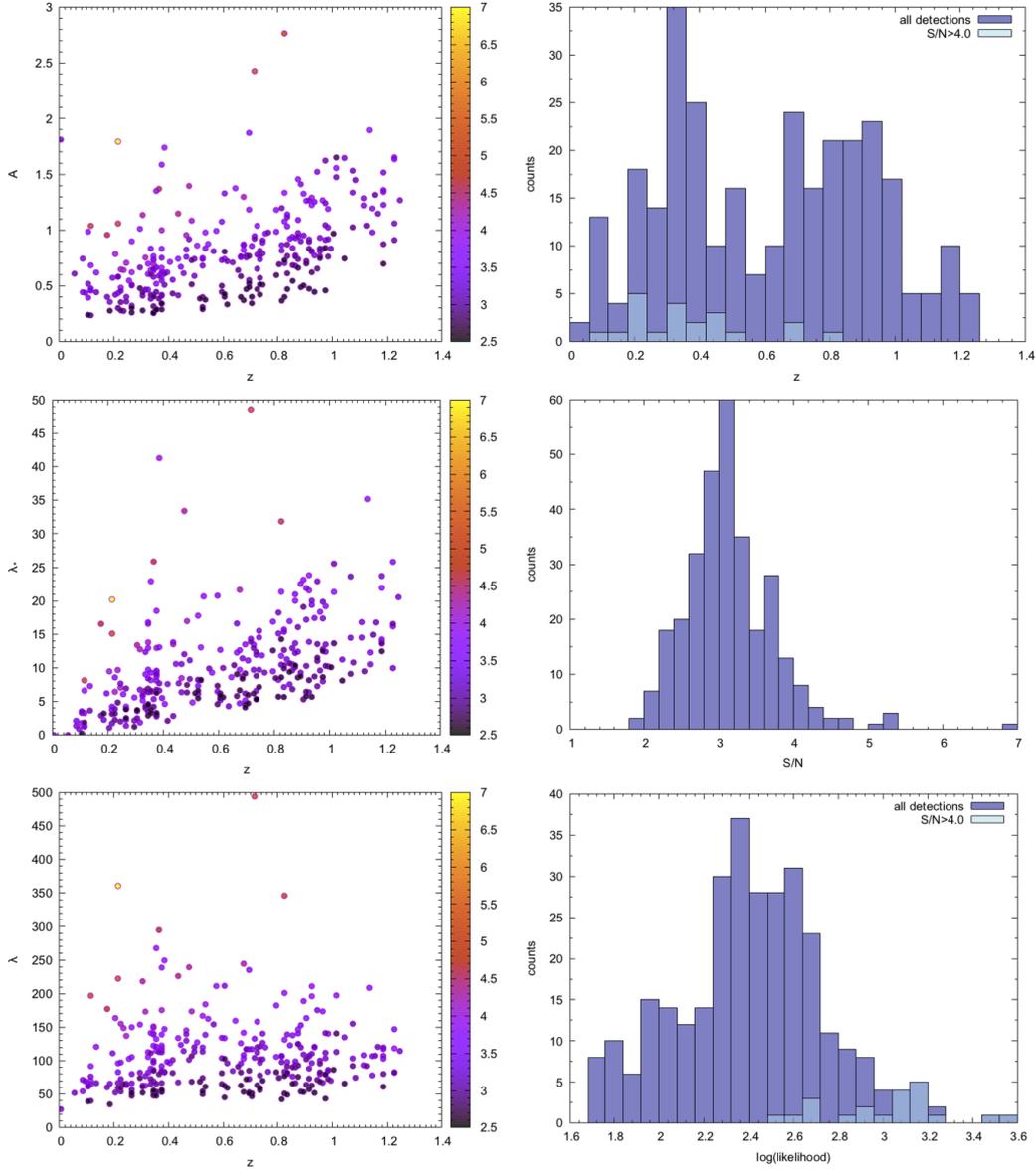


Figure 5.1: The amplitude (top left panel), the intrinsic richness λ_* , the number of visible galaxies restricted to $r < R_{200}$ and $m < m_* + 1.5$ (central left panel) and the apparent richness λ i.e., the visible number of galaxies in total (bottom left panel). These mass proxies provided by AMICO have different evolution trend with redshift due to how they have been defined. Amplitude and intrinsic richness are the most robust quantities to infer mass. The points are colour-coded according to the SNR of the detection as indicated in the side bar. On the right side column panels (from top to bottom) the distribution of detections with respect to redshift, SNR and likelihood. The darker colour histograms represents the full sample of Catalogue r, the light-blue ones the detections with SNR > 4.0.

The richness as defined within AMICO is actually not a simple number count of galaxies associated to the cluster but takes into account how likely

it is that these galaxies belong to the structure.

The apparent richness, λ , is defined as the sum of probabilities of each i -th member galaxy associated to the j -th candidate cluster:

$$\lambda_j = \sum_{i=1}^{N_{gal}} P_{i,j}, \quad (5.1)$$

where $P_{i,j}$ is the membership probability expressed by Equation 2.25.

The intrinsic richness, λ_* , is defined in the exact same way but limiting the sum to galaxies brighter than $m_* + 1.5$ and within R_{200} , where the characteristic magnitude m_* and the virial radius R_{200} are parameters of the cluster model used to define the filter, see Section 4.2:

$$\lambda_{*,j} = \sum_{i=1}^{N_{gal}} P_{i,j} \quad \text{with} \quad \begin{cases} m_i < m_*(z_j) + 1.5 \\ r_i < R_{200}(z_j) \end{cases}, \quad (5.2)$$

being z_j the redshift of the cluster candidate.

Here the characteristic magnitude m_* evolves with z as described in Section 4.2. We have kept the cut radius fixed to avoid the uncertainty associated with the mass, on which R_{200} depends, related for instance to the amplitude of the structure (Maturi et al., 2019). The definition of the cluster richness in Equation 5.2 is similar to the one implemented in redMaPPer (Rykoff et al., 2014), another widely used detection algorithm.

The apparent richness, λ , is a strongly redshift dependent quantity, since the further the cluster, the fewer the visible galaxies. Thus, the magnitude threshold is defined by the survey depth. On the contrary, thanks to how it has been defined, the cut in the intrinsic richness makes it a nearly redshift-independent quantity, providing us with a robust quantity to infer cluster masses (Bellagamba et al., 2019). This clearly holds true as long as the magnitude cut $m_* + 1.5$ is below the limiting magnitude of the survey, which is the case for the COSMOS data-set and the considered redshift range.

The amplitude also provides a mass proxy that is not strongly dependent on redshift. Compared to λ_* , the amplitude is not as closely linked to the direct observable, i.e. the number of galaxies in the cluster and it can be also boosted by bright galaxies contributing to the value of A , due to the magnitude dependence of the filter.

In the left panels of Figure 5.1 it is possible to observe the values of these three mass proxies, A , λ_* and λ (from top to bottom) and how they are

distributed with respect to the redshift of the detected clusters, colour-coded according to the SNR of the detection. Amplitude and intrinsic richness show an expected trend, i.e. the distribution of detections rises with increasing z because at large distances only the intrinsically most massive structures are expected to be detected. For what concerns the apparent richness, a redshift trend is not present, in accordance with what discussed above: this estimate of the richness takes into consideration all visible galaxies associated to the cluster.

5.1.2 Amplitude map and optical image inspection

After investigating the statistical properties of the retrieved sample of clusters, we now examine the amplitude map, in which a value of A is attributed to every point of the selected volume. Investigating the redshift slices of the map can be useful to notice possible problems in the detection procedure, to fine-tune the resolution of the cube bins and to double check the location of the detection centres. The same applies for the examination of the variance map and the cleaned version of the maps. In Figure 5.2, a redshift slice of the amplitude map retrieved from the r -magnitude run is shown on the left panel and the same slice but for the cleaned amplitude map is shown on the right panel. It is possible to appreciate the cleaning procedure described in Section 2.4.2, with the removal of the sequentially detected clusters that were present in the left panel. The cleaning procedure, as already mentioned, takes into consideration the membership probability of the galaxy associations in order to weight their contribution to the cluster signal. This particular redshift slice shows an example of how the imprints of different structures have been removed, starting from the highest amplitude one (small white square in the right panel), then proceeding with smaller and blended structures. Darker regions in the amplitude map trace intrinsic underdensities of galaxies that are expected on these scales, with the exception of the dark circular areas lacking objects due to masking (e.g. bottom left).

Another important step, once the catalogue of candidates has been retrieved, is to visually inspect the optical images in correspondence with the detections to access member galaxies and to look at the significance and the reliability of the detections. Moreover, the image inspection is not only a tool for the detection analysis but also a test for the consistency of the selected input catalogue.

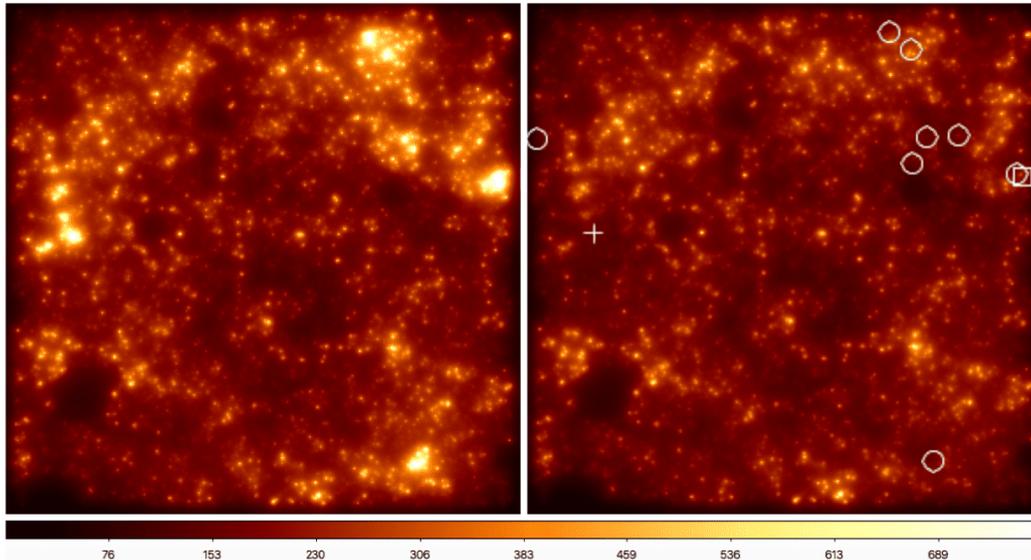


Figure 5.2: Slice of the amplitude map (left) and the cleaned amplitude map (right) computed during the r -band run at $z = 0.465$. AMICO removes the imprint of sequentially detected objects from the map, in order to allow the subsequent detection of smaller and blended structures. The subtraction is performed by exploiting the association probability of member galaxies. The white square in the right panel indicates the position of the first cluster that has been removed, namely the highest amplitude cluster, the cross marks the second one and the circles are placed in correspondence of the the rest of the main removed clusters that were contributing to the amplitude in this slice of the map. To make the comparison consistent the maps are displayed with the same colour scale, reported below the panels.

For this analysis we have used the Subaru Hyper Supreme-Cam (HSC) Data Release 3 (Aihara et al., 2021) images¹ both for consistency with the galaxy catalogue we used and because of the depth that allows to resolve also very distant galaxies, with their morphologies and colours. This data release belongs to the Hyper Supreme-Cam Subaru Strategic Program (HSC-SSP), a multi-band wide-field imaging survey performed with the 8.2 m Subaru telescope on the top of Mauna Kea, Hawaii (Aihara et al., 2018). The total covering is 1470 square degrees, including the COSMOS field, observed in the UltraDeep layer ($r = 27.1$, 5σ depth) (Aihara et al., 2021).

¹<https://hsc-release.mtk.nao.ac.jp/doc/>

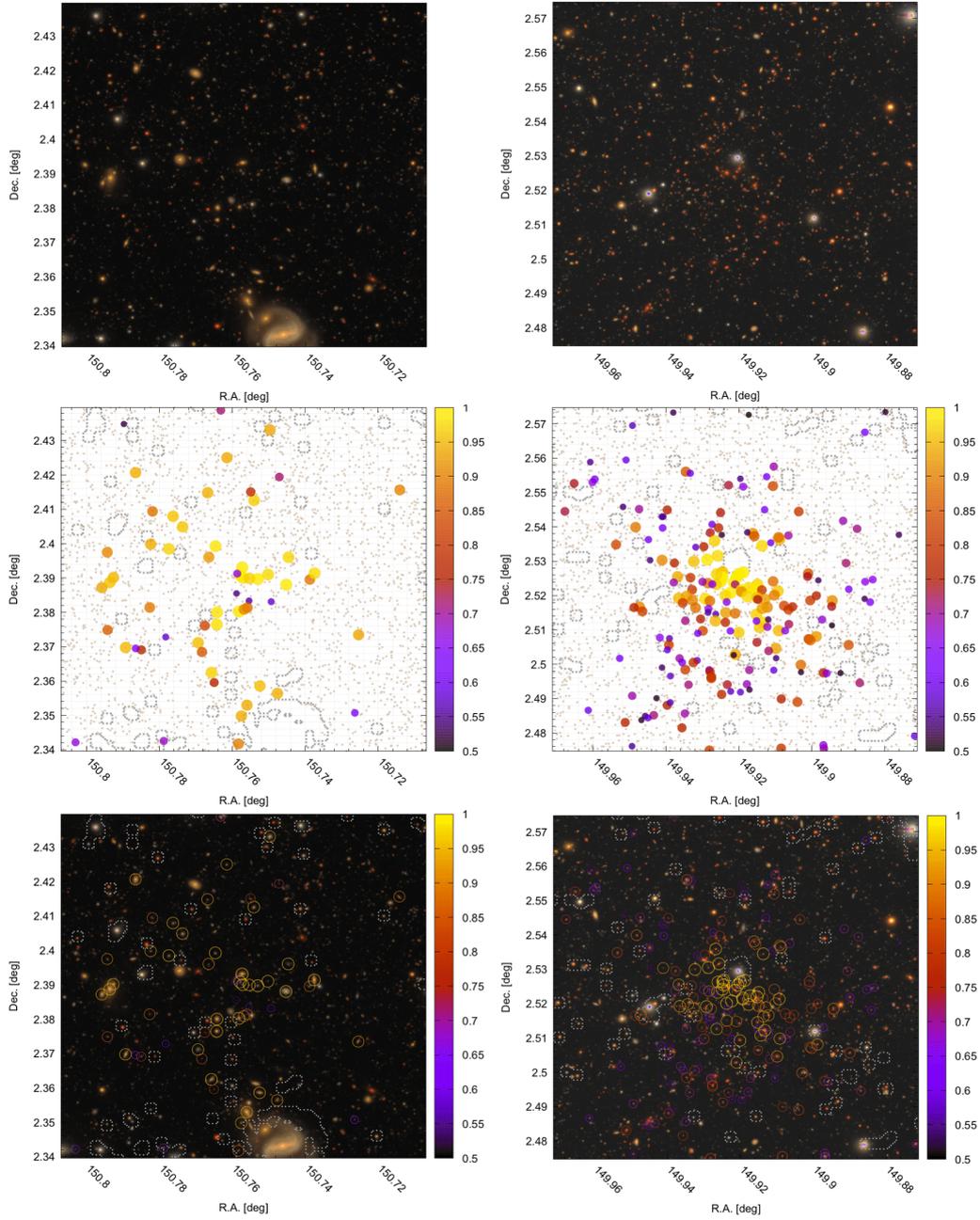


Figure 5.3: Two examples of detections (two columns), one with $\lambda_* \approx 16$ at $z = 0.175$ (left column) and one with $\lambda_* \approx 49$ at $z = 0.715$ (right column). On top, the HSC colour composite (g, r, i) image centered in the cluster candidate centre. The middle panels show how AMICO identified the clusters, with associated galaxies, with colour-coded association probability (according to the side bar). Only galaxies with probability larger than 50% are here shown. Grey dotted contours delimit the masked regions, and beige points in the background indicate the full galaxy catalogue. Especially in the right panel, it is possible to notice how the association probability is higher in the central regions of the cluster. The bottom panels show the HSC image and the associated galaxies with colour-coded probability and mask contours overlapped. All stamps have a side size of 0.1 degrees and are centered in the centre of the cluster as found by AMICO.

In the context of the study described in this Thesis work, during this inspection process, we had the opportunity to notice the absence of a few clear cluster galaxies that were present in the images but absent in the selected galaxy catalogue. In this occasion, we realized that a more relaxed masking procedure was necessary and that we had to include a sample of flagged galaxies due to their relevance within clusters, despite the lower quality of their photometric properties. This procedure has been described in Section 3.3.2 and required a re-examination of the COSMOS2015 catalogue in order to make a more appropriate choice of `FLAG_PETER` values.

Figure 5.3 shows two examples of detections from our final catalogue of clusters for the r magnitude run. The two columns of panels correspond to two different clusters.

5.2 Catalogue H: extending the cluster search up to $z \sim 1.8$

The galaxy catalogue for the analysis in the H -band, described in Section 3.4, has been used as input to extend the cluster detection to higher redshifts. The mask has been corrected to account for the loss of effective area due to the smaller angular coverage of the UltraVISTA data the H -band analysis is based on, as described in Section 4.1.1. The H -band galaxy catalogue covers an area $\sim 23\%$ smaller than the r -band one, and this has to be taken into account when comparing the two cluster catalogues. The noise model was numerically regularized as discussed in Section ?? with a smoothing of the counts of 2 pixels in radius along the redshift direction in order to attenuate signatures generated by the galaxy properties. The final galaxy catalogue used in this run contained 340247 galaxies in total (of which 269985 in the considered redshift interval).

The model describing the galaxy distributions for this band and for the chosen redshift interval is the one described in Section 4.2.2. This model is based on studies performed on a sample with the JKCS041 cluster as the highest- z structure at $z = 1.803$, justifying a possible coverage of the redshift range to look for structures up to ~ 1.8 . At this stage of the analysis, a new feature of the AMICO algorithm has been exploited to make the model suitable to high- z detections: the possibility to cut the profile in the model at a minimum and maximum radius, $R_{min}, R_{max} = (0.05, 1.5)$. The fine-tuning

of the newly introduced cuts has allowed obtaining more stable results.

We searched for clusters in the entire redshift range $0 < z < 2$ with a particular focus on the interval complementary to the r -magnitude search, that is likely to lose reliability at redshift higher than ~ 1.25 . A total of 351 clusters have been detected with $0 < z \lesssim 1.8$ and signal-to-noise ratio between 3 and 8.

The new version of AMICO, used within this run, bases the detection procedure order on SNR and not on likelihood as before. Therefore, the SNR minimum threshold defines the actual minimum SNR of the final catalogue. The SNR threshold has been chosen once again to be $SNR_{min} = 3.0$ and all the detections have $SNR > 3.0$.

In the redshift interval not reliably covered by the r -band analysis, i.e. $1.25 < z \lesssim 1.8$, AMICO detected 31 clusters, mostly found also with different configurations of the initialization parameters.

The main properties of the full sample are shown in Figure 5.4. The same considerations as before hold true for the right column distributions. The full sample of detections is represented by the darker colour, the light-blue colour marks instead only the highest-SNR detections. The redshift distribution is consistent with the one derived in the Catalogue r, with the drops likely to be due to photo- z accumulation points and density fluctuations.

In the left column, the three mass proxies introduced in the previous Section are shown for the full Catalogue H. The expected increasing trend for the amplitude is clearly visible at high redshift. This is not the case for the intrinsic richness which does not show the expected increase of its minimum with increasing redshift for the entire range and at $z > 1$ begins to resemble the tendency of λ . This is due to the fact that the magnitude cut in the definition of λ_* reaches the survey magnitude limit. At high z , the number of faint objects excluded from the count of galaxies in the intrinsic richness gradually approaches the number of galaxies excluded in general from the count, because of the data-set limitations.

The visual inspection of the amplitude map and of the optical Subaru HSC images of the detections have been performed in the same way as for the Catalogue r. Figure 5.5 shows an example of two of the highest- z candidate clusters of Catalogue H with their associated galaxies. Other examples of structures that have been detected within the study described in this Thesis are shown in Appendix C.

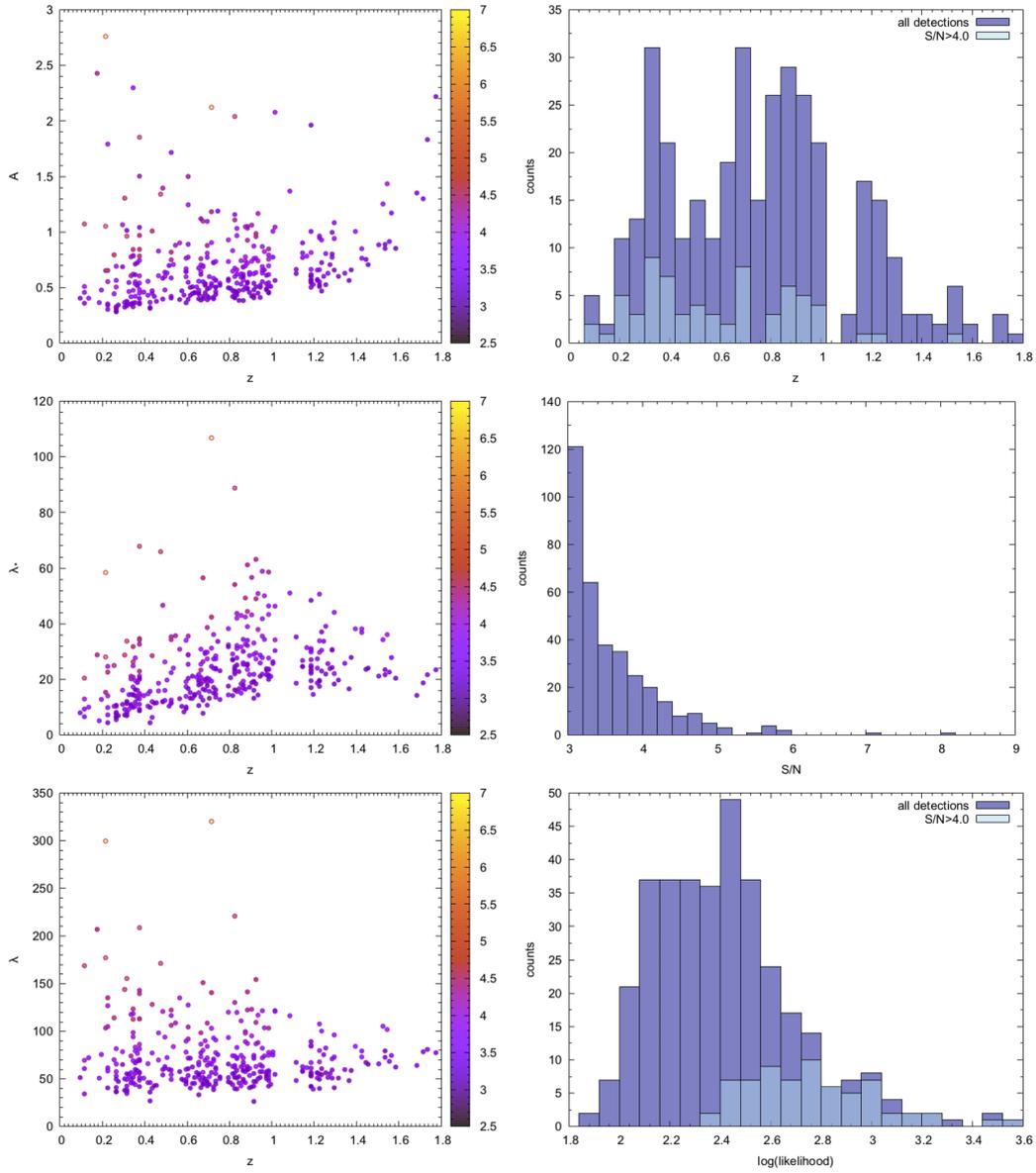


Figure 5.4: The amplitude (top left panel), the intrinsic richness λ_* (central left panel) and the apparent richness λ (bottom left panel) for the full Catalogue H. The points are colour-coded according to the SNR of the detection as indicated in the side bar. On the right side column panels (from top to bottom) the distribution of detections with respect to redshift, SNR and likelihood. The darker colour histograms represent the full sample, the light-blue ones just the detections with SNR > 4.0.

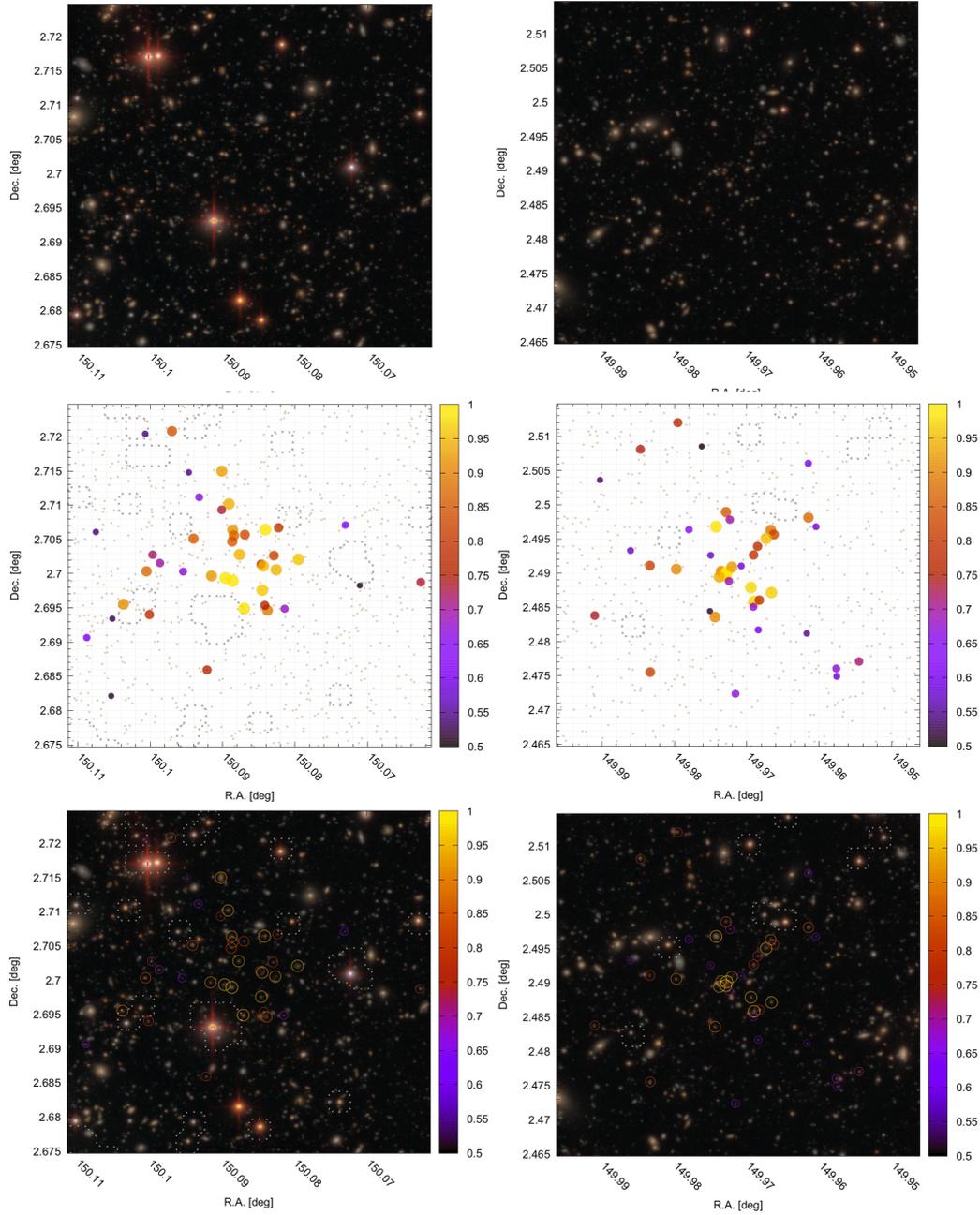


Figure 5.5: Two examples of detections (two columns), one with $\lambda_* \approx 44$ at $z = 1.295$ (left column) and one with $\lambda_* \approx 38$ at $z = 1.425$ (right column). On top, the HSC colour composite (i, z, y) image centered in the cluster candidate center. The middle panels show how AMICO identified the clusters, with associated galaxies, with colour-coded association probability (according to the side bar). Only galaxies with probability larger than the 50% are here shown. Grey dotted contours delimit the masked regions, and beige points in the background indicate the full galaxy catalogue. The bottom panels show the HSC image and the associated galaxies with colour-coded probability and mask contours overlapped. All stamps have a side size of 0.05 degrees, half the size of the stamps shown in Figure 5.3.

5.3 Matching Catalogue r with Catalogue H

In this Section we present the three-dimensional matching performed between Catalogue r and Catalogue H, exploiting the sky position and the assigned redshift of the cluster candidate, using the matching method available within the AMICO code.

The separation in the sky plane has been fixed to $d_{\text{rad}} = 0.3 \text{ Mpc}/h$ where the conversion between angles and Mpc has been based on the detection redshift related to the angular diameter distance expressed by Equation 1.7, given the adopted cosmology (see Section 1.1). The redshift matching separation has been set to $dz = 0.03(1+z)$. Notwithstanding, the two catalogues are not consistent with each other in terms of volumetric coverage. First of all the redshift interval is different, therefore we have restricted the matching to $z < 1.25$ that is the threshold chosen for the Catalogue r. Additionally, we have to take into consideration that the Catalogue H covers a reduced area with respect to Catalogue r. To account for this difference, we rejected all the detections of Catalogue r falling into masked regions according to the mask used for the H -magnitude run. In this way the matching is consistently performed between catalogues covering the same volume. The matching procedure identified the correspondence between 167 clusters in the interval $0 < z < 1.25$, namely $\sim 66\%$ of the selected sub-sample of the Catalogue r.

The deviations in position and redshift between the matched detections are shown in Figure 5.6.

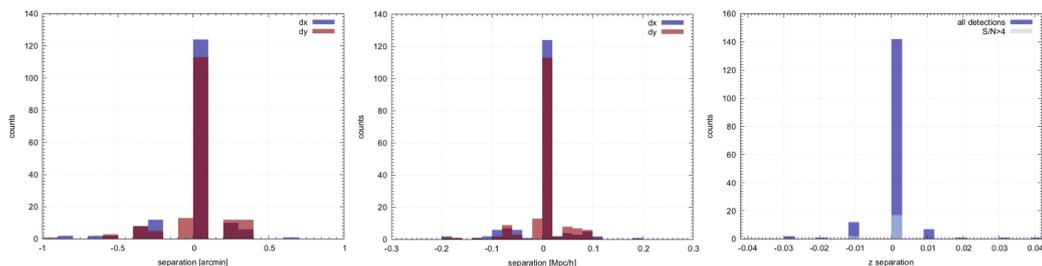


Figure 5.6: Properties of the 167 clusters successfully matched between Catalogue r and H. The left panel shows the separation between matched cluster centers in arcmin and the central panel the separation in Mpc/h. Blue histograms represents separation in RA and red one in DEC. The right panel shows the redshift separation, with blue marking the full distribution and light blue the sub-sample with $\text{SNR} > 4.0$. All distributions are peaked around zero, indicating the centre determination is most of the times consistent.

In the clusters found in both runs the centres have been determined in a consistent way.

Among the 167 successful matches, 106 structures have been detected with $SNR > 3.0$ in both catalogues. We have not restricted the match to structures detected with $SNR > 3.0$ a-priori to not bias the matching, because the signal-to-noise ratio has been found to be not always consistent between the two runs.

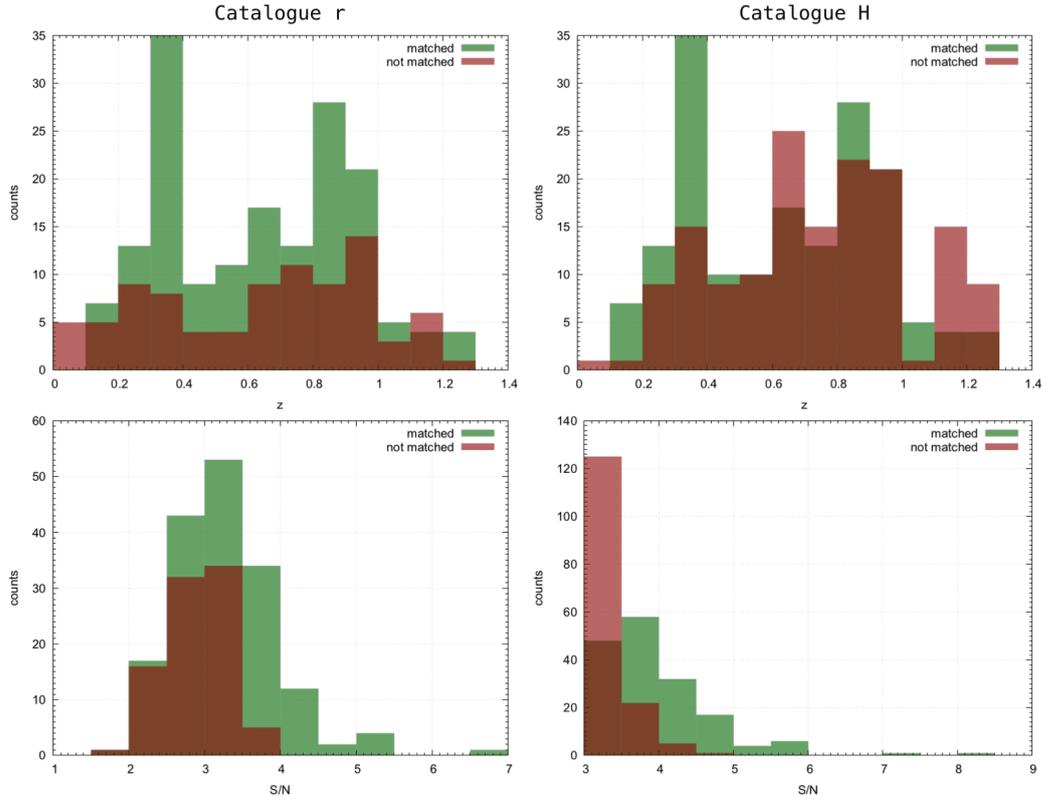


Figure 5.7: The distributions of the clusters detected in the Catalogue r (left column) and the Catalogue H (right column). The green histograms indicate the sample of structures that has been found in both catalogues according to the matching described in the text. The red histograms show the non-matched structures. Top panels show the redshift distributions and bottom panels the SNR ones.

Figure 5.7 shows how the distribution of detections in z and SNR splits into clusters identified in both analyses, represented in green, and cluster identified just during one of the two analyses, represented in red. Both r and H -band catalogues are shown. The non-matched clusters are mostly in the low-SNR regime (bottom panels), but seem to be present over the full redshift range.

It should be noted that among the low- z clusters, 4 have been detected in the r -band run with $\lambda_* < 1$. This gave the chance to consider the possibility to introduce in the AMICO algorithm an adjustable threshold to avoid the detection of structures with unrealistic low richness.

At this point, with the same data described before for the optical image inspection, we present the membership assignment consistency and the properties of the clusters such as center and extension for three examples of matched detections in both catalogues, shown in Figure 5.9. The left side panels show three clusters detected in the r magnitude run, and the right panels show the corresponding clusters found in the H magnitude run. As in the previous Sections, the circles mark the member galaxies with colour expressing the membership probability assigned by the algorithm. All stamps have size 0.1 deg and are based on g, r, i and r, i, z Subaru HSC colour composite images. The corresponding detections in the two catalogues have shown a slight difference in the centering and the absence or the presence of certain galaxies that may have magnitude in one filter available but not in the other, being not included in the respective galaxy samples. The detected cluster shown in the top panels lacks galaxies in the H -band identification, being the structure at the edge of the UltraVISTA field.

Additionally, we examined a sample of non-matched detections both for the Catalogue r and for the Catalogue H. Two examples of detections which appear to be clusters in the HSC images are shown in Figure 5.8. We show here two structures detected in the r -band run but not in the H -band run (left column) and vice versa (right column).

A summary of the different sub-samples of the catalogues produced within this study is shown in Table 5.1.

The total number² of clusters detected within the study presented in this Thesis work, through the application of AMICO to the COSMOS field is 481. We will further discuss the differences between the two catalogues retrieved in our study during the X-ray analysis to take advantage of these additional data to add further information to the comparison. This will be addressed in Section 6.5.

²considering clusters detected within both runs, without counting double the clusters found in both runs and by rejecting the detections with $\lambda_* < 1$.

| Catalogue ID | redshift | area | SNR | # clusters |
|--------------|----------------|------------|-------------|------------|
| Catalogue r | $0 < z < 1.25$ | COSMOS | all | 297* |
| | $0 < z < 1.25$ | COSMOS | $SNR > 3.0$ | 175 |
| | $0 < z < 1.25$ | UltraVISTA | all | 255 |
| | $0 < z < 1.25$ | UltraVISTA | $SNR > 3.0$ | 145 |
| Catalogue H | $0 < z < 1.8$ | UltraVISTA | $SNR > 3.0$ | 351 |
| | $0 < z < 1.25$ | UltraVISTA | $SNR > 3.0$ | 320 |
| | $z > 1.25$ | UltraVISTA | $SNR > 3.0$ | 31 |
| Matches r&H | $0 < z < 1.25$ | UltraVISTA | all | 167 |
| | $0 < z < 1.25$ | UltraVISTA | $SNR > 3.0$ | 106 |

Table 5.1: Summary of the samples of clusters detected in the COSMOS field. Redshift range, covered area of the field, SNR constraint and number of detected structured are shown for the Catalogue r, the Catalogue H and the catalogue of successful matches between the two. The * indicates that clusters with low intrinsic richness have been removed. The total number of detected clusters in the context of this study is 481.

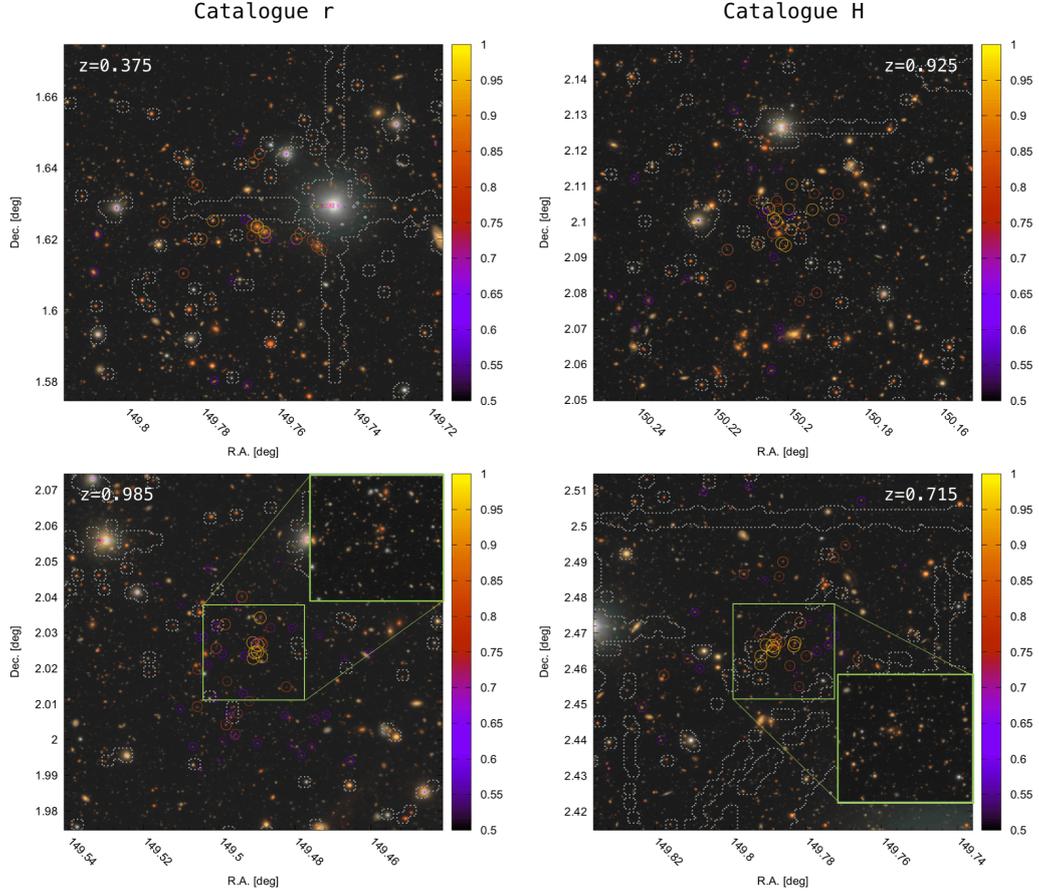


Figure 5.8: Four examples of non-matched detections, present just in one of the two catalogues. The left side column panels show two clusters detected in the r -magnitude run but not in the H one, vice versa in the right column. The circles mark the member galaxies with colour according to the membership probability assigned by the algorithm (colours in the side bar). In the bottom panels, the central regions of the cluster are zoomed in to show the central galaxies. The zoomed stamps have side size of 0.025 deg and are HSC r, i, z colour composites. All the main stamps have side size 0.1 deg and centered in the center of the cluster and based on HSC g, r, i colour composite images. The reshifts of the structures are reported in each image.

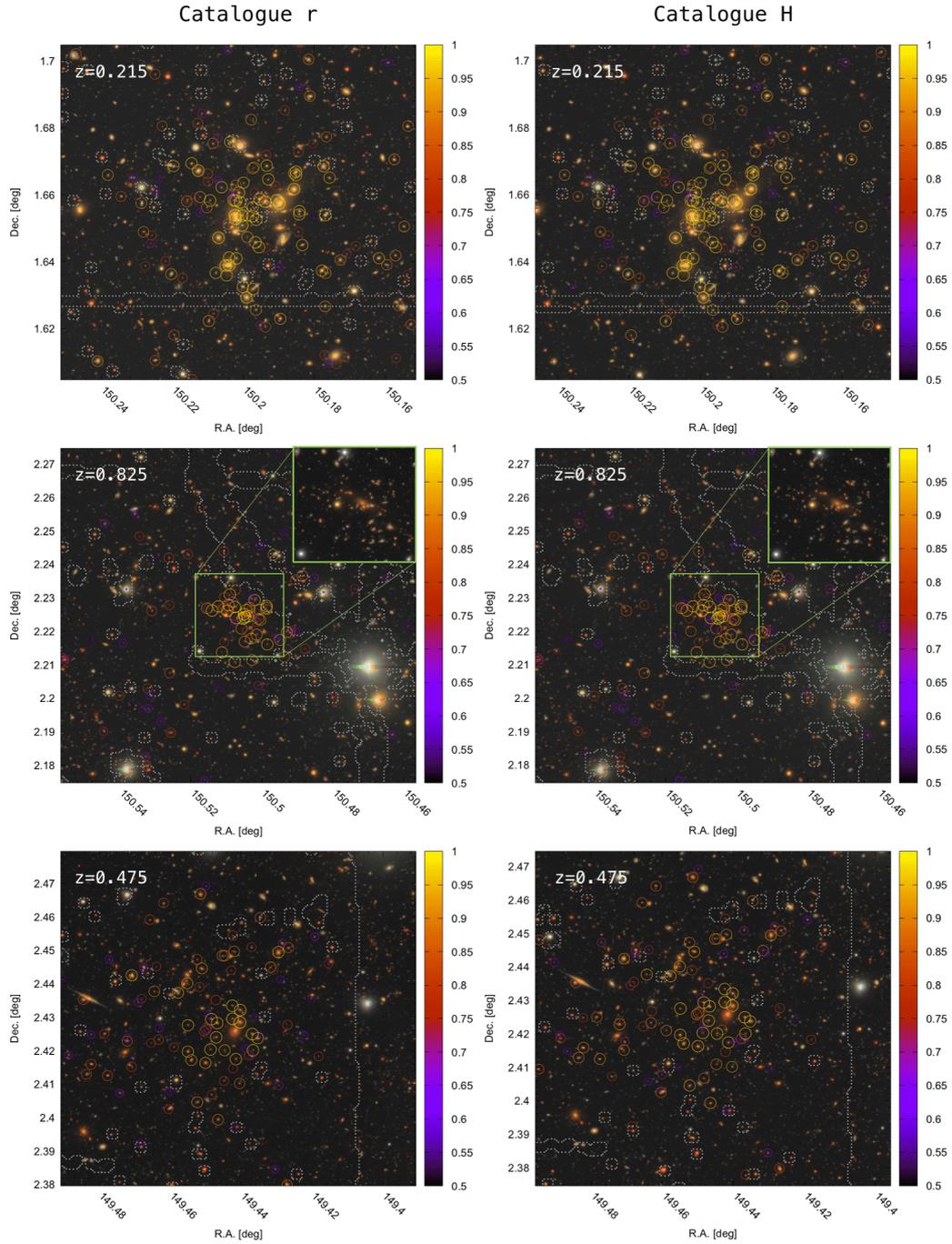


Figure 5.9: Three examples of matched detections, present in both catalogues. The left side column shows three clusters detected in the r -magnitude run, the right column the corresponding detections in the H -magnitude run. The circles mark the member galaxies with colour according to the membership probability assigned by the algorithm (colours in the side bar). In the middle panels, the central regions of the cluster are zoomed in to show the central galaxies. The zoomed stamps have side size of 0.025 deg and are HSC r, i, z colour composites. All the main stamps have side size 0.1 deg, centered in the center of the cluster and based on HSC g, r, i colour composite images. The reshifts of the structures are reported in each image.

Chapter 6

Matching with X-ray catalogues

One of the main sources of discovery of galaxy clusters and sequential study of their properties is the one that exploits the X-ray emission expected from their hot and rarefied gas trapped by the deep potential well of the systems. Successful applications of this method for the search of galaxy clusters are present in literature for the COSMOS field, with publicly available catalogues ([George et al., 2011](#); [Gozaliasl et al., 2019](#)).

The two catalogues of candidate clusters identified within this study have been compared with these two X-ray galaxy group and cluster catalogues made available for the COSMOS field. The analysis has been carried out with the main goal of testing the consistency between the retrieved AMICO catalogues and the X-ray catalogues. In addition to being a significant benchmark for the catalogue obtained within this work and its physical properties it is also an opportunity to compare the reliability of Catalogue r with respect to Catalogue H. Moreover, such a comparison allows to calibrate the mass scaling relations based on the X-rays mass estimates.

The catalogues that have been chosen for this analysis are:

- [George et al. \(2011\)](#): it contains 183 clusters in the range $0 < z < 1$ and, for simplicity, will be referred to from here on as **George+2011**. A catalogue of assigned galaxy members is publicly available together with this release;
- [Gozaliasl et al. \(2019\)](#): it contains 247 groups covering the interval $0 < z < 1.53$, here referred to as **Gozaliasl+2019**.

The catalogues will be here just shortly described. A complete description can be found in the mentioned references.

6.1 COSMOS X-ray cluster catalogues

The catalogue George+2011, presented in [George et al. \(2011\)](#) is based on data extracted from several sources. The X-ray groups and their properties in the catalogue are based on a catalogue of extended sources identified by [Finoguenov et al. \(2007\)](#). X-ray imaging was acquired with *XMM-Newton* with a total exposure of ~ 1.5 Ms and covering 2.13 deg^2 ([Cappelluti et al., 2009](#); [Hasinger et al., 2007](#)). Data from *Chandra* observations were also added and integrated to cover the central region (1.8 Ms over 0.9 deg^2) ([Elvis et al., 2009](#)). The combination of the two data and the derivation of the flux measurements were performed via a wavelet transform method ([Vikhlinin et al., 1998](#)).

The data reduction process followed the procedure presented in [Finoguenov et al., 2009; 2010](#), including a differential cleaning of contaminating point-sources depending on the instrument. After the identification of the extended sources, the optical images inspection allowed for the determination of the redshift via the association of possible known galaxy members. These galaxies were selected within the central projected 0.5 Mpc of the X-ray emission exploiting the red sequence method, a member finder that looks for the presence of red galaxy absolute overdensities.

The photometric redshifts were retrieved from the works of [Ilbert et al. \(2009\)](#) and [Capak et al. \(2007\)](#). In some cases ($\sim 20\%$) the galactic counterpart had also spectroscopic redshifts available extracted from the zCOSMOS survey ([Lilly et al., 2007](#)) along with spectroscopic data from [Prescott et al. \(2006\)](#) and [Capak et al. \(2010\)](#) and a sample retrieved with the FORS2 spectrograph on the VLT¹, allowing a further spectroscopical identification of the candidate groups.

The catalogue contains 183 groups detected up to $z = 1$ with rest frame X-ray luminosities ($0.1\text{-}2.4 \text{ keV}$) spanning the range $10^{41.3} - 10^{44.1} \text{ erg s}^{-1}$ and with typical masses of $M_{200} \simeq 10^{13} - 10^{14} M_{\odot}$.

The second catalogue we investigated, the most recent one, is discussed in [Gozaliasl et al. \(2019\)](#). The sample contains 247 X-ray-selected galaxy groups in a mass range of $M_{200} = 8 \times 10^{12} - 3 \times 10^{14} M_{\odot}$ and with $0.08 \leq z < 1.53$. With respect to the George+2011 catalogue, the improvement for groups identification and redshift assignments reflects the evolution of the observa-

¹for complete description and references we refer the reader to the paper presenting the catalogue: [George et al. \(2011\)](#).

tional campaigns that have taken place in the years since then. An example is the expansion of the Chandra program, with high-resolution imaging of the full COSMOS field (Civano et al., 2016) and the availability of improved photometric catalogues especially at high redshifts. These improvements are reflected in the extension of previous X-ray catalogues to $z \sim 1.53$ and in the improved precision in centres determination of the extended sources (with respect to George et al., 2011, the statistical uncertainty on centre identification improved from $15''$ to $5''$). For this catalogue the combined XMM+Chandra X-ray data described before have been used, with increased sensitivity. For the spectroscopic redshift assignment the used catalogue was based on the work by Hasinger et al. (2018) in an updated version (see Gozaliasl et al., 2019 for details). This latter was carried out with the Deep Imaging Multi-Object Spectrograph (DEIMOS) mounted on the Keck II Telescope (in the range $\sim 550 - 980$ nm). In case no sufficient spectroscopic information was available, the photometric redshifts from the COSMOS2015 were used (Laigle et al., 2016). In order to increase the completeness of the sample for bluer and higher-redshift objects, the data were integrated with the *i*-band selected catalogue data by Ilbert et al. (2009) and McCracken et al. (2012).

Despite the presence of detections flagged as problematic or less secure in both catalogues, in the context of this analysis we used the complete available sample and commented the quality of detections a-posteriori, in order not to bias the comparison and to double check the significance of the flags.

6.1.1 Redshift and membership assignment

The method used to identify and assign optical galactic counterparts to the X-ray detected groups is in both catalogues the red sequence finder. In the Gozaliasl+2019 catalogue, a refined version of the method was used, as described in Finoguenov et al. (2009) and (2010). The technique is based on the identification of overdensities of red galaxies within a certain aperture, centered in the peak of the extended X-ray source. To have an estimate of the overdensity significance with respect to the field, an average number of galaxies in the field was retrieved by extracting equal apertures with random locations within the COSMOS field. Different apertures were chosen for the application of the red sequence finder, such as a fixed radius of 0.5 Mpc and a mass-dependent one based on R_{500} . In case of coexistence of photometric and spectroscopic redshift of the group, the final redshift was estimated via

bi-weight location method (Beers, Flynn, & Gebhardt, 1990).

The catalogue of galaxy members associated with the George+2011 release was based on a simple Bayesian approach to estimate the membership probability to belong to the structure. The chosen selection algorithm assigned the probability by taking into account different observables, including the typical uncertainty of the photo- z at a given magnitude and an estimate of the field galaxy contribution. In case of overlapping volumes surrounding a detected group, a multiple probability was recorded for the galaxies (this was the case for $\sim 3.5\%$ of the galaxies in the George+2011 members catalogue). Further details about the membership assignment can be found in George et al. (2011).

6.2 Matching with AMICO catalogues

Having at hand redshift information, also in this case we performed a three-dimensional matching between our detections and the ones in the George+2011 and Gozaliasl+2019 catalogues. The separation has been considered in physical units and not in angular ones, given the adopted cosmology, in order to make the matching more complete and realistic at different redshifts. We have chosen a separation in redshift computed as $dz = 0.05(1+z)$ and a separation radius of 0.5 Mpc/h. These two input parameters have been fine tuned by testing the obtained successful and not successful matching by comparison with other observables. For instance, when the separation radius was chosen too small a few significant detections that were potentially successfully matched were lost. The identification of this kind of sources has been sometimes possible thanks to a simple optical inspection of the HSC images, sometimes has required a further check of the X-ray maps or of the AMICO amplitude maps.

The adopted criterion is sufficiently restrictive in the sky position and more relaxed in redshift separation to account for possible differences due to the distinct redshift determination methods and uncertainties. Before the matching is computed, the two catalogues can be sorted according to a specific property. The catalogues for this matching have been sorted according to decreasing signal-to-noise ratio for the AMICO catalogues and decreasing mass (M_{200}) for the X-ray catalogues. This was done to prioritize the detections with the largest signal and mass.

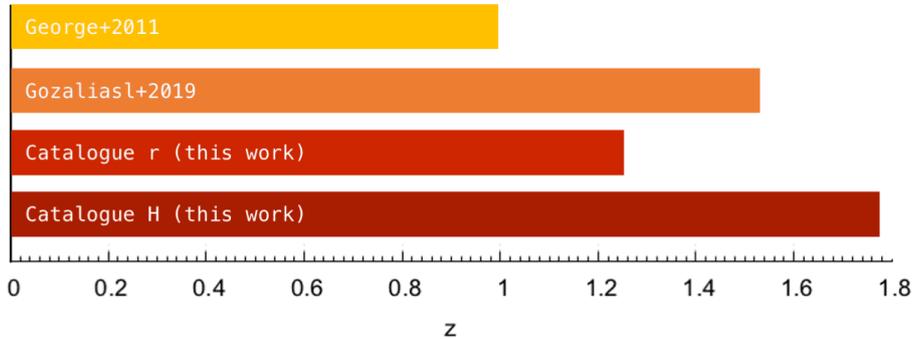


Figure 6.1: Schematic of the different redshift ranges covered by the two X-ray cluster catalogues used for the comparison and the two catalogues produced within this study. From top to bottom: George+2011 (George et al., 2011), $0 < z < 1$; Gozaliasl+2019 (Gozaliasl et al., 2019), $0 < z < 1.53$; the Catalogue r limited to $z = 1.25$; the Catalogue H extended up to $z \sim 1.8$.

The matching has been performed with same parameters both for the George+2011 and the Gozaliasl+2019 catalogue, with respect to both catalogues retrieved within this work. Also for these comparisons, the 4 catalogues do not cover the same volume, given the differences in redshift extent and in effective area. Therefore we restricted the matching procedure to the common area and redshift range solely. In Figure 6.1 we show the redshift coverage of the 4 catalogues. Figure 6.2 gives instead an idea on how the detections of the different catalogues are distributed in the sky. Empty black circles indicate clusters identified by George et al. (2011), green crosses the ones presented by Gozaliasl et al. (2019) and blue triangles and small magenta crosses are the ones from Catalogue r and H, respectively. In Figure 6.2 the displayed detections are limited to $z = 1$ for consistency with the George+2011 catalogue. In the background the masked objects are shown in beige, with the addition of the masked objects for the H -band in grey. This shows the areas in which our detections are absent because of visibility masks.

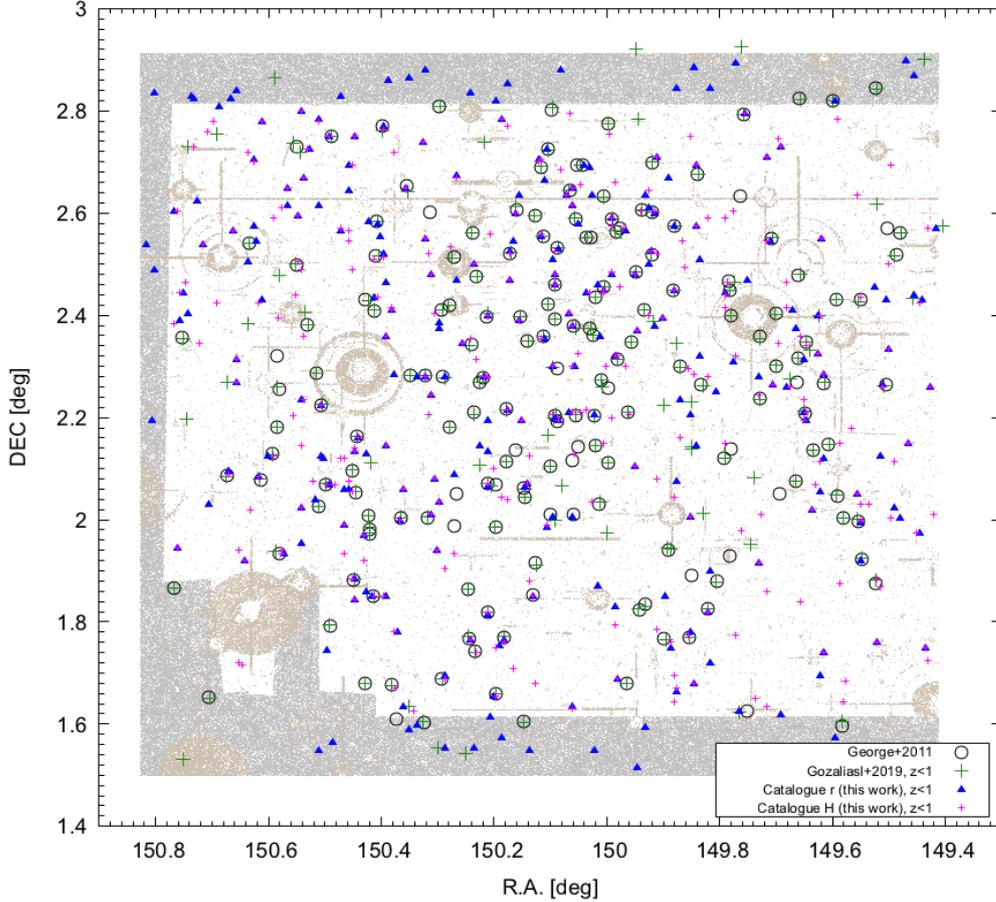


Figure 6.2: Angular positions in the sky plane of the 4 catalogues analyzed in this Chapter. Empty black circles are clusters identified by [George et al. \(2011\)](#), green crosses by [Gozaliasl et al. \(2019\)](#), limited to $z = 1$ and blue triangles and small magenta crosses are the ones from this work, Catalogue r and H, respectively. All the displayed detections are limited to $z = 1$ for consistency with George+2011. In the background the masked objects are marked in beige colour, with masked objects for the H -band in grey. This displays areas in which our detections are absent because of visibility masks.

6.2.1 Catalogue H vs Gozaliasl+2019

The matching between the AMICO Catalogue H and the Gozaliasl+2019 catalogue has been performed by restricting the area to the limits of the former and the redshift extent to the limits of the latter: UltraVISTA area and $0 < z < 1.53$. To account for the effective area of the AMICO Catalogue H, only the X-ray detections falling into unmasked regions according to the H -band mask (described in Section 4.1.1) have been kept.

The comparison has shown correspondence for 99 structures within $0.5 Mpc/h$ and $dz = 0.05(1 + z)$, namely $\sim 46\%$ of the X-ray detections in this effective

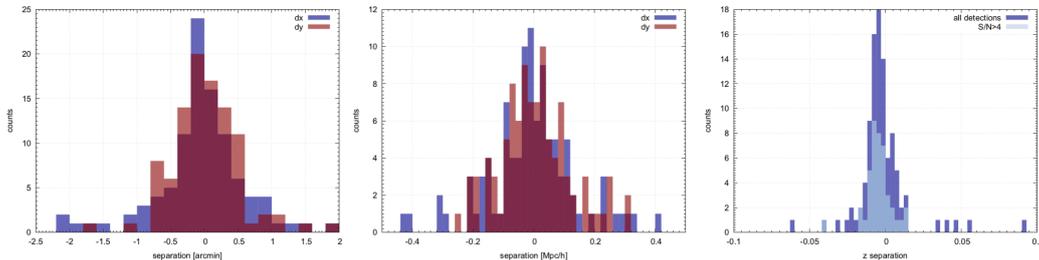


Figure 6.3: The separation distributions of the obtained successful matches between Catalogue H and Gozaliasl+2019. The three distributions show the positional separation both in arcmin (left panel) and Mpc/h (central panel), and the redshift separation (right panel). The distributions obtained for this comparison show quite good consistency between the successfully matched identifications, with most of the clusters having the same centre position within the tolerance range of 0.1 Mpc/h and 0.03 in redshift. In the sky position distributions red stays for DEC separation, blue for RA. In the redshift panel, a slight bias in the distribution is visible, indicating that AMICO redshift tends to be smaller than the redshift reported for X-ray selected clusters.

area.

In Figure 6.3 we show the three distributions of the separation in the sky between successful matches both in arcmin and Mpc/h, and the redshift separation. An ideally good distribution is obviously as peaked as possible, a bad matching displays instead a low and wide distribution of separations. The comparison between the two catalogues has shown quite good consistency between the successfully matched identifications, with most of the clusters having the same centre position within a tolerance range of 0.1 Mpc/h and 0.03 in redshift. The redshift separation distribution shows that the highest-SNR AMICO detections are among the best quality matches in terms of redshift difference. A slight redshift bias is visible in the right panel distribution, namely redshift for the detections in this work tends to be slightly underestimated with respect to the X-ray identifications, i.e. the mean value for this bias is

$$\frac{z_{\text{AMICO}} - z_{\text{X-ray}}}{1 + z_{\text{AMICO}}} \approx -0.0016 \pm 0.0009,$$

which has a dependence on z . Nevertheless, the bias is just minimal and therefore negligible for the purpose of our comparison.

6.2.2 Catalogue r vs Gozaliasl+2019

The matching between the AMICO Catalogue r and the Gozaliasl+2019 catalogue has been performed by restricting the area and the redshift extent to the limits of the former: optical unmasked area and $0 < z < 1.25$. The

optically "masked" X-ray detections have been rejected as before, by making use of the standard visibility masked computed as shown in Section 4.1.

The matching has shown correspondence for 92 structures within $0.5 \text{ Mpc}/h$ and $dz = 0.05(1 + z)$, i.e. $\sim 41\%$ of the X-ray detections in this volume.

In Figure 6.4 we show the three distributions of the separation in the sky between successful matches both in arcmin and Mpc/h , and the redshift separation. The results obtained for this comparison have shown quite good consistency between the successfully matched identifications. The position separation in Mpc/h (central panel) appears slightly better with respect to what has been obtained for the comparison with Catalogue H.

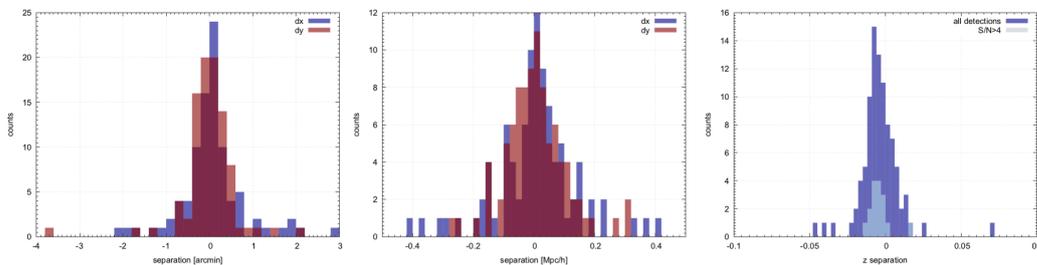


Figure 6.4: The separation distributions of the obtained successful matches between Catalogue r and Gozaliasl+2019. The three distributions show the positional separation both in arcmin (left panel) and Mpc/h (central panel), and the redshift separation (right panel). The distributions obtained for this comparison show quite good consistency between the successfully matched identifications, with slightly better quality than for the Catalogue H. The slight redshift bias mentioned above is present as before and visible in the right panel.

If one restricts the matching limits of the *Catalogue H vs Gozaliasl+2019* comparison to the volume of this matching (i.e. by limiting the analysis to $z = 1.25$), the successful matches are 97, i.e. more than the 47% of X-ray clusters in this volume.

6.2.3 AMICO catalogues vs George+2011

A matching with the same set of parameter has been performed also with the George+2011 catalogue. Despite it may seem outdated, with respect to its successor, it remains a fundamental work with important properties that it is interesting to investigate, and comes with a publicly available member catalogue with membership properties that can be used to test other aspects of our analysis, such as for instance the membership assignment.

89 structures were found to have correspondence with this X-ray detections from the Catalogue H (UltraVISTA area and $0 < z < 1$) and 83 struc-

tures from the Catalogue r (over the full unmasked COSMOS field), namely $\sim 52\%$ and $\sim 46\%$ of the X-ray clusters respectively.

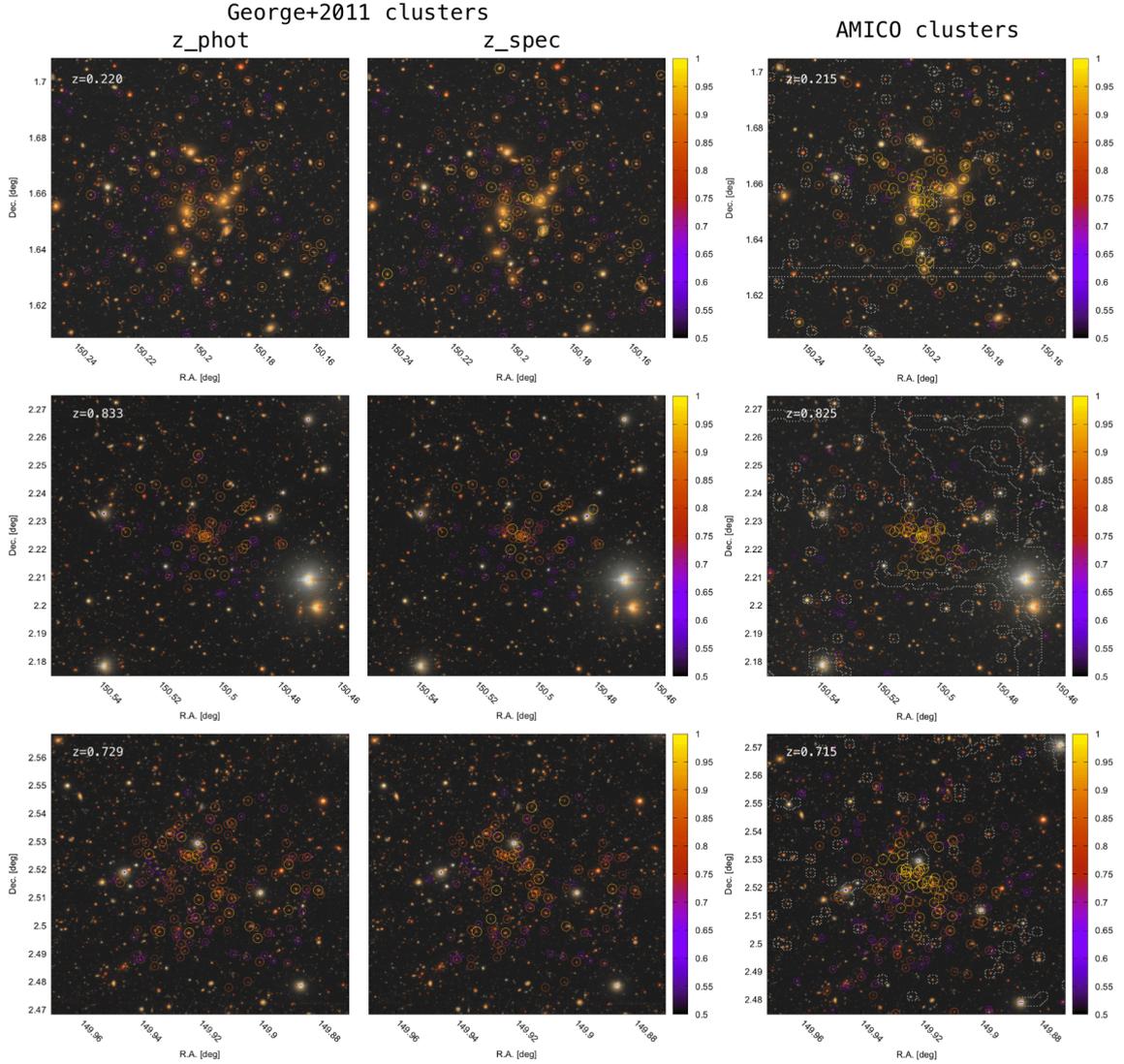


Figure 6.5: Examples of clusters found both in Catalogue r and in George+2011 (each row represents a cluster). The left and central column panels show the clusters as detected in the X-ray, with members assigned using photo- z (left) and spec- z (center) according to [George et al. \(2011\)](#). The right column panels show the correspondent structures as detected by AMICO. Coloured circles always mark the members with their association probability, as in the previous Chapter. Grey dotted lines delimit optically masked region. The redshift of the structure is indicated in the panels. Each stamp is centered in the centre of the cluster as detected in the relevant catalogue and has a side size of 0.1 deg.

For every matched detection that have been checked by looking at the optical HSC images and at the associated members, the membership of

George+2011 has shown good consistency with the membership assigned by AMICO, despite the data were mostly retrieved by different sources. An example of three randomly selected high-SNR detections from the examined sample is shown in Figure 6.5. On the right panels, the detections as identified by AMICO, and on the left and central panels how the same structures are identified in the George+2011 catalogue, both by using photometric (left) and spectroscopic (central) redshifts. Since spectroscopy is not available for the same number of galaxies with photometric information, in the central panels we can see loss of galaxies. Moreover, the spectroscopic membership probabilities are larger because of the much higher quality of the spectroscopic redshift with respect to the photo- z . By considering the two different probability assignments performed in George+2011, the members and their properties have generally shown good correspondence with the AMICO cluster members. Moreover, it should be noted that the AMICO membership assignment does not include only red galaxies in the association of members. Given the differences in terms of optical galaxy catalogues used for the assignment, another discrepancy between the members seems to be related to the different masking that was probably performed in the used galaxy catalogue production. This is visible in the second row panels, where the fraction of masked structure is significant and causes the absence of a fraction of members in the AMICO detection with respect to the X-ray one.

Table 6.1 gives an overview on the number of successful matches between the AMICO and the X-ray catalogues with redshift ranges, covered areas and total number of available objects in the selected volumes.

As mentioned in Section 6.1, X-ray cluster catalogues come with flags indicating the quality and reliability of the detections. Gozaliasl et al. (2019) made use of 4 flag values to describe the cluster identifications.

| Catalogue ID | | redshift | area | # tot | # matches | | |
|--------------|---|----------|---------|------------|-------------------|-------------------|--------|
| r | H | | | | | Geo+11 | Goz+19 |
| ✓ | | ✓ | $z < 1$ | COS | 272 <i>vs</i> 180 | 83 | |
| | ✓ | ✓ | $z < 1$ | UV | 282 <i>vs</i> 170 | 89 | |
| ✓ | | | ✓ | $z < 1.25$ | COS | 301 <i>vs</i> 226 | 92 |
| | ✓ | | ✓ | $z < 1.53$ | UV | 341 <i>vs</i> 216 | 99 |
| | ✓ | | ✓ | $z < 1.25$ | UV | 320 <i>vs</i> 205 | 97 |

Table 6.1: Summary of the successfully matched structures within this comparison with the X-ray available catalogues of clusters. The check marks indicate the possible combinations of catalogues matched with each other. Geo+11 stays for the catalogue by [George et al. \(2011\)](#), Goz+19 for the one from [Gozaliasi et al. \(2019\)](#), r and H for the AMICO catalogues in the respective bands. In the *area* column COS stays for the full unmasked COSMOS field, UV for the unmasked UltraVISTA portion of the field. The total number of objects contained in the matched catalogues according to the selected volume is shown in the *# tot* column, the first value refers to the AMICO catalogue, the second one to the X-ray one. The total number of successful correspondences is shown in the last column.

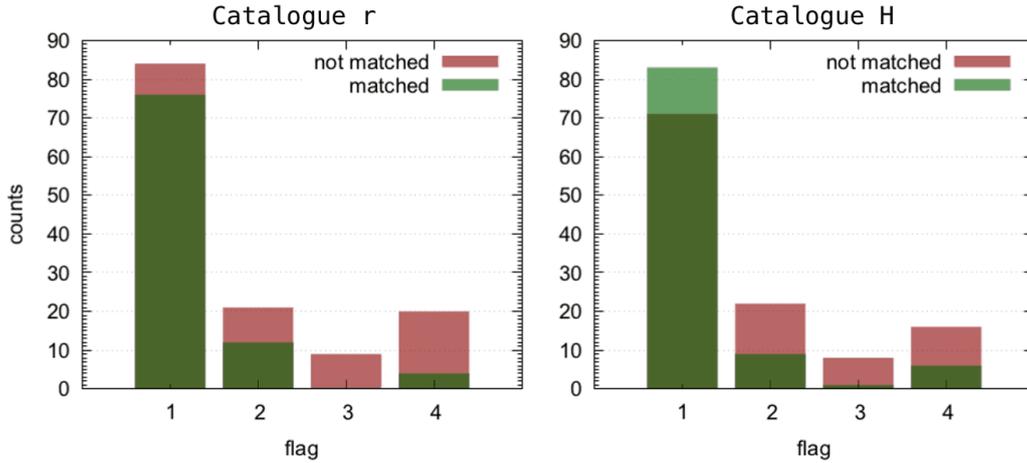


Figure 6.6: The distributions of matches and non-matches with the Gozaliasi+2019 catalogue depending on X-ray flag values. Flag 1 is assigned to the safest sample. The largest fraction of non-matched clusters seems to mainly affect flag 3 and flag 4 structures.

Flag 1 is assigned to secure X-ray detections with emission attributable to the gas of the cluster. Flag 2 indicates a possible foreground/background contamination of the emission. This is the case also for some of the clusters overlapping along the line of sight. Flag 3 is attributed to structures to which only photometric optical counterpart was assigned and therefore the redshift was determined with no spectroscopic information for the member galaxies. Flag 4 marks instead structures with multiple optical counterparts

for which it was not straightforward to make a precise association with the X-ray emission. Figure 6.6 shows how the X-ray flag values correlate with the matched and non-matched detections in the comparison with the AMICO catalogues. For Catalogue H we found the following numbers of flagged objects among the successful matches: flag 1 = 83, flag 2 = 9, flag 3 = 1, flag 4 = 6 out of the total 99 matches. For Catalogue r we found: flag 1 = 76, flag 2 = 12, flag 3 = 0, flag 4 = 4 out of the total 92 matches. The highest fraction of matched clusters concerns the most secure X-ray selected structures (flag 1). Flag 3 and flag 4 display a large fraction of clusters without AMICO counterparts. This might be an indication of the low reliability of these X-ray candidates.

6.3 X-ray counterparts in the 0.5-2 keV emission map

A further analysis has been carried out by making use of the X-ray 0.5-2 keV emission map and X-ray signal-to-noise map of the COSMOS field, which made the comparison more complete and allowed to further investigate the link between the detections and their multi-wavelength observables.²

In this Section, for simplicity and for better consistency with the data of the map we show only the results of the comparison with the Gozaliasl+2019 catalogue. The X-ray emission map is described in Gozaliasl et al. (2019) and it results from the combination of data from XMM-Newton and Chandra observatories, subsequently filtered with a wavelet-transform. We overlapped the list of AMICO detections matched with the X-ray catalogue and the list of non-matched identifications.

Figure 6.7 and Figure 6.8 show the complete Catalogue r and the complete Catalogue H respectively overlapped with the X-ray emission map. The successfully matched clusters are indicated with a yellow circle and the non-matched ones with a cyan circle. The numbers at the centre of each circle indicate the AMICO ID and the redshift of the structure. The radius of the circle scales with the amplitude, A , of the cluster as measured by AMICO. Most of the brightest extended sources in the X-rays have been detected as galaxy clusters by our AMICO detection.

²This analysis has been possible thanks to the collaboration and the material provided by Prof. Dr. Alexis Finoguenov.

By comparing the AMICO detections obtained within our analyses with the X-ray emission map, we have found a sample of 22 new detections with X-ray counterpart not listed in the literature. If we include also these clusters we have obtained a total of 140 AMICO detections with an X-rays counterpart. Among them, 102 have been detected in the r -band analysis and 113 in the H -band analysis.

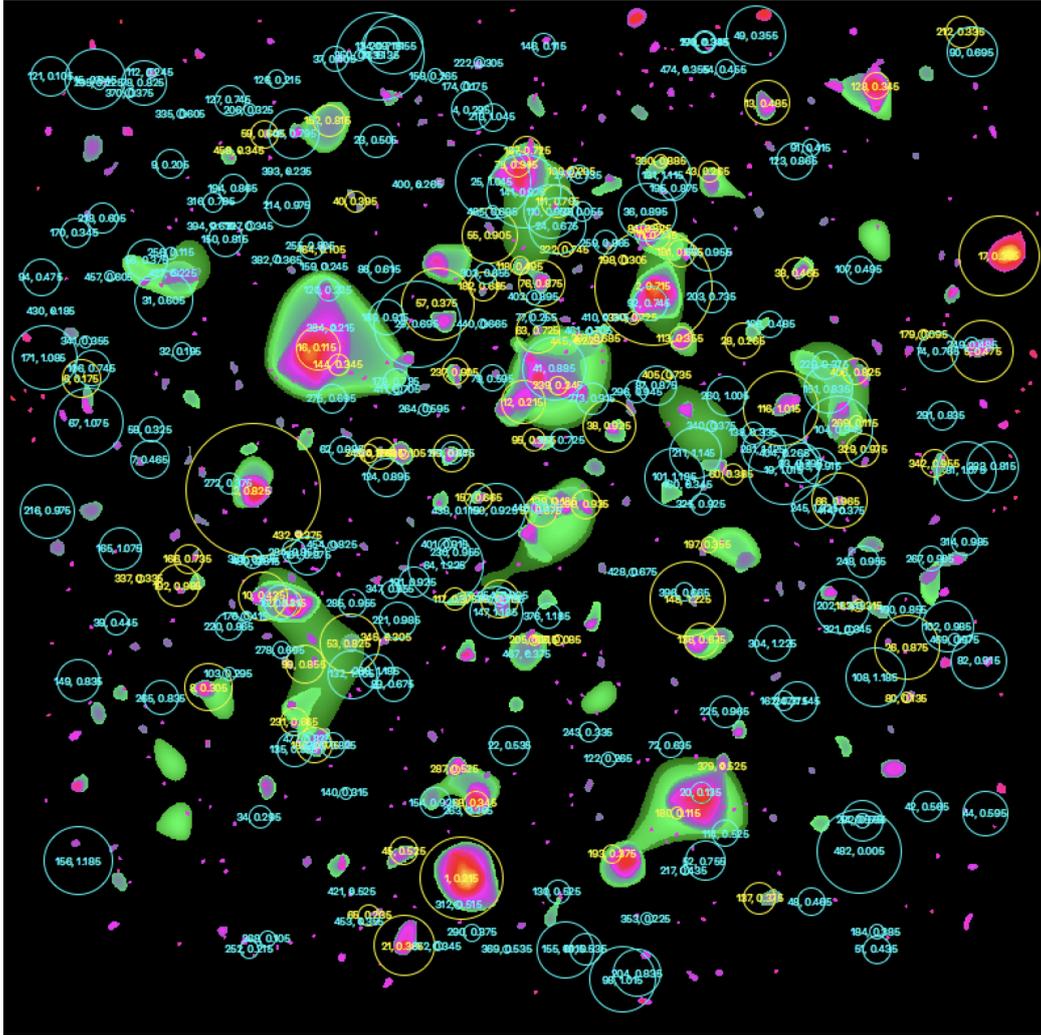


Figure 6.7: The combined XMM-Newton+Chandra 0.5-2 keV wavelet-filtered emission map overlapped with the full Catalogue r . Green areas have low emission, purple and pink medium emission and the brightest regions are highlighted by red and yellow colours. Yellow circles indicate AMICO detections matched in the X-rays using the catalogue by [Gozaliasl et al. \(2019\)](#). Cyan circles mark the non-matched new detections. Each circle has ID and redshift identifying the detection and radius scaled according to the amplitude assigned by AMICO to the structure.



Figure 6.8: The combined XMM-Newton+Chandra 0.5-2 keV wavelet-filtered emission map overlapped with the full Catalogue H. The same of Figure 6.7, but for the Catalogue H.

6.4 Mass-observable scaling relations

We have focused this analysis on the catalogue of cluster candidates presented by [Gozaliasl et al. \(2019\)](#) which includes, among the available properties, the 0.1-2.5 keV rest frame X-ray luminosity (L_X) measured within R_{500} ³ and the virial mass M_{200} . The latter is estimated via the $L_X - M_{200}$ scaling relation presented in [Leauthaud et al. \(2010\)](#), a stacked weak-lensing analysis of X-ray clusters in the COSMOS field.

³radius enclosing a sphere of mean density equal to 500 times the critical density.

In a study carried out by Bellagamba et al. (2019), a stacked weak-lensing mass calibration was instead performed for the clusters of the AMICO-KiDS catalogue (Maturi et al., 2019). The selected cluster sample, including ~ 7000 structures and covering the range $0.1 < z < 0.6$, was used to perform weak-lensing stacked analysis in bins of redshift and AMICO mass proxies. By retrieving in this way estimates of the mean halo mass, they built a scaling relation that links AMICO mass proxies to the mass itself M_{200} . The constructed mass-observable scaling relation reads:

$$\log \frac{M_{200}}{10^{14} M_{\odot} / h} = \alpha + \beta \log \frac{O}{O_{piv}} + \gamma \log \frac{E(z)}{E(z_{piv})}, \quad (6.1)$$

being O the generic observable, O_{piv} and z_{piv} the pivot values, i.e. the typical values representing the chosen sample, $E(z) \equiv H(z)/H_0$ and α , β and γ the parameters of the relation.

We made use of the approach adopted by Bellagamba et al. (2019) to produce a X-ray mass calibration of the AMICO mass proxies, amplitude and intrinsic richness, by fitting the data obtained from the successful X-ray-AMICO matches with the relation expressed by Equation 6.1. We neglected the last term of the relation, accounting for the redshift evolution because of the statistical scarcity of our sample with respect to the one analyzed in Bellagamba et al. (2019). The fit of the sample has been performed with the `fit` function of `GNUPLOT`⁴, that exploits an implementation of the non-linear least squares Marquardt-Levenberg algorithm.

We have chosen the pivot values, O_{piv} , to be the central values of amplitude and richness in the considered sample. For the matches with Catalogue r, these have been set to $A_{piv} = 1.5$ and $\lambda_{*,piv} = 25$. The α and β parameters represent the logarithmic intercept and slope, respectively, of the relations between X-ray virial mass and amplitude $M_{200,X} - A$ and X-ray virial mass and intrinsic richness $M_{200,X} - \lambda_*$. The analysis has produced the following parameters for the scaling relation:

$$A \begin{cases} \alpha = -0.010 \pm 0.025 \\ \beta = 1.21 \pm 0.11 \end{cases} \quad \text{and} \quad \lambda_* \begin{cases} \alpha = 0.024 \pm 0.021 \\ \beta = 1.17 \pm 0.09. \end{cases}$$

⁴http://www.gnuplot.info/docs_4.0/gnuplot.html

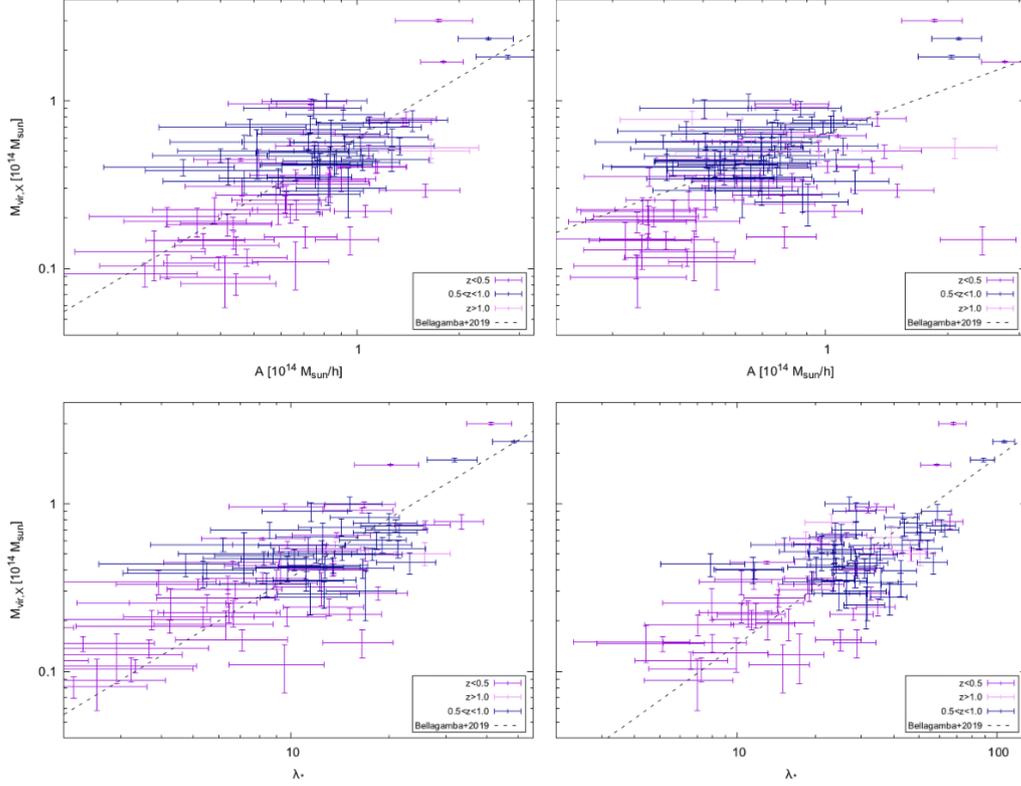


Figure 6.9: The X-ray virial mass $M_{vir,X}$ relation with amplitude and intrinsic richness for the Catalogue r (left) and the Catalogue H (right) and the Gozaliasl+2019 corresponding clusters. Different colours mark different intervals of redshift. The dashed line is the best fit of the data according to the approach followed in Bellagamba et al. (2019), on which we have based the scaling relations. The error bars indicate the 1σ error on mass as provided in the X-ray catalogue, the square root of λ_* and the amplitude variance, i.e., $\sigma = A/SNR$.

For the comparison between Catalogue H and Gozaliasl+2019, the analysis has shown consistent results, namely:

$$A \begin{cases} \alpha = -0.04 \pm 0.03 \\ \beta = 0.89 \pm 0.11 \end{cases} \quad \text{and} \quad \lambda_* \begin{cases} \alpha = -0.037 \pm 0.027 \\ \beta = 0.99 \pm 0.12, \end{cases}$$

where we have used $A_{piv} = 1.5$ and $\lambda_{*,piv} = 55$.

The result of the fit for the comparison between Catalogue r and Catalogue H with Gozaliasl+2019 is shown in Figure 6.9, where the different colours indicate different redshift intervals.

The overall mass range covered by the AMICO detected clusters in the COSMOS field is the interval $8.09 \times 10^{12} M_\odot \lesssim M_{200} \lesssim 2.99 \times 10^{14} M_\odot$.

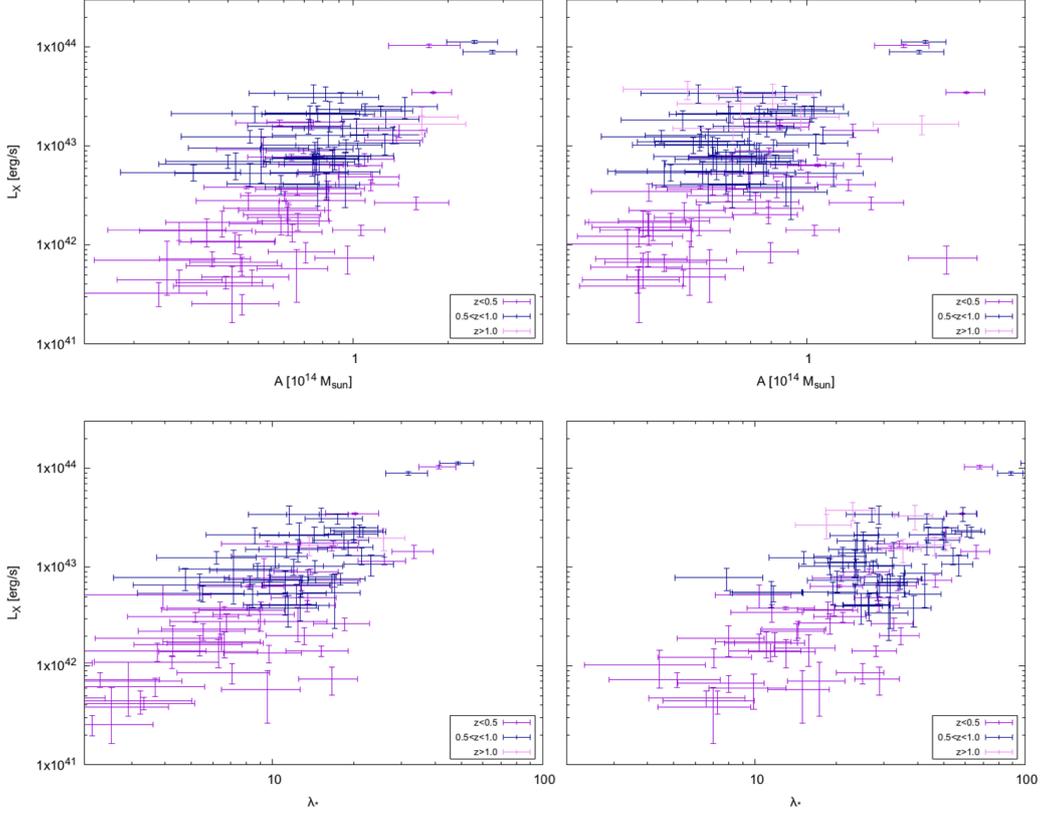


Figure 6.10: The 0.1-2.5 keV rest frame X-ray luminosity (L_X) in relation with amplitude and intrinsic richness for the Catalogue r (left) and the Catalogue H (right) for the matches with the Gozaliasl+2019 catalogue. Different colours mark different intervals of redshift. The error bars indicate the 1σ error on luminosity as provided in the X-ray catalogue, the square root of λ_* and the amplitude variance, i.e., $\sigma = A/SNR$.

In Figure 6.10, we show the relation that has been found between the two AMICO mass proxies and the X-ray luminosity of the clusters with X-ray correspondence. A clear correlation is visible in all four cases, both for Catalogue r and H detections. The scatter of the scaling relation is slightly larger for the amplitude with respect to the intrinsic richness, especially in the context of Catalogue H matched identifications (top right panel).

6.5 Catalogue r *vs* Catalogue H in the X-rays

After having matched and compared the AMICO catalogues with the available X-ray catalogues, we introduced a new set of observables, that come into play in the verification of the reliability of our catalogues, also in

comparison with each other.

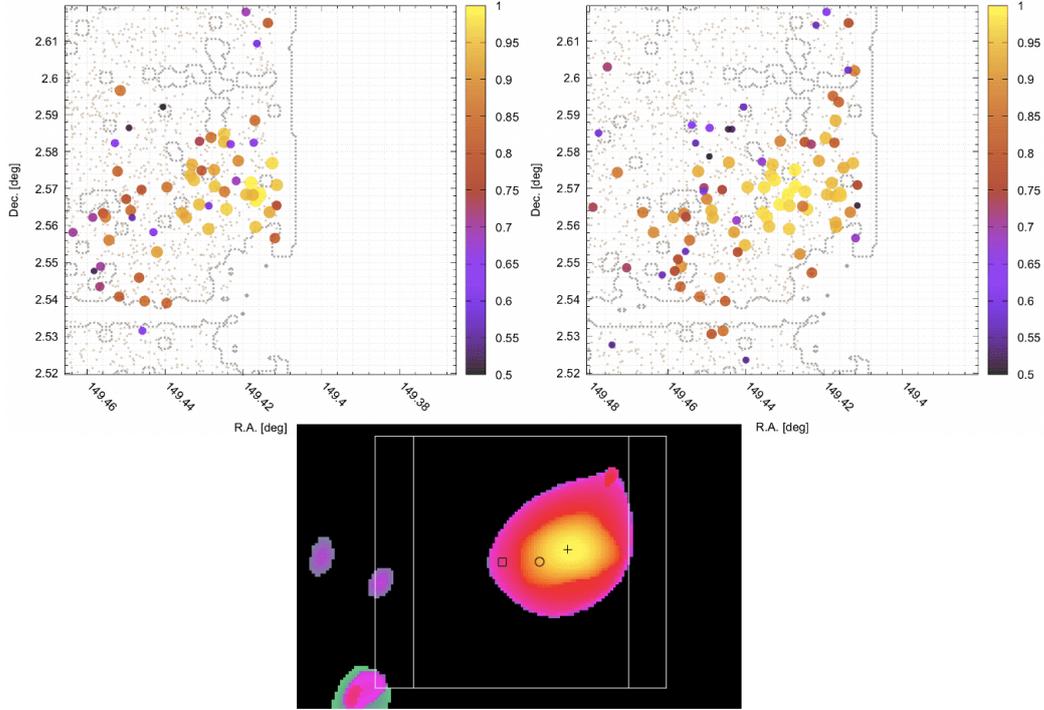


Figure 6.11: The detection of a massive structure at the edge of the COSMOS field, as identified during the r magnitude run (top left) and during the H magnitude run (top right) ($z=0.385$). The clusters are centered in different locations and therefore show discrepancies in the association of member galaxies. The bottom panel shows the X-ray emission in correspondence with the centre of the cluster in H -band (square) and in r -band (circle). The latter is more consistent with the X-ray peak marked with the black cross. White contours indicate the 0.1 deg side boxes represented on the top panels.

Due to the different way in which cluster model and noise model were computed for the two magnitude bands, and therefore to differences in the computation of the amplitude map, one of the main discrepancies in the matches between Catalogue r and H concerns the determination of cluster centres. The top panels of Figure 6.11 shows an example of a clear clustered structure, already known in literature (Gozaliasl et al., 2019), as detected by AMICO in the r (left) and in the H band (right). The cluster is unfortunately at the edge of the COSMOS field, and probably for this reason is not present in the George+2011 catalogue. Nevertheless, the detection is present in the Gozaliasl+2019 catalogue. We can see how the centres of the AMICO detections produced in the two different runs differ and how this has an impact on the membership determination and therefore on the richness. The high fraction of field masked for this detection makes it hard to draw

conclusions, but thanks to the X-ray map (see bottom panel of Figure 6.11) it is possible to see that the centre according to the r -band magnitude run (black circle) is more accurate if we take the peak of the X-ray emission as reference (black cross).

Moreover, a certain number of detections that were not found both in Catalogue r and in Catalogue H have an X-ray counterpart, in the cluster lists and/or in the X-ray emission map.

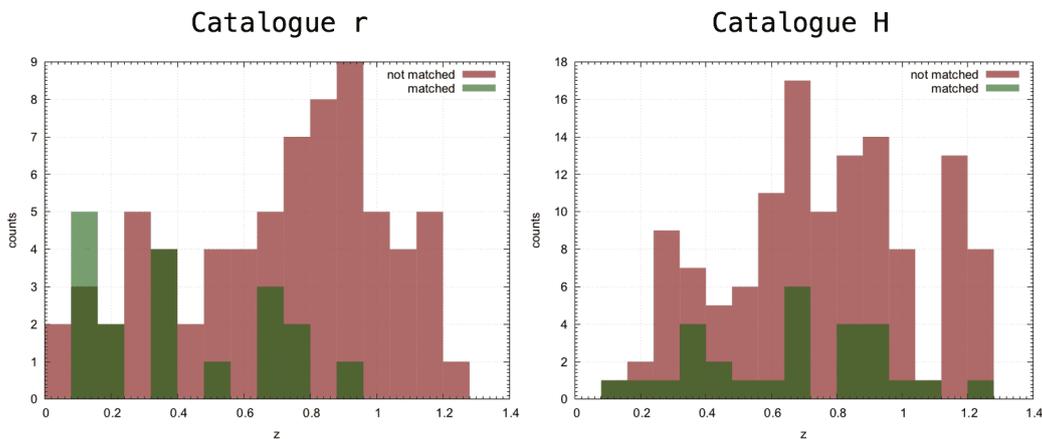


Figure 6.12: The redshift distribution of the non-matched clusters from the comparison performed in Section 5.3. Green indicates clusters that have been matched successfully with the Gozaliasl+2019 catalogue. The redshift distributions do not show particular features, other than a larger fraction of matches for the r -band run at low redshift ($z < 0.2$).

Among the structures shown in Figure 5.8, for instance, the top left one and the bottom right one display X-ray counterparts in the emission map.

By restricting the comparison to the UltraVISTA area and to $z < 1.25$, we found X-ray correspondence (with Gozaliasl+2019) between 18 clusters detected in the r -band run and not in the H -band, and 28 clusters detected in the H -band run and not in the r -band, over 88 and 153 total clusters, respectively. The distribution in redshift of the X-ray matched detections is shown in green in Figure 6.12. The same analysis has been performed once again in the interval $0 < z < 1$ by making use of the George+2011 catalogue: this has shown again 18 matched for the Catalogue r and 26 for the Catalogue H for the clusters not found in both of the AMICO catalogues. From this analysis a few considerations can be made about the different reliability of the two catalogues produced within this work.

The X-ray analysis confirms what has been shown with the optical HSC image inspection and the study of the statistical properties, namely that

Catalogue H appears to be more complete, in the sense that the number of detected clusters on the same volume is larger than for the Catalogue r, both on absolute numbers and in relation to the identifications confirmed in the X-rays. Relative to the total number of detections in the respective catalogues, the reliability of the two is comparable (by taking the Gozaliasl+2019 as reference catalogue). Namely, the Catalogue r "new" detections, not found in the H -band are 88, of which 18 have X-ray correspondence, i.e. $\sim 20.5\%$. Vice versa, we have 28 H -band detections matched in the X-ray among the 153 not found in the r -band, i.e. $\sim 18.3\%$. This shows how the Catalogue H is more complete but the reliability of the two catalogues in terms of verification with previously released catalogues is consistent or slightly higher for the Catalogue r.

Conclusions and future outlook

In the study presented in this work, we applied the AMICO algorithm to the photometric galaxy catalogue COSMOS2015 (Laigle et al., 2016) to detect galaxy clusters in the COSMOS 2 deg². We have taken into account the area of the field which is not significantly affected by image artifacts and generated suited masks to this end.

We performed our cluster detection in two different ways: one analysis was based on the r -band photometry and one on the H -band photometry. The first analysis resulted in 301 detections in the range $0 < z < 1.25$, including 175 detections with $SNR > 3.0$ over an effective area of 1.759 deg². The second analysis, based on the H -band photometry has been performed on the COSMOS-UltraVISTA area (1.351 deg² of effective area) and allowed us to extend the cluster search up to $z \sim 1.8$. In this way 351 clusters have been detected with $SNR > 3.0$, of which 31 structures in the interval $1.25 < z \lesssim 1.8$.

We identified 167 structures in common between the two catalogues over the overlapping area and redshift range. A total of 481 galaxy clusters over the range $0 < z \lesssim 1.8$, with SNR up to 8 have been detected in the context of this study. Both analyses provided us with newly identified structures. The obtained catalogues of clusters we derived include identification number, sky position, redshift, signal-to-noise ratio, amplitude A , apparent richness λ , intrinsic richness λ_* , likelihood \mathcal{L} , masked fraction and full redshift probability distribution. Along with the cluster catalogue, AMICO provides the full galaxy member catalogue with association probabilities and field probabilities. The presence of a membership catalogue allows to improve the detection efficiency through the cleaning procedure and characterize the galaxy populations opening the chance to investigate environmental evolution and galaxy formation. Moreover, the catalogue of member galaxies can be for instance exploited to improve strong lensing measurements from photometric based

algorithms (e.g. Stapelberg, Carrasco, and Maturi, 2019) and can be a helpful tool to remove foreground contribution for weak lensing analysis (Maturi et al., 2019).

We performed a multi-wavelength analysis of the resulting detections by executing a 3-dimensional matching with the two publicly available X-ray cluster catalogues (George et al., 2011; Gozaliasl et al., 2019). The comparison between our detections and the X-ray detections has shown good consistency between successfully matched detections in the two catalogues. By matching over the common area and redshift range, we found 99 and 89 successful matches with the catalogues by Gozaliasl et al. (2019) and George et al. (2011) respectively, for our catalogue based on the H -band and 92 and 83 respectively for the r -band one. Most of the matched identifications were flagged as secure detections by Gozaliasl et al. (2019).

We then calibrated the AMICO mass proxies, A and λ_* , with the X-rays mass estimates and examined their relation with the X-rays luminosity. The derived relations display a small scatter. The expression of the scaling relations follows the relation by Bellagamba et al. (2019),

$$\log \frac{M_{200}}{10^{14} M_{\odot} / h} = \alpha + \beta \log \frac{O}{O_{piv}},$$

where O is the observable. The obtained parameters are, for the r analysis, $\alpha = -0.010 \pm 0.025$, $\beta = 1.21 \pm 0.11$ considering the amplitude as observable, $\alpha = 0.024 \pm 0.021$, $\beta = 1.17 \pm 0.09$ considering λ_* and for the H analysis $\alpha = -0.04 \pm 0.03$, $\beta = 0.89 \pm 0.11$ for the amplitude, $\alpha = -0.037 \pm 0.027$, $\beta = 0.99 \pm 0.12$ for λ_* . Thus, the mass range covered by the AMICO detected clusters in the COSMOS field is the interval $8.09 \times 10^{12} M_{\odot} \lesssim M_{200} \lesssim 2.99 \times 10^{14} M_{\odot}$.

The AMICO algorithm is constantly under improvement and testing and is subject to modification and integration. During the cluster search, from a technical point of view, we tested the usability of AMICO for the case of the high redshifts available in COSMOS and of its small covered area. We have introduced some new methods within AMICO, to allow its application on new kinds of data-sets. We introduced a new way to generate from the data-set a complete visibility mask, we tested a regularization method for the noise and we introduced a new way to attenuate cluster contribution to the noise model.

An interesting future development of our analysis may be the integration of a spectroscopic confirmation of the detected structures by adding infor-

mation from the spectroscopic samples available for the COSMOS field (e.g. DEIMOS, [Hasinger et al., 2018](#)).

Moreover, for what concerns the data-set we used in this study, the COSMOS2015 catalogue ([Laigle et al., 2016](#)), which is near-infrared selected, it would be interesting to further investigate the approach followed for instance by [Gozaliasl et al. \(2019\)](#), who integrated the data-set with *i*-band selected catalogues (e.g. [Ilbert et al., 2009](#)), in order to improve the completeness for blue and high-redshift galaxies.

The already available possibility within AMICO to neglect the magnitude dependence when computing the average redshift probability distribution has been shortly explored. It will be interesting to further analyze the role this feature plays in the case of surveys with small covered area and poor statistics.

The possibility to improve and include a multi-magnitude study with AMICO will be also taken into account. In addition, the influence of the regulating parameters on the noise model briefly discussed in Appendix B will be also further investigated.

In the direction of the multi-wavelength comparison, the X-ray analysis will be further expanded in view of the important role that calibrations and scaling relations between observables will play in future challenges of galaxy cluster cosmology. Additionally, the X-ray emission analysis will soon allow the integration of X-ray observables in the cluster catalogue, not only for matched identifications but also for the new detections resulted from this work that have been shown to have X-ray counterparts in the XMM-Newton+Chandra X-ray emission map ([Gozaliasl et al., 2019](#)).

Appendix A

Further analyses on the galaxy catalogue

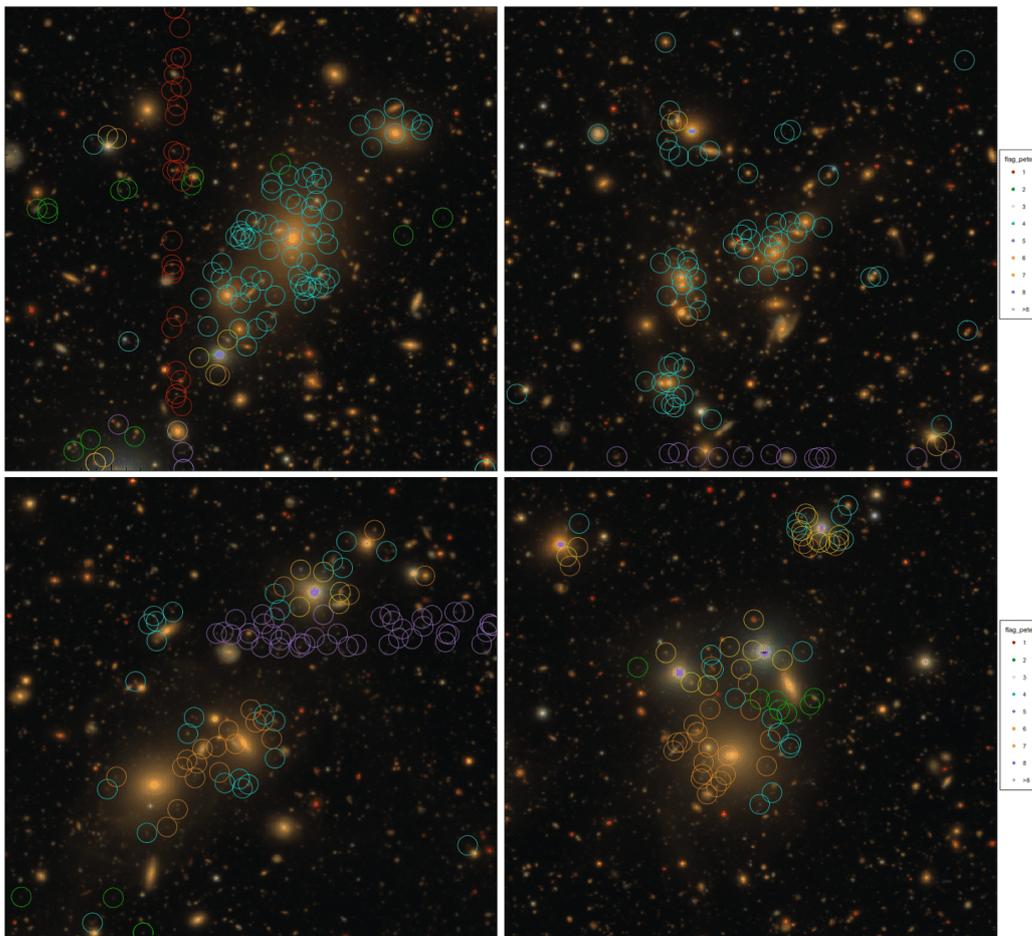


Figure A.1: The presence of potential cluster galaxies among the masked objects according to the flag `FLAG_PETER`. Galaxies are marked with circles and colour-coded according to flag values, in accordance with Figure 3.4. These examples show how important it was to recover galaxies flagged with `FLAG_PETER` = 4 (cyan) or `FLAG_PETER` = 6 (orange), mostly potential cluster galaxies. Images on background are HSC colour composites (g, r, i).

As discussed in Section 3.3.2, most of the cluster members are rejected when adopting standard selection criteria (`FLAG_PETER = 0`) in the COSMOS2015 catalogue.

In Figure A.1 we indicated highlighted with different colours the flag values associated to galaxies as in Figure 3.4. It is possible to see the impact that excluding galaxies marked (especially) with orange and cyan circles would have on cluster detection. Most of the galaxies masked because blended with other galaxies are indeed cluster members and had to be included in the final galaxy sample used to identify clusters.

Appendix B

Minimum noise value

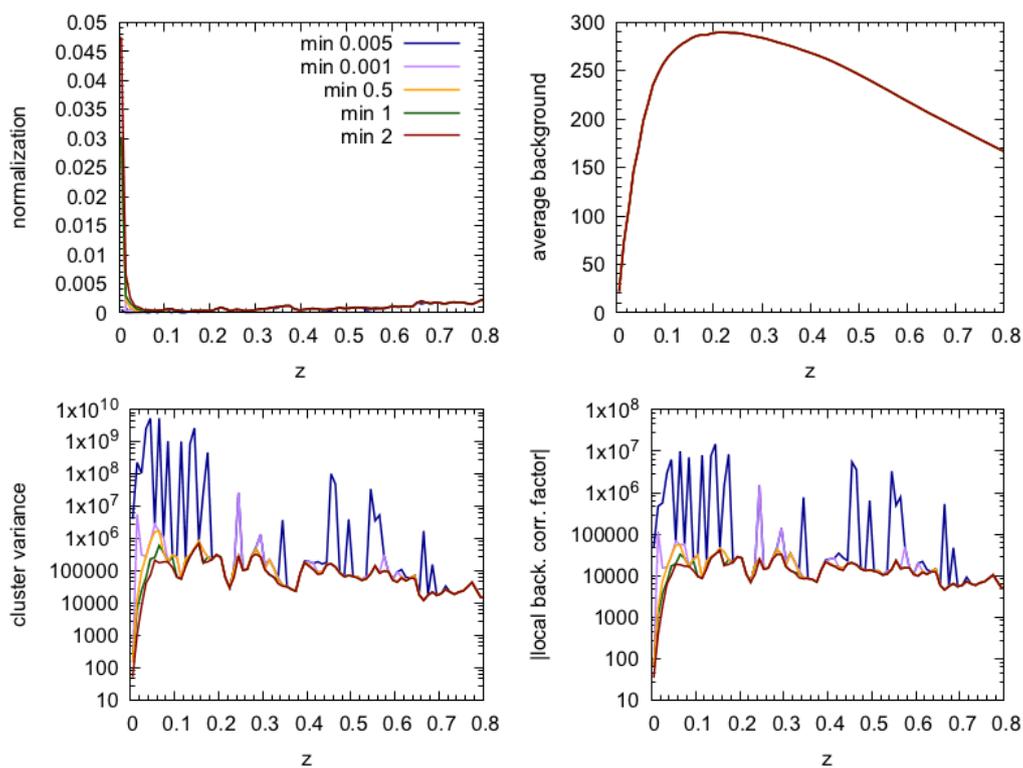


Figure B.1: Effect of the values chosen for the minimum threshold of the noise (`dummy_value_min`) on the 4 filter constants at $z < 0.8$. The data are from 5 different runs performed with r -band magnitude information, but the result is analogous for the H -band runs. Different colours represent different values of the minimum.

As discussed in Section 4.3, the choice of a minimum threshold for the noise model values is not trivial. This is mainly due to the normalization

of the noise model and the introduction of redshift dependence through the redshift probability distribution of galaxies performed by AMICO in building the noise.

In the context of the study presented in this Thesis, to tune the filter response we studied the impact that this threshold value has on the filter constants. Figure B.1 shows the effect that different values of `dummy_value_min` (2, 1, 0.5, 0.001, 0.005) have on the 4 filter constants for the *r*-band magnitude runs. The value `dummy_value_min=2` seems to provide the smoothest filter constants and the most realistic results. Larger values may result in a loss of information.

Appendix C

Examples of detected clusters

Figure [C.1](#) and [C.2](#) show a few examples of optical images in correspondence of randomly chosen detected clusters for the r -band and H -band, respectively. The intrinsic richness and redshift of the structures are reported in caption.

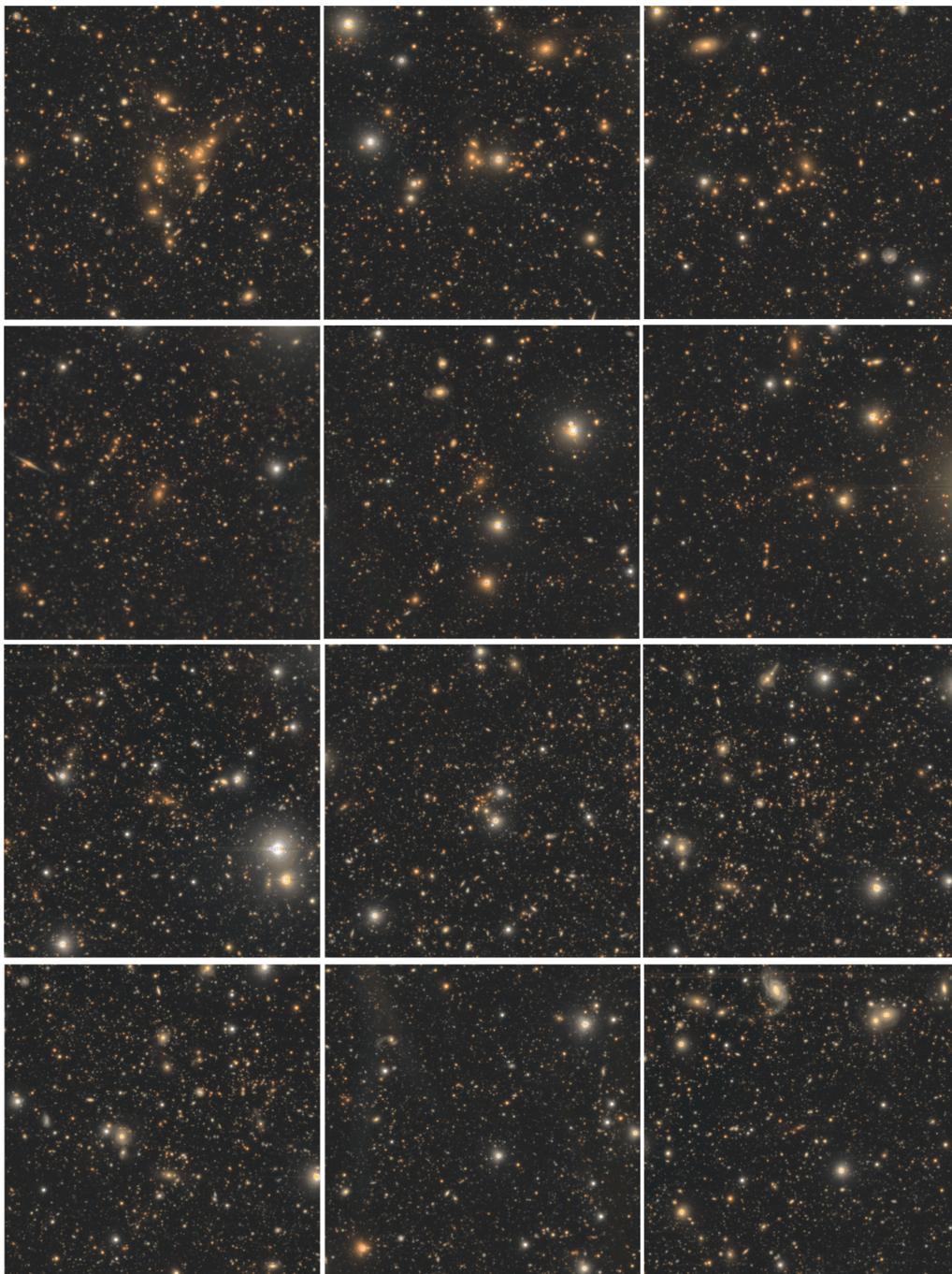


Figure C.1: A sample of 12 detections extracted from the resulting catalogue of the analysis performed with r -band photometry. The clusters are located at $z \approx 0.2 - 0.3, 0.5, 0.8$ and 1.0 (rows from top to bottom) and with $\lambda_* \approx 20 - 50, 10 - 20$ and < 10 (columns from left to right). Every image is an HSC colour composite, g, r, i for first two rows and r, i, z for the last two, and it is centered in the center of the cluster as detected by AMICO with side size of 0.1 deg.



Figure C.2: A sample of 12 detections extracted from the resulting catalogue of the analysis performed with H -band photometry. The clusters are located at $z \approx 0.8, 1.0, 1.3$ and > 1.5 (rows from top to bottom) and with $\lambda_* \approx 30+, 20-30$ and < 20 (columns from left to right). Every image is an HSC colour composite, r, i, z for first two rows and i, z, y for the last two, and it is centered in the center of the cluster as detected by AMICO. Every image of the first 3 rows has side size 0.1 deg, every image of the last row has side size 0.05 deg.

Acknowledgments

I would like to thank Professor Lauro Moscardini and Professor Matteo Maturi for their kind availability, their professionalism and their patience.

I express my gratitude to all the collaborators of AMICO, for giving me the opportunity to work with the code.

I would like to thank Professor Alexis Finoguenov (Helsinki University) for the collaboration and availability and for providing access to the X-ray emission map and signal-to-noise map used within this work.

A thank you goes to all the members of the *Galaxy Clusters* group of the Institute of Theoretical Astrophysics, Heidelberg.

Based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under ESO programme ID 179.A-2005 and on data products produced by TERAPIX and the Cambridge Astronomy Survey Unit on behalf of the UltraVISTA consortium.

Based in part on data collected at the Subaru Telescope and retrieved from the HSC data archive system, which is operated by Subaru Telescope and Astronomy Data Center at National Astronomical Observatory of Japan.

Bibliography

- Abdullah, M. H., Klypin, A., & Wilson, G. (2020). Cosmological Constraints on Ω_m and σ_8 from Cluster Abundances Using the GalWCat19 Optical-spectroscopic SDSS Catalog., *901*(2), Article 90, 90. <https://doi.org/10.3847/1538-4357/aba619>
- Abell, G. O. (1958). The Distribution of Rich Clusters of Galaxies., *3*, 211. <https://doi.org/10.1086/190036>
- Abell, G. O. (1959). The National Geographic Society-Palomar Observatory Sky Survey. *Leaflet of the Astronomical Society of the Pacific*, *8*(366), 121.
- Aihara, H., Arimoto, N., Armstrong, R., Arnouts, S., Bahcall, N. A., Bickerton, S., Bosch, J., Bundy, K., Capak, P. L., Chan, J. H. H., Chiba, M., Coupon, J., Egami, E., Enoki, M., Finet, F., Fujimori, H., Fujimoto, S., Furusawa, H., Furusawa, J., ... Yuma, S. (2018). The Hyper Suprime-Cam SSP Survey: Overview and survey design., *70*, Article S4, S4. <https://doi.org/10.1093/pasj/psx066>
- Aihara, H., AlSayyad, Y., Ando, M., Armstrong, R., Bosch, J., Egami, E., Furusawa, H., Furusawa, J., Harasawa, S., Harikane, Y., Hsieh, B.-C., Ikeda, H., Ito, K., Iwata, I., Kodama, T., Koike, M., Kokubo, M., Komiyama, Y., Li, X., ... Yamashita, T. (2021). Third Data Release of the Hyper Suprime-Cam Subaru Strategic Program. *arXiv e-prints*, Article arXiv:2108.13045, arXiv:2108.13045.
- Alberts, S., Pope, A., Brodwin, M., Chung, S. M., Cybulski, R., Dey, A., Eisenhardt, P. R. M., Galametz, A., Gonzalez, A. H., Jannuzi, B. T., Stanford, S. A., Snyder, G. F., Stern, D., & Zeimann, G. R. (2016). Star Formation and AGN Activity in Galaxy Clusters from $z=1-2$: a Multi-Wavelength Analysis Featuring Herschel/PACS., *825*(1), Article 72, 72. <https://doi.org/10.3847/0004-637X/825/1/72>

- Allen, S. W., Evrard, A. E., & Mantz, A. B. (2011). Cosmological Parameters from Observations of Galaxy Clusters., *49*(1), 409–470. <https://doi.org/10.1146/annurev-astro-081710-102514>
- Amendola, L., Appleby, S., Bacon, D., Baker, T., Baldi, M., Bartolo, N., Blanchard, A., Bonvin, C., Borgani, S., Branchini, E., Burrage, C., Camera, S., Carbone, C., Casarini, L., Cropper, M., de Rham, C., Di Porto, C., Ealet, A., Ferreira, P. G., . . . Zlosnik, T. (2013). Cosmology and Fundamental Physics with the Euclid Satellite. *Living Reviews in Relativity*, *16*(1), Article 6, 6. <https://doi.org/10.12942/lrr-2013-6>
- Andreon, S. (2008). The history of mass assembly of faint red galaxies in 28 galaxy clusters since $z = 1.3$., *386*(2), 1045–1052. <https://doi.org/10.1111/j.1365-2966.2008.13100.x>
- Andreon, S., Newman, A. B., Trinchieri, G., Raichoor, A., Ellis, R. S., & Treu, T. (2014). JKCS 041: a Coma cluster progenitor at $z = 1.803$., *565*, Article A120, A120. <https://doi.org/10.1051/0004-6361/201323077>
- Arnouts, S., Moscardini, L., Vanzella, E., Colombi, S., Cristiani, S., Fontana, A., Giallongo, E., Matarrese, S., & Saracco, P. (2002). Measuring the redshift evolution of clustering: the Hubble Deep Field South., *329*(2), 355–366. <https://doi.org/10.1046/j.1365-8711.2002.04988.x>
- Ashby, M. L. N., Willner, S. P., Fazio, G. G., Huang, J. -S., Arendt, R., Barmby, P., Barro, G., Bell, E. F., Bouwens, R., Cattaneo, A., Croton, D., Davé, R., Dunlop, J. S., Egami, E., Faber, S., Finlator, K., Grogin, N. A., Guhathakurta, P., Hernquist, L., . . . Yan, H. (2013). SEDS: The Spitzer Extended Deep Survey. Survey Design, Photometry, and Deep IRAC Source Counts., *769*(1), Article 80, 80. <https://doi.org/10.1088/0004-637X/769/1/80>
- Ashby, M. L. N., Willner, S. P., Fazio, G. G., Dunlop, J. S., Egami, E., Faber, S. M., Ferguson, H. C., Grogin, N. A., Hora, J. L., Huang, J. -S., Koekemoer, A. M., Labbé, I., & Wang, Z. (2015). S-CANDELS: The Spitzer-Cosmic Assembly Near-Infrared Deep Extragalactic Survey. Survey Design, Photometry, and Deep IRAC Source Counts., *218*(2), Article 33, 33. <https://doi.org/10.1088/0067-0049/218/2/33>
- Bartelmann, M. (1996). Arcs from a universal dark-matter halo profile., *313*, 697–702.
- Bartelmann, M. (2010). TOPICAL REVIEW Gravitational lensing. *Classical and Quantum Gravity*, *27*(23), Article 233001, 233001. <https://doi.org/10.1088/0264-9381/27/23/233001>

- Baumann, D., & Peiris, H. V. (2008). Cosmological Inflation: Theory and Observations. *arXiv e-prints*, Article arXiv:0810.3022, arXiv:0810.3022.
- Beers, T. C., Flynn, K., & Gebhardt, K. (1990). Measures of Location and Scale for Velocities in Clusters of Galaxies—A Robust Approach., *100*, 32. <https://doi.org/10.1086/115487>
- Bellagamba, F., Maturi, M., Hamana, T., Meneghetti, M., Miyazaki, S., & Moscardini, L. (2011). Optimal filtering of optical and weak lensing data to search for galaxy clusters: application to the COSMOS field., *413*(2), 1145–1157. <https://doi.org/10.1111/j.1365-2966.2011.18202.x>
- Bellagamba, F., Roncarelli, M., Maturi, M., & Moscardini, L. (2018). AMICO: optimized detection of galaxy clusters in photometric surveys., *473*(4), 5221–5236. <https://doi.org/10.1093/mnras/stx2701>
- Bellagamba, F., Sereno, M., Roncarelli, M., Maturi, M., Radovich, M., Bardelli, S., Puddu, E., Moscardini, L., Getman, F., Hildebrandt, H., & Napolitano, N. (2019). AMICO galaxy clusters in KiDS-DR3: weak lensing mass calibration., *484*(2), 1598–1615. <https://doi.org/10.1093/mnras/stz090>
- Bertin, E., & Arnouts, S. (1996). SExtractor: Software for source extraction., *117*, 393–404. <https://doi.org/10.1051/aas:1996164>
- Blanton, M. R., Hogg, D. W., Bahcall, N. A., Brinkmann, J., Britton, M., Connolly, A. J., Csabai, I., Fukugita, M., Loveday, J., Meiksin, A., Munn, J. A., Nichol, R. C., Okamura, S., Quinn, T., Schneider, D. P., Shimasaku, K., Strauss, M. A., Tegmark, M., Vogeley, M. S., & Weinberg, D. H. (2003). The Galaxy Luminosity Function and Luminosity Density at Redshift $z = 0.1$., *592*(2), 819–838. <https://doi.org/10.1086/375776>
- Blanton, M. R., Bershady, M. A., Abolfathi, B., Albareti, F. D., Allende Prieto, C., Almeida, A., Alonso-García, J., Anders, F., Anderson, S. F., Andrews, B., Aquino-Ortíz, E., Aragón-Salamanca, A., Argudo-Fernández, M., Armengaud, E., Aubourg, E., Avila-Reese, V., Badenes, C., Bailey, S., Barger, K. A., ... Zou, H. (2017). Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe., *154*(1), Article 28, 28. <https://doi.org/10.3847/1538-3881/aa7567>
- Bleem, L. E., Stalder, B., de Haan, T., Aird, K. A., Allen, S. W., Applegate, D. E., Ashby, M. L. N., Bautz, M., Bayliss, M., Benson, B. A., Bocquet, S., Brodwin, M., Carlstrom, J. E., Chang, C. L., Chiu, I., Cho,

- H. M., Clocchiatti, A., Crawford, T. M., Crites, A. T., . . . Zenteno, A. (2015). Galaxy Clusters Discovered via the Sunyaev-Zel'dovich Effect in the 2500-Square-Degree SPT-SZ Survey., *216*(2), Article 27, 27. <https://doi.org/10.1088/0067-0049/216/2/27>
- Bocquet, S., Saro, A., Mohr, J. J., Aird, K. A., Ashby, M. L. N., Bautz, M., Bayliss, M., Bazin, G., Benson, B. A., Bleem, L. E., Brodwin, M., Carlstrom, J. E., Chang, C. L., Chiu, I., Cho, H. M., Clocchiatti, A., Crawford, T. M., Crites, A. T., Desai, S., . . . Zenteno, A. (2015). Mass Calibration and Cosmological Analysis of the SPT-SZ Galaxy Cluster Sample Using Velocity Dispersion σ_v and X-Ray Y_x Measurements., *799*(2), Article 214, 214. <https://doi.org/10.1088/0004-637X/799/2/214>
- Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. (1991). Excursion Set Mass Functions for Hierarchical Gaussian Fluctuations., *379*, 440. <https://doi.org/10.1086/170520>
- Bond, J. R., Kofman, L., & Pogosyan, D. (1996). How filaments of galaxies are woven into the cosmic web., *380*(6575), 603–606. <https://doi.org/10.1038/380603a0>
- Brodwin, M., Stanford, S. A., Gonzalez, A. H., Zeimann, G. R., Snyder, G. F., Mancone, C. L., Pope, A., Eisenhardt, P. R., Stern, D., Alberts, S., Ashby, M. L. N., Brown, M. J. I., Chary, R. -R., Dey, A., Galametz, A., Gettings, D. P., Jannuzi, B. T., Miller, E. D., Moustakas, J., & Moustakas, L. A. (2013). The Era of Star Formation in Galaxy Clusters., *779*(2), Article 138, 138. <https://doi.org/10.1088/0004-637X/779/2/138>
- Bruzual, G., & Charlot, S. (2003). Stellar population synthesis at the resolution of 2003., *344*(4), 1000–1028. <https://doi.org/10.1046/j.1365-8711.2003.06897.x>
- Capak, P., Aussel, H., Ajiki, M., McCracken, H. J., Mobasher, B., Scoville, N., Shopbell, P., Taniguchi, Y., Thompson, D., Tribiano, S., Sasaki, S., Blain, A. W., Brusa, M., Carilli, C., Comastri, A., Carollo, C. M., Casata, P., Colbert, J., Ellis, R. S., . . . Zamorani, G. (2007). The First Release COSMOS Optical and Near-IR Data and Catalog., *172*(1), 99–116. <https://doi.org/10.1086/519081>
- Capak, P., Scoville, N. Z., Sanders, D. B., Mobasher, B., Masters, D., & Salvato, M. (2010). A Large Sample Of $z_i > 3.5$ Galaxies With Keck

- Deimos Spectroscopy. *American Astronomical Society Meeting Abstracts #215, 215*, Article 410.05, 410.05.
- Cappelluti, N., Brusa, M., Hasinger, G., Comastri, A., Zamorani, G., Finoguenov, A., Gilli, R., Puccetti, S., Miyaji, T., Salvato, M., Vignali, C., Aldcroft, T., Böhringer, H., Brunner, H., Civano, F., Elvis, M., Fiore, F., Fruscione, A., Griffiths, R. E., ... Urry, C. M. (2009). The XMM-Newton wide-field survey in the COSMOS field. The point-like X-ray source catalogue., *497*(2), 635–648. <https://doi.org/10.1051/0004-6361/200810794>
- Chabrier, G. (2003). Galactic Stellar and Substellar Initial Mass Function., *115*(809), 763–795. <https://doi.org/10.1086/376392>
- Cimatti, A., Fraternali, F., & Nipoti, C. (2019). *Introduction to Galaxy Formation and Evolution: From Primordial Gas to Present-Day Galaxies*. <https://doi.org/10.1017/9781316471180>
- Civano, F., Marchesi, S., Comastri, A., Urry, M. C., Elvis, M., Cappelluti, N., Puccetti, S., Brusa, M., Zamorani, G., Hasinger, G., Aldcroft, T., Alexander, D. M., Allevato, V., Brunner, H., Capak, P., Finoguenov, A., Fiore, F., Fruscione, A., Gilli, R., ... Vignali, C. (2016). The Chandra Cosmos Legacy Survey: Overview and Point Source Catalog., *819*(1), Article 62, 62. <https://doi.org/10.3847/0004-637X/819/1/62>
- Coupon, J., Czakon, N., Bosch, J., Komiyama, Y., Medezinski, E., Miyazaki, S., & Oguri, M. (2018). The bright-star masks for the HSC-SSP survey., *70*, Article S7, S7. <https://doi.org/10.1093/pasj/psx047>
- Crawford, S. M., Bershady, M. A., & Hoessel, J. G. (2009). The Red-Sequence Luminosity Function in Massive Intermediate-Redshift Galaxy Clusters., *690*(2), 1158–1180. <https://doi.org/10.1088/0004-637X/690/2/1158>
- de Jong, J. T. A., Verdoes Kleijn, G. A., Erben, T., Hildebrandt, H., Kuijken, K., Sikkema, G., Brescia, M., Bilicki, M., Napolitano, N. R., Amaro, V., Begeman, K. G., Boxhoorn, D. R., Buddelmeijer, H., Cavuoti, S., Getman, F., Grado, A., Helmich, E., Huang, Z., Irisarri, N., ... Viola, M. (2017). The third data release of the Kilo-Degree Survey and associated data products., *604*, Article A134, A134. <https://doi.org/10.1051/0004-6361/201730747>
- De Propris, R., Phillipps, S., & Bremer, M. N. (2013). Deep luminosity functions and colour-magnitude relations for cluster galaxies at $0.2 < z < 0.6$., *434*(4), 3469–3486. <https://doi.org/10.1093/mnras/stt1262>

- de Propriis, R., Stanford, S. A., Eisenhardt, P. R., Dickinson, M., & Elston, R. (1999). The K-Band Luminosity Function in Galaxy Clusters to $Z \sim 1$, *118*(2), 719–729. <https://doi.org/10.1086/300978>
- Delvecchio, I., Smolčić, V., Zamorani, G., Lagos, C. D. P., Berta, S., Delhaize, J., Baran, N., Alexander, D. M., Rosario, D. J., Gonzalez-Perez, V., Ilbert, O., Lacey, C. G., Le Fèvre, O., Miettinen, O., Aravena, M., Bondi, M., Carilli, C., Ciliegi, P., Mooley, K., . . . Tasca, L. (2017). The VLA-COSMOS 3 GHz Large Project: AGN and host-galaxy properties out to $z \lesssim 6$, *602*, Article A3, A3. <https://doi.org/10.1051/0004-6361/201629367>
- Desjacques, V., Jeong, D., & Schmidt, F. (2018). Large-scale galaxy bias., *733*, 1–193. <https://doi.org/10.1016/j.physrep.2017.12.002>
- Despali, G., Giocoli, C., Angulo, R. E., Tormen, G., Sheth, R. K., Baso, G., & Moscardini, L. (2016). The universality of the virial halo mass function and models for non-universality of other halo definitions., *456*(3), 2486–2504. <https://doi.org/10.1093/mnras/stv2842>
- Dodelson, S., & Schmidt, F. (2020). *Modern cosmology*. Academic Press.
- Dressler, A. (1980). Galaxy morphology in rich clusters: implications for the formation and evolution of galaxies., *236*, 351–365. <https://doi.org/10.1086/157753>
- Dunkley, J., Komatsu, E., Nolta, M. R., Spergel, D. N., Larson, D., Hinshaw, G., Page, L., Bennett, C. L., Gold, B., Jarosik, N., Weiland, J. L., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Tucker, G. S., Wollack, E., & Wright, E. L. (2009). Five-Year Wilkinson Microwave Anisotropy Probe Observations: Likelihoods and Parameters from the WMAP Data., *180*(2), 306–329. <https://doi.org/10.1088/0067-0049/180/2/306>
- Einstein, A. (1916). Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik*, *354*(7), 769–822. <https://doi.org/10.1002/andp.19163540702>
- Einstein, A. (1917). Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften (Berlin)*, 142–152.
- Elvis, M., Civano, F., Vignali, C., Puccetti, S., Fiore, F., Cappelluti, N., Aldcroft, T. L., Fruscione, A., Zamorani, G., Comastri, A., Brusa, M., Gilli, R., Miyaji, T., Damiani, F., Koekemoer, A. M., Finoguenov, A., Brunner, H., Urry, C. M., Silverman, J., . . . Volonteri, M. (2009). The

- Chandra COSMOS Survey. I. Overview and Point Source Catalog., *184*(1), 158–171. <https://doi.org/10.1088/0067-0049/184/1/158>
- Euclid Collaboration, Adam, R., Vannier, M., Maurogordato, S., Biviano, A., Adami, C., Ascaso, B., Bellagamba, F., Benoist, C., Cappi, A., Díaz-Sánchez, A., Durret, F., Farrens, S., Gonzalez, A. H., Iovino, A., Licitra, R., Maturi, M., Mei, S., Merson, A., ... Zamorani, G. (2019). Euclid preparation. III. Galaxy cluster detection in the wide photometric survey, performance and algorithm selection., *627*, Article A23, A23. <https://doi.org/10.1051/0004-6361/201935088>
- Finoguenov, A., Guzzo, L., Hasinger, G., Scoville, N. Z., Aussel, H., Böhringer, H., Brusa, M., Capak, P., Cappelluti, N., Comastri, A., Giodini, S., Griffiths, R. E., Impey, C., Koekemoer, A. M., Kneib, J. -P., Leauthaud, A., Le Fèvre, O., Lilly, S., Mainieri, V., ... Zamorani, G. (2007). The XMM-Newton Wide-Field Survey in the COSMOS Field: Statistical Properties of Clusters of Galaxies., *172*(1), 182–195. <https://doi.org/10.1086/516577>
- Finoguenov, A., Connelly, J. L., Parker, L. C., Wilman, D. J., Mulchaey, J. S., Saglia, R. P., Balogh, M. L., Bower, R. G., & McGee, S. L. (2009). The Roadmap for Unification in Galaxy Group Selection. I. A Search for Extended X-ray Emission in the CNOC2 Survey., *704*(1), 564–575. <https://doi.org/10.1088/0004-637X/704/1/564>
- Finoguenov, A., Watson, M. G., Tanaka, M., Simpson, C., Cirasuolo, M., Dunlop, J. S., Peacock, J. A., Farrah, D., Akiyama, M., Ueda, Y., Smolčić, V., Stewart, G., Rawlings, S., van Breukelen, C., Almaini, O., Clewley, L., Bonfield, D. G., Jarvis, M. J., Barr, J. M., ... Egami, E. (2010). X-ray groups and clusters of galaxies in the Subaru-XMM Deep Field., *403*(4), 2063–2076. <https://doi.org/10.1111/j.1365-2966.2010.16256.x>
- Fixsen, D. J. (2009). The Temperature of the Cosmic Microwave Background., *707*(2), 916–920. <https://doi.org/10.1088/0004-637X/707/2/916>
- Freedman, W. L., & Madore, B. F. (2010). The Hubble Constant., *48*, 673–710. <https://doi.org/10.1146/annurev-astro-082708-101829>
- Friedmann, A. (1922). Über die Krümmung des Raumes. *Zeitschrift für Physik*, *10*, 377–386. <https://doi.org/10.1007/BF01332580>

- Friedmann, A. (1924). Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes. *Zeitschrift für Physik*, *21*(1), 326–332. <https://doi.org/10.1007/BF01328280>
- George, M. R., Leauthaud, A., Bundy, K., Finoguenov, A., Tinker, J., Lin, Y.-T., Mei, S., Kneib, J.-P., Aussel, H., Behroozi, P. S., Busha, M. T., Capak, P., Coccato, L., Covone, G., Faure, C., Fiorenza, S. L., Ilbert, O., Le Floch, E., Koekemoer, A. M., ... Wolk, M. (2011). Galaxies in X-Ray Groups. I. Robust Membership Assignment and the Impact of Group Environments on Quenching., *742*(2), Article 125, 125. <https://doi.org/10.1088/0004-637X/742/2/125>
- Giacconi, R., Murray, S., Gursky, H., Kellogg, E., Schreier, E., & Tananbaum, H. (1972). The Uhuru catalog of X-ray sources., *178*, 281–308. <https://doi.org/10.1086/151790>
- Gonzalez, A. (2014). Cluster detection via wavelets. *Building the Euclid Cluster Survey - Scientific Program*, Article 7, 7.
- Gozaliasl, G., Finoguenov, A., Tanaka, M., Dolag, K., Montanari, F., Kirkpatrick, C. C., Vardoulaki, E., Khosroshahi, H. G., Salvato, M., Laigle, C., McCracken, H. J., Ilbert, O., Cappelluti, N., Daddi, E., Hasinger, G., Capak, P., Scoville, N. Z., Toft, S., Civano, F., ... Erfanianfar, G. (2019). Chandra centres for COSMOS X-ray galaxy groups: differences in stellar properties between central dominant and offset brightest group galaxies., *483*(3), 3545–3565. <https://doi.org/10.1093/mnras/sty3203>
- Harrison, E. R. (1970). Fluctuations at the Threshold of Classical Cosmology., *1*(10), 2726–2730. <https://doi.org/10.1103/PhysRevD.1.2726>
- Hasinger, G., Cappelluti, N., Brunner, H., Brusa, M., Comastri, A., Elvis, M., Finoguenov, A., Fiore, F., Franceschini, A., Gilli, R., Griffiths, R. E., Lehmann, I., Mainieri, V., Matt, G., Matute, I., Miyaji, T., Molendi, S., Paltani, S., Sanders, D. B., ... Zamorani, G. (2007). The XMM-Newton Wide-Field Survey in the COSMOS Field. I. Survey Description., *172*(1), 29–37. <https://doi.org/10.1086/516576>
- Hasinger, G., Capak, P., Salvato, M., Barger, A. J., Cowie, L. L., Faisst, A., Hemmati, S., Kakazu, Y., Kartaltepe, J., Masters, D., Mobasher, B., Nayyeri, H., Sanders, D., Scoville, N. Z., Suh, H., Steinhardt, C., & Yang, F. (2018). The DEIMOS 10K Spectroscopic Survey Catalog of the COSMOS Field., *858*(2), Article 77, 77. <https://doi.org/10.3847/1538-4357/aabacf>

- Hennig, C., Mohr, J. J., Zenteno, A., Desai, S., Dietrich, J. P., Bocquet, S., Strazzullo, V., Saro, A., Abbott, T. M. C., Abdalla, F. B., Bayliss, M., Benoit-Lévy, A., Bernstein, R. A., Bertin, E., Brooks, D., Capasso, R., Capozzi, D., Carnero, A., Carrasco Kind, M., . . . Zhang, Y. (2017). Galaxy populations in massive galaxy clusters to $z = 1.1$: colour distribution, concentration, halo occupation number and red sequence fraction., *467*(4), 4015–4035. <https://doi.org/10.1093/mnras/stx175>
- Hubble, E. (1929). A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, *15*(3), 168–173. <https://doi.org/10.1073/pnas.15.3.168>
- Ilbert, O., Arnouts, S., McCracken, H. J., Bolzonella, M., Bertin, E., Le Fèvre, O., Mellier, Y., Zamorani, G., Pellò, R., Iovino, A., Tresse, L., Le Brun, V., Bottini, D., Garilli, B., Maccagni, D., Picat, J. P., Scaramella, R., Scodreggio, M., Vettolani, G., . . . Vergani, D. (2006). Accurate photometric redshifts for the CFHT legacy survey calibrated using the VIMOS VLT deep survey., *457*(3), 841–856. <https://doi.org/10.1051/0004-6361:20065138>
- Ilbert, O., Capak, P., Salvato, M., Aussel, H., McCracken, H. J., Sanders, D. B., Scoville, N., Kartaltepe, J., Arnouts, S., Le Floch, E., Mobasher, B., Taniguchi, Y., Lamareille, F., Leauthaud, A., Sasaki, S., Thompson, D., Zamojski, M., Zamorani, G., Bardelli, S., . . . Zucca, E. (2009). Cosmos Photometric Redshifts with 30-Bands for 2-deg²., *690*(2), 1236–1249. <https://doi.org/10.1088/0004-637X/690/2/1236>
- Ilbert, O., McCracken, H. J., Le Fèvre, O., Capak, P., Dunlop, J., Karim, A., Renzini, M. A., Caputi, K., Boissier, S., Arnouts, S., Aussel, H., Comparat, J., Guo, Q., Hudelot, P., Kartaltepe, J., Kneib, J. P., Krogager, J. K., Le Floch, E., Lilly, S., . . . Zirm, A. (2013). Mass assembly in quiescent and star-forming galaxies since $z \hat{=} 4$ from UltraVISTA., *556*, Article A55, A55. <https://doi.org/10.1051/0004-6361/201321100>
- Kaiser, N. (1984). On the spatial correlations of Abell clusters., *284*, L9–L12. <https://doi.org/10.1086/184341>
- Koekemoer, A. M., Aussel, H., Calzetti, D., Capak, P., Giavalisco, M., Kneib, J. -P., Leauthaud, A., Le Fèvre, O., McCracken, H. J., Massey, R., Mobasher, B., Rhodes, J., Scoville, N., & Shopbell, P. L. (2007). The COSMOS Survey: Hubble Space Telescope Advanced Camera for Surveys Observations and Data Processing., *172*(1), 196–202. <https://doi.org/10.1086/520086>

- Koester, B. P., McKay, T. A., Annis, J., Wechsler, R. H., Evrard, A., Bleem, L., Becker, M., Johnston, D., Sheldon, E., Nichol, R., Miller, C., Scranton, R., Bahcall, N., Barentine, J., Brewington, H., Brinkmann, J., Harvanek, M., Kleinman, S., Krzesinski, J., ... York, D. (2007). A MaxBCG Catalog of 13,823 Galaxy Clusters from the Sloan Digital Sky Survey., *660*(1), 239–255. <https://doi.org/10.1086/509599>
- Komatsu, E., Smith, K. M., Dunkley, J., Bennett, C. L., Gold, B., Hinshaw, G., Jarosik, N., Larson, D., Nolta, M. R., Page, L., Spergel, D. N., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Tucker, G. S., Weiland, J. L., ... Wright, E. L. (2011). Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation., *192*(2), Article 18, 18. <https://doi.org/10.1088/0067-0049/192/2/18>
- Laigle, C., McCracken, H. J., Ilbert, O., Hsieh, B. C., Davidzon, I., Capak, P., Hasinger, G., Silverman, J. D., Pichon, C., Coupon, J., Aussel, H., Le Borgne, D., Caputi, K., Cassata, P., Chang, Y. -Y., Civano, F., Dunlop, J., Fynbo, J., Kartaltepe, J. S., ... Zabl, J. (2016). The COSMOS2015 Catalog: Exploring the $1 < z < 6$ Universe with Half a Million Galaxies., *224*(2), Article 24, 24. <https://doi.org/10.3847/0067-0049/224/2/24>
- Laureijs, R., Amiaux, J., Arduini, S., Auguères, J. -L., Brinchmann, J., Cole, R., Cropper, M., Dabin, C., Duvet, L., Ealet, A., Garilli, B., Gondoin, P., Guzzo, L., Hoar, J., Hoekstra, H., Holmes, R., Kitching, T., Maciaszek, T., Mellier, Y., ... Zucca, E. (2011). Euclid Definition Study Report. *arXiv e-prints*, Article arXiv:1110.3193, arXiv:1110.3193.
- Leauthaud, A., Finoguenov, A., Kneib, J.-P., Taylor, J. E., Massey, R., Rhodes, J., Ilbert, O., Bundy, K., Tinker, J., George, M. R., Capak, P., Koekemoer, A. M., Johnston, D. E., Zhang, Y.-Y., Cappelluti, N., Ellis, R. S., Elvis, M., Giodini, S., Heymans, C., ... Wolk, M. (2010). A Weak Lensing Study of X-ray Groups in the Cosmos Survey: Form and Evolution of the Mass-Luminosity Relation., *709*(1), 97–114. <https://doi.org/10.1088/0004-637X/709/1/97>
- Lemaître, G. (1927). Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extragalactiques. *Annales de la Société; Scientifique de Bruxelles*, *47*, 49–59.

- Lesci, G. F., Marulli, F., Moscardini, L., Sereno, M., Veropalumbo, A., Maturi, M., Giocoli, C., Radovich, M., Bellagamba, F., Roncarelli, M., Bardelli, S., Contarini, S., Covone, G., Ingoglia, L., Nanni, L., & Puddu, E. (2020). AMICO galaxy clusters in KiDS-DR3: cosmological constraints from counts and stacked weak-lensing. *arXiv e-prints*, Article arXiv:2012.12273, arXiv:2012.12273.
- Lilly, S. J., Le Fèvre, O., Renzini, A., Zamorani, G., Scodreggio, M., Contini, T., Carollo, C. M., Hasinger, G., Kneib, J. -P., Iovino, A., Le Brun, V., Maier, C., Mainieri, V., Mignoli, M., Silverman, J., Tasca, L. A. M., Bolzonella, M., Bongiorno, A., Bottini, D., ... Zucca, E. (2007). zCOSMOS: A Large VLT/VIMOS Redshift Survey Covering $0 < z < 3$ in the COSMOS Field., *172*(1), 70–85. <https://doi.org/10.1086/516589>
- Lutz, D., Poglitsch, A., Altieri, B., Andreani, P., Aussel, H., Berta, S., Bongiovanni, A., Brisbin, D., Cava, A., Cepa, J., Cimatti, A., Daddi, E., Dominguez-Sanchez, H., Elbaz, D., Förster Schreiber, N. M., Genzel, R., Grazian, A., Gruppioni, C., Harwit, M., ... Wieprecht, E. (2011). PACS Evolutionary Probe (PEP) - A Herschel key program., *532*, Article A90, A90. <https://doi.org/10.1051/0004-6361/201117107>
- Mancone, C. L., & Gonzalez, A. H. (2012). EzGal: A Flexible Interface for Stellar Population Synthesis Models., *124*(916), 606. <https://doi.org/10.1086/666502>
- Mancone, C. L., Gonzalez, A. H., Brodwin, M., Stanford, S. A., Eisenhardt, P. R. M., Stern, D., & Jones, C. (2010). The Formation of Massive Cluster Galaxies., *720*(1), 284–298. <https://doi.org/10.1088/0004-637X/720/1/284>
- Mantz, A., Allen, S. W., Ebeling, H., Rapetti, D., & Drlica-Wagner, A. (2010). The observed growth of massive galaxy clusters - II. X-ray scaling relations., *406*(3), 1773–1795. <https://doi.org/10.1111/j.1365-2966.2010.16993.x>
- Marchesi, S., Civano, F., Elvis, M., Salvato, M., Brusa, M., Comastri, A., Gilli, R., Hasinger, G., Lanzuisi, G., Miyaji, T., Treister, E., Urry, C. M., Vignali, C., Zamorani, G., Alleinato, V., Cappelluti, N., Cardamone, C., Finoguenov, A., Griffiths, R. E., ... Trakhtenbrot, B. (2016). The Chandra COSMOS Legacy survey: optical/IR identifications., *817*(1), Article 34, 34. <https://doi.org/10.3847/0004-637X/817/1/34>

- Maturi, M., Meneghetti, M., Bartelmann, M., Dolag, K., & Moscardini, L. (2005). An optimal filter for the detection of galaxy clusters through weak lensing., *442*(3), 851–860. <https://doi.org/10.1051/0004-6361:20042600>
- Maturi, M., Bellagamba, F., Radovich, M., Roncarelli, M., Sereno, M., Moscardini, L., Bardelli, S., & Puddu, E. (2019). AMICO galaxy clusters in KiDS-DR3: sample properties and selection function., *485*(1), 498–512. <https://doi.org/10.1093/mnras/stz294>
- McCracken, H. J., Capak, P., Salvato, M., Aussel, H., Thompson, D., Daddi, E., Sanders, D. B., Kneib, J. -P., Willott, C. J., Mancini, C., Renzini, A., Cook, R., Le Fèvre, O., Ilbert, O., Kartaltepe, J., Koekemoer, A. M., Mellier, Y., Murayama, T., Scoville, N. Z., ... Tanaguchi, Y. (2010). The COSMOS-WIRCam Near-Infrared Imaging Survey. I. BzK-Selected Passive and Star-Forming Galaxy Candidates at $z_{\text{gsim}} 1.4$., *708*(1), 202–217. <https://doi.org/10.1088/0004-637X/708/1/202>
- McCracken, H. J., Milvang-Jensen, B., Dunlop, J., Franx, M., Fynbo, J. P. U., Le Fèvre, O., Holt, J., Caputi, K. I., Goranova, Y., Buitrago, F., Emerson, J. P., Freudling, W., Hudelot, P., López-Sanjuan, C., Magnard, F., Mellier, Y., Møller, P., Nilsson, K. K., Sutherland, W., ... Zabl, J. (2012). UltraVISTA: a new ultra-deep near-infrared survey in COSMOS., *544*, Article A156, A156. <https://doi.org/10.1051/0004-6361/201219507>
- Meneghetti, M., Bartelmann, M., & Moscardini, L. (2003). Cluster cross-sections for strong lensing: analytic and numerical lens models., *340*(1), 105–114. <https://doi.org/10.1046/j.1365-8711.2003.06276.x>
- Miyazaki, S., Komiyama, Y., Nakaya, H., Kamata, Y., Doi, Y., Hamana, T., Karoji, H., Furusawa, H., Kawanomoto, S., Morokuma, T., Ishizuka, Y., Nariai, K., Tanaka, Y., Uraguchi, F., Utsumi, Y., Obuchi, Y., Okura, Y., Oguri, M., Takata, T., ... Takeshi, K. (2012). Hyper Suprime-Cam. In I. S. McLean, S. K. Ramsay, & H. Takami (Eds.), *Ground-based and Airborne Instrumentation for Astronomy IV* (84460Z). <https://doi.org/10.1117/12.926844>
- Molendi, S. (2004). The Intra-Cluster Medium. In G. Bertin, D. Farina, & R. Pozzoli (Eds.), *Plasmas in the Laboratory and in the Universe: New Insights and New Challenges* (pp. 345–354). <https://doi.org/10.1063/1.1718479>

- Muldrew, S. I., Hatch, N. A., & Cooke, E. A. (2015). What are protoclusters? - Defining high-redshift galaxy clusters and protoclusters., *452*(3), 2528–2539. <https://doi.org/10.1093/mnras/stv1449>
- Navarro, J. F., Frenk, C. S., & White, S. D. M. (1997). A Universal Density Profile from Hierarchical Clustering., *490*(2), 493–508. <https://doi.org/10.1086/304888>
- Oke, J. B. (1974). Absolute Spectral Energy Distributions for White Dwarfs., *27*, 21. <https://doi.org/10.1086/190287>
- Overzier, R. A. (2016). The realm of the galaxy protoclusters. A review., *24*(1), Article 14, 14. <https://doi.org/10.1007/s00159-016-0100-3>
- Penzias, A. A., & Wilson, R. W. (1965). A Measurement of Excess Antenna Temperature at 4080 Mc/s., *142*, 419–421. <https://doi.org/10.1086/148307>
- Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R. A., Nugent, P., Castro, P. G., Deustua, S., Fabbro, S., Goobar, A., Groom, D. E., Hook, I. M., Kim, A. G., Kim, M. Y., Lee, J. C., Nunes, N. J., Pain, R., Pennypacker, C. R., Quimby, R., Lidman, C., ... Project, T. S. C. (1999). Measurements of Ω and Λ from 42 High-Redshift Supernovae., *517*(2), 565–586. <https://doi.org/10.1086/307221>
- Planck Collaboration, Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Battye, R., Benabed, K., Bernard, J. -P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bond, J. R., Borrill, J., ... Zonca, A. (2020). Planck 2018 results. VI. Cosmological parameters., *641*, Article A6, A6. <https://doi.org/10.1051/0004-6361/201833910>
- Postman, M., Lubin, L. M., Gunn, J. E., Oke, J. B., Hoessel, J. G., Schneider, D. P., & Christensen, J. A. (1996). The Palomar Distant Clusters Survey. I. The Cluster Catalog., *111*, 615. <https://doi.org/10.1086/117811>
- Pratt, G. W., Arnaud, M., Biviano, A., Eckert, D., Ettori, S., Nagai, D., Okabe, N., & Reiprich, T. H. (2019). The Galaxy Cluster Mass Scale and Its Impact on Cosmological Constraints from the Cluster Population., *215*(2), Article 25, 25. <https://doi.org/10.1007/s11214-019-0591-0>
- Prescott, M. K. M., Impey, C. D., Cool, R. J., & Scoville, N. Z. (2006). Quasars in the COSMOS Field., *644*(1), 100–115. <https://doi.org/10.1086/503325>

- Press, W. H., & Schechter, P. (1974). Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation., *187*, 425–438. <https://doi.org/10.1086/152650>
- Ramella, M., Boschin, W., Fadda, D., & Nonino, M. (2001). Finding galaxy clusters using Voronoi tessellations., *368*, 776–786. <https://doi.org/10.1051/0004-6361:20010071>
- Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan, C. J., Jha, S., Kirshner, R. P., Leibundgut, B., Phillips, M. M., Reiss, D., Schmidt, B. P., Schommer, R. A., Smith, R. C., Spyromilio, J., Stubbs, C., Suntzeff, N. B., & Tonry, J. (1998). Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant., *116*(3), 1009–1038. <https://doi.org/10.1086/300499>
- Robertson, H. P. (1935). Kinematics and World-Structure., *82*, 284. <https://doi.org/10.1086/143681>
- Rosati, P., Borgani, S., & Norman, C. (2002). The Evolution of X-ray Clusters of Galaxies., *40*, 539–577. <https://doi.org/10.1146/annurev.astro.40.120401.150547>
- Rozo, E., Wechsler, R. H., Rykoff, E. S., Annis, J. T., Becker, M. R., Evrard, A. E., Frieman, J. A., Hansen, S. M., Hao, J., Johnston, D. E., Koester, B. P., McKay, T. A., Sheldon, E. S., & Weinberg, D. H. (2010). Cosmological Constraints from the Sloan Digital Sky Survey maxBCG Cluster Catalog., *708*(1), 645–660. <https://doi.org/10.1088/0004-637X/708/1/645>
- Rykoff, E. S., Rozo, E., Busha, M. T., Cunha, C. E., Finoguenov, A., Evrard, A., Hao, J., Koester, B. P., Leauthaud, A., Nord, B., Pierre, M., Reddick, R., Sadibekova, T., Sheldon, E. S., & Wechsler, R. H. (2014). redMaPPer. I. Algorithm and SDSS DR8 Catalog., *785*(2), Article 104, 104. <https://doi.org/10.1088/0004-637X/785/2/104>
- Salpeter, E. E. (1955). The Luminosity Function and Stellar Evolution., *121*, 161. <https://doi.org/10.1086/145971>
- Salvato, M., Ilbert, O., Hasinger, G., Rau, A., Civano, F., Zamorani, G., Brusa, M., Elvis, M., Vignali, C., Aussel, H., Comastri, A., Fiore, F., Le Floch, E., Mainieri, V., Bardelli, S., Bolzonella, M., Bongiorno, A., Capak, P., Caputi, K., . . . Zucca, E. (2011). Dissecting Photometric Redshift for Active Galactic Nucleus Using XMM- and Chandra-

- COSMOS Samples., *742*(2), Article 61, 61. <https://doi.org/10.1088/0004-637X/742/2/61>
- Sanders, D. B., Salvato, M., Aussel, H., Ilbert, O., Scoville, N., Surace, J. A., Frayer, D. T., Sheth, K., Helou, G., Brooke, T., Bhattacharya, B., Yan, L., Kartaltepe, J. S., Barnes, J. E., Blain, A. W., Calzetti, D., Capak, P., Carilli, C., Carollo, C. M., ... Williams, J. P. (2007). S-COSMOS: The Spitzer Legacy Survey of the Hubble Space Telescope ACS 2 deg² COSMOS Field I: Survey Strategy and First Analysis., *172*(1), 86–98. <https://doi.org/10.1086/517885>
- Sarazin, C. L. (1986). X-ray emission from clusters of galaxies. *Reviews of Modern Physics*, *58*(1), 1–115. <https://doi.org/10.1103/RevModPhys.58.1>
- Sartoris, B., Biviano, A., Fedeli, C., Bartlett, J. G., Borgani, S., Costanzi, M., Giocoli, C., Moscardini, L., Weller, J., Ascaso, B., Bardelli, S., Maurogordato, S., & Viana, P. T. P. (2016). Next generation cosmology: constraints from the Euclid galaxy cluster survey., *459*(2), 1764–1780. <https://doi.org/10.1093/mnras/stw630>
- Schechter, P. (1976). An analytic expression for the luminosity function for galaxies., *203*, 297–306. <https://doi.org/10.1086/154079>
- Scoville, N., Aussel, H., Brusa, M., Capak, P., Carollo, C. M., Elvis, M., Giavalisco, M., Guzzo, L., Hasinger, G., Impey, C., Kneib, J. -P., LeFevre, O., Lilly, S. J., Mobasher, B., Renzini, A., Rich, R. M., Sanders, D. B., Schinnerer, E., Schminovich, D., ... Tyson, N. D. (2007). The Cosmic Evolution Survey (COSMOS): Overview., *172*(1), 1–8. <https://doi.org/10.1086/516585>
- Shan, H., Kneib, J.-P., Tao, C., Fan, Z., Jauzac, M., Limousin, M., Massey, R., Rhodes, J., Thanjavur, K., & McCracken, H. J. (2012). Weak Lensing Measurement of Galaxy Clusters in the CFHTLS-Wide Survey., *748*(1), Article 56, 56. <https://doi.org/10.1088/0004-637X/748/1/56>
- Sheth, R. K., & Tormen, G. (1999). Large-scale bias and the peak background split., *308*(1), 119–126. <https://doi.org/10.1046/j.1365-8711.1999.02692.x>
- Shirasaki, M., Ishiyama, T., & Ando, S. (2021). Virial Halo Mass Function in the Planck Cosmology., *922*(1), Article 89, 89. <https://doi.org/10.3847/1538-4357/ac214b>

- Singh, P., Saro, A., Costanzi, M., & Dolag, K. (2020). Cosmology dependence of galaxy cluster scaling relations., *494*(3), 3728–3740. <https://doi.org/10.1093/mnras/staa1004>
- Smolčić, V., Schinnerer, E., Finoguenov, A., Sakelliou, I., Carilli, C. L., Botzler, C. S., Brusa, M., Scoville, N., Ajiki, M., Capak, P., Guzzo, L., Hasinger, G., Impey, C., Jahnke, K., Kartaltepe, J. S., McCracken, H. J., Mobasher, B., Murayama, T., Sasaki, S. S., ... Trump, J. R. (2007). A Wide-Angle Tail Radio Galaxy in the COSMOS Field: Evidence for Cluster Formation., *172*(1), 295–313. <https://doi.org/10.1086/516583>
- Smolčić, V., Novak, M., Bondi, M., Ciliegi, P., Mooley, K. P., Schinnerer, E., Zamorani, G., Navarrete, F., Bourke, S., Karim, A., Vardoulaki, E., Leslie, S., Delhaize, J., Carilli, C. L., Myers, S. T., Baran, N., Delvecchio, I., Miettinen, O., Banfield, J., ... Sheth, K. (2017). The VLA-COSMOS 3 GHz Large Project: Continuum data and source catalog release., *602*, Article A1, A1. <https://doi.org/10.1051/0004-6361/201628704>
- Springel, V., Frenk, C. S., & White, S. D. M. (2006). The large-scale structure of the Universe., *440*(7088), 1137–1144. <https://doi.org/10.1038/nature04805>
- Stapelberg, S., Carrasco, M., & Maturi, M. (2019). EasyCritics - I. Efficient detection of strongly lensing galaxy groups and clusters in wide-field surveys., *482*(2), 1824–1839. <https://doi.org/10.1093/mnras/sty2784>
- Story, K. T., Reichardt, C. L., Hou, Z., Keisler, R., Aird, K. A., Benson, B. A., Bleem, L. E., Carlstrom, J. E., Chang, C. L., Cho, H. -M., Crawford, T. M., Crites, A. T., de Haan, T., Dobbs, M. A., Dudley, J., Follin, B., George, E. M., Halverson, N. W., Holder, G. P., ... Zahn, O. (2013). A Measurement of the Cosmic Microwave Background Damping Tail from the 2500-Square-Degree SPT-SZ Survey., *779*(1), Article 86, 86. <https://doi.org/10.1088/0004-637X/779/1/86>
- Stoughton, C., Lupton, R. H., Bernardi, M., Blanton, M. R., Burles, S., Castander, F. J., Connolly, A. J., Eisenstein, D. J., Frieman, J. A., Hennessy, G. S., Hindsley, R. B., Ivezić, Ž., Kent, S., Kunszt, P. Z., Lee, B. C., Meiksin, A., Munn, J. A., Newberg, H. J., Nichol, R. C., ... Zheng, W. (2002). Sloan Digital Sky Survey: Early Data Release., *123*(1), 485–548. <https://doi.org/10.1086/324741>

- Strazzullo, V., Daddi, E., Gobat, R., Valentino, F., Pannella, M., Dickinson, M., Renzini, A., Brammer, G., Onodera, M., Finoguenov, A., Cimatti, A., Carollo, C. M., & Arimoto, N. (2016). The Red Sequence at Birth in the Galaxy Cluster Cl J1449+0856 at $z = 2$. *833*(2), Article L20, L20. <https://doi.org/10.3847/2041-8213/833/2/L20>
- Sunyaev, R. A., & Zeldovich, Y. B. (1970). The Spectrum of Primordial Radiation, its Distortions and their Significance. *Comments on Astrophysics and Space Physics*, *2*, 66.
- Taniguchi, Y., Scoville, N., Murayama, T., Sanders, D. B., Mobasher, B., Aussel, H., Capak, P., Ajiki, M., Miyazaki, S., Komiyama, Y., Shioya, Y., Nagao, T., Sasaki, S. S., Koda, J., Carilli, C., Giavalisco, M., Guzzo, L., Hasinger, G., Impey, C., ... Ohta, K. (2007). The Cosmic Evolution Survey (COSMOS): Subaru Observations of the HST Cosmos Field., *172*(1), 9–28. <https://doi.org/10.1086/516596>
- Taniguchi, Y., Kajisawa, M., Kobayashi, M. A. R., Shioya, Y., Nagao, T., Capak, P. L., Aussel, H., Ichikawa, A., Murayama, T., Scoville, N. Z., Ilbert, O., Salvato, M., Sanders, D. B. B., Mobasher, B., Miyazaki, S., Komiyama, Y., Le Fèvre, O., Tasca, L., Lilly, S., ... Hayashino, T. (2015). The Subaru COSMOS 20: Subaru optical imaging of the HST COSMOS field with 20 filters*, *67*(6), Article 104, 104. <https://doi.org/10.1093/pasj/psv106>
- The Dark Energy Survey Collaboration. (2005). The Dark Energy Survey. *arXiv e-prints*, Article astro-ph/0510346, astro-ph/0510346.
- Tinker, J. L., Robertson, B. E., Kravtsov, A. V., Klypin, A., Warren, M. S., Yepes, G., & Gottlöber, S. (2010). The Large-scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests., *724*(2), 878–886. <https://doi.org/10.1088/0004-637X/724/2/878>
- van Weeren, R. J., de Gasperin, F., Akamatsu, H., Brügger, M., Feretti, L., Kang, H., Stroe, A., & Zandanel, F. (2019). Diffuse Radio Emission from Galaxy Clusters., *215*(1), Article 16, 16. <https://doi.org/10.1007/s11214-019-0584-z>
- Vikhlinin, A., McNamara, B. R., Forman, W., Jones, C., Quintana, H., & Hornstrup, A. (1998). A Catalog of 200 Galaxy Clusters Serendipitously Detected in the ROSAT PSPC Pointed Observations., *502*(2), 558–581. <https://doi.org/10.1086/305951>
- Vikhlinin, A., Kravtsov, A. V., Burenin, R. A., Ebeling, H., Forman, W. R., Hornstrup, A., Jones, C., Murray, S. S., Nagai, D., Quintana, H., &

- Voevodkin, A. (2009). Chandra Cluster Cosmology Project III: Cosmological Parameter Constraints., *692*(2), 1060–1074. <https://doi.org/10.1088/0004-637X/692/2/1060>
- Weaver, J. R., Kauffmann, O. B., Ilbert, O., McCracken, H. J., Moneti, A., Toft, S., Brammer, G., Shuntov, M., Davidzon, I., Hsieh, B. C., Laigle, C., Anastasiou, A., Jespersen, C. K., Vinther, J., Capak, P., Casey, C. M., McPartland, C. J. R., Milvang-Jensen, B., Mobasher, B., ... Zamorani, G. (2022). COSMOS2020: A Panchromatic View of the Universe to z 10 from Two Complementary Catalogs., *258*(1), Article 11, 11. <https://doi.org/10.3847/1538-4365/ac3078>
- Wen, Z. L., Han, J. L., & Liu, F. S. (2012). A Catalog of 132,684 Clusters of Galaxies Identified from Sloan Digital Sky Survey III., *199*(2), Article 34, 34. <https://doi.org/10.1088/0067-0049/199/2/34>
- Zamojski, M. A., Schiminovich, D., Rich, R. M., Mobasher, B., Koekemoer, A. M., Capak, P., Taniguchi, Y., Sasaki, S. S., McCracken, H. J., Mellier, Y., Bertin, E., Aussel, H., Sanders, D. B., Le Fèvre, O., Ilbert, O., Salvato, M., Thompson, D. J., Kartaltepe, J. S., Scoville, N., ... Yi, S. K. (2007). Deep GALEX Imaging of the COSMOS HST Field: A First Look at the Morphology of $z \sim 0.7$ Star-forming Galaxies., *172*(1), 468–493. <https://doi.org/10.1086/516593>
- Zehavi, I., Zheng, Z., Weinberg, D. H., Blanton, M. R., Bahcall, N. A., Berlind, A. A., Brinkmann, J., Frieman, J. A., Gunn, J. E., Lupton, R. H., Nichol, R. C., Percival, W. J., Schneider, D. P., Skibba, R. A., Strauss, M. A., Tegmark, M., & York, D. G. (2011). Galaxy Clustering in the Completed SDSS Redshift Survey: The Dependence on Color and Luminosity., *736*(1), Article 59, 59. <https://doi.org/10.1088/0004-637X/736/1/59>
- Zeldovich, Y. B. (1972). A hypothesis, unifying the structure and the entropy of the Universe., *160*, 1P. <https://doi.org/10.1093/mnras/160.1.1P>
- Zenteno, A., Mohr, J. J., Desai, S., Stalder, B., Saro, A., Dietrich, J. P., Bayliss, M., Bocquet, S., Chiu, I., Gonzalez, A. H., Gangkofner, C., Gupta, N., Hlavacek-Larrondo, J., McDonald, M., Reichardt, C., & Rest, A. (2016). Galaxy populations in the 26 most massive galaxy clusters in the South Pole Telescope SPT-SZ survey., *462*(1), 830–843. <https://doi.org/10.1093/mnras/stw1649>
- Zwicky, F. (1933). Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, *6*, 110–127.

Zwicky, F. (1937). On the Masses of Nebulae and of Clusters of Nebulae., *86*, 217. <https://doi.org/10.1086/143864>