UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI FISICA E ASTRONOMIA "GALILEO GALILEI"

Master Degree in Astrophysics and Cosmology

Final dissertation

Candidates (proto-)clusters in the near-infrared wide field SHARKS survey: preparing the path to Euclid

Supervisor:

Prof. Giulia Rodighiero

Co-supervisor:

Dr. Francesco Sinigaglia

Candidate:

Giuliano Lorenzon

Academic year 2021/2022

Contents

Abstract

I have constructed a catalog of galaxy clusters and proto-clusters using data from the Southern H-ATLAS Regions Ks-band Survey (SHARKS). This is an ESO public survey in Ks photometric band, covering almost the entire GAlaxy Mass Assembly (GAMA) fields for a total of ∼300 square degrees in the sky. I used data from the first public release (dr1), that includes an area of ∼20 square degrees including parts of GAMA-12, GAMA-15 and South Galactic Pole (SGP) regions. I removed spurious sources from the public catalog, calculated the completeness level and removed Galactic sources to have a complete catalog of galaxies. The sensitivity level at 5σ in AB magnitudes as described in the public release is 22.7 mag. I calculated an average completeness level of 22.1 mag, varying along the different areas of the survey. I performed the search for clusters in the catalog using a Friends-of-Friends (FoF) algorithm from part of the Python library nbodykit. The algorithm was originally developed to perform analysis on N-body cosmological simulations, however I applied it to a two-dimensional sky-projection of the distribution of galaxies. This required a new definition of the parameters governing the algorithm, that was tested and discussed for the entire dr1 catalog. To construct the catalog, instead, I have combined SHARKS data to data from the photometric Y band of the Dark Energy Survey (DES). I used observations in Y and Ks band to develop a color-redshift relation calibrated for the galaxy red sequence. I have then assigned redshift values for each galaxy and divided the catalog in bins in the redshift space of width Δz=0.3 from z=0 to z=4. The FoF algorithm was applied to each bin. This analysis was performed in the overlapping area between SHARKS and DES: ∼8 squared degrees in the SGP region. The resulting catalog contains 186 galaxy clusters and proto-clusters between redshift z∼0.6 and z∼3 in such an area, for an average of 23 clusters per squared degree. This seems consistent by order of magnitude with analysis from Kilo Degree Survey (KiDS), having 11 clusters per squared degree, and DES, having 28 clusters per squared degree at the same photometric limit of SHARKS.

Chapter 1

Introduction

Infrared (IR) astronomy has undoubtedly gained importance in the last decades, thanks to new powerful technologies employed in ground and space observatories.
As a true protagonist of the decade, NASA James Webb Space Telescope (JWST) (Katz et al. 2019) is already providing stunning photometric and spectroscopic data, exposing the secrets of the local and very distant infrared Universe. Recent missions as Spitzer(Werner et al. 2004), Hershel (Pilbratt et al. 2010) and WISE (Wright et al. 2010) provided essential data on both planetary formation, stellar evolution and galaxy evolution. In the immediate future, among others, ESA Euclid space observatory (Scaramella et al. 2014) will soon bring its own wonderful contribution, exploring in detail almost the entire sky to show the appearance of a Universe only few billion years old.
Telescopes of this kind are essential for the study of infrared phenomena in the local Universe, ranging from characterization of exoplanets and planetary systems formation, to mapping dust content and precisely measuring stellar masses in nearby galaxies.
In this work, however, I will focus on the importance of observation of the distant Universe, for which Near InfraRed (NIR) and IR observation are mainly required for the two following purposes.
Many photometric and spectral features of local galaxies are observed in the UV and optical part of the electromagnetic spectrum. Nevertheless when looking at very distant objects, cosmological redshift [1] maps the entire spectrum of a galaxy to lower frequencies, and such features are no longer observable at their rest frame frequencies.
Another important aspect is the increasing amount of cold gas and dust found observing towards higher redshifts (Casey, Narayanan, and Cooray 2014; Manning et al. 2022; Chapman et al. 2004): these largely absorb (and scatter) ultraviolet and optical radiation, re-emitting them at lower frequencies. For this reason a galaxy rich in dust content will appear practically optically invisible, while it could shine in the IR.
In the following section I will briefly summarize the standard cosmological theory arising from studies of the matter distribution of the Universe. I will also try to focus on some key information we can draw from mapping how galaxies and gas are distributed on the large cosmological scale.

[1] Light traveling through expanding space undergoes a progressive decrease in energy and thus lowers its frequency. This is a consequence of the use of a Friedmann–Lemaître–Robertson–Walker metric describing the behavior of space-time in a Universe in which the Cosmological Principle holds.

1.1 A brief description of the cosmic structure

The Universe appears to have very different structure depending on the spatial scale we are observing it.
An effective way to understand this is starting from our familiar reference frame, Earth, and zooming out observing bigger and bigger volumes centered on this frame. This is what pointing our telescopes at the sky corresponds to, allowing us to reconstruct a three dimensional map of what we normally see projected on the two-dimensional sky plane.
Zooming out at a scale of about 1 Mpc, the observed cosmic volume encompasses a system of at least 35 confirmed galaxies: this is the local galactic group (Bergh 1999). As shown in *Figure 1.1*, we can distinguish well single galaxies, including our Milky Way and its famous neighbor Andromeda. Here we can notice a strong inhomogeneity in the matter distribution. Matter appears in fact to be clumped up in galaxies occupying a rather small fraction of the neighborhood volume.

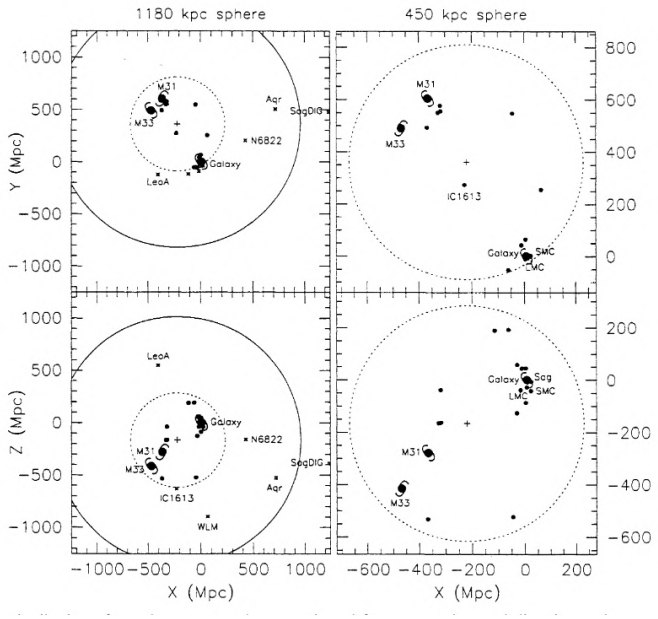
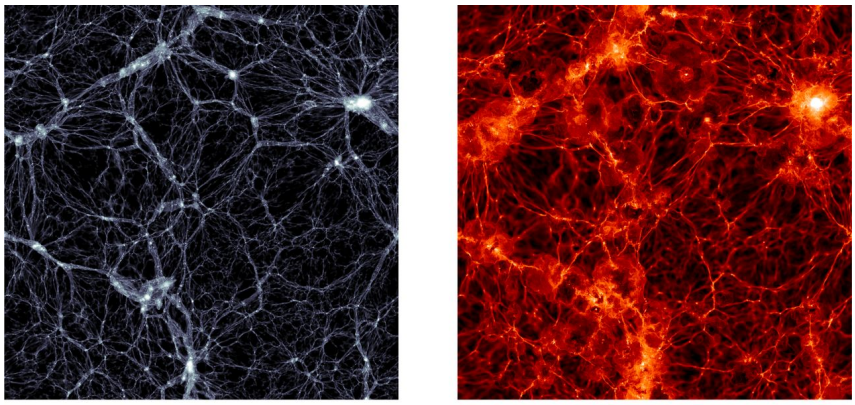


Figure 1.1. Projection on different planes of the local galactic group. Image from Bergh 1999

Zooming out further to enclose volumes of the order of few to some hundreds of megaparsecs, the Universe appears to be rather structured. Galaxies are as bits of luminous matter arranged to form walls, filaments and sheets, intertwined to regions of very low matter density called voids (Bond, Kofman, and Pogosyan 1996; Shectman et al. 1996; Libeskind et al. 2018). The ensemble of all such structures forms the so called Cosmic Web, well visible both on simulations as in *Figure 1.2* and real observations as in *Figure 1.3*.
Different structures are characterized by different densities and origins, with nodes of the web being the densest regions and acting as junction points of filaments or sheets. These regions exceptionally dense in matter are called galaxy clusters and are probably the first large structures to be formed in the Universe (Cautun et al. 2014).
Fairly close to our reference frame, at a distance of only $\sim$ 17 Mpc from us we can find the Virgo cluster (Bruno Binggeli, Tammann, and Sandage 1987), a collection of thousands of galaxies inside a volume of few Mpc in radius [2]. It represents one of the typical clusters that we can observe

[2]It is really a complex structure not completely virialized, with 6 components identified until now (Mei et al.

in the local universe, containing many different species of galaxies: from very faint dwarfs (B. Binggeli and Cameron 1991) to extremely bright giant ellipticals (Mei et al. 2007), all moving in a dynamically complex environment.
Matter distribution at these intermediate scales is still very inhomogeneous. This leads to creation of tidal forces and torques related to flows of matter that shape the structure itself. Such enormous formations originate from the non-linear collapse of density perturbations: these, initially expanding with the Universe, grow further by incorporating new matter. When they reach a critical density the volume of Universe they occupy stops expanding and starts collapsing (Eisenstein, Seo, and M. White 2007).



(a) dark matter (b) baryons

Figure 1.2. A slice of the simulated Universe from the Illustris simulation at z=0. This shows the entire simulated volume 106.5 Mpc^2 for a thickness of 104 Kpc. On the left we see the underlined dark matter structure and on the right the distribution of baryonic matter. Image from Haider et al. 2016

Going to the largest scales, encompassing volumes of many hundreds if not thousands of megaparsecs, we find ourselves in an exceptionally smooth distribution of matter. At this scale the immense structures of the Cosmic Web are not distinguishable for they are overlaid in a random pattern along the line of sight.
Many clues point to the idea of a cosmological principle manifesting at this scale. This states that matter is distributed homogeneously and isotropically, every point and every direction indistinguishable from the other (Mandolesi et al. 1986). Here galaxies are as particles in an almost perfect fluid guided and shaped by gravity. Gravitational effects can be linearized, since tidal forces, torques and higher order effects tend to average out as the density field (traced by the structures) do.
The Cosmic Microwave Background (CMB) (Dicke et al. 1965) is a strong argument in favor of a cosmological principle. It represents the moment in which light decoupled from baryonic matter, exiting the phase of thermodynamic equilibrium. Thus it shows the matter distribution at an epoch of 380,000 yrs after the big-bang. As shown in *Figure 1.4* photon (and thus matter) distribution is generally highly consistent with the cosmological principle. Recent discoveries from Planck data shows some deviations from isotropy in the CMB but their origin is still unclear (Copi et al. 2010).

2007), so it is difficult to give a precise delimitation. The gravitational center of the system seems to be close to the giant luminous galaxy M87.

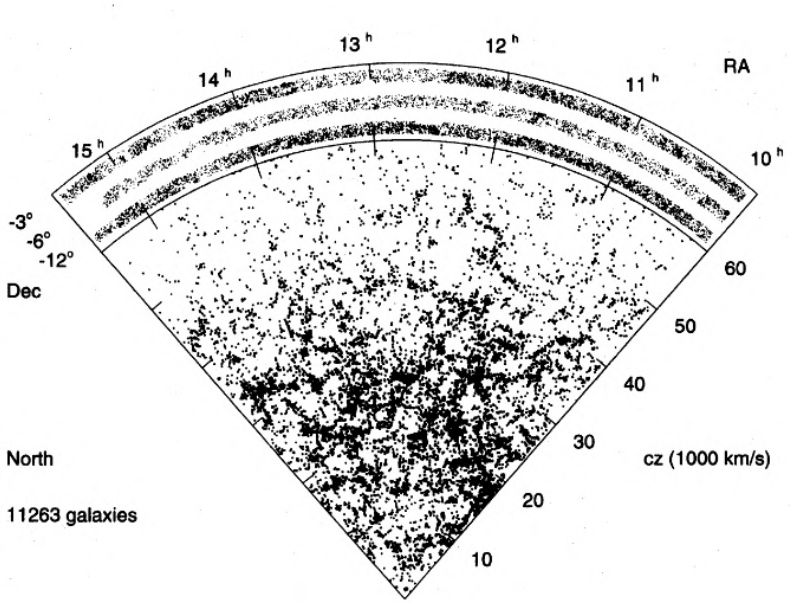


Figure 1.3. Galaxies distribution from the Las Campanas Redshift survey. Three images at different declinations but same redshift depth and right ascension are overlapped. It is evident in the large-scale structure of the local Universe. Image from Shectman et al. 1996

The discovery of Baryonic Acoustic Oscillations [3] (BAO) (Eisenstein, Zehavi, et al. 2005; Busca et al. 2013) paved the way for a direct connection between density anisotropies of the CMB and the resulting spatial distribution of galaxies in cosmic structures.

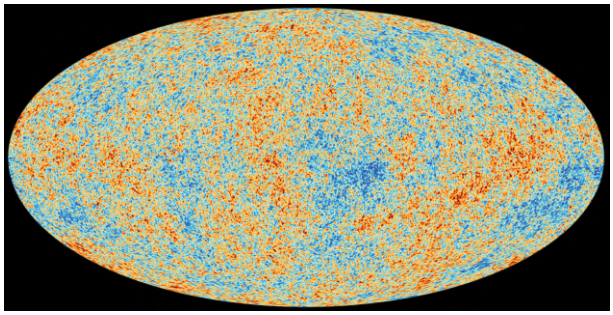


Figure 1.4. Cosmic Microwave Background from Planck satellite. The entirety of the sky is seen in one shot, showing information measured from light coming from the further we can currently observe the Universe both in space and time. The map shows the temperature distribution of the CMB and this derives from the distribution of photons at each point. The Universe here appears to be strikingly homogeneous, with temperature fluctuations of the order of $10^{-5}K$. Image from ESA.

The standard cosmological model adopted to describe the structure of the Universe is a 6 parameters model called flat Lambda Cold Dark Matter model (ΛCDM) (Aghanim et al. 2020).

[3]These are density perturbations arising from the forced oscillation of primordial plasma. This plasma contained baryonic matter and photons in thermodynamic equilibrium, plus a dark matter component. Dark matter is thought as a type of matter with low or nonexistent interactions with normal matter besides gravitational force. Thus already in this early phase is considered not in equilibrium with the plasma and can only interact with it through continuous gravitational attraction. This is countered in the plasma by the radiation pressure of photons, creating oscillations of the plasma density field. These oscillations stopped once photons decoupled from the plasma generating the CMB radiation. BAO are thus imprinted both in the CMB radiation and in the matter distribution (gas, galaxies and dark matter halos).

This allows for the existence of a cosmological constant Λ with an associated density parameter Ω_Λ, which is responsible for the acceleration of cosmic expansion dominating at redshift z<1 (Kowalski and Rubin 2008). The model is flat in the sense that is flat the geometry of space-time, ruled by the amount of energy contained in the Universe. The model assumes the existence of a dark matter component, that is matter that interacts almost exclusively with the gravitational field. This is "cold" since it must have decoupled in early times in the primordial Universe from the primordial plasma. In this way we ensure dark matter to be able to quickly lose energy and form structures by gravitational collapse even at early times. Once baryonic matter decoupled in turn from primordial plasma, it collapsed on the already-forming dark matter structure.
ΛCDM model predicts what is called a hierarchical clustering of matter (Navarro, Frenk, and S. D. M. White 1997). We can consider a fluctuation density field injected somewhat in the primordial Universe. In this scenario matter collapsed first on the peaks of the primordial density field, forming dense halos of dark matter. Then these halos merged together forming bigger structures. In a second time baryonic matter fell into dark matter halos condensing in stars and galaxies.
According to this model we can expect clusters to form in the peaks of the injected density fluctuation field from a collection of smaller sub-halos drawn to merge by gravity. Thus they must have formed as the first large scale structures, before filaments and sheets.
To have an observed counterpart for the way in which high redshift structures collapse requires precise dynamical information. Precise spectroscopy, combined with deep photometric surveys, would allow to see if clusters really form by merging of smaller structures or by other means. Moreover this would allow to measure the timescale of the formation and thus to shed light on the process of galaxy formation and evolution.

1.2 Why studying the large-scale structure

In order to confirm or exclude the standard flat ΛCDM model it is essential to perform precise observations of the large scale structure at different cosmic epochs. This requires complete and consistent surveys looking at very large patches of the sky and going very deep, to redshifts significantly greater than $z = 1$.
The spatial distribution of clusters is in fact fundamentally linked to the cosmological parameters regulating not only the geometry of the Universe, but also its evolution (Allen, Evrard, and Mantz 2011).
Galaxy clusters are in fact considered tracers of the peaks of the underlined density field which permeates the Universe and which primordial form is imprinted in the CMB. This means that they can provide information on the origin of the matter distribution and on how this is shaped through gravitational collapse.
Clusters also serve as standard candles to measure distances and thus to give information on the nature of dark energy [4] and provides insights on the nature of Universe expansion. Moreover dynamical studies of galaxy clusters provide clues for the presence of a dark matter component, but it is yet unclear if its distribution in these structures follows predictions from standard ΛCDM model (Sand et al. 2004).

[4]The elusive cause at the base of the accelerated expansion of the universe is called dark energy. This acceleration has been observed from studies on the distribution of Supernovae Ia and it is significant for redshifts lower than $z = 1$. The nature of this acceleration is not yet clear, but it is mainly regarded as a consequence of a very small non-zero vacuum energy (Copeland, Sami, and Tsujikawa 2006)

Another way in which we can use these extremely massive environments comes from the fact that they act as powerful gravitational lensing systems. This allows to use them to observe very distant galaxies unreachable from direct observation. Having a large map of galaxy clusters can thus open the possibility to have a significant number of extremely powerful natural telescopes (Smith et al. 2005) (Jauzac et al. 2021).
According to the standard cosmological model, galaxy clusters are the first structures to form in the Universe. We can then use them to study the evolution of single galaxies and to explore how this process depends on the cluster environment (Zavala et al. 2019). Here in fact we can observe how galaxies and intergalactic medium interacts providing insights to processes that can increase or quench star formation events (Casey 2016). We can also see how galaxies interact between themselves both through tidal forces, stellar winds and relativistic jets.
So having a complete and statistically consistent catalog of galaxy clusters can pave the way for new discoveries and constraints on different aspects of the Universe. These include formation and evolution of both the largest-scale structures and of single galaxies but also characterization of the gas dynamics, thus how this interacts with galaxies and stars.
In this frame the Southern H-ATLAS Regions Ks-band Survey (SHARKS) plays an important role, providing the possibility to construct a large catalog of galaxy clusters observed in IR light. The use of IR observation ensures that it is not biased by those dust-rich galaxies that would result invisible in optical surveys. The resulting catalog of candidate clusters will cover a redshift range from $z \sim 0$ to $z \sim 3$ and a sky area of $\sim$300 square degrees, possibly delivering hundreds of candidate high-density structures all within a consistent observation.

The structure of this work is divided as such:
in *Chapter 1* I present the characteristics of clusters and proto-clusters and the different methods used to find them in wide surveys.
Chapter 2 describes the SHARKS survey and the content of the first data release.
In *Chapter 3* the process of catalog cleaning is described, starting from the removal of flux-contaminated sources, then calculating the catalog completeness and finally removing the foreground stars and obtaining a complete catalog of galaxies.
Chapter 4 describes the methodology used to analyze the cleaned catalog and focus on the characterization of the algorithm used to search for the presence of candidate clusters in the catalog. Some preliminary statistics is presented here underlining advantages and disadvantages of using this type of algorithm.
In *Chapter 5* I describe an improved method for the cluster search. This combines multi-band photometric information to select the galaxy red sequence at different redshift intervals. Selecting galaxies in this way allows for enhanced contrast between the number density of galaxies in a cluster and in the field. This in turn increases the accuracy of the search algorithm.
In *Chapter 7* a summary of the analysis process is presented with its results. Here I try to show possible interpretations of the results together with the known problems and the possible future solutions. Attention is drawn to some possible implementations of this method of analysis and to the many applications that SHARKS could provide to the study of galaxy evolution.
A list of the most interesting candidate clusters found in the analysis, together with a summary of their characteristics can be found in the Appendix.

.

1.3 What are clusters

Figure 1.5. A classic example of the different matter components forming galaxy clusters. In the image two clusters are observed after a head-on collision in which the Bullet Cluster (on the right) passed through a larger structure. The galaxy component of both clusters is visible in the optical observation from Hubble Space Telescope. The Intra Cluster Medium component, instead, being constituted of diffuse ionized gas, interacted during the encounter and thus was stripped out from both structures. This material heated up in the process emitting strong fluxes in X-ray, here traced in pink from Chandra space telescope observations. The total matter distribution (in blue) is mapped using the gravitational lensing effect exerted by the cluster on background sources. Measurements show that the mass of the hot gas component is of the same order of magnitude as the luminous mass of stars in the galaxy component. Dark matter, however, is the dominant mass component in the system. Image from ESA.

Galaxy clusters represent the densest environments in the local Universe. They are majestic structures containing hundreds of galaxies packed in volumes of only a few cubic megaparsecs. This is rather in contrast with the field environment, constituted by a diffuse component of galaxies that are not arranged (at the same scale of clusters) in gravitationally bound structures.
From simulations it appears that the largest clusters tend to be connected by cosmic filaments, which feed matter to them through gravitational accretion (Kuchner et al. 2020). These volumes are usually virialized or near virialization, meaning that their kinetic and potential energy balance in a way that they do not contract or expand. The radius which comprises the virialized part of the structure is the so-called r200. It is defined as the radius including an energy density 200 times the critical one (ρ_c) required to have a geometrically flat Universe (Hansen et al. 2005).
Galaxies in a cluster are usually characterized by a very small amount of star formation, especially along the central region of the structure, spanning a volume of radius usually a fraction of Mpc called core.
It seems that galaxy clusters start to become virialized around $z \sim 1$. At $z \leq 1$ galaxies in clusters appear to have on average very low star formation rates[5], while at $z \geq 1$ star formation rates tend to increase reaching values order of magnitudes greater than in the field (Madau and Dickinson

[5]Star formation rate is defined as the amount of star produced per year. It is measured in units of $[\frac{M_\odot}{yr}]$. If this quantity is weighted by volume of comoving space, it is called star formation rate density.

2014).
There is still a strong debate on the reality and magnitude of this inversion. This in fact can be due to biases in the observational methods, to poor statistical sampling, or to lack of complete observations at higher redshift instead of having physical meaning(Santos et al. 2015).
There are typically three main components we can identify in a cluster.

- An **halo** of dark matter that embodies the core region and accounts for the majority of the cluster's mass (of the order of $\sim 10^{14} M_{\odot}$). Secondary halos may or may not be present and are typically visible only through precise spectroscopic analysis. These halos have much smaller masses and tend to fall into the potential well of the main halo, merging with it.

- A large amount of **diffuse** and extremely **hot gas** reaching densities of 10^{-4}cm^{-3} and temperatures of T $10^{8} K$.
  This ionized gas is responsible for strong X-ray emissions and is called Intra Cluster Medium (ICM).
  Such high temperatures are reached probably through different mechanisms like tidal forces, friction between ICM and galaxies, collective stellar winds or supernovae shock waves.
  The nature and efficiency of the cooling mechanisms of the medium are also unclear and may depend on the considered region (Peterson and Fabian 2006).
  The core of the cluster, where the gas density is very high, may experience strong cooling flows combined with efficient radiative cooling by emission lines.
  In the outskirts, instead, due to the much lower plasma density, cooling seems to be inefficient, driven mainly by continuous emission.

- A luminous component of **galaxies** of different morphology.
  Inside the cluster core, the innermost and densest region of the structure, we usually find early type galaxies, with little to no content of dust and practically no star formation.
  Among them it is possible to find giant elliptical galaxies of extremely high luminosity. These are characterized by brilliant halos of stars stripped from other galaxies by means of strong tidal forces.
  Going further radially from the core, the fraction of late type galaxies increases (Dressler 1980, e.g.). Galaxies seem to generally have stable orbits inside the cluster, having time to evolve during each crossing of its volume.

1.4 What are proto-clusters

Proto-clusters are generally considered the progenitors of the galaxy clusters we can find in the local Universe.
Literature generally refers to proto-clusters when talking about structures that are found at $z \gtrsim 1.5$. However, an ultimate definition for these objects is still to be developed due to the lack of a sufficiently large statistical sample of detailed observations.
The leading characteristic distinguishing local clusters from distant proto-clusters is in the morphology. Differently from their local counterpart, in fact, proto-clusters are not virialized structures, being instead still in a phase of gravitational collapse. For this reason their density is much lower and their spatial extension is way bigger than in more recent structures. Thus we can expect the absence of a main halo, observing instead a structure which is an ensemble of smaller sub-halos falling in the common gravitational potential well.

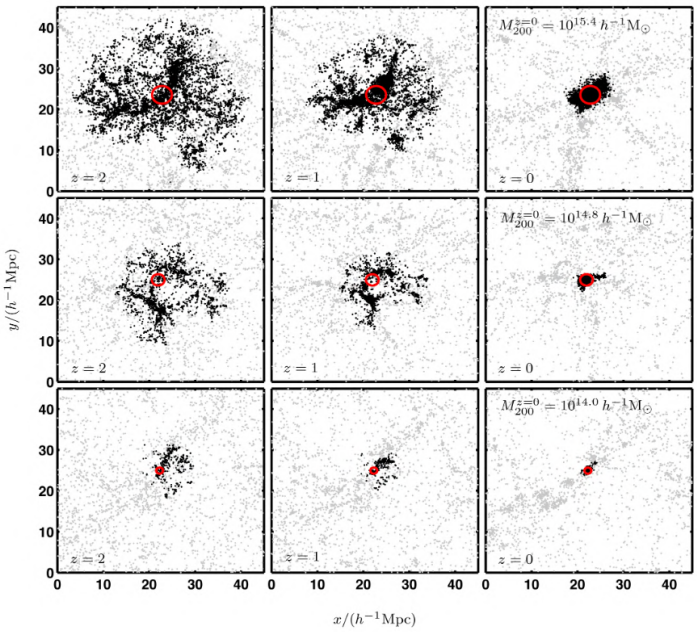


Figure 1.6. Evolution of proto-clusters of three different final masses. At each mass (changing from top to bottom) proto-clusters are shown in different phases of their evolution, respectively at redshift z=2, z=1, z=0 from left to right. The red circle represents the final virial radius reached by the structure. Each box has a fixed comoving size of 45 Mpc/h (Muldrew, Hatch, and Cooke 2015).

It is not clear what are the conditions that allow a proto-cluster to become a massive cluster at z=0. These structures can in fact be so sparse that the sub-halos system does not have time to merge completely before reaching our cosmic epoch. As a result, it is not obvious that a denser environment will form in the future, since the action of accelerating cosmological expansion must be taken into account.

Figure 1.6 shows the typical evolution of a proto-cluster (Muldrew, Hatch, and Cooke 2015) for different choices of the mass it will reach at the present time. Fixed the mass, the younger the structure the larger and rarefied it is. As the Universe evolves, new matter is attracted to the proto-cluster, increasing its mass and steepening the potential well. This, in turn, causes a larger tendency of matter to collapse, increasing the density. This is observed largely around $z \sim 1$. However the most dense structures can collapse at higher redshifts, while less dense structures collapse at $z < 1$ or have yet to collapse. At more distant cosmic times, instead, proto-clusters tend to expand with the Universe, their size being thus mostly independent from redshift, but strongly dependent on the mass they will reach at z=0.

These structures are exceptionally elusive since, due to low number density of galaxies, they tend to be easily confused with field and background galaxies. Another difficulty for the detection of proto-clusters resides in the large amount of dust and cold gas that they contain, which heavily obscures any optical radiation. This seems connected to a large amount of star formation, that makes galaxies residing in such ancient structures much different from what we find in their local descendants. Often these star formation events appear to be quick, happening in order of millions of years and are thus called starbursts (Vieira et al. 2013).

Since proto-clusters do not have reached virialization yet, they also have not faced relaxation processes yet. These are phases in which the cluster collapses and re-expands in a violent and rapid way until it reaches an equilibrium point. Such a violent relaxation is apparently triggered when the structure reaches a critical mass and its expansion stops. This process of rapid mass redistribution contributes to the heating up of IGM and thus proto-clusters do not show the hot diffuse baryonic component emitting in X-ray typical of clusters.

1.5 How to catch giants: method for (proto-) clusters detection

Identification methods must be different for clusters and proto-clusters since they have very different properties. Here I present some of the many methods used to find them in wide surveys.

Spectroscopic surveys: they are the most accurate way to find clusters, tracing almost directly their three-dimensional structure, but require an enormous amount of observational time. Moreover, they are strongly subjected to deformation effects related to the peculiar velocity field of galaxies.

X-ray: Using space-based observatories it is possible to look for the X-ray emission from hot intra-cluster medium (Böhringer et al. 2004).

Sunyaev-Zel'dovich (SZ) effect: hot electrons from ICM scatter CMB photons transferring energy to them. The effect generates a deformation of the black body distribution of photons in the CMB, decreasing the number of low energy photons and increasing the number of high energy ones. So the method consists of observation of the CMB, searching for anomalies in the almost perfect Planck distribution of photons (Bleem et al. 2015).

However, both X-ray and SZ methods are not efficient when looking for distant clusters since they

tend to lack an hot ICM component.
A more efficient method to find proto-clusters consists in identifying peculiar galaxies that tend to reside largely in dense environments.
Lyman-break galaxies: These are galaxies presenting a strong drop in observed flux $\lambda < 912$Å. This is caused by absorption of UV photons by ionized gas clouds around young stars in the galaxy itself and by large assemblies of neutral hydrogen along the line of sight (Ouchi et al. 2004).
It is not necessary to use spectroscopy to identify such galaxies, but it is sufficient to use photometric observations in at least 3 wide bands. The bluer filter of the trio and the intermediate one are close in frequency and are used to find the spectral drop. The redder filter is used to ensure that the identified drop is in the spectral region in which we expect the Lyman forest to manifest.
Lyman-Alpha galaxies: Another kind of peculiar galaxies are Lyman-Alpha (Ly-α) emitters. These objects presents a bright rest frame UV emission ($\lambda = 1215.67$Å) due to intense ongoing star formation.
This search requires narrow-band filters in order to identify the emission line. Again, both methods are not really effective for ancient proto-clusters, since galaxies in these environment can be so rich in cold gas and dust to be almost totally obscured in the short wavelengths required.
H alpha emitters: The Balmer H_α line ($\lambda = 6563$Å) is less affected by dust absorption than Ly_α so can be used to select strongly star forming galaxies in dense environments.
Photometric surveys are able to go much deeper in redshift than spectroscopic observations in the same area, since they require less observational time.
Red sequence selection: local clusters are observed to contain larger fraction of red sequence galaxies [6] than the field. Thus it is possible to use multi-band photometric surveys to search for clustering of red sequence galaxies tracing galaxy clusters (Muzzin et al. 2008). **Single band photometric surveys:** The search for clusters in a single band survey is performed by algorithms that look for regions where the number density of galaxies is significantly higher than the average of the survey.
The downside of these methods is that the three dimensional structure of the universe is collapsed in a two dimensional image erasing distance information and thus blending together all structures along the line of sight.

[6] Galaxies in color-magnitude diagrams tend to arrange themselves in two regions: red sequence and blue cloud. The former is characterized by galaxies with redder colors and are usually associated with bright early type galaxies. The latter contains galaxies with bluer color that are typically late type (Strateva et al. 2001).

Chapter 2

The SHARKS survey

Figure 2.1. Logo of the SHARKS survey from http://research.iac.es/proyecto/sharks/pages/en/home.php

The Southern H-ATLAS Regions Ks-band Survey (SHARKS) is an European Southern Observatory (ESO) public survey performed with the ground telescope VISTA. It covers 300 squared degrees in the sky comprising all equatorial and southern Galaxy Mass Assembly (GAMA)[1] fields. This for exception of GAMA19, which is a field particularly contaminated by Galactic cirrus emission.
SHARKS is performed using a single photometric band (Ks) centered at ($\lambda = 2.2\mu m$). It was granted 1200h of observing time, allowing it to reach a magnitude limit of ∼22.7 AB magnitude. SHARKS was granted 300h observational time in the equatorial region and 900h in the larger southern region, which is situated in the South Galactic Pole (SGP) region (Saavedra Esquivel et al. 2021).

2.1 The SHARKS science case

The project has three key goals:

[1]GAMA is a project involving observations with multiple instruments both from ground and from space (Driver et al. 2011).

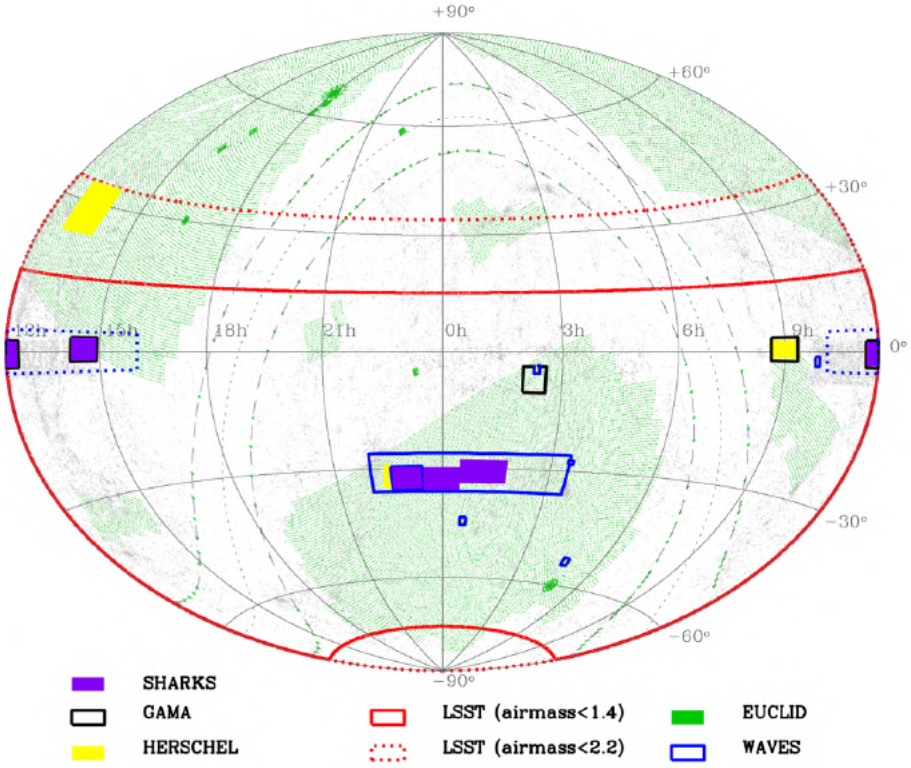


Figure 2.2. SHARKS coverage on the sky plane from ESO SHARKS management plan https://www.eso.org/sci/observing/PublicSurveys/docs/SHARKS_SMP_10012017.pdf

- To provide counterpart high resolution observations for ∼90% of infrared and sub-millimeter sources detected up to redshift z 3 by previous observations (Hershel atlas (H-ATLAS), South Kilometer Array (SKA), Australian SKA Pathfinder Telescope ASKAP, LOw Frequency Array (LOFAR)).
- To produce a map of strong gravitational lenses in a fairly large region of the sky.
- To provide a large catalog of galaxies in order to study the evolution of cosmic structures.

To fulfill these objectives it is clearly necessary to observe over a large patch of the sky and simultaneously reach a fairly high depth in magnitude.
The survey area and position were chosen mainly to cover a large portion of the Herschel Astrophysical Terahertz Large Area Survey (H-ATLAS). H-ATLAS is a ∼600 square degrees survey performed with HERSHEL space observatory with the aim of providing information on dust masses and on star formation rates for a large number of galaxies. The survey covers the sub-millimetric spectral range with photometric bands centered at 100 μm, 160 μm, 250 μm, 350 μm, and 500 μm (Valiante et al. 2016). The observed area covers regions observed with the Sloan Digital Sky Survey (SDSS), the 2-Degree-Field Galaxy Redshift Survey (2dF), the GAlaxy Mass Assembly project (GAMA), the Wide-field Infrared Survey Explorer (WISE), the VISTA Kilo-Degree Infrared Galaxy Survey VIKING and the VST Kilo-Degree Survey (KIDS).
Ks band was established to be the most performing photometric band for observation of infrared H-ATLAS sources counterparts and radio sources counterparts.
Furthermore Ks band is situated in the last infrared window observable from Earth. The far infrared region of the spectrum is in fact unreachable from ground due to absorption by water molecules and CO_2 molecule in the atmosphere.
We can study the rest frame optical properties of very distant galaxies using observations on the infrared region of the spectrum due to cosmological redshift.
In the particular case of Ks observations a further advantage comes from the boost effect due to K-correction [2]. This depends on the slope of the galaxy continuum, that typically decreases in its rest frame IR region and thus it receives a positive contribution from the correction.
Stellar population studies and star formation models tell us that the dominant portion of star by number and by mass is constituted by dwarf stars. These, with masses lower than the Sun, have low effective temperature and so, due to Wien's law, their quasi black body emission peaks in the NIR around ∼1.6μm. The Ks band, centered at 2.2μm is then able to observe the emission peak of low mass stars for galaxies at $z \sim 0.4$. The full width at half maximum (FWHM)[3] of the response function[4] for VIRcam Ks photometric filter is 3075.13Å, meaning that the band includes approximately a wavelength range of $1.9\mu\text{m} \leq \lambda \leq 2.5\mu\text{m}$. The very peak at ∼1.6$\mu$m is then detected in the redshift range $0.2 \leq z \leq 0.6$. However we do not need to be precisely in the peak

[2]The frequency of radiation coming from a distant galaxy is lowered by effect of cosmic expansion. This means that if we observe the infrared portion of the spectrum of a galaxy at $z \sim 2$, we are really observing the optical region of its rest frame spectrum. This, however, is modulated by the instrumental function comprising all information about telescope, camera and IR filter. To know the real rest frame emission of the galaxy it is necessary to have a theoretical model of its spectrum, the instrumental function and a cosmological model describing the effect of cosmological redshift. Combining all this information we can, starting from the observed portion of a galaxy's spectrum, derive its rest frame spectrum.

[3]This is the width of a distribution calculated at half of its maximum amplitude

[4]Here for response function we indicate the function modulating the output signal of the instrumentation. This function includes information, among other things, on the optics of the telescope and of the mounted camera, on the filters and on the sensitivity of the CCD.

to measure the mass, the intensity of the peak can be estimated from points of the distribution close enough to λ ∼1.6 Å. For this reason the Ks band is fundamental to measure the luminous mass component of galaxies in a quite broad redshift interval.
The total exposure time was chosen to reach a photometric depth of ∼ 22.7 AB magnitude. This was calculated as the minimum depth required to observe a complete sampling of H-ATLAS galaxies up to $z \sim 3$.

2.1.1 On the way to EUCLID

The SHARKS survey does not serve just as a follow up observation of past missions, but was designed with the intent of providing a considerable amount of complementary data for future generation telescopes.
A list of the surveys that had covered or will cover part of SHARKS area can be found in *Table 2.1*.
Among these, the Euclid space survey is of particular interest since it will deliver NIR spectroscopy in a range between 0.92 μm and 1.85 μm. This will provide fundamental information for the characterization of many SHARKS sources.
In the context of this work, Euclid spectroscopic data will serve as a way to confirm or reject each candidate cluster and proto-cluster found in the SHARKS survey.
Subsequently its data will be used to study different characteristics of the confirmed candidates, for example constraining mass and star formation rate.
Euclid will in fact provide both photometric and spectroscopic measurements thanks to the instrumentation it is equipped with.
The VISible Instrument (VIS) is a CCD camera covering a photometric band from 550 nm to 900 nm, with resolution of 0.23 arcsec.
VIS will be able to do precision photometry to a depth of about 25 mag at 10σ, so it will provide complementary photometric data for each SHARKS source.
A near infrared camera is available in the Near Infrared Spectrometer and Photometer (NISP) instrument. NISP will have 3 photometric filters: *Y (900-1192 nm)*, *J (1192-1544 nm)* and *H (1544–2000 nm)* with a resolution of ∼ 0.3 arcsec.
The spectrometer in NISP is slitless and contains 4 low resolution (R=380 for a 0.5 arcsecond diameter source) grisms covering the two ranges (1250 nm - 1850 nm) and (920 nm - 1250 nm).
Despite the low resolution, NISP will be able to observe fundamental spectral features as H_α line, which can be used to constrain star formation rate. According to the NISP spectroscopic range, the H_α would be visible in the range $0.4 \leq z \leq 0.9$. However there are techniques that are able to detect the presence of H_α line in photometric observations, leading to its detection with NISP up to z∼2.
Data from SHARKS will serve as ancillary information for the Euclid space mission.

2.2 The VISTA instrumentation

VISTA (Will Sutherland et al. 2015; Emerson, McPherson, and W. Sutherland 2006) is a 4 meters ESO telescope located at Cerro Paranal Observatory in Chile.
The telescope is characterized by a wide field for observation of the southern sky, coupled with

Survey	Frequency	Overlapping Area	Depth (5σ)	Status	Telescope
H-ATLAS	100-500 μm	SGP and GAMA	$\sim$ 30 mJy at 250 μm	E	Herschel
EMU	$\sim$ 1.3 GHz	SGP and GAMA	$\sim$ 50 μJy at 20 cm	F	ASKAP
LOFAR Tier 2	$\sim$ 150 MHz	GAMA	$\sim$ 0.1 mJy at 150 MHz	E	LOFAR
WALLABY	HI survey	SGP	MHI >108M	E	ASKAP
DINGO	HI survey	SGP and GAMA	MHI >107M	E	ASKAP
GALEX MIS	UV phot.	SGP and GAMA	$\sim$ 22.7 AB mag	E	GALEX
Deep-WAVES	optical spec.	SGP	r <22 ABmag	F	VISTA
WEAVE	optical spec.	in GAMA		F	WHT
LSST	u,g,R,I,z phot.	SGP and GAMA	$\sim$ 26-27 AB mag	F	Rubin Obs.
EUCLID	Y, J, H phot.	SGP	$\sim$ 24 AB mag	F	EUCLID

Table 2.1. An overview of the main surveys that have covered or will cover part of the SHARKS area. The *Status* column expresses if the survey is already existent (E) or if it is scheduled for the future (F).

an adaptive optic system that brings the point spread function to be PSF$\sim$0.5 arcsec[5] for an expected average seeing of $\sim$0.8.
The mirror operates on an azimuth-altitude mount and with a quasi-Ritchey-Chretien configuration.

Figure 2.3. The VISTA telescope facility.

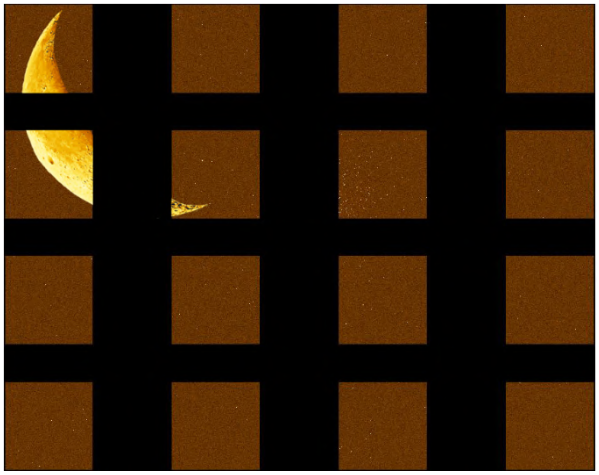


Figure 2.4. Simulated pawprint from VIRCAM. Black spaces are caused by the separation between CCDs in the mosaic from the VISTA consortioum https://vista.maths.qmul.ac.uk/index.html.

The telescope is equipped with Vista InfraRed CAMera (VIRCAM) (Dalton et al. 2006) a multi-band near-infrared camera covering Z,Y,J,H,Ks broad photometric bands and a 1.18 micron narrow band.
The diameter of one VIRcam FOV is 1.65 degrees. To cover such a large field of view VIRcam's CCD is not a single solid piece of semiconductor, but a mosaic of 16 plates.
This means that the gaps between plates will result in blind stripes on the final image as can be

[5]As a comparison this is similar to the resolution that Euclid's NISP camera can reach in NIR from space.

seen in *Figure 2.4*.
For this reason the dithering method was used during observation for the SHARKS survey.
This consists in observing the same patch of sky repeatedly shifting each time the image center. Images are then stacked and, due to their different centers, the blind spots of one image are covered by the other images.
One single pointing of VIRcam is called a pawprint. The minimum number of pawprints required to completely cover the field of view is 6, that, when stacked together, form a tile.
One tile is constructed by observing the same sky at three slightly different positions (shifted by half of a detector) along the Y direction and for each of them observing two times (shifted by 0.9 of a detector) along the X direction. A schematic view of the dithering method can be found in *Figure 2.5*
Each tile is observed for a total of ∼7.8 min. A mosaic is formed by summing 7 tiles observed at the same position for a total observation time of ∼55 min.

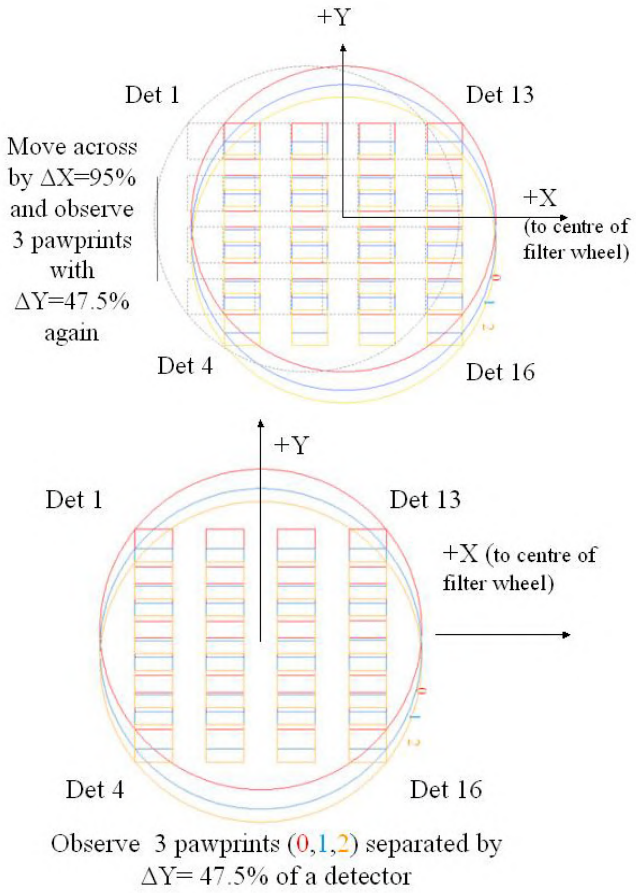


Figure 2.5. VISTA pawprint from the VISTA consortioum web page https://vista.maths.qmul.ac.uk/index.html

2.3 First data release: dr1

SHARKS first Data Release (dr1) is comprehensive of 10 mosaics each of ∼ 2 squared degrees for a total of ∼20deg^2 over the ∼300deg^2 of the complete survey. The ten mosaics sample both equatorial and southern regions of GAMA: two mosaics for each GAMA12 and GAMA15 fields, and 6 mosaics in the SGP region. This ensures to have an overview, although minimal, of the

entire field.
Of the 6 SGP mosaics, 4 are contiguous and slightly overlapped. They cover part of the Dark Energy Survey (DES) area, which has a flux depth of ∼24 AB magnitudes and uses 5 different photometric bands (g,r,i,z,Y) complementary to SHARKS Ks.
Dr1 consists of a series of pre-reduced image files and calibrated catalogs. The images were acquired over 141 nights between 3 March 2017 and 18 January 2019. The criteria for observation are: airmass[6] <1.7 and seeing < 1.2" for SGP images and <1" for GAMA images. Images are still considered good if these parameters fall within a 20% threshold from these optimal limits. Photometric contamination from the presence of the moon is negligible in Ks band.
Images are pre-reduced and corrected for sky background. Photometric calibration is performed with respect to 2MASS [7] in the AB magnitude system [8]. Each mosaic is calibrated independently. Catalogs are constructed from the calibrated images using the source extractor algorithm SExtractor. Flux is calculated inside 13 photometric apertures for each source, listed in *Table 2.2*. Kron [9] and Petrosian [10] magnitudes are also present for extended sources. Pixel size corresponds to 0.34″.
Astrometric calibration is also performed using 2MASS, but it is planned to add much more precise information on position from the GAIA space observatory. The effective average seeing effect on the corrected and calibrated images is around 1″, larger than the average seeing affecting raw images. This is caused by the image reduction process generating a degradation of the image up to 15%. Background removal, in fact, was performed prioritizing completeness over purity and image quality.
As it will be described in *Chapter 3* completeness is of ∼22.0 mag, varying slightly with catalogs. This is close to the calculated 5σ flux depth limit of 22.7 mag.

[6]In astronomy it is a measure of the amount of air along the line of sight and it is a function of the zenith angle, increasing towards the horizon. Higher its value, stronger are the effects of absorption and refraction.

[7]Two Micron All Sky Survey is a photometric survey observing the whole sky in J, H and Ks band between 1997 and 2001. It thus covered the spectral region from 1.2μm to 2.2μm using two 1.3 m telescopes for a photometric depth of 14 magnitudes.

[8]In AB (for *ABsolute*) system the zero point for flux-magnitude conversion is independent on the observed frequency.

[9]Kron radius (R_k) is obtained as the first moment of the surface brightness light profile. It is found empirically that a choice of an aperture of $2.5R_k$ would encompass > 90% of a galaxy flux (Kron 1980).

[10]Given a radius r′ from the center of an extended source, the ratio between the local surface brightness inside an annulus at r′ and the mean surface brightness within r′ shows how much the total flux increases with the aperture. Petrosian radius is the value of r′ at which the ratio stops changing significantly with r′.

	Diameter [arcseconds]
Aper 1	1″
Aper 2	1.4″
Aper 3	2″
Aper 4	2.8″
Aper 5	4″
Aper 6	5.7″
Aper 7	8″
Aper 8	10″
Aper 9	12″
Aper 10	14″
Aper 11	16″
Aper 12	20″
Aper 13	24″

Table 2.2. The size of each photometric aperture as angular diameter is units of arcseconds.

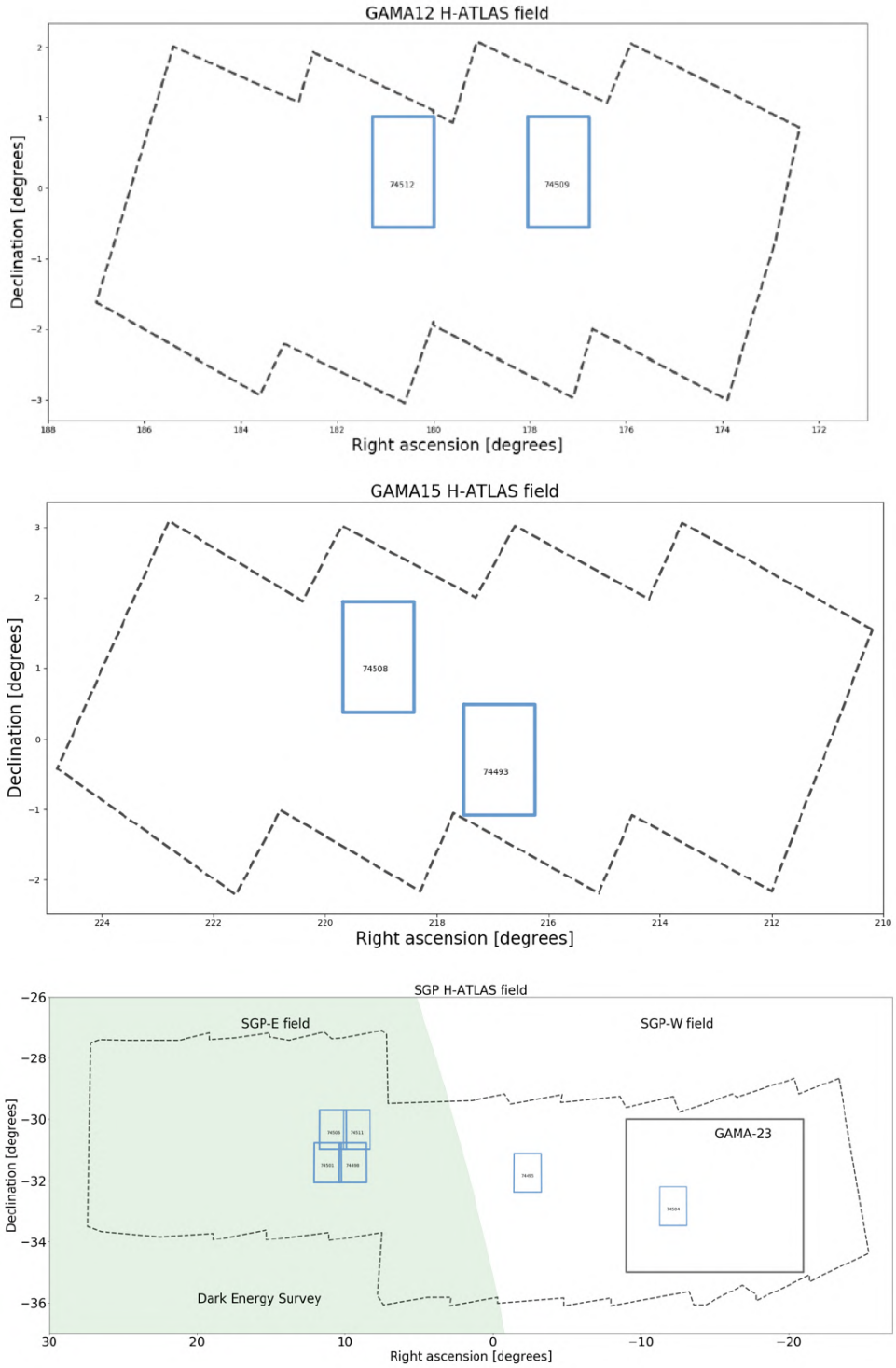


Figure 2.6. SHARKS coverage (in *Blue*) confronted with GAMA fields (GAMA-23 in *Black* and SGP in *dashed*) and DES field (the *Green* area) from SHARKS dr1 web page http://research.iac.es/proyecto/sharks/pages/en/data-releases/dr1.php

Chapter 3

Cleaning SHARKS catalog

3.1 Spurious sources

Catalogs from SHARKS dr1 contain flags for each source stating the quality of the corresponding data. Here I describe these flags and how they were applied to the dr1 for the removal of the spurious sources. I also provide a description of the problems of spurious sources removal using these flags and, in the section that follows, how I have constructed an automatic algorithm to help solving the problem
. Dr1 flags are derived from the SExtractor output ERRBITS. Flags are integers generated by a binary number for which each bit is associated to one parameter. A parameter can be True or False for a given source, corresponding respectively to a 1 or 0 in the bit position. Quality parameters identify spurious sources generated by errors in the application of SExtractor or issues on the photometry due to bad pixels or flux contamination. A list of the parameters present in ERRBIT is shown in *Table 3.1.*
Selection of spurious sources, however, is not efficient in removing artifacts caused by diffraction effects due to saturated objects. If a source is exceptionally bright, in fact, multiple reflections in the optical apparatus generates visible patterns in the image.
I divided such patterns into 3 main typologies, highlighted in *Figure 3.1.* Many other effects exist, as described in Paillassa, Bertin, and Bouy 2020, however they are generally well treated by SExtractor. The three artifacts treated in this work are:

- **Halo:** large deviation from the point-like behavior of a star or star-like source. It is a bright region enlarging the angular diameter of the object.
- **Ghost:** is typically composed of multiple nested faint disks with shifted centers and bright borders. The position of the ghost's center relative to the star depends on the relative position of the ghost center with respect to the star.
- **Spikes:** are narrow lines of light departing from a source in a fixed pattern that depends on the telescope. They originate from light diffraction happening on the structure that supports the secondary mirror.

In the case of infrared observations, it is expected such artifacts to be much less prominent and present in negligible number with respect to optical observations.
Spurious sources in the published catalog are shown in red, together with non-spurious ones (highlighted in green), in *Figure 3.2.* Such sources are identified by the ERRBITS flag corresponding

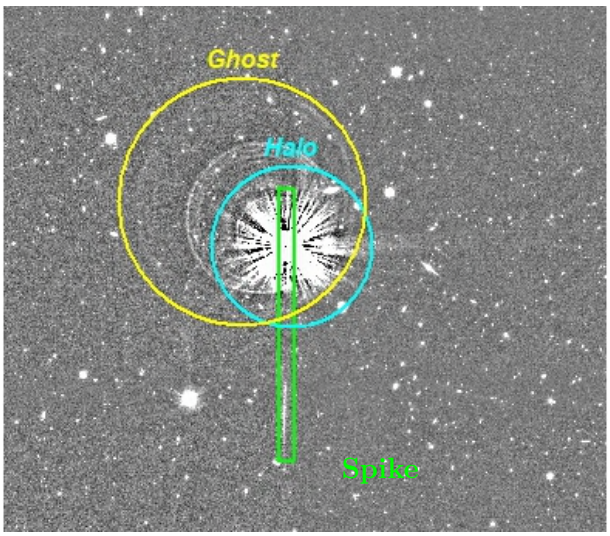


Figure 3.1. A typical example of diffraction artifacts due to a saturated star SHARKS. The star comes from dr1 reduced image 1000000074511.

to flux contamination or pixel corruption. It is possible to see how a non-negligible amount of sources follows diffraction patterns around a saturated star without being identified as spurious. Such sources are by nature not randomly arranged, but follows patterns that could be identified as cosmic structures by the algorithm used for cluster search.
It was decided then to completely remove any source that resides inside regions dominated by halo's or ghost's light. This can be done without affecting the catalog completeness too much. Saturated stars, in fact, are not frequent, with an average of ∼3 for squared degrees within the ∼10deg^2 of dr1.

Value	Meaning
1	aperture photometry is likely to be biased by neighboring sources or by more than 10% of bad pixels in any aperture
2	the object has been deblended
4	at least one object pixel is saturated
8	the isophotal footprint of the detected object is truncated (too close to an image boundary)
16	at least one photometric aperture is incomplete or corrupted (hitting buffer or memory limits)
32	the isophotal footprint is incomplete or corrupted (hitting buffer or memory limits)
64	a memory overflow occurred during deblending
128	a memory overflow occurred during extraction

Table 3.1. List of parameters that constitute SExtractor's flags.

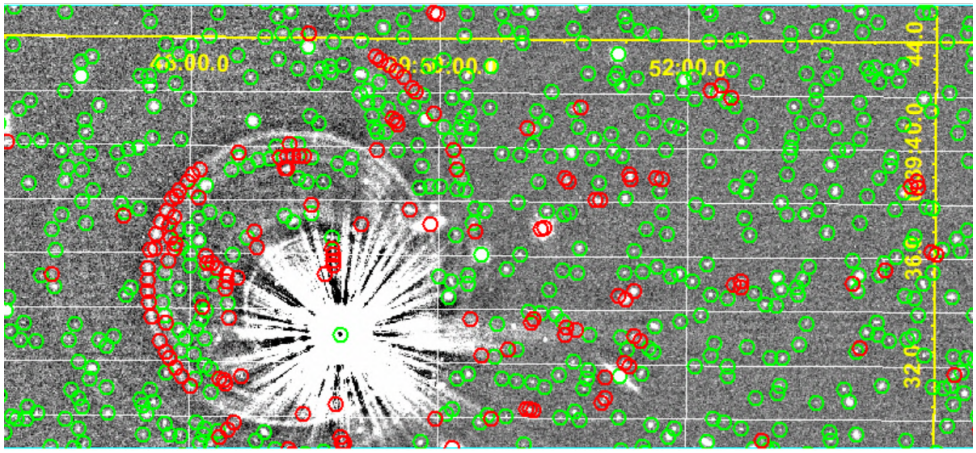


Figure 3.2. Diffraction effect on a saturated star. In *green* are shown non-spurious sources from the SHARKS catalog. Spurious sources selected by the SExtractor flag ERRBITS are highlighted in *red*. Many catalog non-spurious sources are clustered along the borders of artifacts in a clearly visible pattern. This could result in identification of candidate clusters among these clustered fake sources.

3.1.1 Automatic identification of diffraction patterns

Here I describe the algorithm I have constructed based on the work of (Estrada et al., in preparation; Amata Mercurio and Team 2018) to identify and automatically remove diffraction artifacts from SHARKS catalog.

For this task I developed a python code to automatically look for saturated stars in SHARKS catalogs and remove sources residing within the corresponding halos and ghosts. Saturated stars that generate diffraction artifacts will be hereafter called *parent stars*. The code uses empirical linear relations between magnitude and position of saturated stars to find position and size of the corresponding halos and ghosts.
Calibration of the linear relations was performed using the 5 images listed in *Table 3.3*. A similar simple relation for spikes was not adopted, since the vast majority of spurious sources around spikes are well identified by the SExtractor ERRBITS flag.
To construct parametric relations independently on SHARKS photometry, it was decided to make use of GAIA stellar catalogs. Halos and ghosts were selected manually from the 5 calibration images and the corresponding parent stars were matched to Gaia[1] dr2 catalog.
Then relations are constructed using the position of Gaia sources in pixels and fluxes from Gaia's photometric G band in magnitudes. Positions and distances were calculated in pixel space so to avoid distortion effects due to projecting spherical coordinates to flat space.

- **Magnitude-size for halos:** The position of the center of a halo is considered to coincide with the center of the parent star (the star generating the diffraction pattern). The linear relation between the stellar magnitude in Gaia G band and the halo size in SHARKS image is plotted in *Figure 3.3*. Given the magnitude of the parent star, the radius of the halo (R_h) is:

  $$R_h = b_1 \times magnitude + b_2$$

  . Only stars with magnitude < 12 mag are considered to form large halos. Coefficients b_1 and b_2 are listed in *Table 3.2*.

- **Magnitude-size for ghosts:** As for halos, ghosts are visually selected and radius and position are traced by hand. A plot of the magnitude-size relation can be found in *Figure 3.4*. Given the magnitude of the parent star, the radius of the ghost (R_g) is:

  $$R_g = a_1 \times magnitude + a_2$$

  . Only sources with magnitude < 10 mag are empirically considered to form bright ghosts. Coefficients a_1 and a_2 are listed in *Table 3.2*.

- **Ghost position:** Ghosts are not generally centered on the parent star. Instead, they tend to be radially shifted from the optical axis (which usually coincide with the center of the image) outward. The more distant the star is from the optical axis, the larger is the distance between the ghost center and the star.
  A linear relation can thus be found between the relative ghost-star and the star-center distance. Star and ghost are assumed to be aligned along the radial direction.
  SHARKS' reduced images are, however, constructed by stacking at least 7 multiple compositions of 6 dithered images. This means that there are 6 optical axes on the final image, all misaligned. Each image will generate its own ghost pattern, based on the relative position of sources with respect to its own center. Thus, with the stacking process, a star should end

[1]Gaia is an ESA space mission with the aim of constructing a 3D map of the Galaxy using parallaxes and radial velocity observations. It will measure precisely positions for ∼1% of the Galaxy's population.

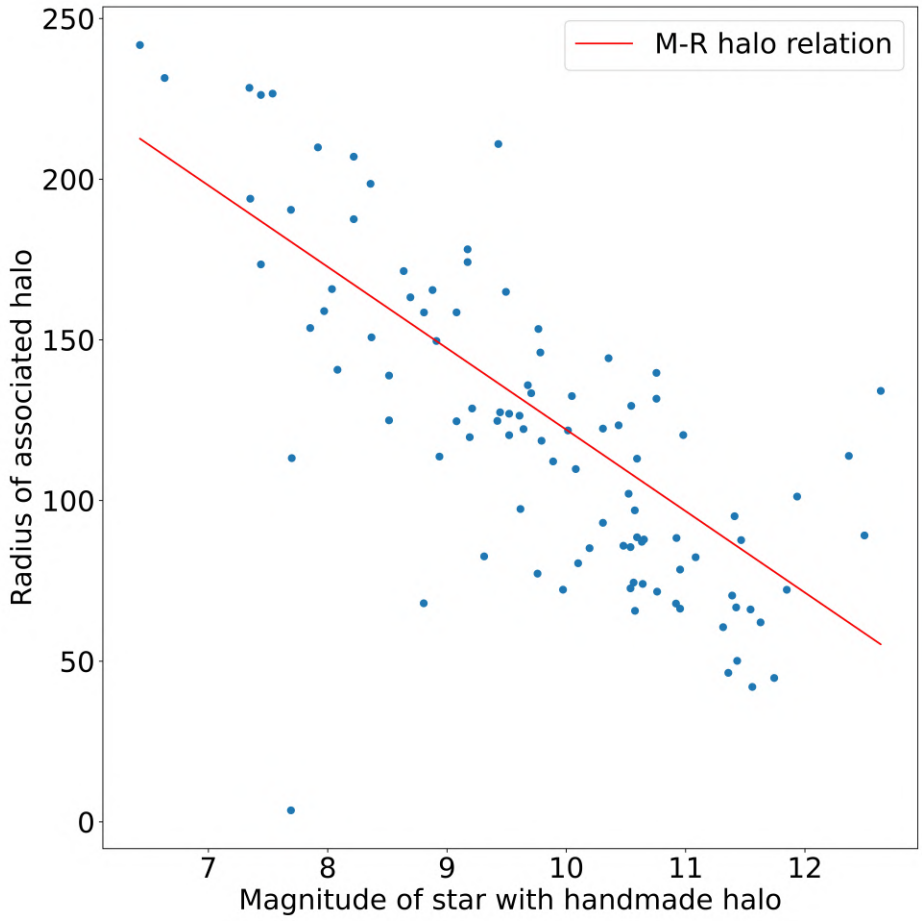


Figure 3.3. Linear relation between Gaia G band stellar magnitude and halo radius in pixel (M-R halo relation). Halos are visually selected and traced by hand on the 5 calibration images. After various tests it was decided to define sources with magnitude < 12 mag as those capable of generating an halo of significantly large size.

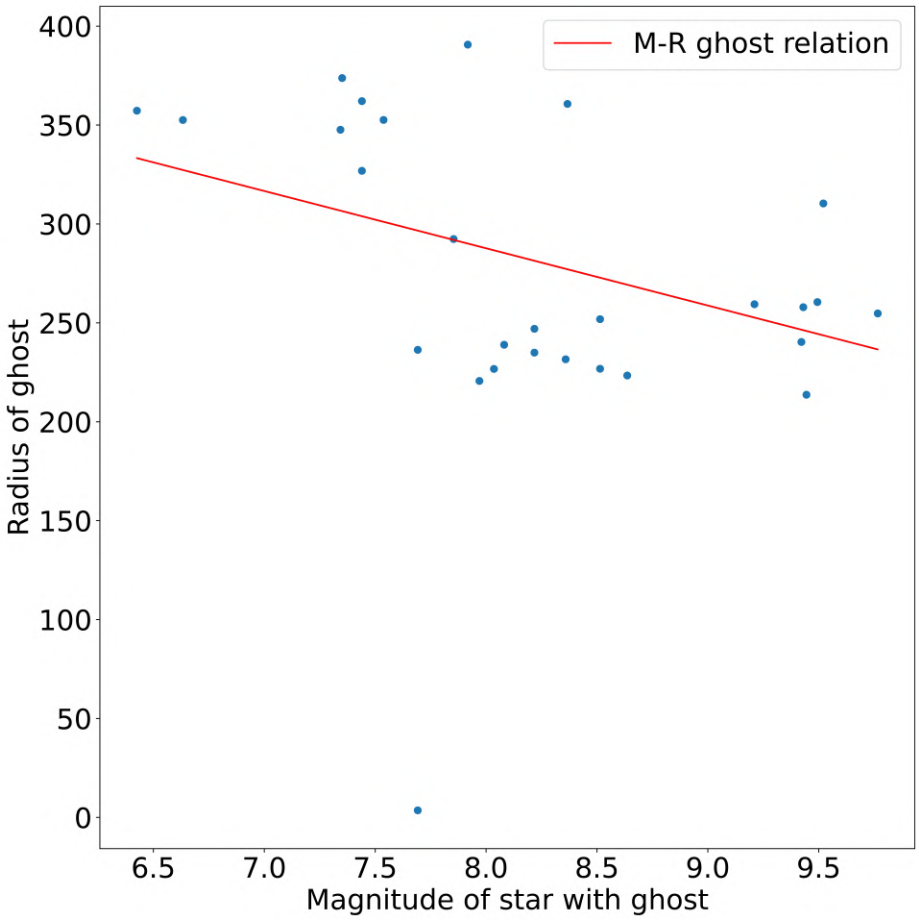


Figure 3.4. Linear relation between star magnitude in Gaia G band and ghost size in pixel. Ghosts are visually selected and traced by hand on the 5 calibration images. Sources capable of producing significantly bright ghosts are empirically defined to have magnitude < 10 mag.

up having multiple misaligned ghosts in the final image.
In reality the appearance of multiple ghosts depends on different factors during the observation of a pawprint and during stacking and reduction processes. A given star could manifest a ghost only in one pawprint and thus ending up with a single ghost in the final stacked image.
One simple example of this is the situation in which a ghost ends up on the blind spot of the camera mosaic for multiple pawprints, being then detected only on the remaining ones. In such a case not all six ghosts would be visible.
Figure 3.5 shows 4 examples of ghosts that represent exceptions to the rules described above.

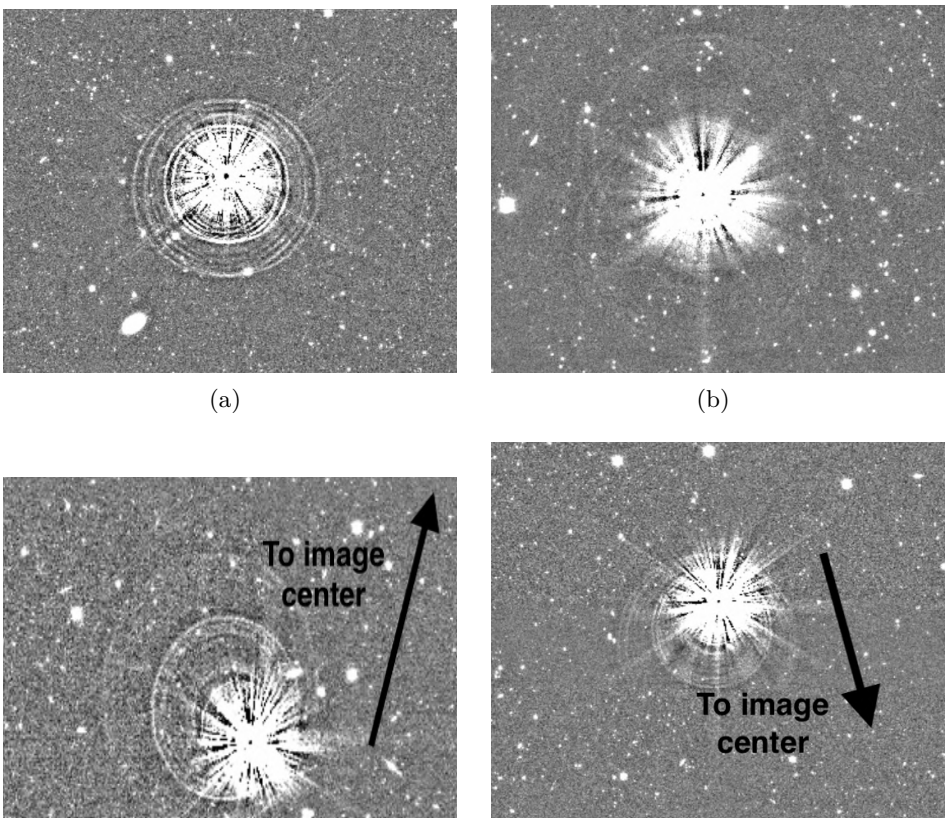


Figure 3.5. Different ways in which ghosts manifest in SHARKS images. (a) Near the center of the image ghosts tend to be centered on the parent star. (b) Sometimes multiple ghosts can form, probably due to dithering effects. Here two faint artifacts are visible on top and bottom of the parent star. In (c) and (d) the center of the ghost does not lay on the straight line connecting the parent star to the image center (direction represented by the arrow) nor are directed radially outwards as we would expect.

Multiple factors are responsible for the position and orientation of a ghost. This makes it impossible to uniquely fit a linear relation between the relative position of a ghost and its parent star.
However, it is assumed that a linear relation can be obtained in order to find and remove all ghosts visually detected in the images. Multiple ghosts are in fact often faint, while ghosts that are not aligned with the expected direction are rare. For this reason it was decided to still make use of a linear relation to approximately find ghosts in an automated way for a given image.
After a visual inspection of the 5 calibration images, many anomalous ghosts seemed to be

found at large radial distances from the image center. It was then selected a "safe radius" on the image, within which a linear relation can be used. Outside the safe radius, ghosts are assumed by the algorithm to be centered on the parent star and their size is calculated as 1.5 times the one asserted from the linear relation. This ensures that anomalous ghosts are still engulfed in the selected area independently on their orientation.
This choice increases the number of sources that are removed from the final catalog. Still, since the number of total sources is very large, of the order of ∼200.000, this loss was considered acceptable and found to be of ∼0.7%.
Figure 3.6 shows the relation linking the relative star-ghost distance to the radial star-image center distance. It can be noticed the large scatter of points, especially for large distances of the parent star from the image center. Here many ghosts are considered anomalous, having centers that do not lay along the radial direction image center-to-star.
Given the distance of the parent star from the center (D_*), the relative distance between ghost and parent star (Δ_g) is:

$$\Delta_g = c_1 \times D_* + c_2$$

Coefficients c_1 and c_2 are extracted from the linear fitting once Δ_g and D_* are calculated using the positions of the visually detected ghosts and of the corresponding parent stars from Gaia DR2 catalog. They are listed in *Table 3.2*.

		Coefficient Value
Ghost size	a_1	-28.9
	a_2	519.3
Halo size	b_1	-25.3
	b_2	375.5
Ghost position	c_1	0.01
	c_2	25.1

Table 3.2. Coefficients for linear relations between radius and position of ghosts/halos and magnitude of parent stars. With i=1,2 coefficients a_i relate magnitude of parent star to ghost radius, while b_i relate magnitude of parent star to halo radius. Coefficients c_i link distance of parent star from the center of the image to the relative distance between parent star and ghost center.

Parameters from *Table 3.2* were used to automatically find halos and ghosts on all dr1 SHARKS image. Any source residing in the area automatically selected was removed from the catalog. An example of the result of automatic selection is shown in *Figure 3.7*.

3.2 Completeness

A perfect photometric image would show any existing source present in the observed patch of the sky for each possible flux value. In reality, however, many effects contribute to generating a background noise permeating each pixel of the image. This noise consists, among other things, of random fluctuations of the number of photons received by a pixel and it is not connected to a specific source.

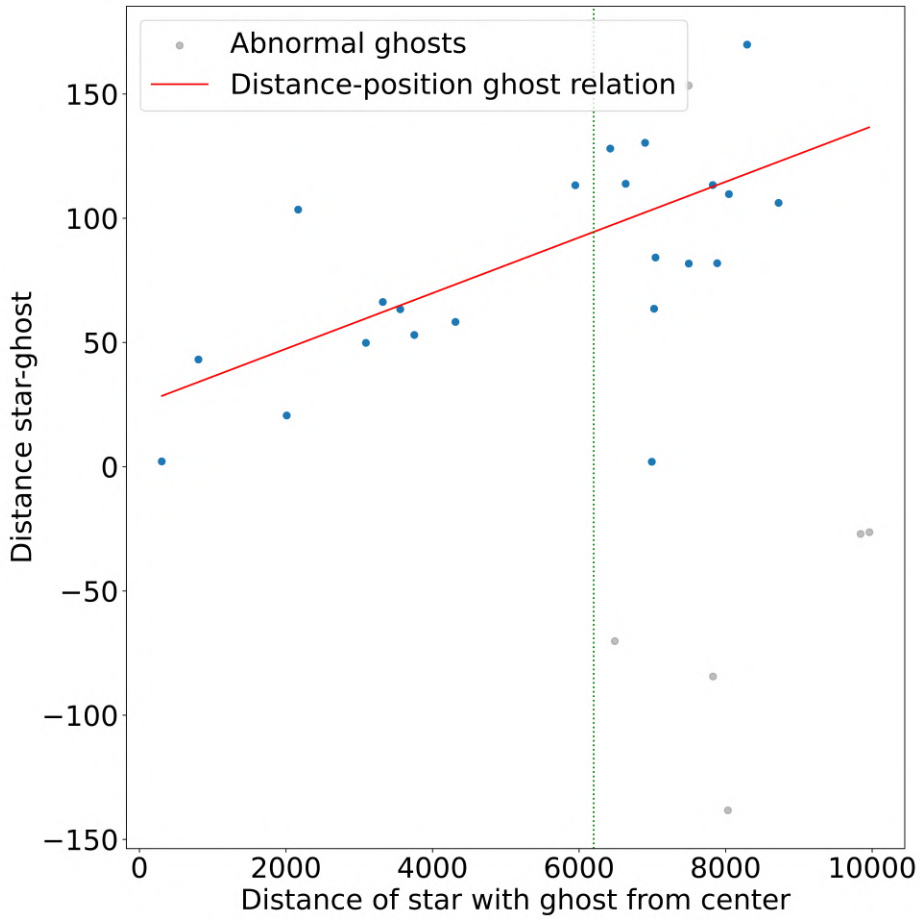


Figure 3.6. Linear relation between Gaia G band stellar magnitude and ghost size in pixel. Ghosts are visually selected and traced by hand on the 5 calibration images. Dashed line represents the limit over which ghosts are found to be anomalous and so for the artificial masking of the catalog, only the size relation is considered. Position is centered on the parent star and the radius is 1.5 times the radius found from such a relation.

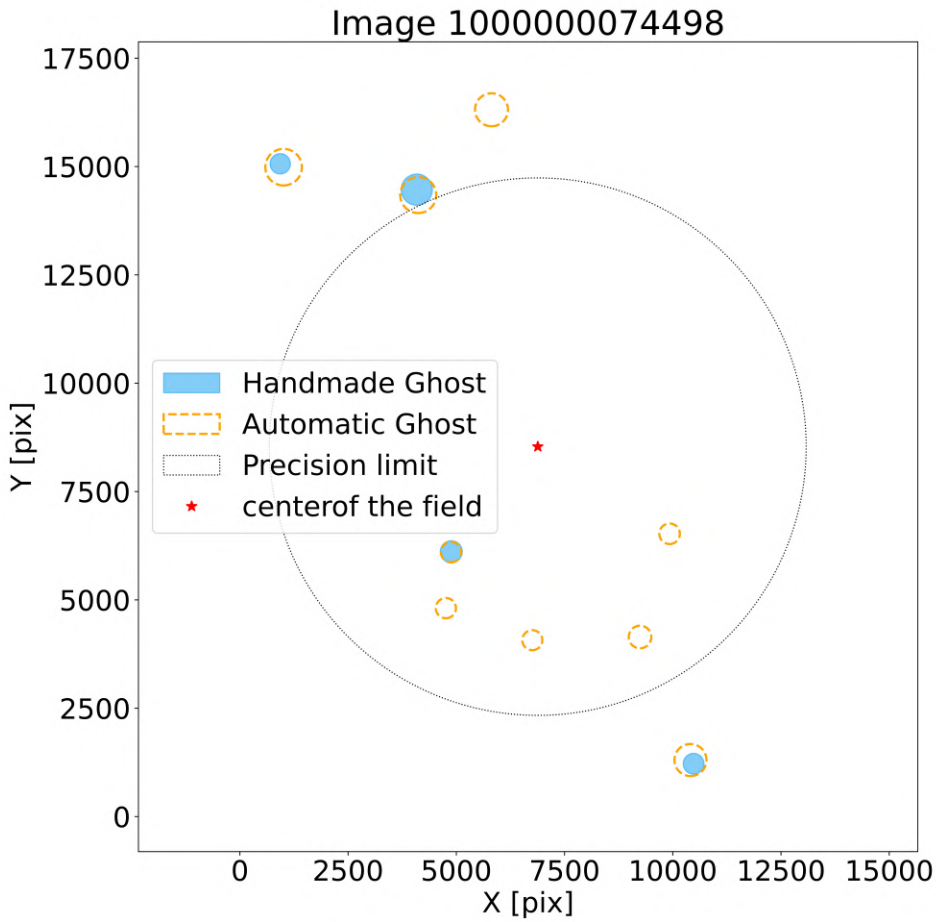


Figure 3.7. Comparison between manually-selected ghosts from the image (filled cyan disks) and automatically-generated ghosts from the algorithm (dashed orange circles). The dotted circle represents the region inside which the linear relation is directly used. Outside this region, calibration images showed the presence of anomalous ghosts.

Any source with a flux lower than the average flux of the noise cannot be distinguished from the background. In reality, noise flux fluctuates (both in time and in space) due to its random nature. Thus it is possible to observe sources with fluxes lower than the average noise level if they happen to be in a region for which, at the time of the observation, the random noise is particularly low. Looking only at sources with fluxes much higher than the noise, we expect them to be all completely visible. Going to fainter flux levels, instead, there is an increasing probability that some sources are indistinguishable from noise fluctuations.

Completeness is then defined as the flux level under which it is possible that some existing sources are not detectable.

Completeness is usually calculated by means of simulations. Once a catalog is constructed, sources with known position and flux are simulated with the same noise conditions of the original image. Then a catalog is extracted from the simulated image. The number of simulated sources for each bin of flux is known and it is confronted with the number of extracted simulated sources for the flux bin. This is repeated for a number of times in order to associate to each flux interval a probability of being completely detected.In this way it is possible to establish the flux limit under which not all the simulated sources are extracted.

A simpler method was chosen for this analysis, following what was done by A. Mercurio et al. 2015a. The method consists in searching for the appearance of sources strongly degraded by the

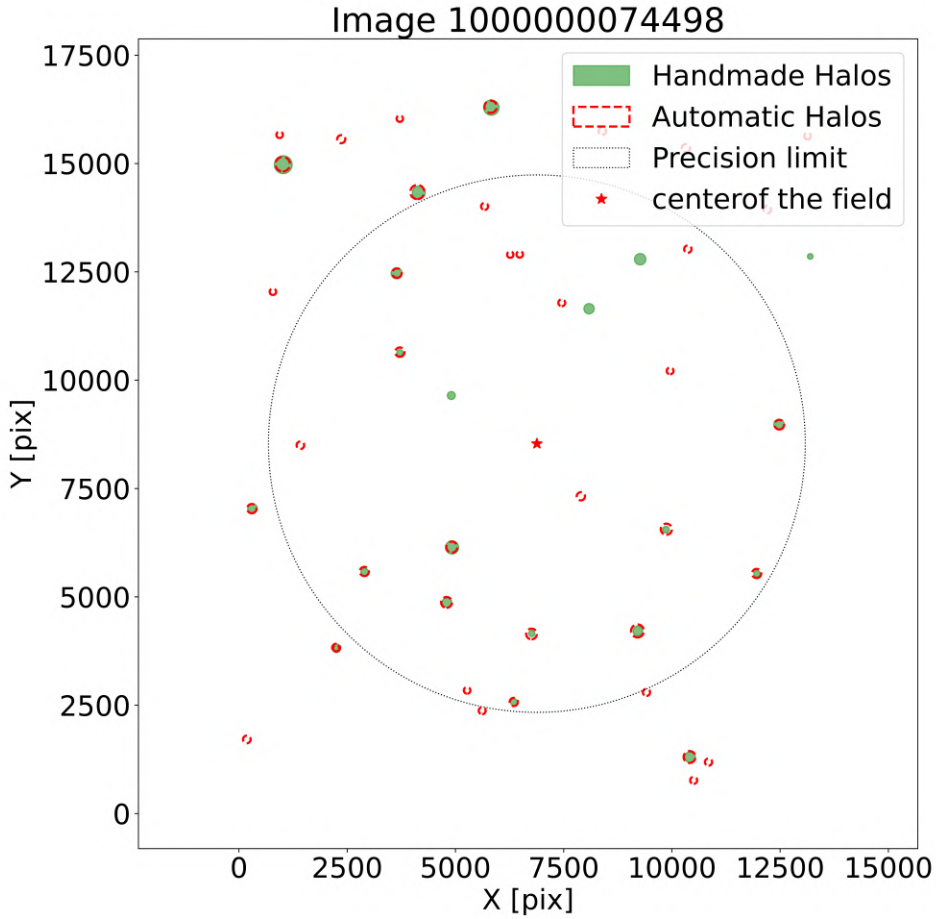


Figure 3.8. Comparison between manually-selected halos from the image (filled green disks) and automatically-generated halos from the algorithm (dashed red circles). The dotted circle represents the region inside which the linear relation is directly used. Outside this region, calibration images showed the presence of anomalous ghosts.

noise fluctuations. This is done by monitoring the relation between the flux calculated inside two different apertures for the same source. It is in fact expected for a source emitting a fairly strong flux to have a somewhat well defined shape, even taking diffraction and seeing effects into account. This means that flux calculated within apertures with different diameter increases almost linearly as the aperture diameter increases.

If the flux of a source is close to the noise level, instead, increasing the aperture size will include a significant part of the random fluctuating noise, thus making the relation non linear and subjected to a random component. The selected aperture for SHARKS catalogs have a diameter of 2″ and 8″. The former is chosen since it is the aperture used for calculating the noise level in dr1. The latter, instead, is chosen to be much larger than the first aperture, but still not large enough to be always dominated by noise random flux.

The relation between flux from the two aperture for each source is showed in *Figure 3.9*. Flux here is in units of magnitudes. Bright sources correspond to lower magnitudes and are associated with a small scatter around the linear relation function (dashed line). Faint sources have much larger scatter. However, as shown in the color map, the vast majority of sources tend to be located in a pattern very close to the linear relation up to magnitudes around 22 mag.

Completeness level is calculated as the intersection between the lower limit of the scatter and the 5σ sensitivity of the catalog (dotted line).

	Saturated stars	Halos (hand made)	Ghosts (hand made)	Ghosts (automated)
10000000 74498	136	21	4	9
10000000 74501	112	17	6	10
10000000 74506	90	22	7	10
10000000 74511	100	25	5	6
10000000 74512	111	14	6	14

Table 3.3. Properties for each calibration image. The first column shows the ID of each calibration image. The second column contains the number of saturated stars, defined as stars with magnitude m < 12. The third and fourth columns show instead the number of visually selected halos and ghosts respectively. On the fifth column can be found the number of ghosts identified with the automatic algorithms using linear relations calibrated with these same images. The number of automated halos corresponds to the number of saturated stars by construction. It is evident that the algorithm assigns ghosts to bright stars that present none to a second visual inspection. On average, approximately $\sim$60% of ghosts automatically found on calibration images have a visual confirmation. Sources in the remaining $\sim$40% are probably not contaminated and are nevertheless discarded. However these sources account for $< 0.1\%$ of the total catalog and are considered an acceptable loss.

Completeness values for each image of dr1 are listed in *Table 3.4* together with the percentage of sources outside completeness and some statistics on photometrically contaminated sources.

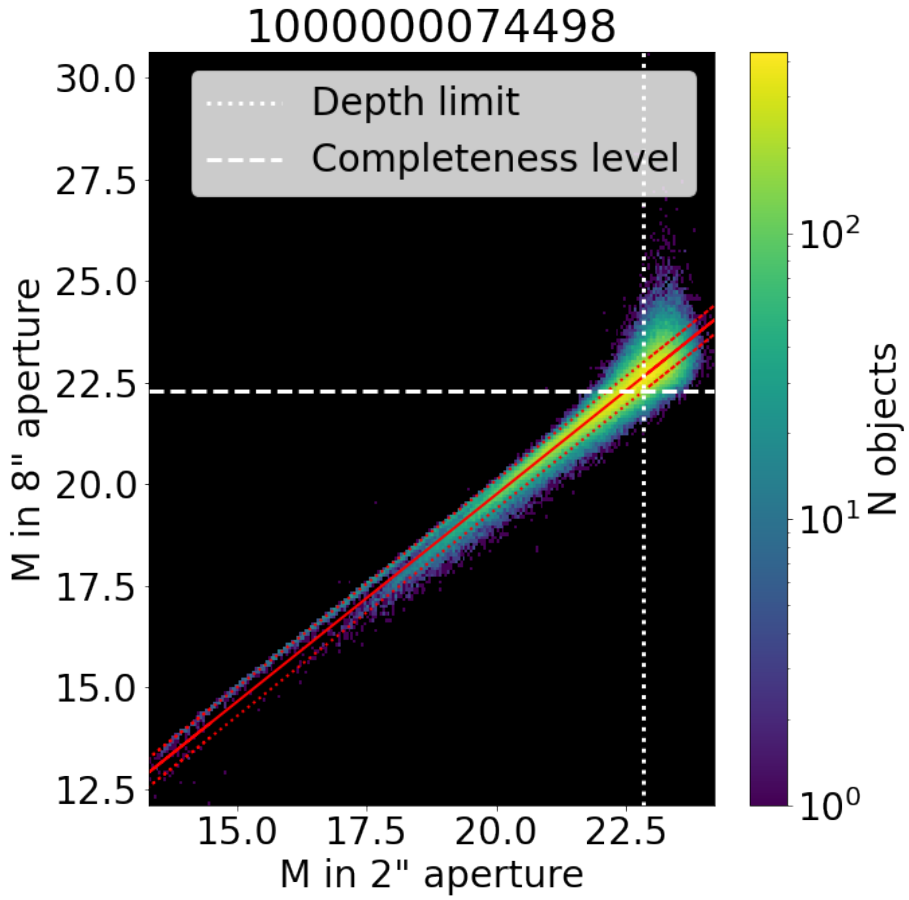


Figure 3.9. Density plot of the relation between fluxes from 2 different diameter apertures. On abscissa the flux from 2" aperture, in ordinate flux from 8" aperture. The continuous red line represents the linear relation and generally follows the peak of the distribution. Dashed red lines represent the 1σ scattering of the relation. The dotted white line shows the 5σ sensitivity level, while the dashed white line indicates the completeness level. This is given as the intersection between the lower limit of the scatter and the sensitivity. The plot refers to image 1000000074498. Calculated level of completeness is 22.3

	Completeness level (mag)	**Contaminated sources (%)**	**Sources in artifacts (%)**	**Sources Outside Completness (%)**	**Contaminated sources outside completeness (%)**
74493	21.93	9.2	2.2	40.4	94.6
74495	22.17	10.0	2.0	43.1	93.5
74498	22.3	10.6	1.6	42.4	92.9
74501	22.19	9.1	2.1	43.8	94.3
74504	22.05	8.4	1.0	45.0	94.3
74506	22.20	9.7	2.4	42.3	93.9
74508	22.03	9.3	2.2	40.5	94.3
74509	21.97	8.2	1.6	41.7	94.6
74511	22.26	9.5	1.3	45.5	92.8
74512	21.87	8.7	2.0	41.3	95.0

Table 3.4. Completeness limits of each dr1 catalog. The first column shows the ID if each image. The third and fourth columns show respectively the fraction of sources flagged by the ghost/halos searching algorithm, and the one of sources with photometric contamination (from SExtractor analysis) that are not inside ghosts/halos. The fifth column shows the percentage of sources that have fluxes lower than the completeness level. The last column on the right, instead, tells how many of the sources outside completeness are contaminated by background light or diffraction artifacts.
A significant number of sources are above the completeness limit. However the vast majority of these sources are largely contaminated as we can expect due to noise effects. Still, contaminated sources that are found below completeness limit constitute a non-negligible fraction of the complete catalog. The amount of sources removed due to diffraction artifacts is comparably low.

3.3 Dealing with catalogs intersections

The 4 mosaics residing in the SGP region slightly overlap with each other.
The presence of regions of the sky observed with different mosaics, independently corrected and calibrated, gives the opportunity to highlight possible biases in the photometric calibration.
Testing this requires to have flux measurements from each overlapping catalog for each source.
Source matching was performed using the software TOPCAT. A source is considered matched if the distance between its position in the two catalog is less than 1".
The test for photometric calibration is shown in 3.10.
In each plot are presented in ordinate the Kron magnitude differences for matched sources and in abscissa Kron magnitudes measured from one of the 2 catalog.
Bright sources have low poissonian noise [2] due to the large number of photons the detector receives from them. Thus the measured fluxes tend to be stable in time, inducing significant fluctuations in the measured flux value.
For faint sources, however, noise has non-negligible contribution and, due to its random nature, it influences the source flux differently at different times.
Following this reasoning is to be expected a close-to-zero difference between the flux measured in different images for bright sources.
For faint sources, instead, flux difference deviates from zero and can assume random values due to strong random noise fluctuations.
A straight line can be fitted to the flux difference distribution in order to find an estimator of its mean value.
The mean flux difference between overlapping images is significantly different from zero for 3 out of the 4 overlapping images.
The most suitable explanation relies on calibration uncertainty. Somewhat different mosaics ended up having systematically different photometric calibrations.

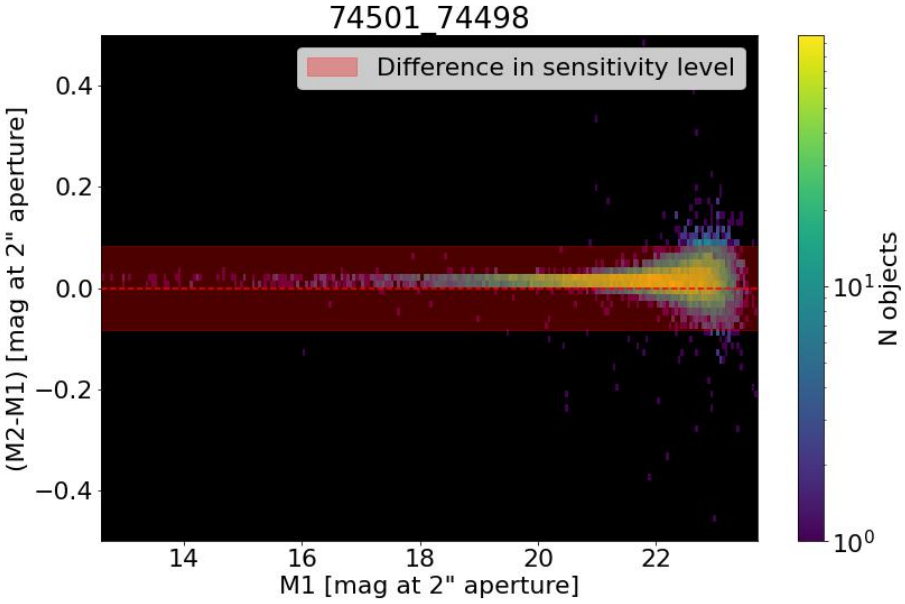
3.4 Star-galaxy separation

SHARKS catalog contains all sources identified by SExtractor. However we need to accurately remove foreground stars belonging to the Milky Way.
Distant galaxies in fact appear very similar to stars: both are almost point-like and shaped by seeing effects [3].
Having multi-band photometry information it is a relatively easy task to separate stars from galaxies, since they generate two well separated sequences in a color-color diagram [4].

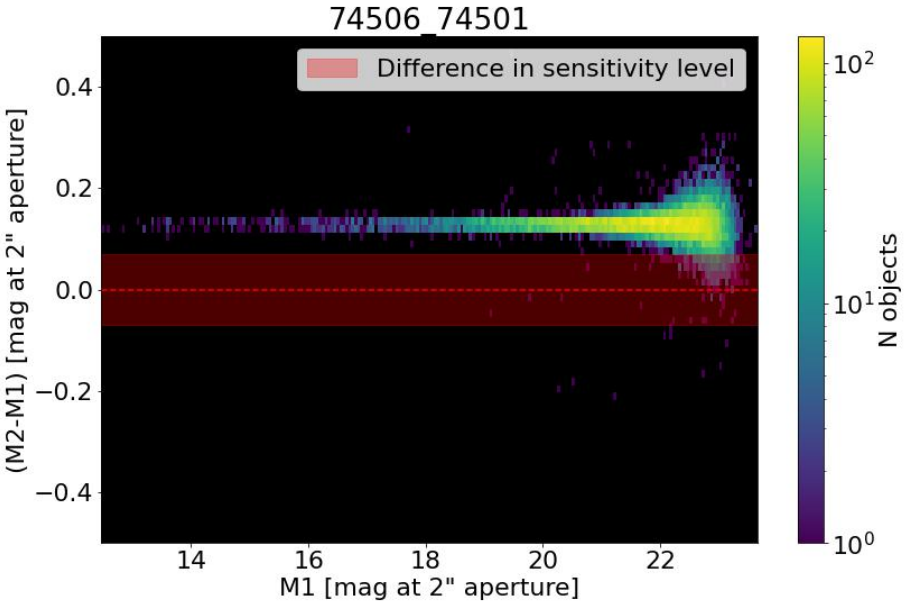
[2]Poisson distribution is a discrete probability distribution for independent events happening with an average constant rate. It describes the distribution in time of photons arriving on a pixel. From this distribution it is possible to estimate that the variance associated to the rate of photons decreases with the square root of the number of detected photons.

[3]These are effects influencing the shape of the source's image on the instrument's CCD. They are due to the presence of atmospheric turbulence, so mixing strata of air along the line of sight with different refraction indexes.

[4]Given three photometric band it is possible to construct a diagram showing a different color index for each axis. One could have, for example, J, H, K bands and form J-H and H-K colors. This diagram is useful to distinguish populations of different objects, one example being the galactic extreme red objects, later confirmed to be proto-planetary disks.



(a)



(b)

Figure 3.10. *(a)* difference between magnitude in a 2″ aperture for sources in both image 1000000074501 and 1000000074498. *(b)* same as (a), but for images 1000000074506 and 1000000074501. Plots are density plots color-coded as a function of the number of objects within a given bin of magnitude. The distribution of magnitude difference has a tight relation for bright sources and a large dispersion for faint sources, due to the effect of random noise. The dashed line marks the position at which magnitude difference is zero. Red band shows the difference of the sensitivity level at 5σ between the two overlapping images. It is possible to see that for the common sources of images 1000000074506 and 1000000074501 there is a systematic magnitude difference. This is probably due to systematic errors in the photometric calibration.

In a single band photometric survey as SHARKS, however, this is not possible: we need to base ourselves only on the shape of the source.
Stars act as point sources and are regularly shaped in an Airy disk by diffraction effects, and their light is even more smeared out by seeing effect.
On the other hand galaxies are extended objects and their shape on the sky plane depends on both morphological type and orientation. This provides a large variety of possible projected shapes even for relatively distant galaxies.
Observing at very high magnitude values, we see faint objects or extremely distant ones and the ability to distinguish their nature decrease strongly.
It is not possible to completely identify every source with certainty: the applied method can overestimate the number of galaxies interpreted as stars or vice versa.
In this work the final goal is to construct a catalog of galaxy clusters. Thus it is necessary to ensure that the galaxy sample is complete. For this reason it is then preferable to overestimate the number of stars interpreted as galaxies rather than lose extra-galactic objects.
A progressive approach is adopted, selecting stars through 3 different parameters applied in different magnitude ranges. The parameters are present in SHARKS dr1 catalogs and derive from SExtractor outputs.

- **Stellarity index:** this is a value provided by SExstractor parameter $CLASS\_STAR$, that expresses the probability of a source to be a galaxy or a star.
  The classifier is based on a multi layer feed-forward neural network.
  It comes from the study of 8 isophotal areas of the source and the highest pixel value, combined to the highest pixel value with respect to the background. A seeing parameter is also considered.
  The closer $CLASS\_STAR$ is to 1, the more likely it is for the source to be a star, while the closer it is to 0, the more probable it is to be a galaxy.
  This classifier, however, is reliable only for the extreme values; any intermediate between almost 1 and almost 0 does not provide sufficient information.

- **Half-light radius:** this is a parameter expressing, for a source, the angular distance (from its center) at which it emits half of its total flux.
  This correlates with the total magnitude of the source if the source is a star.
  This is due to the fact that the image of a star tends to have circular symmetry due to a combination of diffraction and seeing effects.
  A galaxy, instead, is an extended source that can have many different shapes on the sky plane and hardly a circular one.

- **Spread model:** this parameter is the result of the application of a parameter estimator that indicates the best PSF model fitting the observed source.
  The estimator makes a choice between the local PSF of the image, resampled on the source itself, and a convolution between this and a circular exponential model.
  The former model better suits point like-sources affected by seeing, while the latter one fits sources that are more smeared out, like extended sources.
  SExtractor's $SPREAD\_MODEL$ is close to zero for point sources, while it is positive for extended sources.

- μ **max:** This is not used as a parameter to separate stars from galaxies, but as a way to

check the selection performed with the other methods.
This parameter represents the flux measured on the brightest pixel of a source.
For symmetry reasons it is expected stars to arrange themselves in a narrow sequence along a relation between μ and total flux.

3.4.1 Application of the progressive method

I constructed distributions of each of the parameters described in the previous section with the source's magnitude. The tight sequence of points in such distributions represents objects that have the characteristics of the typical stellar component. I selected such sequences and removed the corresponding sources from the catalog of galaxies.
Figure 3.11 shows a density plot for the distribution of `Half-light radius`, μ `max` and `Spread model` parameters in the mosaic 1000000074498.
The color map indicates the number of sources within a given interval of magnitude and parameter value. There is a clearly visible narrow sequence of points in each plot: those are assumed to be stars for the reasons already stated for each parameter.

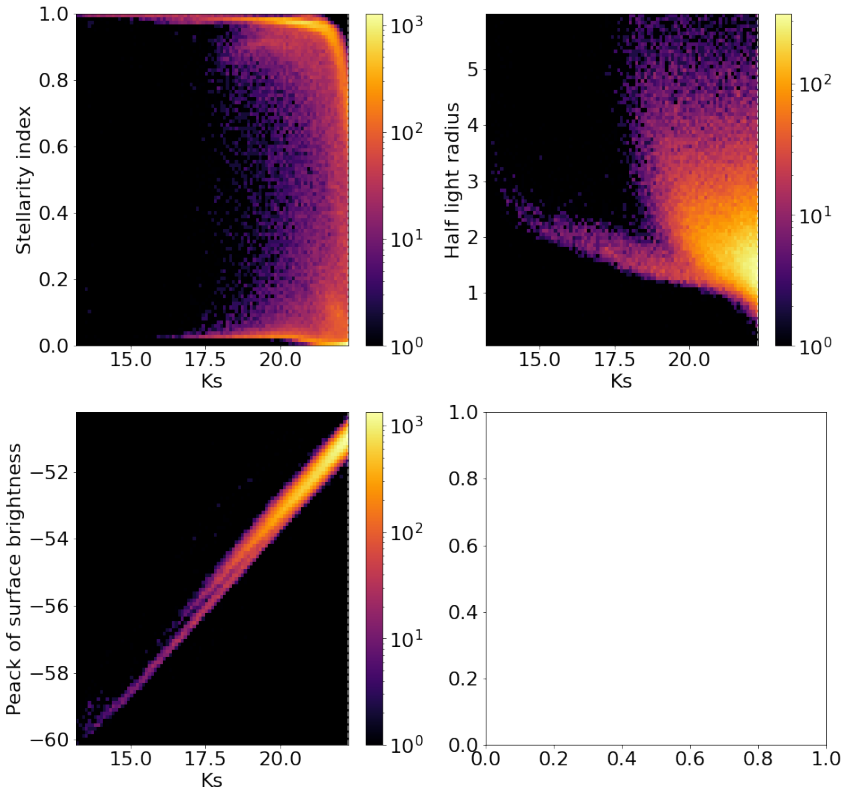


Figure 3.11. Density plot for distribution of parameters `Stellarity`, `half-light radius` and μ_{max}. Tight sequences of points are visible in each image: they represent the stellar component. The higher density of points seems to appear in regions with dispersion of one order of magnitude larger, showing that the majority of sources in the catalog is constituted by extra-galactic extended objects.

For all catalogs of SHARKS dr1, this work used only *Stellarity index* and *Half-light radius* in order to separate the stellar component from the galactic one.
However the *Spread model* parameter was tested for the 4 contiguous images in SGP region.
The test highlighted the possibility of performing the separation down to much fainter sources, increasing significantly the method with only two parameters.

In *Figure 3.12* it is shown a selection performed with the *Stellarity index*.
Stellarity parameter is used only for bright sources, defined as having a flux M < 19.2 mag from the the dr1 release team. A bright source is considered a star if STELLARITY > 0.99 following a more restrictive requirement than in A. Mercurio et al. 2015b.
It is possible to appreciate how this conservative selection includes a small fraction of points in the assumed star sequences visible on the distribution of the other parameters.

The sequence of points for $14 <$ M < 21 in the *Half-light radius* plot (in cyan) was visually selected. The largest magnitude included in the selection is found to be M $= 20.56$, far from the completeness level having a mean value of $\sim 22 mag$.
The selected points are clearly arranged on a narrow linear sequence in the μ_{max} plot.
Results are shown in *Figure 3.13*.

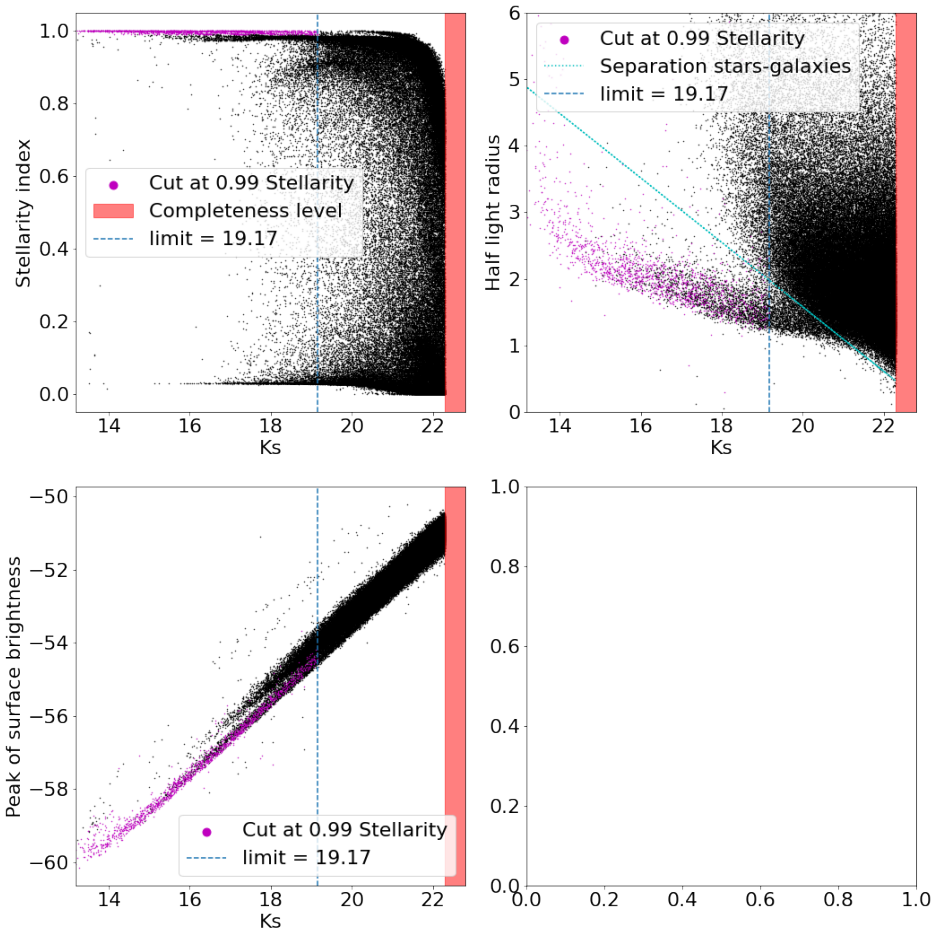


Figure 3.12. The upper right plot shows (magenta) the adopted selection of bright stars. They are objects with magnitude < 19.2 and stellarity > 0.99. The same objects are reported in the half-light radius plot and in the μ_{max} plots highlighting how they arrange themselves again in tight sequences.

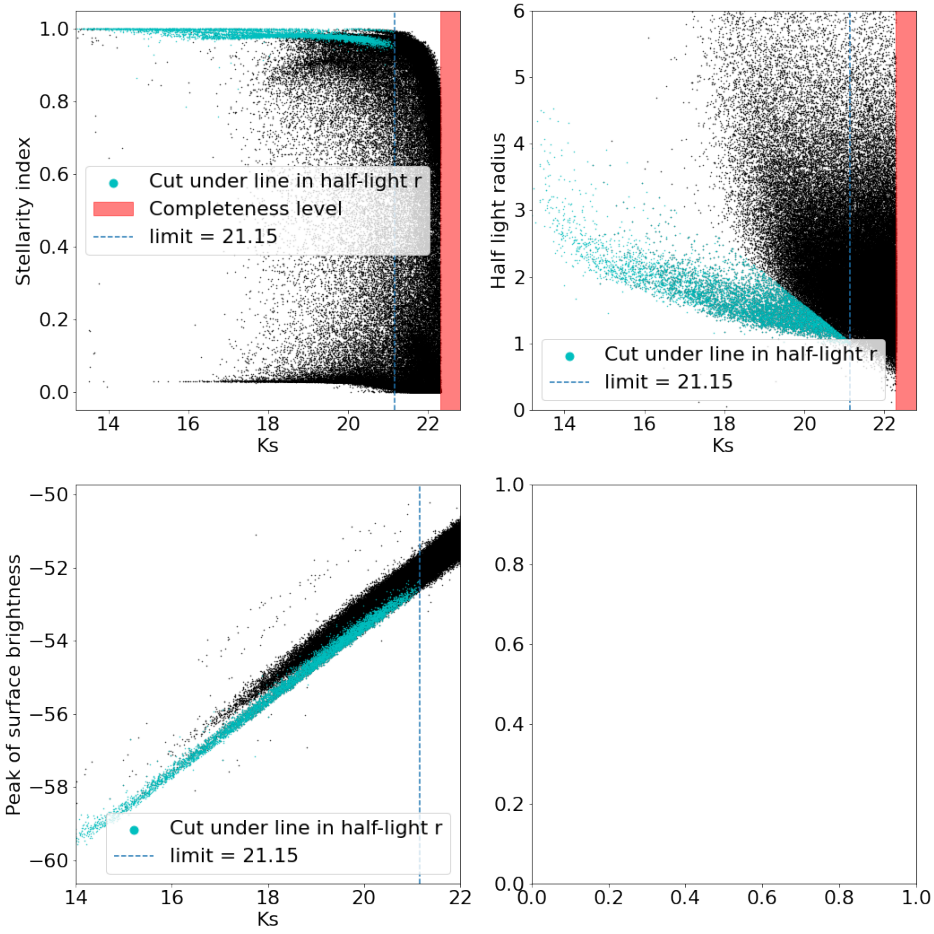


Figure 3.13. The upper right plot shows the selection performed on the half-light radius parameter. The narrow sequence of points at low magnitude is interpreted as a sequence of the stellar component. The selected points (cyan) are shown also in the stellarity and μ_{max} plots, highlighting how they arrange themselves again in tight sequences. The largest magnitude reached with this relation is 20.56 mag.

3.4.2 Adding the spread model parameter

I performed a test adding the *Spread model* parameter for the 4 contiguous fields 1000000074498, 1000000074501, 1000000074506 and 1000000074511. The aim is to understand if the addition of this new parameter is able to provide a more in-depth removal of the stellar component, bringing the selection of stellar objects to higher magnitudes.
Result for the catalog 1000000074498 is shown in *Figure 3.14*. The sequence in the *Spread model* parameter distribution is selected visually. The largest magnitude included in the selection is M = 21.69 mag.
Again, the spread-model-selected points are found to have a small scatter in the μ *max* plot around the sequences of both stellarity and half-light radius selected sources.

The test shows that adding a third parameter increased the depth of the selection process of more than 1 mag.

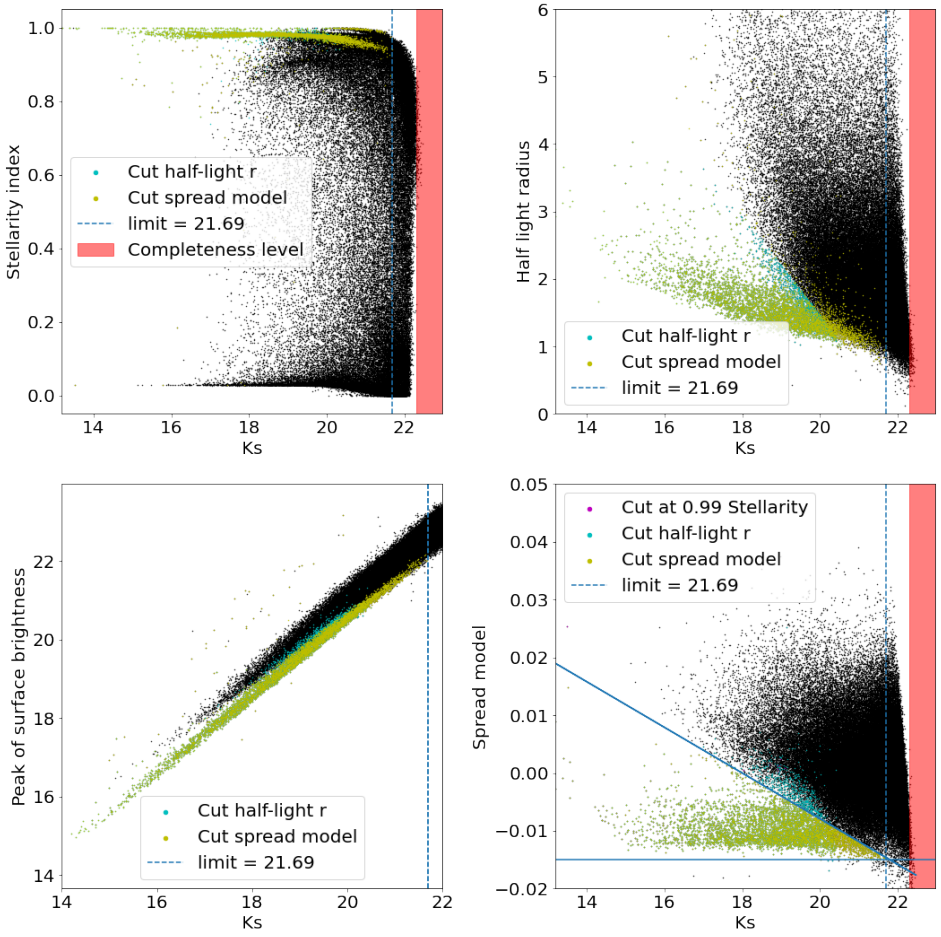


Figure 3.14. Stellar component selected using the spread model parameter from SExtractor (in green). In the bottom right plot, for values of the spread model around -0.01, the spread model's tight sequence of points is visible. The selection was performed visually, leading to a maximum magnitude of 21.69 mag. Spread-model-selected points show tight relations on each plot, confirming the validity of the selection.

All the analysis that follows is obtained using the cleaned catalogs. These catalogs contain sources with AB magnitude down to each respective completeness limit (that average out to $M = 22.1$mag). Stellar component is removed down to AB magnitude $M = 20.56$. Spurious sources are removed using both SExtractor flags contained in SHARKS data release one and using the self-made algorithm for the removal of the diffraction artifacts.

3.5 Some statistics of cleaned catalogs

I applied a simple counting function to the 4 contiguous fields in SGP in order to understand the effects of removing contaminated sources on the catalogs statistics. Galaxies are counted for bins of Kron magnitudes.

The results of the counting functions for the 4 cleaned catalogs are shown in *Figure 3.15* together with completeness levels. It is clear that all 4 images are consistent, having roughly the same statistics other than almost the same completeness level.

In *Figure 3.16* instead, it is shown a comparison between different counting functions from the same image 1000000074498. The counting function for the raw non-cleaned catalog is plotted in blue. The green line represents the counting function for the catalog cleaned from contaminated

sources with both ERRBITS flag and halo/ghost. The counting function for the contaminated sources only is shown in orange. The amount of contaminated sources do not influence much the amount of sources in the resulting catalog, as expected.
The largest deviation between the cleaned and non-cleaned catalogs is found for the faintest sources. This behavior is again expected since fainter sources are more affected by noise and so more easily contaminated by background light. However this region is not considered in the final version of the catalog since it is at magnitudes far higher than the completeness level. In fact, in the region under the completeness condition, we find that such a difference is $< 10\%$, while is $\sim 20\%$ in the region outside completeness. Moreover, all three the counting functions behave in the same way.
Furthermore, it can be noticed that removing sources around bright stars does not seem to induce significant biases. This can be seen in *Table 3.4*, which shows that more than 90% of sources with magnitudes larger than the completeness level are flux-contaminated.

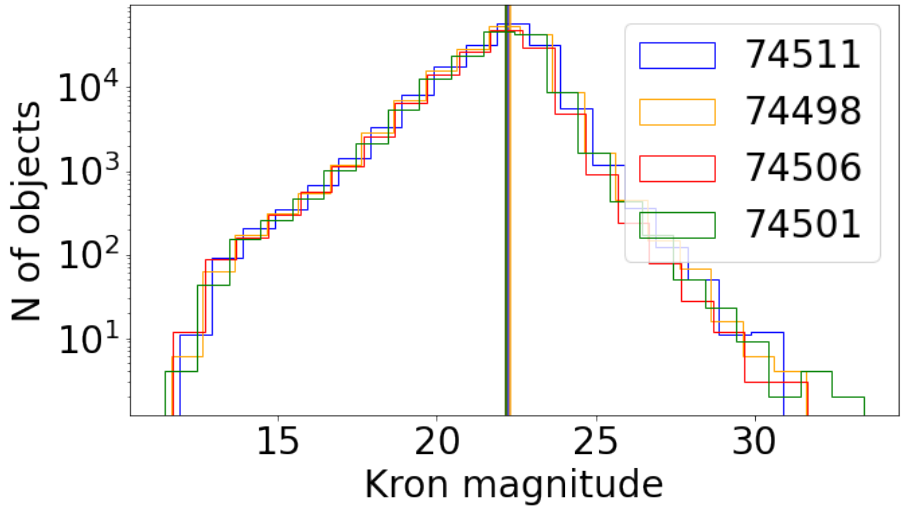


Figure 3.15. Comparison between counting functions for each of the 4 contiguous images on the SGP. Dashed lines indicate the completeness level. The images have consistent statistics, without any substantial bias.

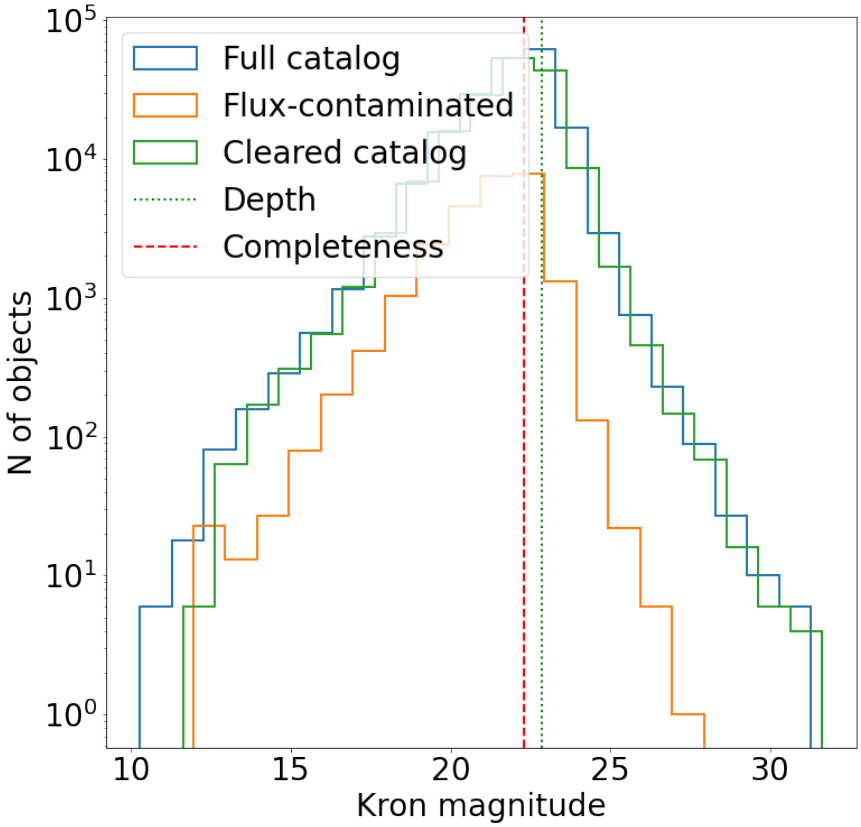


Figure 3.16. Comparison between counting function of the non-cleaned catalog (*blue*) and the cleaned one (*green*). The former contains all sources of the dr1 catalog, while the latter have all contaminated sources removed. Counting function for all contaminated sources is in *orange*. Plot is logarithmic. The shape of the counting function is practically the same in all 3 cases, indicating that flux-contaminated sources do not originate from biases, but from random occurrences as closeness to other sources of similar magnitude. 5σ limit flux (green dotted line) is compared to completeness level (dashed line) as expected completeness is at slight lower magnitudes.

Chapter 4

The Friends-of-Friends algorithm

In order to have a preliminary catalog of clusters from a single band observation, we can use algorithms that search regions of the survey with a number density of galaxies significantly higher than the average.
Once an overdense region is identified, multi-band photometric observations and spectroscopic observations can be used to tentatively separate all the structures projected along the line of sight, to eventually confirm the presence of a real gravitationally bound structure, and to constrain its parameters.

Here I briefly present some algorithms that can be used to search clusters from number density or spatial distribution of galaxies.
Match filter: the image is smoothed (eg. using a Gaussian filter) in order to have a continuous density field. Then a filter function is drawn onto the image to find matching signals.
This method is strongly dependent on the assumption used to construct the filter. These assumptions are based on previous observations so the method is likely to find objects with characteristics close to what we already know to exist. Thus it is poorly sensitive to non-canonical objects (Radovich et al. 2020).

Voronoi tessellation: Applying a Voronoi tessellation on a two dimensional survey means to partition the area of the survey in elemental polygons each one enclosing one and only one galaxy, called a nucleus. Points inside a polygon are such to be the ones closer to the nucleus than to any other galaxy.
Candidate clusters are considered to be sets of adjacent polygons having an area lower than a given threshold, meaning that their nuclei are all very close together (Ramella et al. 2001).

Friends-of-Friends (FoF): A distance is defined in the space on which we apply the FoF. This could be any parameter space, but in tracing galaxy clusters it is a physical three dimensional space, or an angular space if galaxies are projected in the sky plane.
The algorithm groups galaxies together based on this defined distance. If the relative distance among two galaxies is less than a given threshold called linking length (b), those galaxies are called friends.
For each galaxy in a group, there is at least a friend defined in this way. This happens since being friends is a transitive property, meaning that having 3 galaxies A, B and C, if A and B are friends, and B and C are also friends, then A and C are considered friends (Farrens et al. 2011).
In principle the algorithm process can be thought as this:

- We start with a point **p** from the set S we want to explore.

- We define a metric in the space of our set. Then we calculate the distance between **p** and any other point of the set using that metric.
- We decide a threshold distance, a so-called linking length b, defining the relation between the points of the group we want to build.
- Every point x whose distance from **p** is lower than b is then considered part of the same group of **p**, that we can call group G.
- We can now repeat the search for every point x in G: if a point y in the set of points that are in S, but not in G (S-G) has a distance from x which is lower than b, y is part of the group G.
- The same strategy is applied now iteratively for each of the newly acquired points.
- If there is no point in S-G which is closer than b to any point in G, then the group is complete.
- We then pass scan another point of S-G creating a new group K in the same manner.
- It is clear that if b is chosen to be too large, then every point of S would be part of the same group G. For any point p in S would exist at least another point in S which is closer to p than a threshold b.

4.0.1 Algorithm choice

It was decided to use a Friends-of-Friends algorithm to search for possible overdensities in SHARKS catalogs.
The match filtering algorithm was discarded since it requires a series of assumptions on density or luminosity profile of the clusters. The structures that are possible to find with this method are in fact limited to the ones that strictly follow what is already known and well studied. This typically includes local clusters. At redshift sufficiently larger than $z \sim 1.5$, structures as proto-clusters are instead more irregular and unpredictable in shape due to not being yet virialized. On the contrary they consists of many halos collapsing together, allowing for many different configurations of proto-clusters' shapes.
Moreover statistics for these structure is not complete enough so that any filter constructed on its base could be biased by the available observations.
As regards Voronoi tessellation algorithms, they were valid candidates for the search in two dimensional space, however, they were ruled out since they do not allow for a precise definition of the cluster's shape.

4.1 nbodykit package description

The FoF algorithm used is part of the python library *nbodykit* (Hand et al. 2018), which is object oriented and can work with multiple CPU in parallel. nbodykit was originally developed to carry out analysis of the results of N-body cosmological simulations in a fast, efficient and pythonic fashion. However, in this work I apply it to the two-dimensional projection on the sky of the distribution of galaxies. I discuss the implications, the caveats and the possible limitations of this choice in what follows.

Parallel computation was not necessary for dr1 analysis since the amount of data in the catalog was easily manageable.
Nbodykit FoF algorithm, in principle, needs only two basic parameters: the Linking Length and the minimum number of members per group. The latter describes the minimum number of galaxies that must be present in a group for it to be considered as a galaxy cluster. If a group is too small it is ignored. The former, instead, describes the maximum distance between two points for them to be part of the same group. The linking length can be a fixed spatial distance or can be expressed as a fraction of the average distance between the points in the sample.
As said in the general description of the algorithm in chapter 4, it exploits transitivity property. Thus, considering points A, B and C, if A and B are closer together than a linking length and so are B and C, then A, B and C all belong to the same group. This even if A and C are not directly friends (meaning they are further apart than a linking length).

In the following section a rough statistical overview of the application of the FoF to dr1 catalogs is described. Then two methods for a more robust application of the algorithm are presented.

4.2 First statistics with FoF

I performed a first blind search exploring the two-parameter space used as input for the Friends-of-Friends algorithm. It is important to note that this is a non-conventional application of a Friends-of-Friends algorithm, since all galaxies are projected on the sky plane independently of their distance. Hence, apparent cosmic structures might actually be artificial, caused by projection effects. Therefore, it is not possible to adopt linking length values usually used in simulations and extensively tabulated in the literature. It is then necessary to derive my own linking length parameter tailored to the way the problem is formulated in this study. Once one takes this into account, the FoF algorithm can here be applied without loss of generality.
In this phase it was important to establish lower and upper limits for each parameter. I tested parameters independently running the FoF on the 5 calibration images. Results were very similar for all images, therefore in this chapter only the application of the method to image 1000000074498 will be shown.
The linking length (b) was first tested for a value of mingal=3, since no smaller collection of galaxies can logically been considered a cluster. The lower limit for b is then defined as the value for which no candidate cluster can be found, which was calculated to be b = 0.03. This means that b must be so small that no triplets of galaxies can be found being closer together than a fraction b of the mean distance between galaxies in the whole image. The upper limit, instead, is considered the minimum value of b needed to find a single candidate cluster containing a number of galaxies of the order of the entire catalog itself, which is b = 2.
The behavior of the two parameters was tested again in a range further from such extremes and trivial values. These new tested ranges were then $5 < \text{mingal} < 15$ (at steps of 1) and $0.07 < b < 0.2$ (at steps of b = 0.01). FoF was applied for all combinations of parameters in these intervals. We can then study how the number of detected candidate clusters changes for each parameter choice, together with the number of galaxies contained in the larger cluster. We can also study the expected behavior of the distribution of the number of clusters for each parameter choice.

4.2.1 Comparison with a random field of galaxies

As already stated, this is a peculiar application of a Friends-of-Friends algorithm since it makes use of galaxies projected on a two-dimensional plane. The linking length parameter must be redefined in this context, however, there are no means of confrontations and no pre-existent physical or mathematical descriptions were found to provide a meaningful value for it.
I then decided to compare the application of the FoF to a simulated catalog of galaxies, created by randomly reassign the coordinates of the sources in the original catalog. A confrontation between real and simulated catalogs would highlight similar patterns and differences between the results of the FoF application for different linking length. This would allow to study the behavior of the linking length on a two-dimensional space and to constrain its significance in this new application.
The template catalog selected for the simulation was again 1000000074498. I assigned new spatial coordinates to each galaxy in the original catalog. These coordinates were sampled from a uniform distribution in the spatial region spanned by the catalog itself. In this new arrangement, then, no prior is present for the distribution of galaxies, every clustered structure is completely happening by chance.

Distribution of total number of found clusters

At fixed values for mingal, choosing small values of b implies asking the algorithm to find galaxies that are very close to each other.
In *Figure 4.1* it is possible to see that with this condition a low number of clusters can be found, associated to the most compact structures in the image. Increasing b, the number of found clusters increases as we are allowing identification of less compact structures. This is clearly not a linear increase. Such behavior could be explained if at first, requiring a structure to be very compact means that is more likely for it to be a bound structure instead of a random occurrence.
Increasing b could, however, increase the probability for an identified candidate cluster to be just a collection of galaxies that appear close together due to projection effects. These fake structures appear to populate almost any region of the image due to the large amount of galaxies populating the sky plane.
By applying the FoF with a value b' of the linking length it is possible to find close by groups of sources each one identified as a candidate cluster. With a new application of the FoF with a linking length of $b'' > b'$ it is possible that the close-by candidates previously selected will now be considered as a single large cluster.
For this reason increasing the value of b there can not only yield an increase of found candidates, but also a decrease. The increase dominates at first, but then, being the image filled with identified clusters, the union of close-by candidates inverts the trend and the number of found clusters decreases and converges rapidly to one. This single cluster contains all sources in the image and there is no further evolution with increasing b over this point.
The trend of the number of clusters found using different b has the same shape even for different mingal. Changing the minimum number of galaxies required to form a cluster, in fact, seems only to shift the entire curve and to change its normalization. Low values of mingal means that even small collections of galaxies can be considered as clusters.
Considering the 2D image as a random distribution of points it is more likely to have galaxies more frequently arranged in small groups than in big ones. This translates into more clusters found at fixed b as we lower the value of mingal. Moreover the peak of the function will move towards lower values of b, since these are sufficient to find large amounts of very small clusters

before the merging effect takes place.
The same analysis was performed also on the simulated random redistribution of galaxies. The described behaviour can be observed for both real and simulated catalog in *Figure 4.1.*
It is evident that the distribution of the number of clusters at the variation of b and mingal parameters is almost indistinguishable in the two cases.
This means that for a blind application of the FoF algorithm, the distribution of galaxies in the actual SHARKS catalog is similar to a random distribution of galaxies. For this reason any clustering effect could be interpreted as fictitious, without a real physical reason binding galaxies together.
This is probably caused by projection effects, since the single band catalog is showing a three dimensional structure collapsed in a two dimensional one, increasing the average number density of the image close to the typical number density of a cluster.

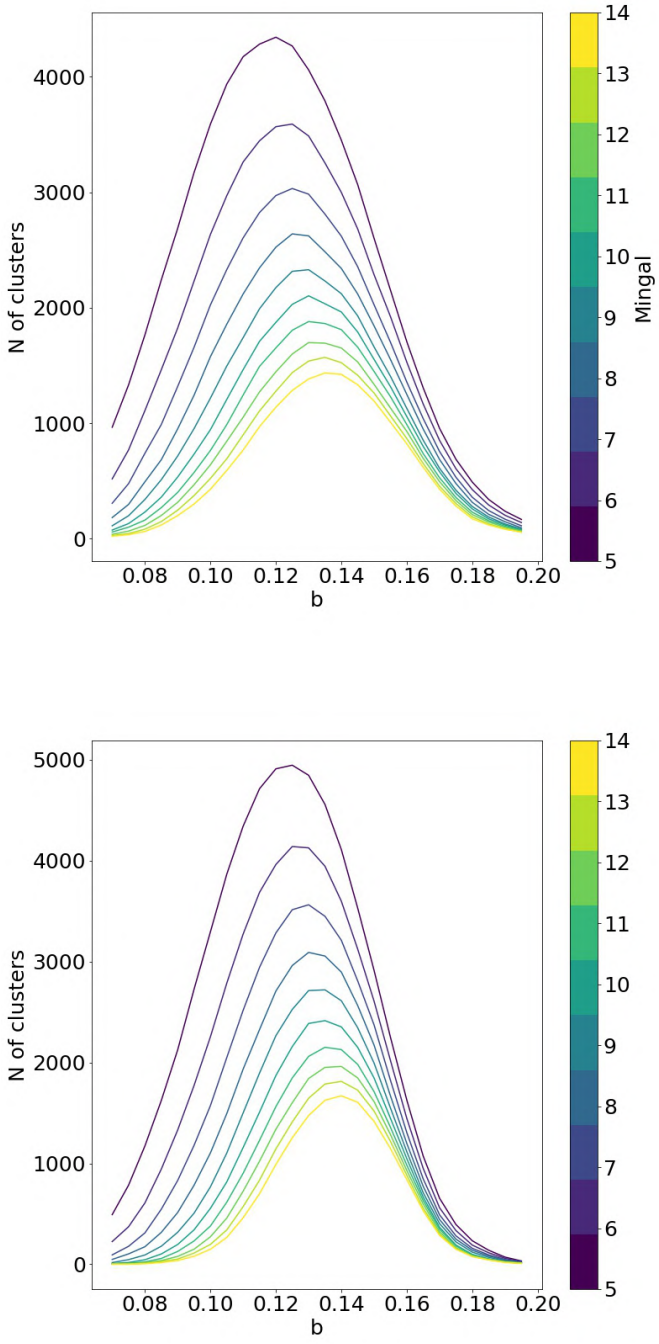
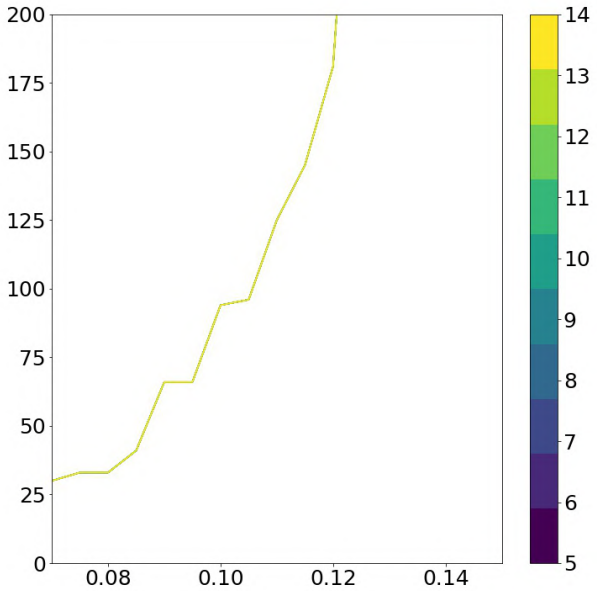


Figure 4.1. *(Upper)* distribution of the number of clusters found in catalog 1000000074498 varying the Linking Length b and minimum number of members mingal. *(Bottom)* distribution of the number of clusters found in random spatial redistribution of catalog 1000000074498. Variation of b and mingal is the same as for the application on the real image. The two distributions are almost identical. It seems not possible then to distinguish a real cluster from a statistical fluctuation of the distribution of galaxies in the 2D image using a FoF algorithm.

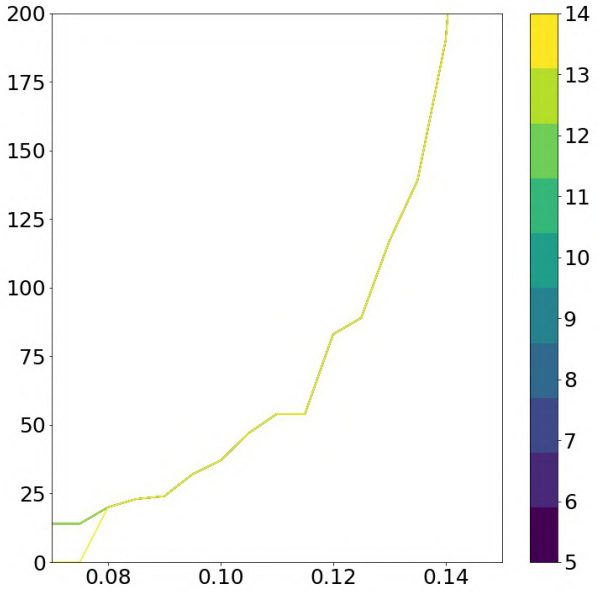
Distribution of the number of galaxies in the larger cluster

At fixed mingal, it is expected that increasing b, the average number of galaxies in a cluster will increase. This is an obvious consequence of the definition of linking length. Moreover, as b increases, the cluster-finding algorithm saturates when a cluster is found to be formed from all

galaxies in the image.
However this distribution is less sensitive to the parameter mingal than the total number of clusters. The reason is that mingal is a lower limit so it does not influence the maximum number of galaxies a cluster can contain.
A small difference in the distributions at different mingal can be observed for the lowest values of b. In this case it is possible that the candidate clusters contain less galaxies than the required mingal, so no cluster is actually found until b is sufficiently large.
A very similar distribution is obtained by repeating the analysis on a random redistribution of galaxies in the catalog.
The two distributions are shown in *Figure 4.2*. Again, this is a clue for the fact that projection effects seem to erase any information on gravitational clustering. This is in fact overwhelmed by the effect of random clustering, that is galaxies with very large three-dimensional distances projecting along the line of sight and showing very small bi-dimensional distances on the sky plane.



(a)



(b)

Figure 4.2. *(Upper)* distribution of the number of cluster members found in the largest candidate cluster for catalog 1000000074498, varying the Linking Length b and minimum number of members mingal. *(Bottom)* distribution of the number of cluster members found in the largest candidate cluster found in random spatial redistribution of catalog 1000000074498. Variation of b and mingal is the same as for the application on the real image. The two distributions are almost identical. It seems not possible then to distinguish a real cluster from a statistical fluctuation of the distribution of galaxies in the 2D image using a FoF algorithm.

4.3 Constraining the minimum number of galaxies

The problem with a blind test of the FoF algorithm is that there are no obvious ways to select the best parameter couple (b, mingal) to use.
Reasonable intervals for each parameter can be found, but the behavior of the algorithm is smooth and regular in such intervals, and no feature favors a parameter value more effectively than others. In other words: we do not know what to search for, nor how to search for it. For this reason I decided to link together the two parameters, so as to move the parameter choice from a two-dimensional to a one-dimensional parameter space. This choice simplifies the process of selecting the more effective setup for the FoF algorithm, since now only one parameter must be studied and user selected.
The mingal parameter was thus chosen to be dependent from the linking length. It was decided to force a deeper meaning on mingal, defining it as *the typical dimension of the largest galaxy cluster that can form by chance in a random distribution of points.*
A random spatial redistribution of the SHARKS catalog is then constructed assigning randomly a position to each source. Then the FoF was run with a mingal=3 and for different values of b. The number of galaxy members for each cluster is calculated and the cluster with the largest amount of galaxies is selected. This is repeated 1000 times and an average value of the number of members for the largest cluster over the 1000 iterations is calculated for each b.
The result is plotted in *Figure 4.3* and showed in *Table 4.1* together with the standard deviation associated to each mean value. The value of mingal to associate to each choice of b is then computed accounting for the standard deviation σ calculated from the 1000 simulations. The formula is

$$mingal = \mathrm{M}_{\mathrm{max}} + 3 \times \sigma$$

With $\mathrm{M}_{\mathrm{max}}$ the number of galaxies in the largest cluster found in the simulation. In this way a lower limit for the parameter mingal is constructed for each value of linking length b. This is now not a simple definition of a cluster, but has a deeper meaning since it represents the largest most probable collection of galaxies that can occur in a SHARKS image and be confused for a cluster. Now it is possible to explore the parameter space for b and having from this relation the corresponding limit mingal.

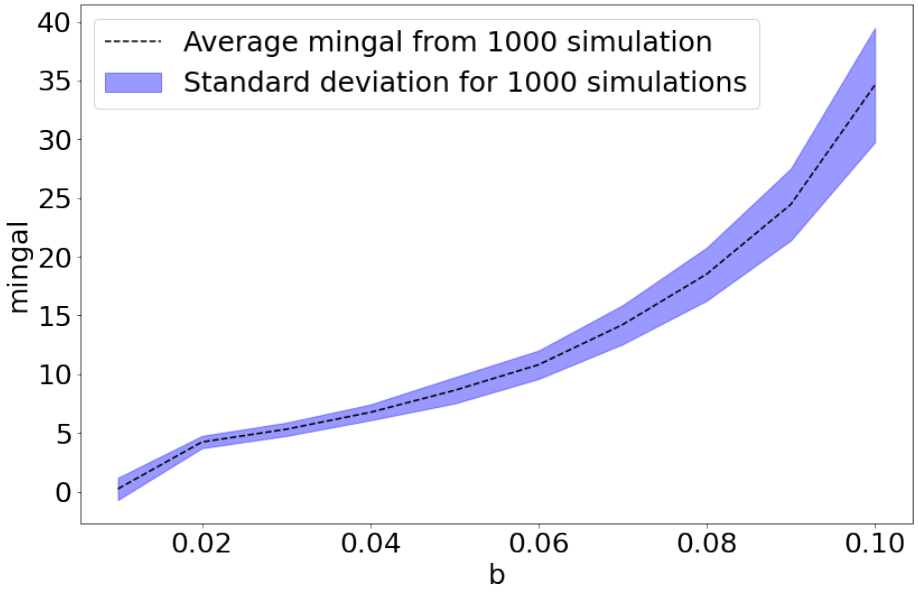


Figure 4.3. Relation between the parameters b and mingal. Mingal here is intended as the minimum number of galaxies that a candidate cluster must have to lower the chance of being a statistical fluctuation effect on the distribution of galaxies. We simulated 1000 random spatial redistribution of galaxies in catalog 1000000074498. The FoF algorithm was applied for different b and a mingal = 3 to each simulated catalog. The number of cluster members for the largest cluster was extracted at each iteration for each b value. The average of this number over all 1000 iterations for each b value was calculated, together with its standard deviation. The resulting relation is plotted. Increasing the value of b, the corresponding mingal increase in a non-linear way.

Linking length	**Mean n galaxies (1000 iter)**	**Standard deviation (σ) (1000 iter)**
0.01	0.28	1.02
0.02	4.29	0.63
0.03	5.35	0.59
0.04	6.78	0.95
0.05	8.66	1.00
0.06	11.05	1.35
0.07	14.34	1.91
0.08	18.15	2.29
0.09	24.34	3.44
0.10	34.73	4.74

Table 4.1. b-mingal relation. For each value of b, 1000 simulated spatial redistribution of SHARKS' image 1000000074498 were generated. The FoF algorithm was applied to each simulation with the given b and a mingal=3. The number of members in the largest candidate cluster is saved. An average of this value over the 1000 simulations is obtained together with its standard deviation.
We decided to use $Mean + 3 \times \sigma$ as value of mingal to apply to the FoF.

4.4 Testing FoF with dependent parameters

There are different properties associated with the candidate clusters found using FoF.
Three macro-proprieties show the general characteristics coming from a given choice of b and mingal parameters. These are:

- The number N of total cluster candidates found in one catalog.
- The number M of galaxies contained in the largest found cluster.
- The radius R of the largest found cluster.

They provide some insights on the goodness of the parameter choice for the FoF algorithm. For example, if R results to be of the order of the size of the image itself, then the parameter b used in the FoF is too large. If, on the other hand, N is much larger than expected, for example showing thousands of candidate clusters per squared degree, then maybe the value of mingal used is not restrictive enough.

Other important characteristics are to be found in the distribution of properties associated to the single candidate clusters detected with the FoF. These properties are:

- The distribution of the number n of galaxies contained in each cluster.
- The distribution of cluster radii r in angular units.
- The distribution of number density ρ for each cluster, defined as the number of galaxies per unit area:
$$n = \pi \times radius^2$$

In *Figure 4.4* it is possible to see the distribution of these properties for a choice of b and mingal respecting the relation obtained from simulated catalogs (*Figure 4.3*).
These distributions are compared with distributions from the same analysis applied to 10 simulated images in *Figure 4.5*. Parameter distribution appears quite different between the real and the simulated catalogs, highlighting how the choice of the new definition for mingal change the characteristics of found candidates, now larger and denser, with a broader variety of proprieties.
Figure 4.6 and *Figure 4.7* show respectively the number of clusters and the number of galaxies in the largest clusters found with different values of b (and corresponding mingal) for real and for simulated catalogs.

The best b choice seems to be b = 0.08, with corresponding mingal = 25.
This comes for the fact that this parameter corresponds to a peak in the number of clusters found and to a region of modest increase in the mingal value, right before its strong increase.
Further reasons to use b = 0.08 comes from the exploration of plots, showing a confrontation between the proprieties N, M, R for the different catalogs and 10 simulated ones.
An example can be found on *Figure 4.11* where the property N is shown. Here the FoF is performed using b = 0.08 for the corresponding mingal = 25 and for a much smaller value mingal = 15.
In the former there are no clusters from the simulated images as expected, while in the latter, N is generally consistent between simulated and real images.
For the proprieties M and R the real and simulated distributions are still well separated.

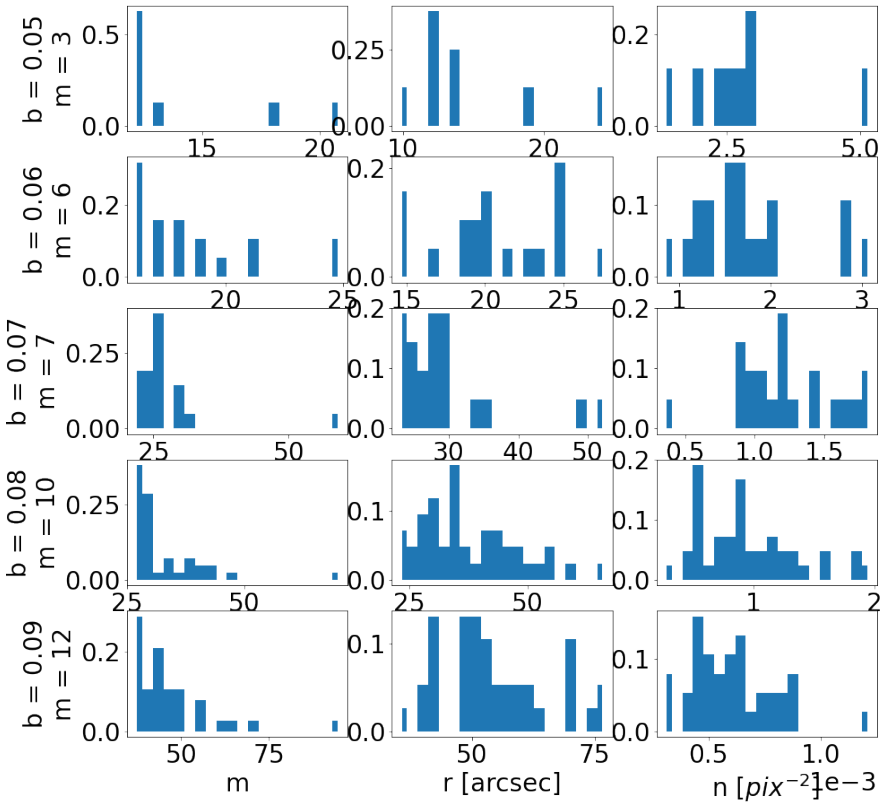


Figure 4.4. Distribution of the m, r and n parameters for linking length in the range 0.05 $\leq$ b $\leq$ 0.09. and corresponding mingal from the b-mingal relation. The analysis is applied to all 10 catalogs from dr1.

Visual inspection of the candidate clusters found with this method confirms the possibility of finding at least the most concentrated structures, that is candidate clusters with a large numbers of galaxies within a very small radius.

An example of candidate clusters is shown in *Figure 4.14* coming from the analysis of catalog 1000000074501 and 1000000074498.

Eight candidates are identified in each catalog. Three candidates in catalog 1000000074501 are very close together in the sky plane.

Their radius is of the order of 1 arcminute encompassing at least 25 galaxies.

These structures are much smaller than one could expect from cosmological simulations, but it is probable that they do not represent the entirety of a cluster. Such dense regions can in fact represent the main halo of a distant cluster, or even some dense sub halos that have still to merge together to form a larger structure.

A strong limitation of this method comes from the large value of mingal required for a given value of b.

This requirement comes from simulations and it is needed to have a justification or reference point for the choice of mingal.

The aim is to reduce ambiguity on the nature of the identified structure, which then have low probability of being caused by statistical fluctuations.

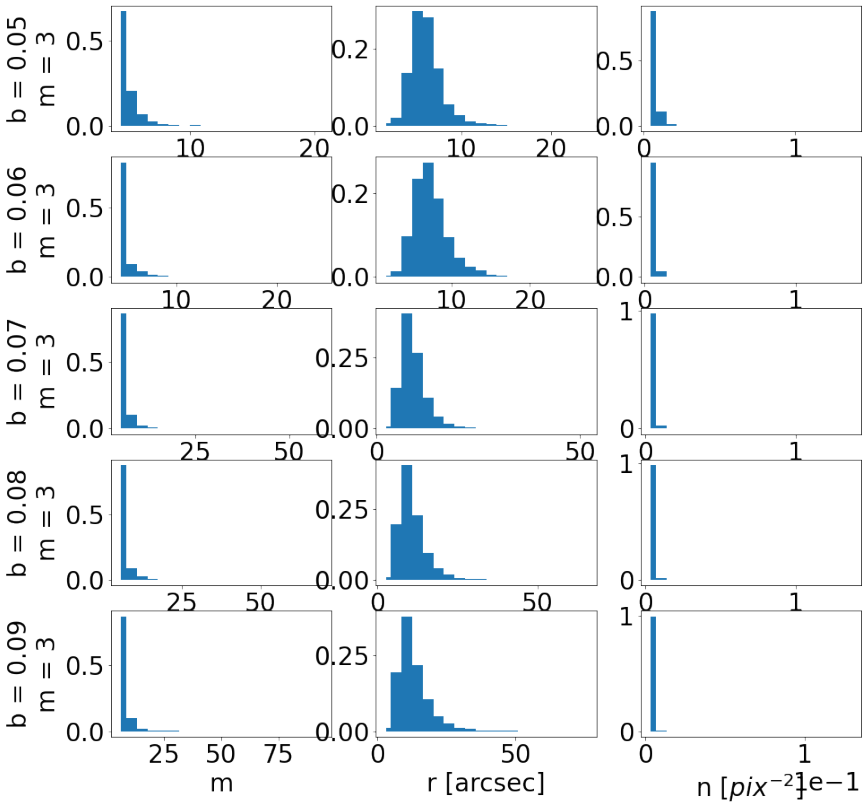


Figure 4.5. Distribution of the m, r and n parameters for linking length in the range $0.05 \leq b \leq 0.09$. Mingal is chosen as the minimum possible value: mingal = 3. The analysis is applied to 10 simulated images, obtained by randomly reassigning spatial coordinates to catalogs' sources.

This however limits the search of candidate clusters to only the densest environments.
Such a choice appears to be rather in contrast with what we can expect from distant clusters: structures with very large angular sizes and with very low surface density.
Furthermore it is important to notice that the simulations employed assume to have a two-dimensional distribution of sources for each any position is equiprobable.
In reality what we observe is a structured three-dimensional volume projected onto a two-dimensional plane. This means that the overlap of low density structures present in the 3D distribution could form high density regions in the 2D one. These would be interpreted as candidate clusters from the FoF algorithm, increasing the probability of having false signals.
However results from the 2D simulation seem to be even excessively restrictive for the choice of mingal.
Deeper insights on the argument are left for future works and improvements of the simulated approach.

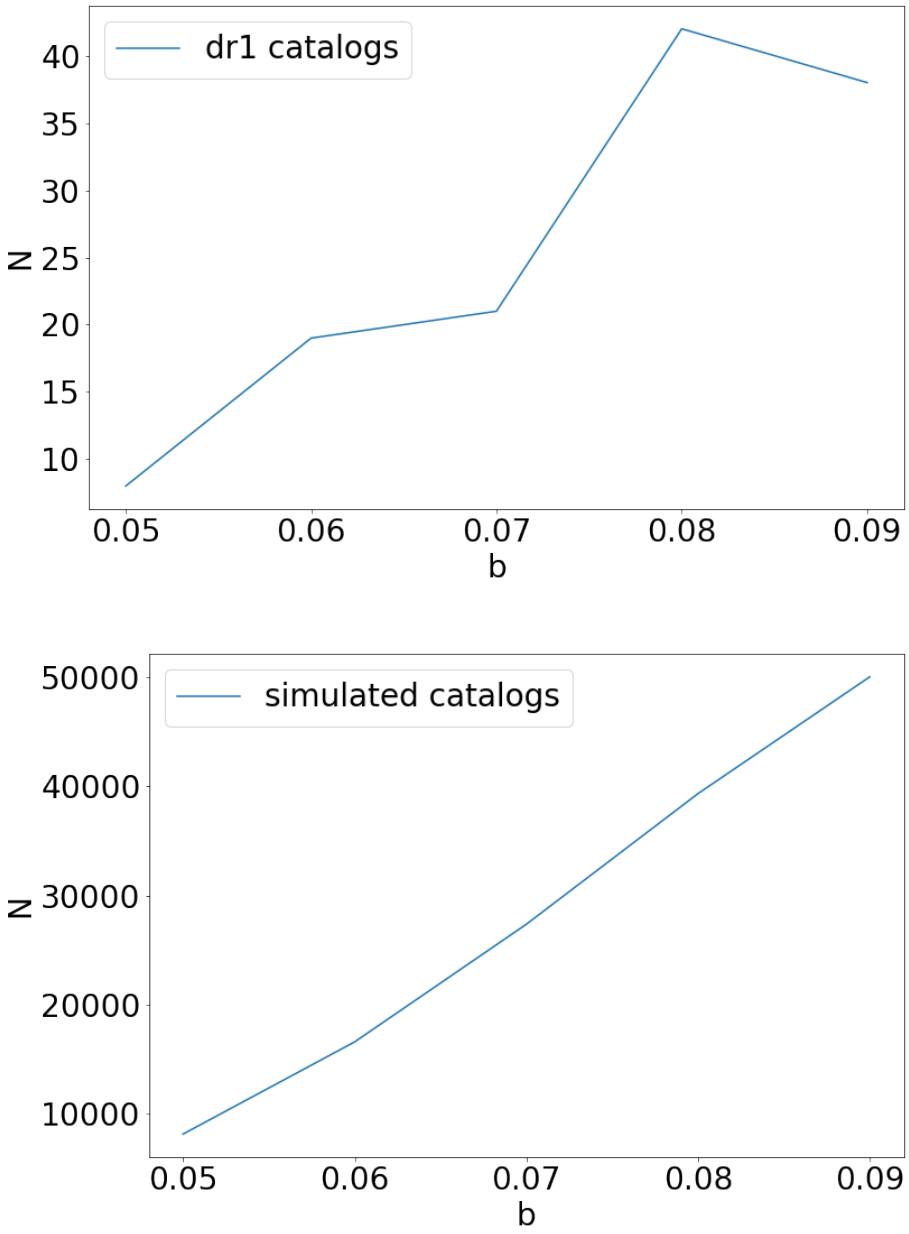


Figure 4.6. Distribution of the N parameter for linking length in the range $0.05 \leq b \leq 0.09$. The corresponding value of the mingal parameter is found from the b-mingal relation. It is shown the result for all 10 real images (*on top*) and for 10 simulated images (*on bottom*)

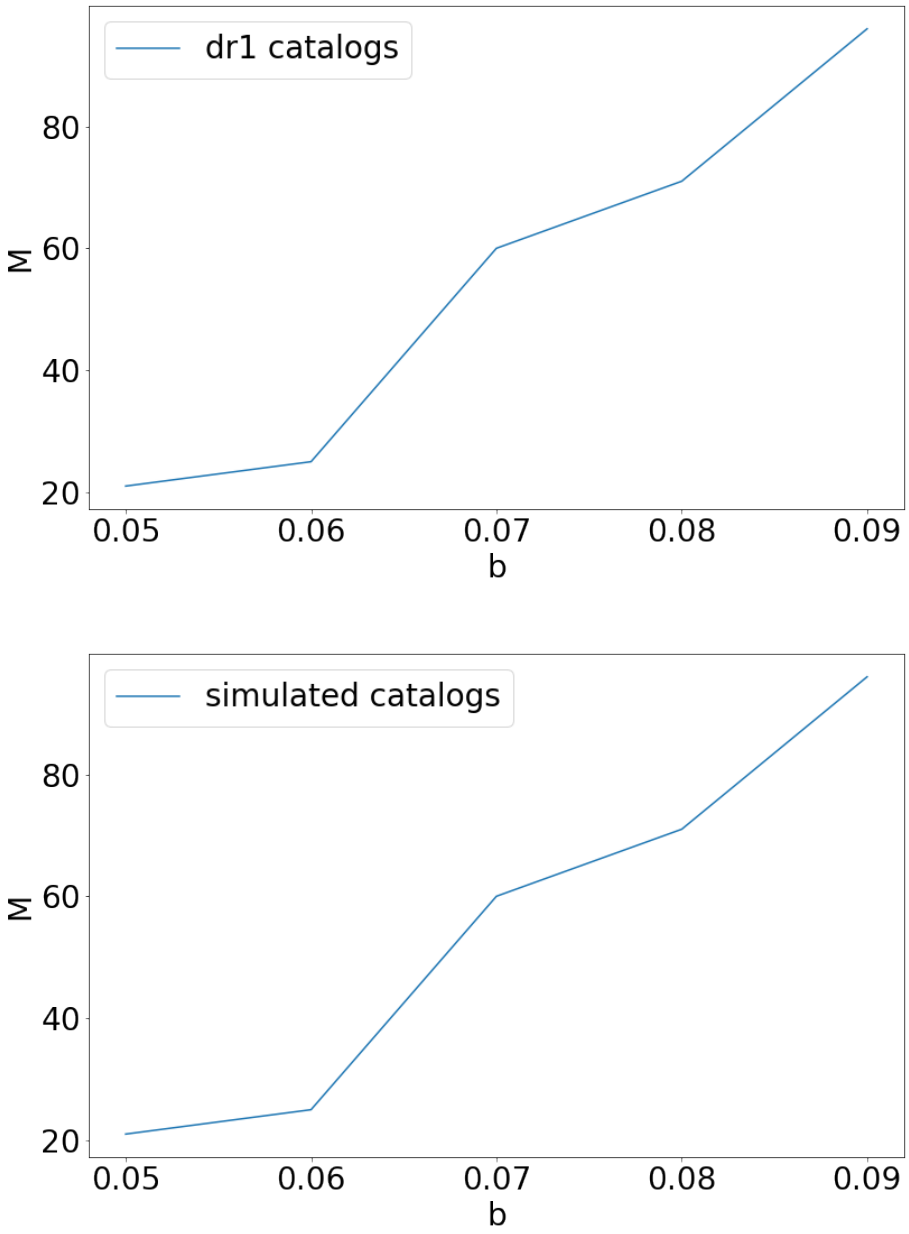


Figure 4.7. Distribution of the M parameter for linking length in the range $0.05 \leq b \leq 0.09$. The corresponding value of the mingal parameter is found from the b-mingal relation. It is shown the result for all 10 real images (*on top*) and for 10 simulated images (*on bottom*)

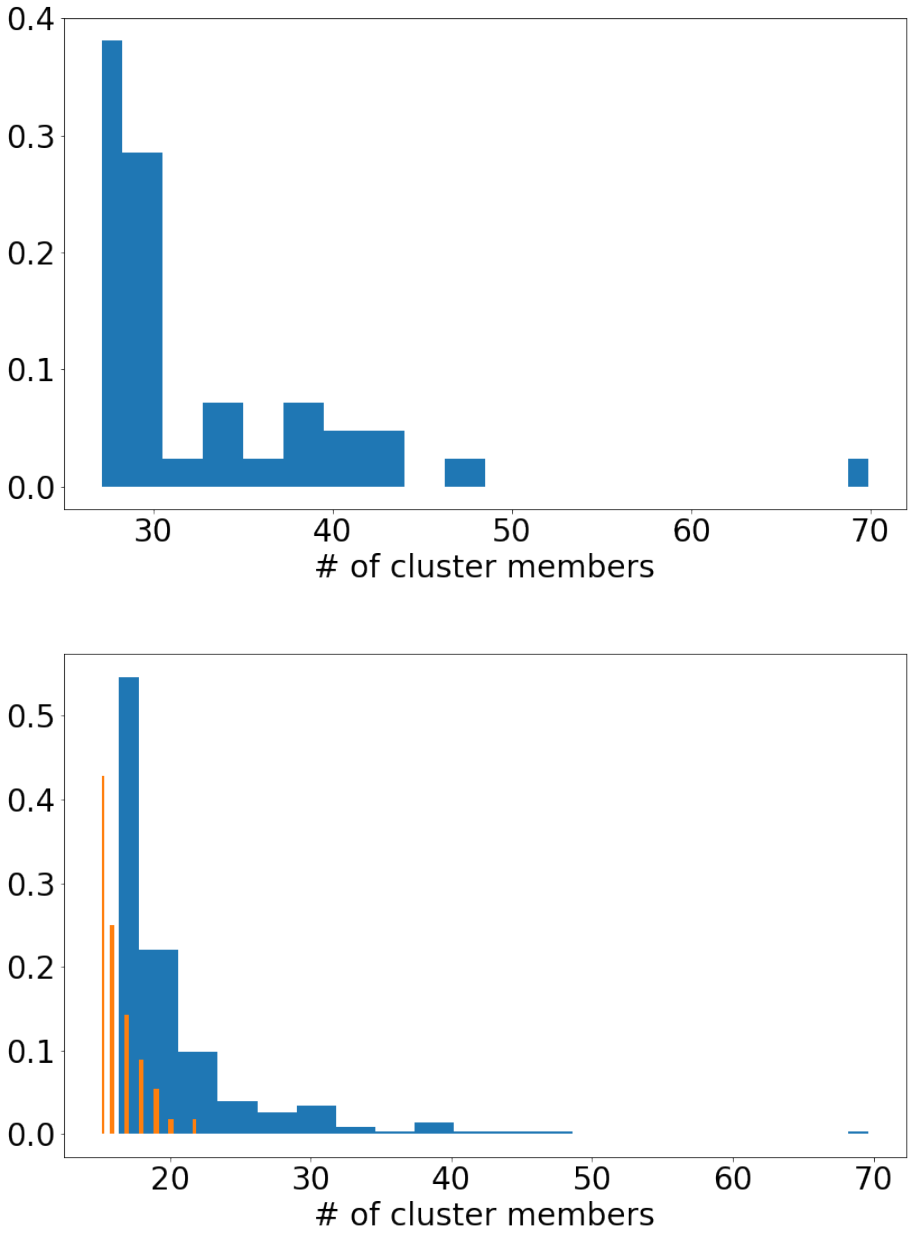


Figure 4.8. *Top:* (In cyan) distribution of the number of cluster members found for every cluster in all 10 dr1 images. The FoF algorithm was applied using b=0.08 and mingal=25, coming from the b-mingal relation. (In orange) the same distribution is shown for an application of the FoF to 10 simulated catalogs. No clusters are found in the simulated images as expected from the relation. The distribution is broad, however it seems to be peaked around 30-40 arcsecs, with a long tail engulfing larger radii. The peak is less prominent than the one in the distribution of the number of cluster members. *Bottom:* (In cyan) distribution of the number of cluster members found for every cluster in all 10 dr1 images. The FoF algorithm was applied using b=0.08 and an arbitrarily selected mingal=3. (In orange) the same distribution is shown for an application of the FoF to 10 simulated catalogs. The two distributions appear to be almost identical. The distribution for the simulated catalogs is shifted to lower numbers of members. It appears the smallest clusters to be systematically larger on the real catalog with respect to the simulated one.

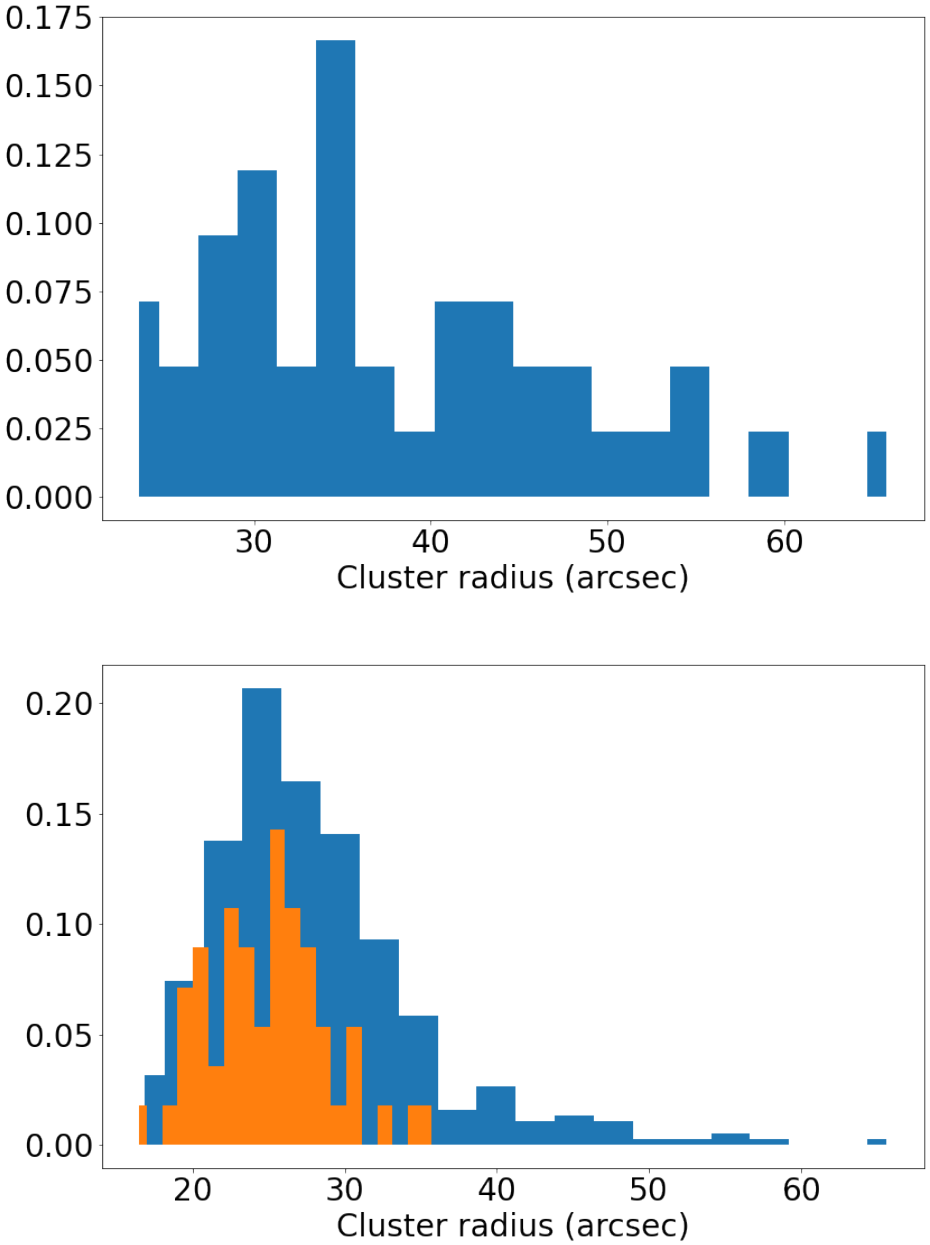


Figure 4.9. *Top:* (In cyan) distribution of the radius found for every cluster in all 10 dr1 images. The FoF algorithm was applied using b=0.08 and mingal=25, coming from the b-mingal relation. (In orange) the same distribution is shown for an application of the FoF to 10 simulated catalogs. No clusters are found in the simulated images. *Bottom:* (In cyan) distribution of the radius found for every cluster in all 10 dr1 images. The FoF algorithm was applied using b=0.08 and the minimum possible value of mingal=3. (In orange) the same distribution is shown for an application of the FoF to 10 simulated catalogs. The two distributions appear to be almost identical.

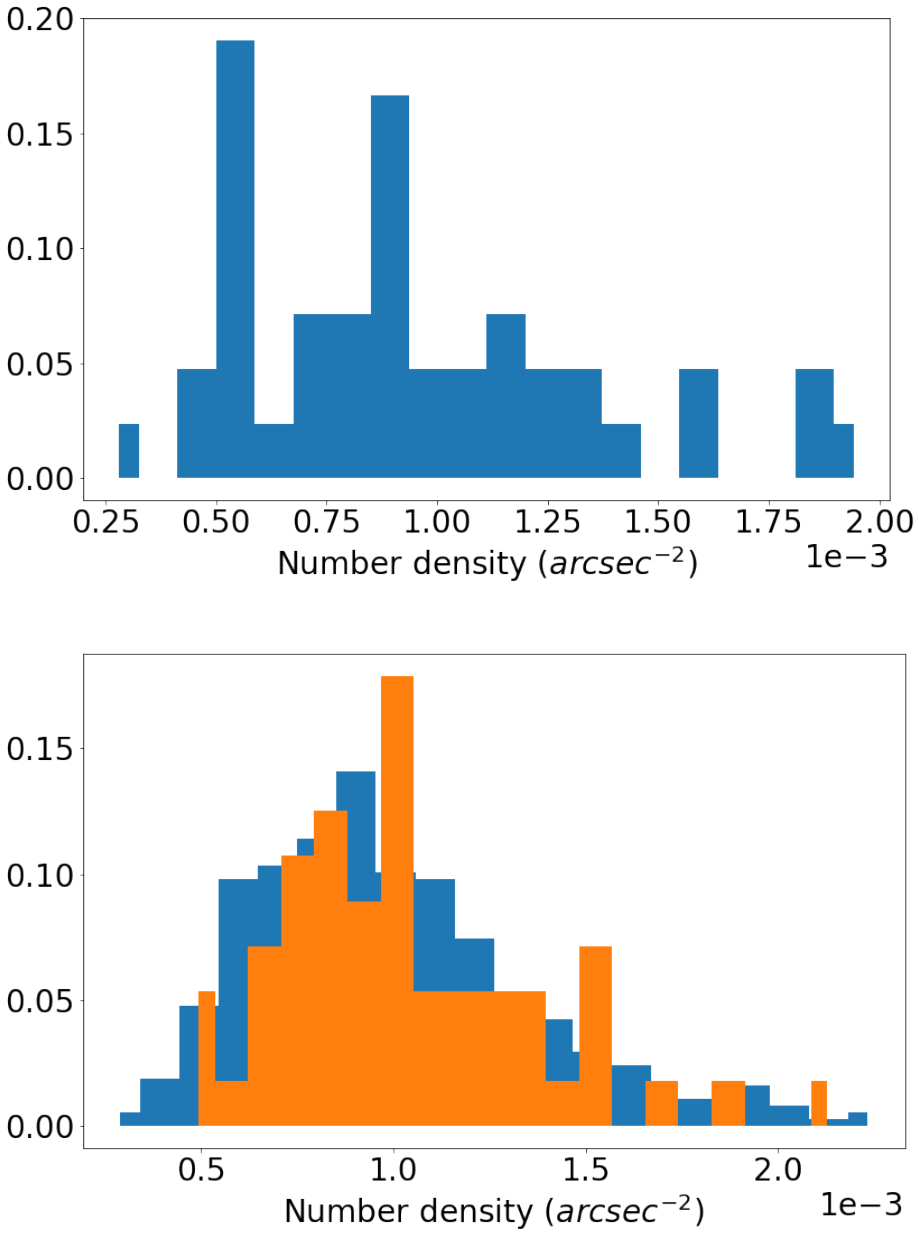


Figure 4.10. *Top:* (In cyan) distribution of the 2D galaxy number density for every cluster in all 10 dr1 images. This value is obtained dividing the number of members in the cluster for the projected area of the cluster. The FoF algorithm was applied using b=0.08 and mingal=25, coming from the b-mingal relation. (In orange) the same distribution is shown for an application of the FoF to 10 simulated catalogs. No clusters are found in the simulated images. Distribution is broad, reflecting the vast dishomegeneity of the candidate cluster radii. *Bottom:* (In cyan) distribution of the 2D galaxy number density for every cluster in all 10 dr1 images. The FoF algorithm was applied using b=0.08 and the minimum mingal=3. (In orange) the same distribution is shown for an application of the FoF to 10 simulated catalogs. The two distributions appear to be almost identical. Since both the radius and number of members have strongly peaked distributions, the same is true also for the 2D number density. Clusters found with the minimum possible requirement for mingal are then almost identical, very small random fluctuations of the spatial distribution of sources.

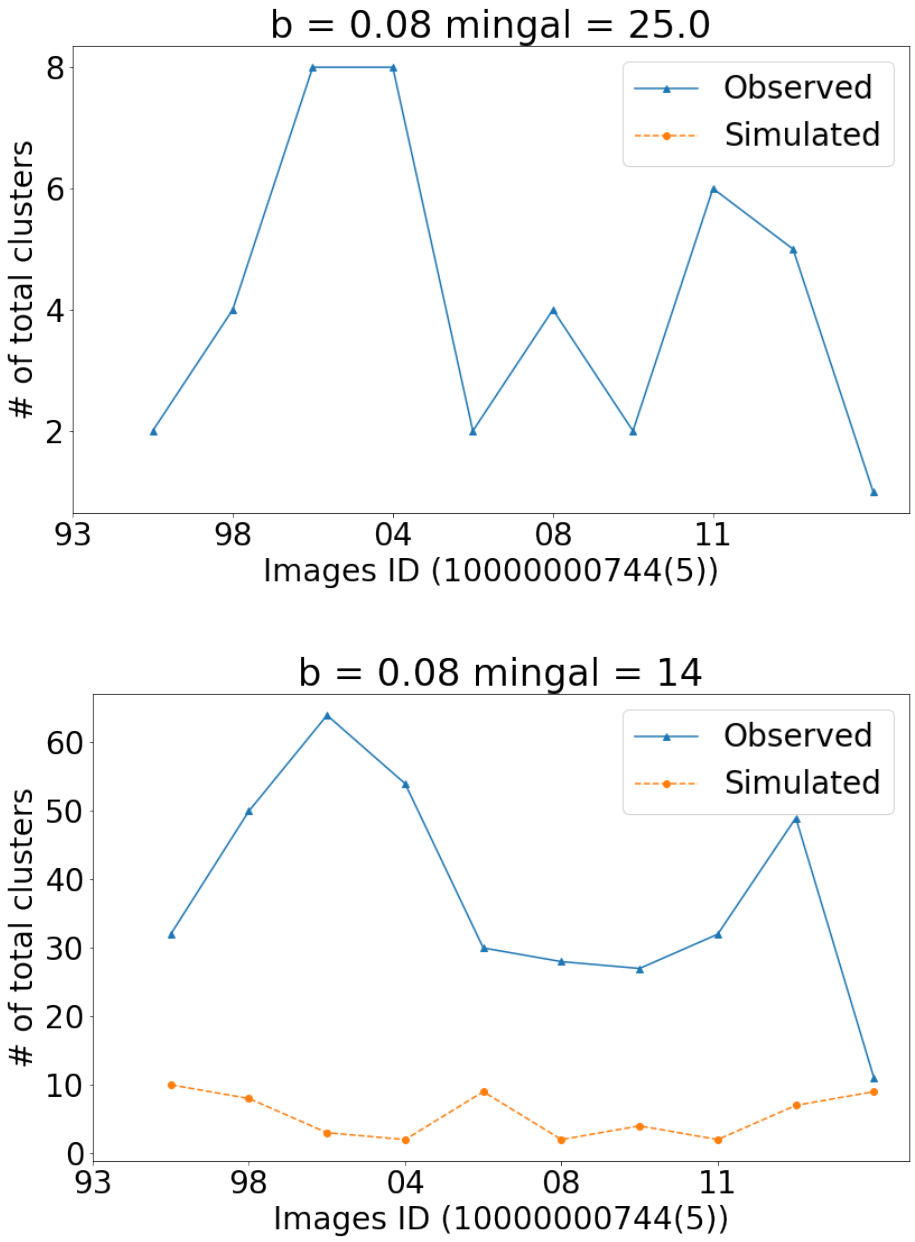


Figure 4.11. *Top:* (In cyan) distribution of the number of clusters found for each dr1 catalog applying the FoF using b=0.08. The mingal parameter is selected through the b-mingal relation to be mingal = 25. (dashed orange line) represents the distribution of the number of clusters found for 10 simulated images. No clusters are found using these parameters in the simulated images. *Bottom:* (In cyan) distribution of the number of clusters found for each image applying the FoF using b=0.08 and the minimum mingal=5. (dashed orange line) shows the same distribution obtained for an application of FoF to 10 simulated images. The two distributions seem to be very different. The number of candidates is extremely high for both . This is probably an effect of the value of b selected.

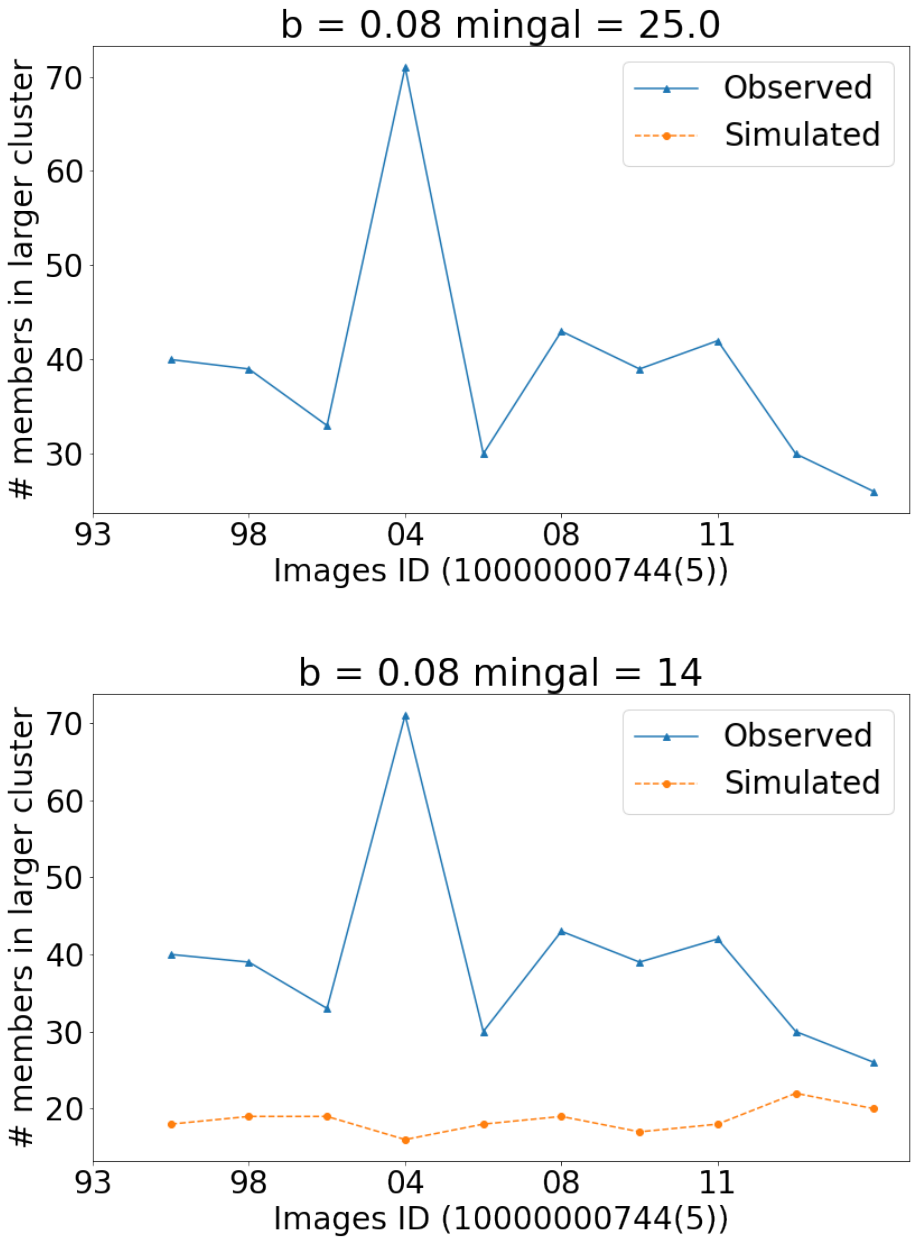


Figure 4.12. *Top:* (In cyan) distribution of the number of cluster members for the largest cluster found for each image applying the FoF using b=0.08. The mingal parameter is selected through the b-mingal relation to be mingal = 33. (dashed orange line) represents the distribution of the number of clusters found for 10 simulated images. No clusters are found using these parameters in the simulated images. *Bottom:* (In cyan) distribution of the number of cluster members for the largest cluster found for each image applying the FoF using b=0.08 and an arbitrarily selected mingal=5. (dashed orange line) shows the same distribution obtained for an application of FoF to 10 simulated images. The two distributions seem to be different from each other. This is probably an effect of the value of b selected.

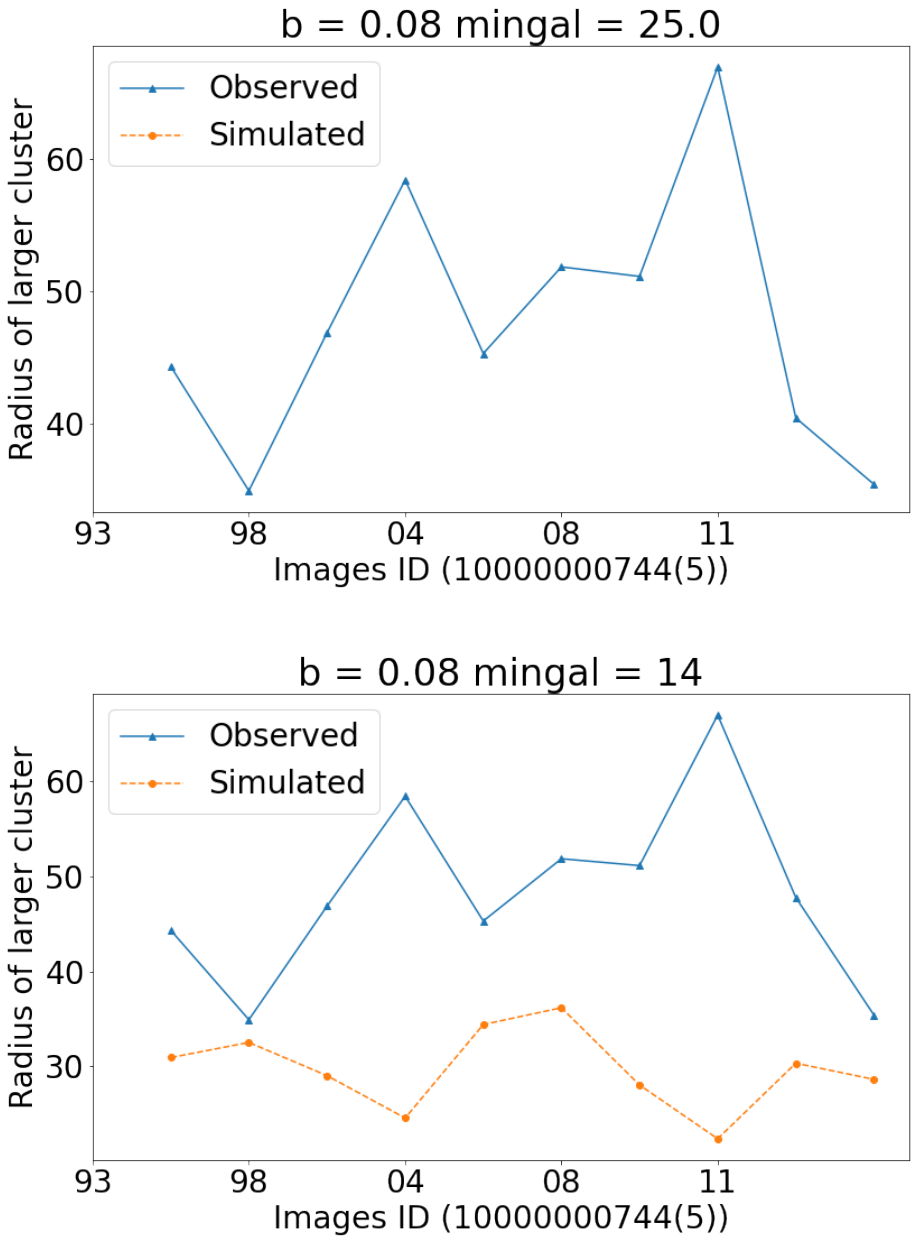


Figure 4.13. *Top:* (In cyan) distribution of the radius (in arcsec) of the largest cluster found for each image applying the FoF using b=0.08. The mingal parameter is selected through the b-mingal relation to be mingal = 33. (dashed orange line) represents the distribution of the number of clusters found for 10 simulated images. No clusters are found using these parameters in the simulated images. *Bottom:* (In cyan) distribution of the radius (in arcsec) of the largest cluster found for each image applying the FoF using b=0.08 and an arbitrarily selected mingal=5. (dashed orange line) shows the same distribution obtained for an application of FoF to 10 simulated images. The two distributions seem to be different from each other. This is probably an effect of the value of b selected.

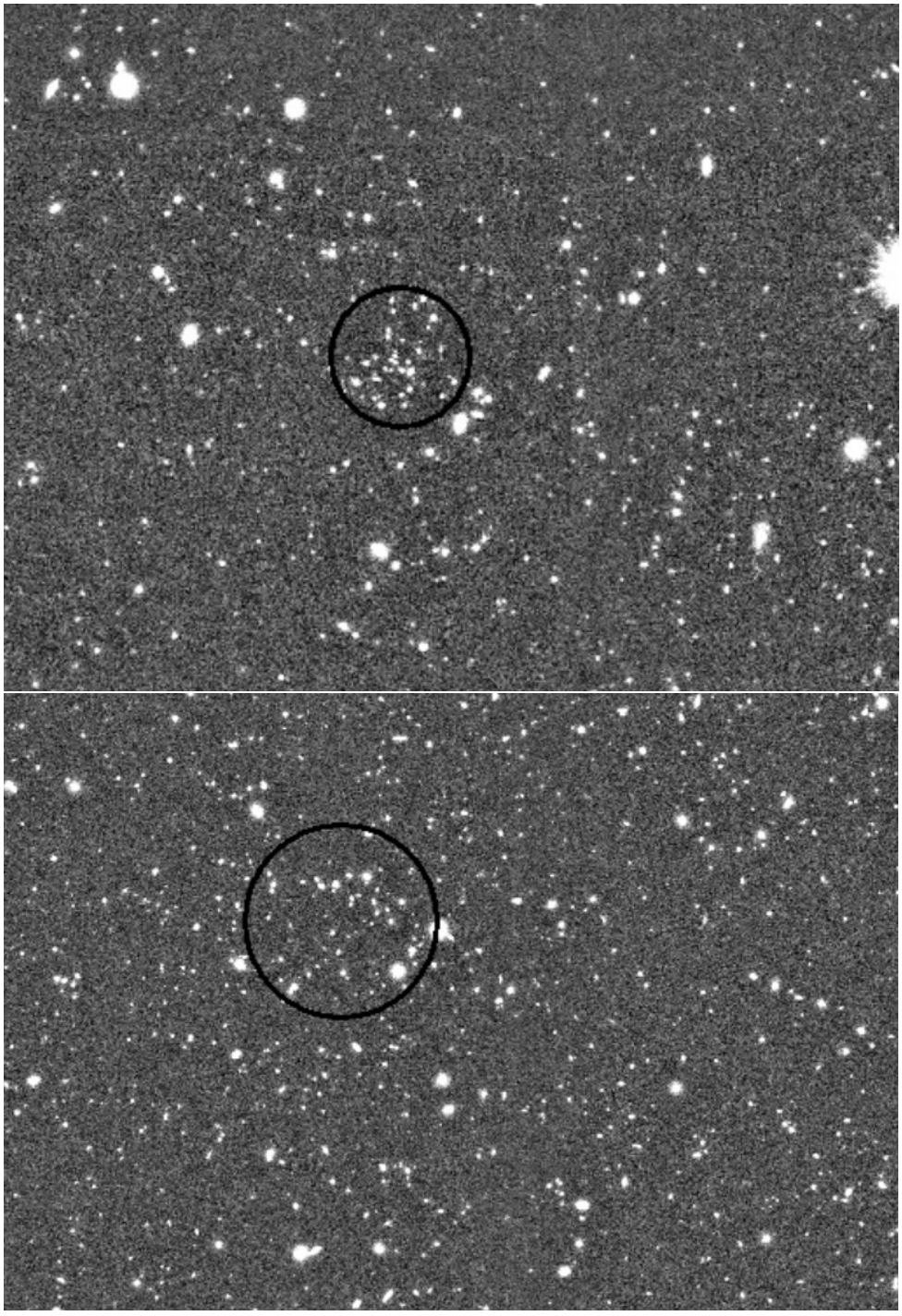


Figure 4.14. Example of candidate clusters from the application of the FoF algorithm using b = 0.08 and mingal=25 on catalog 1000000074498.

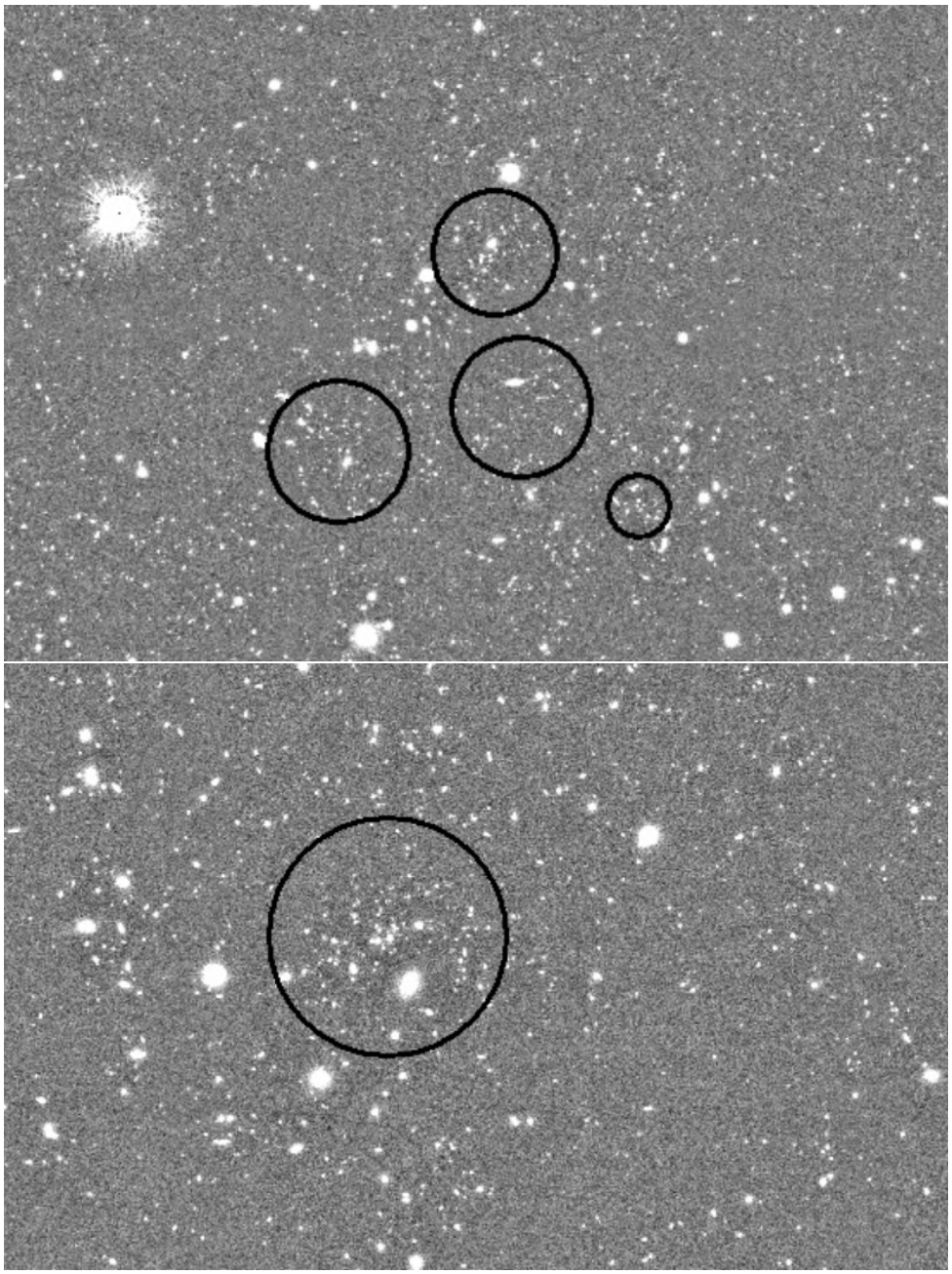


Figure 4.15. Example of candidate clusters from the application of the FoF algorithm using b = 0.08 and mingal=33 on catalog 1000000074501. *On top:* 3 very close candidates, probably part of the same larger structure. *Bottom:* the fourth very dense structure is presented.

Chapter 5

Increasing contrast: color selection of galaxy red sequence

As stated in the previous chapter, applications of the FoF algorithms to a single band survey can be constrained in order to obtain somewhat interesting results.
To constrain by trial the two parameters b and mingal is almost impossible: the projected distribution of galaxies have, at first approximation, similar properties to a uniform random distribution of the same number of sources in the same area.
Simulations can be used to constrain the parameter mingal to given values of b. This however strongly restricts the characteristics of the possible candidate clusters, requiring for them very high galaxy number densities.
More complex simulations can be constructed to explore the effects of two dimensional projection of the three dimensional cosmic structure within the observational limits of SHARKS.
However this is left for future works. A different approach is instead explored in this chapter.

To increase the efficiency of the FoF algorithm it is necessary to deproject the cosmic structure from the sky plane to a three-dimensional space. This can be done if the distance of each object from the observer is known. Distance can be determined once the redshift of the source is known. This requires spectroscopic or photometric multi-wavelength observations. Many different instruments have observed SHARKS sky area, so photometric measurements at different bands are already available.
Photometric redshifts [1] can then be constructed using templates of galaxies Spectral Energy Distribution (SED) derived from a combination of stellar evolution theory and observations.
In this work, however, photometric redshifts were not calculated. It was instead decided to estimate the redshift of each galaxy using a color-redshift relation. Once the spectrum of a galaxy is known, in fact, it is possible to apply to it the effects of cosmological redshift, simulating how the spectrum would appear if the corresponding galaxy was placed at different distances from us. Simulating the observation of the galaxy at different distances with the same photometric filter, the photometric measurement changes due to the effect of cosmological redshift on the observed light. The resulting color will also change. This means that it is possible to associate, for the same

[1] Flux measurements for different photometric bands can be performed for a given source. These fluxes can be fitted with a template model of the spectral energy distribution of a galaxy. The positions of characteristic features of the model depends, among other things, on the effects of cosmological redshift. Thus, the redshift associated with the observed object can be estimated. This will be influenced mainly by metallicity and age of the stellar population used for the modeling.

known spectrum, a color to each new distance of the galaxy and vice versa.
Having the color of a known galaxy, then, it is possible to estimate its redshift. This relation is useful since it requires information on only two photometric bands in order to calculate the color of an object. However, the object must manifest a known SED.
The solution is then to apply this relation to a class of galaxies for which we assume to know the typical spectral energy distribution.
As stated in chapter 1, the densest regions of clusters appear from literature to contain a larger fraction of early type galaxies with respect to the field. These galaxies are characterized by having a prevalent old stellar population, with low to no star formation. This, combined with the fact that they form in the densest region of the cluster (usually the core) indicates that these galaxies might be very old, forming together with the cluster, or at even earlier times.
Thus the spectrum of such galaxies have a characteristic slope due to the dominating old stellar population and can be used to establish a color-redshift relation even if the spectrum of each single object is different. Age and metallicity of the galaxy can influence its color significantly, however for the scope of this work the relation just needs to be indicative rather than extremely precise.
The color-redshift relation for red sequence galaxies was obtained using galaxy evolution models from professor Claudia Maraston (private communication). Models were generated using a dominant fraction of metal-rich stars together with a small fraction of metal-poor stars instead of assuming an on-going star formation. These models are calibrated on observations of luminous red galaxies from the Sloan Digital Sky Survey (SDSS) and the 2dF-SDSS LRG and QSO (2SLAQ) survey at z 0.4. Such empirical models seem to better reproduce observed data that theoretical ones, especially in r and i bands (Maraston et al. 2009).

Using the color-redshift relation it is possible to group galaxies according to the calculated redshift. The FoF algorithm can be then applied independently to each group.
The relation was calibrated for galaxies belonging to the red sequence. These galaxies are of the kind typically found inside cluster cores, so can be used as a tracer for clusters. Selecting galaxies in a narrow redshift bin, the number of sources to which the FoF is applied is much smaller than in a blind search.
However, in each redshift bin a mix of galaxy types can be found, since the relation is calibrated for the red sequence galaxies, but is applied to the whole survey, independently from the galaxy proprieties.
Two main advantages can be found sustaining this relation:

- The number of projected galaxies is reduced, in favor of red sequence galaxies, thus increasing the contrast between density of a cluster and density of the field.
- If a cluster is present in the image, then its red sequence galaxies are highlighted. Since they tend to reside in large fractions in dense environments, the chance of detecting the cluster increases.

5.1 Color-redshift relation

To calculate the color of SHARKS sources it was decided to use complementary photometric bands from the Dark Energy Survey (DES), which has a larger depth and so it is capable of providing fluxes for much of SHARKS sources.

DES area covers only partially SHARKS area. The only catalogs available in dr1 that contain information from both surveys are the four contiguous fields in SGP. Thus, analysis was performed only on ~ 8 squared degrees out of the 20 available.
DES provides photometry in five broad optical bands (g,r,i,z,Y), while SHARKS use infrared Ks band. The Y-Ks color was used for the color-redshift relation, since it is the combination that better shows the red sequence of galaxies in a color-magnitude diagram (Y-Ks over Ks). In *Figure 5.1* it is shown the distribution of SHARKS galaxies in a Y-Ks over Ks color-magnitude diagram for the image 1000000074498. A narrow sequence of galaxies is evident around the color Y-Ks = 0, which represents the red sequence of local galaxies. The color map of the image represents the number of galaxies for each bin of magnitude and color.

Figure 5.2 shows the color-redshift relation overlapped to the color-magnitude diagram. Each horizontal line represents a redshift value and it is associated with a color value. Redshift values are calculated with a separation of z = 0.3 in redshift space. It is important to notice that the selected bins are equally spaced in redshift space, but not in color space. A reason for this is the fact that the effect of cosmological redshift, responsible for the color-redshift relation, is dependent on the slope of the spectral energy distribution of the galaxy, which is not constant. Another effect that could change the spectral slope is the evolution of the stellar populations with redshift. Due to the evolution, in fact, as the stellar population grows older, the blue part of the spectrum tends to decrease in intensity in favour of the redder regions.

5.2 Further stellar component removal with two-color diagram

A brief discussion is to be made on the removal of stellar components from the catalog of galaxies. As seen in chapter 3, foreground stars are removed from SHARKS catalog using a procedural method that involves different parameters to highlight the symmetry and concentration of a source. However this process removes stars up to a magnitude level of Ks = 20.56, far from the completeness level of Ks $\sim$ 22.
Many faint foreground objects are thus not completely removed from the survey catalogs. This is confirmed by the color-magnitude diagram in *Figure 5.3*. In the figure it is present a very narrow sequence of points at negative values of Y-Ks color. This is assumed to represent a sequence of bright stars, which increases its dispersion going towards larger magnitudes.
The two color diagram in *Figure 5.4* confirms the presence of a tight sequence of sources. The diagram is constructed using Y-Ks and I-Y colors. A very distinct sequence of sources appears for negative Y-Ks color, together with a wider and redder distribution. This narrow sequence is confirmed to contain the same objects residing in the narrow negative sequence on the color-magnitude diagram.
Using the two-color diagram is then possible to further remove stars up to the completeness level.

5.3 The cluster catalog

5.3.1 Application of FoF to color selected galaxies

For each redshift bin I applied the algorithm using the same parameters b and mingal. It is important here to remember that b is not an absolute value, but represents a percentage of the average

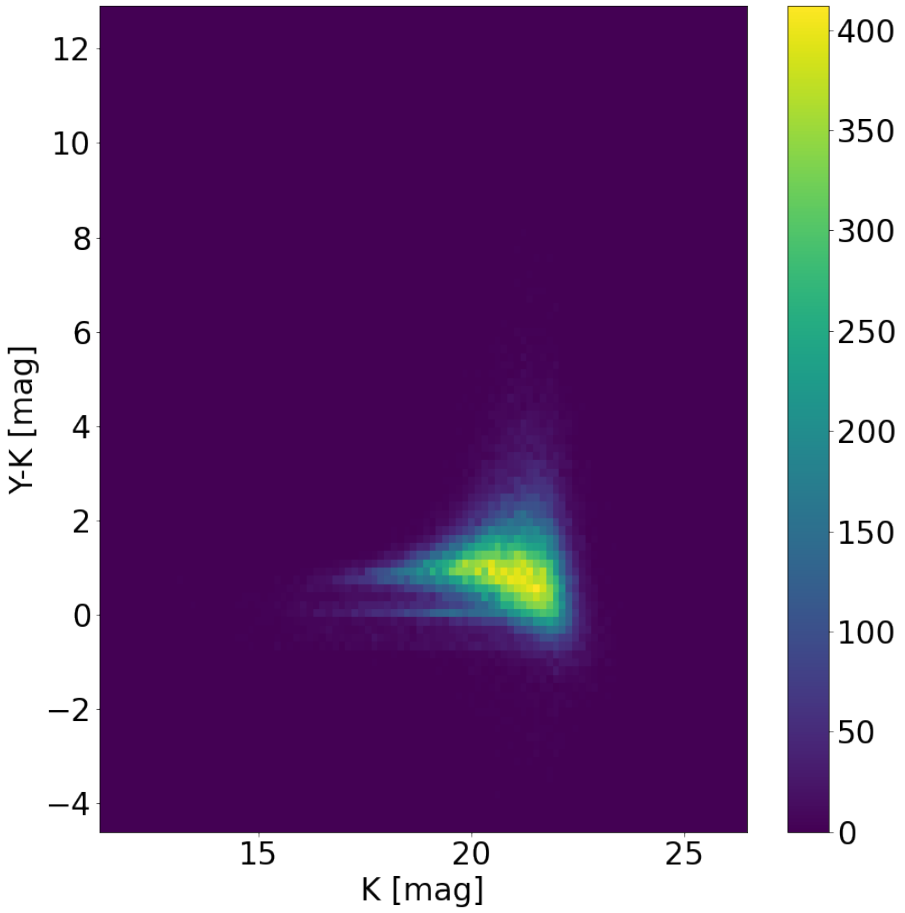


Figure 5.1. Color magnitude diagram of catalog 1000000074498 constructed from SHARKS Ks band and DES Y band. The narrow sequence around Y-Ks = 1 is probably the red sequence of local galaxies. At lower color a very tight line of sources can be observed, representing faint stellar component that was not removed with previous methods.

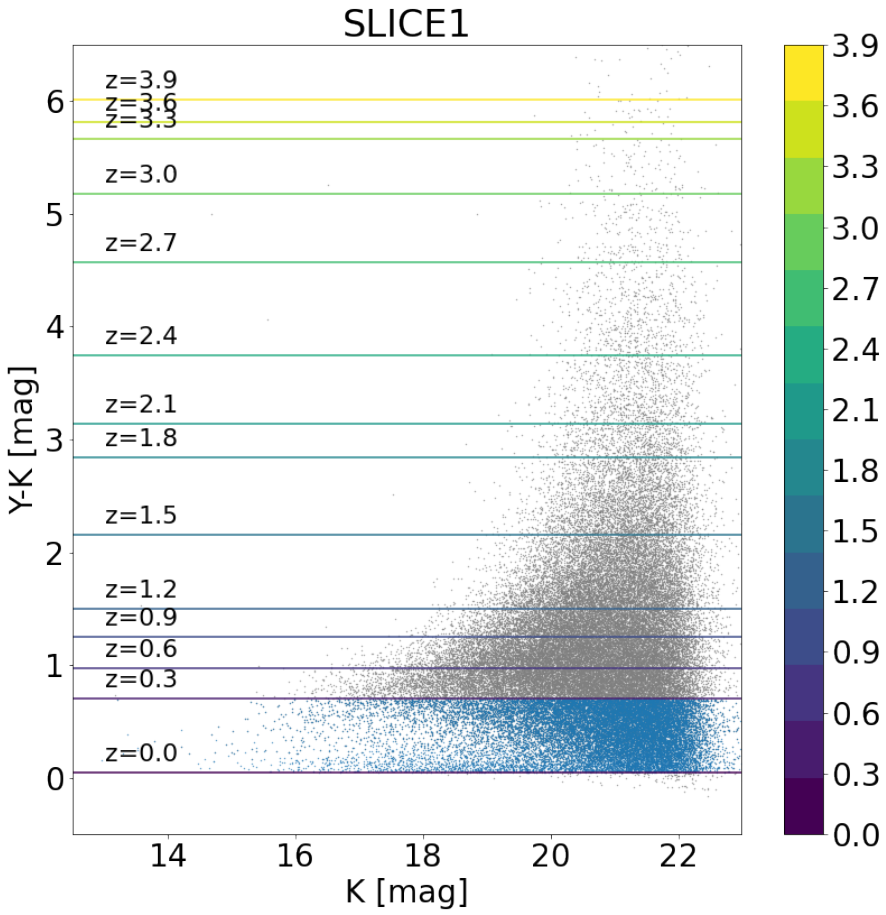


Figure 5.2. The color-redshift relation calibrated using spectral energy distribution models for red sequence galaxies. The relation was applied up to z = 4. The width of redshift bins is of Δz = 0.3, but due to the non constant slope of the galaxy spectrum, color spacing appears irregular. Points in blue represent galaxies falling in the redshift range $0 < z < 0.3$.

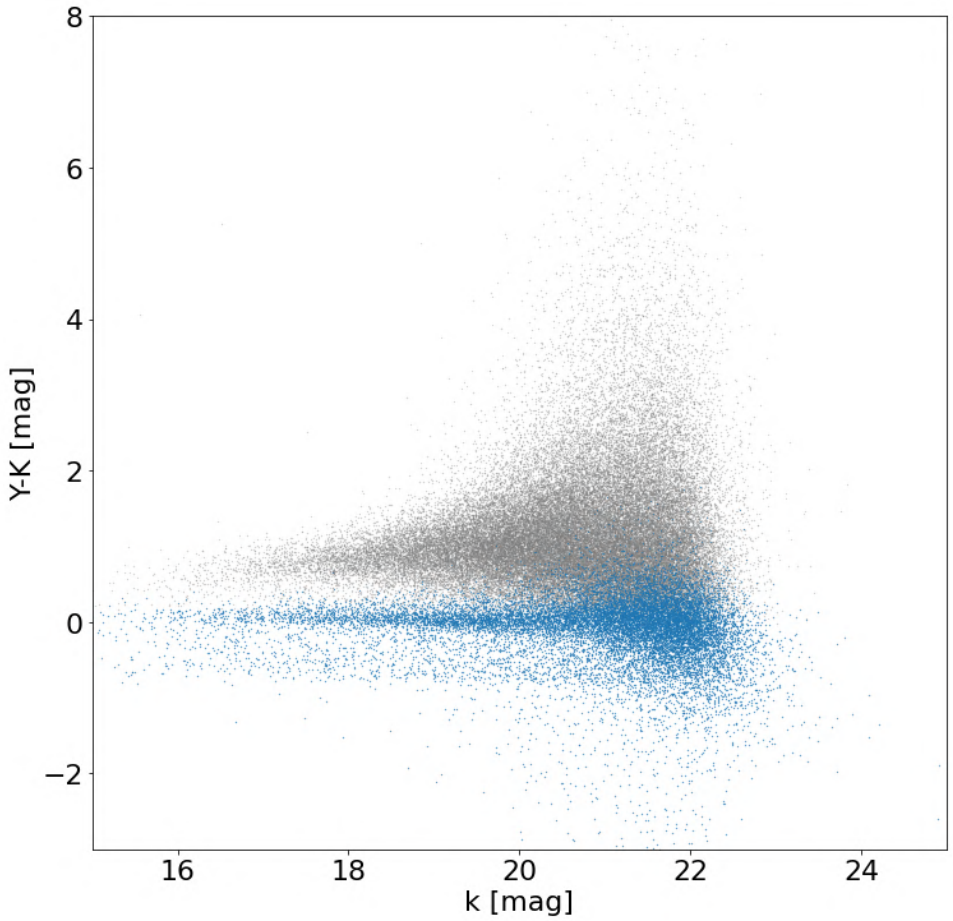


Figure 5.3. Color-magnitude diagram. Color Y-Ks is constructed using Y band from DES and Ks band from SHARKS. In blue are shown points corresponding to the tight sequence of sources in the two-color diagram. These sources have been interpreted as residual stellar components which were not removed with the combination of stellarity index and half-light radius.

distance between galaxies in the catalog. This means that the definition of clusters adapts with the redshift selection. At high redshifts in fact the number of galaxies in a bin is usually smaller than at lower redshifts, however galaxies are required to be less close together to form a candidate cluster.

For the redshift selected galaxies is not possible to generate a relation between the linking length (b) and the minimum number of galaxies (mingal) as it was done for a blind single band search. The selection of galaxies, in fact, applies new information on the catalog that cannot be reproduced by just randomly rearrange the position of the original sources. More complicated simulations would be needed to account for the color-redshift selection on the simulated data.

A blind application of the FoF algorithm for each color bin of SHARKS catalog was then performed. Variations of b and mingal parameters produce distributions of the number of clusters found very similar to the ones obtained from the unbinned catalog. As already verified in chapter 4 it is then of little use to search for the right combination of parameter using these distributions. Calibration of the FoF on the binned catalogs was then performed visually, searching for the combination of parameters that provide the most convincing results.

The visual calibration took advantages of some collection of galaxies in image 1000000074498 that were identified from the much more restrictive application of the FOF that used simulations of the catalogs to estimate the mingal parameter for each value of b. Such structures were visually

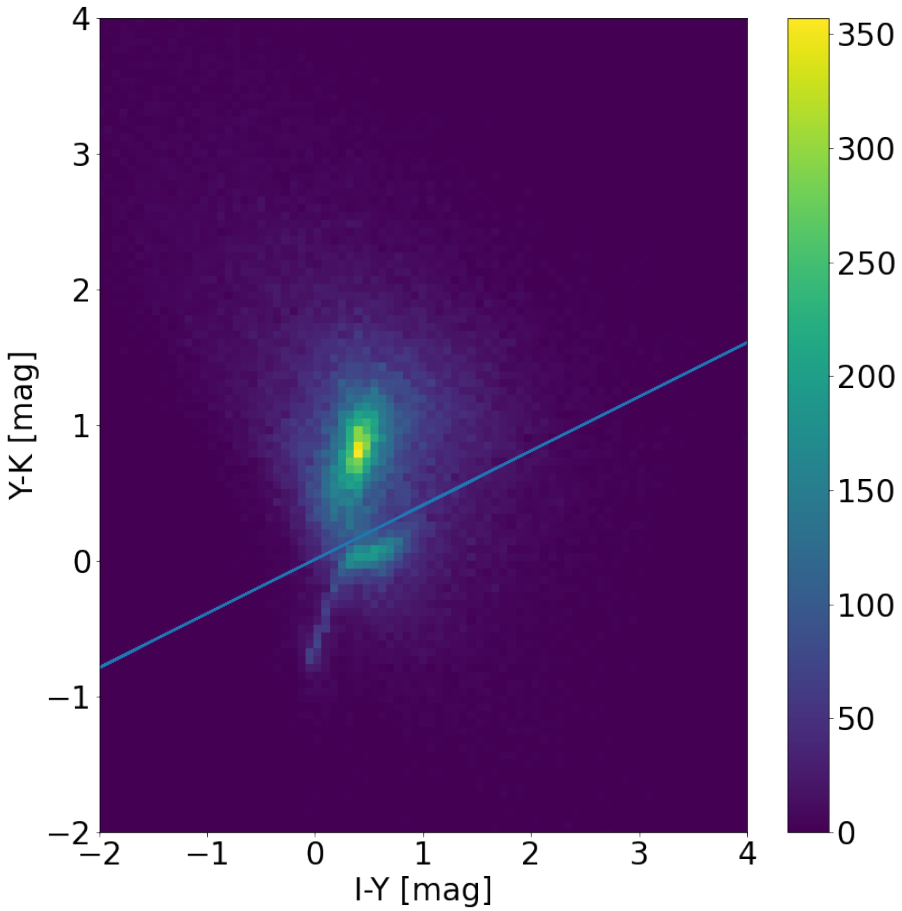


Figure 5.4. Two-color-diagram obtained combining SHARKS Ks band with DES Y and I bands. A narrow distribution of points is observed for negative Y-K values: this is the faint stellar component that was not removed with previous methods.

confirmed following two main criteria:

- a large number of sources in a restricted area
- a similar angular dimension of the sources, suggesting they might be at a similar distance from the observer.

Examples of such evident structures found on SHARKS images are shown in *Figure 5.5*. The best values considered for linking length and minimum number of cluster members are the ones for which the algorithm is able to find all the clusters found with the more restrictive application and visual inspection.

A good choice of b and mingal parameters must be reasonable in terms of what already explored in the blind search.
This means that it is expected for mingal to be much smaller than what obtained applying the FoF to the unbinned catalogs together with the b-mingal relation.
The reason for this is that in the color binned catalogs the average number density of sources is much smaller than in the complete catalog since there are less galaxies for the same area covered by the catalog.
For the same reason b parameter is expected to be larger than in a search over the entire catalog.
Moreover, a good parameter choice is expected (but not necessarily required) to identify the visual

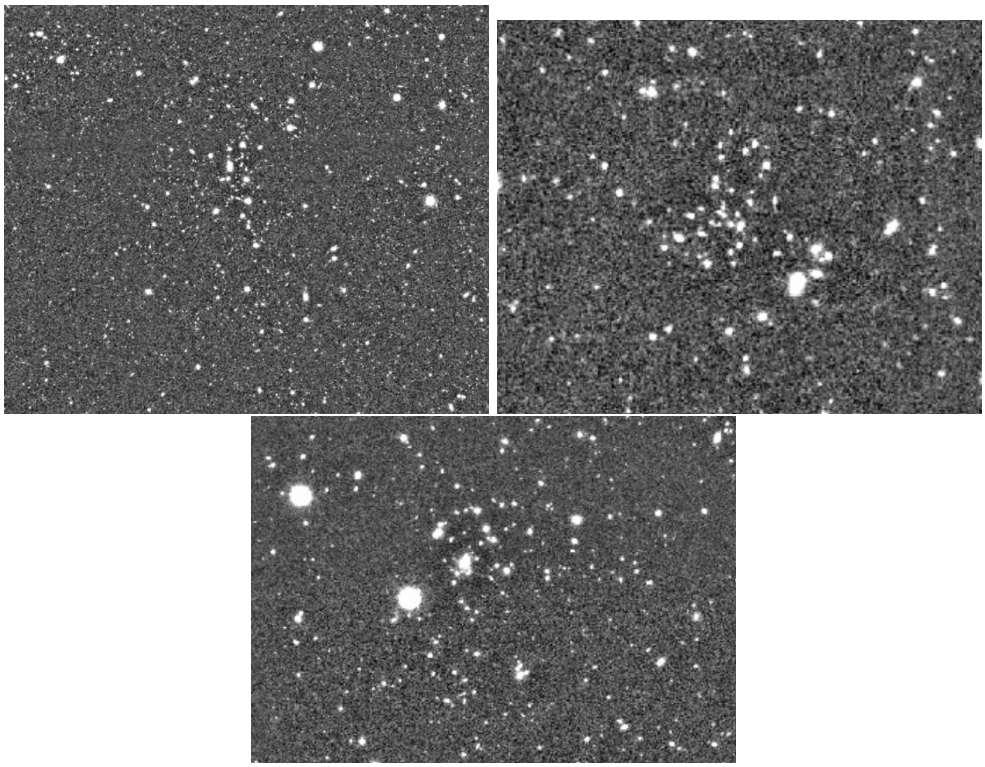


Figure 5.5. Calibration structures used for the selection of the best b and mingal parameters for the FoF algorithm in the redshift binned catalog 1000000074498. Such structures were found upon visual inspection of the original image. They consist of a large number of objects in a rather small area (of the order of 1 arcmin). For these reasons are considered to be possible candidate clusters.

observed structures.

The best observed combinations of parameters is then found to be:

$$b = 0.1$$

and

$$mingal = 1.0$$

Results however depend on the spacing adopted for the redshift binning of red sequence galaxies. Larger the spacing, larger the number of clusters found.
Upon visual inspection of the candidate clusters detected using different Δz as redshift binning of the catalog, it was decided to use a spacing of Δz=0.3. For lower values of Δz, due to the low precision of the color-redshift relation, large and convincing structures are not individuated by the algorithm. They are instead fragmented in various redshift bins and the minimum number of galaxies to form a structure is not reached for a given redshift.
Larger values of Δz increases the number of found clusters, but their radius become much larger and in some cases engulfs the entirety of the survey area, meaning that the average number density of the selected sources is small enough to be compared with the linking length.
All found clusters are listed in *Table 6.2*. Additive information and images can be found in *Chapter A*.

5.4 Other cluster catalogs in the literature

In the following chapter I will describe the catalog of candidate galaxy clusters and proto-clusters obtained from analysis of SHARKS data. This will be confronted with 3 other catalogs of clusters from the literature.

- The Kilo-Degree Survey (KiDS) was performed at the VLT Survey Telescope (VST), a 2.6 m ESO telescope at Paranal, with the instrument OmegaCAM. The survey covered 2 areas of 750 square degrees each in u,g,r,i bands reaching a median redshift of 0.7. It reached a 5σ depth in AB magnitude in i band of 24.2 mag over an aperture of $2''$. Unfortunately there is no overlap between the area covered by KiDS and the one observed by SHARKS.
  A catalog of galaxy clusters was obtained over an area of 114 square degrees for a redshift interval $0 \leq z \leq 0.7$.

- The Dark Energy Survey (DES) was performed from Cerro Tololo Inter-American Observatory (CTIO) using the Dark Energy Camera (DECam), which is mounted on the Victor M. Blanco 4-meter Telescope. The survey covers a contiguous area of 5000 square degrees using g,r,i,z,Y photometric filters. The catalog of candidate clusters contains $\sim$100,000 objects in a redshift interval $0.1 \leq z \leq 1.5$

- The Planck space mission's main goal was to study in detail the Cosmic Microwave Background. The satellite used a Low Frequency Instrument (LFI) spanning the frequency region between 30 GHz and 70 GHz ($\sim$4 mm to $\sim$10 mm), a High Frequency Instrument (HFI), which operated between 100 GHz and 857 GHz ($\sim$3 mm to $\sim$0.3 mm). The sub-millimeter region observed by HFI is useful for the discovery of strongly dust-obscured objects. These

include distant young galaxies with large content of gas and cold dust, which is heated by the UV radiation coming from massive newly formed stars. These galaxies are of the type we expect to find in dense and active regions as distant proto-clusters, which seem to manifest high fractions of star bursting objects almost invisible to optical light. Due to the limited size of Planck primary mirror (1.5 m) and to the large wavelength observed, the resulting resolution of Planck's survey is quite modest, about $\sim$5 arcmin. Planck High redshift sources that show a strong flux in sub-millimeter light are called Cold Extragalactic Sub-Millimiter Sources. They seem to be associated with emissions from dense and dusty environments. In Negrello et al. 2017 it is explored the possibility for these sources to have originated from proto-clusters at different redshifts projected along the line of sight. The collection of projected smaller structures is smoothed out by the low resolution of the instrument and appears as a larger single source. Different methods can be used to disentangle these sources, decreasing the confusion effect. However a resolved observation of such regions is always the best way to determine their nature. With its large survey area, SHARKS is capable of observing many of Plank's cold extragalactic sub-millimeter sources. The Ks band observed by SHARKS is practically not affected by dust absorption nor by dust infrared emission, which starts to be consistent around $\sim 8\mu$m. For this reason SHARKS survey is capable of piercing through dust, revealing the light of heavily obscured galaxies. This, together with the good resolution of the telescope, makes the survey capable of producing important insights on Planck sources, possibly resolving the structures responsible for the Cold-Sub-Millimeter sources. Studies of these objects within a large and consistent survey as SHARKS would provide further statistical information for a general comprehension of their nature and origin.

Chapter 6

The SHARKS Cluster Catalog

In this chapter I present a description of the resulting SHARKS candidate clusters catalog.
The catalog covers an area of ∼8 square degrees in the SGP region. Candidate clusters are found through the application of a Friends-of-Friends algorithm on redshift-selected volumes of the observed sky. Redshift for each source is calculated from a color-redshift relation using Y-Ks color. Y band comes from the Dark Energy Survey, while Ks band comes from SHARKS.
The catalog contains 186 clusters in an area of ∼8 square degrees, meaning an average of 23 candidate clusters for each square degree. We perform here a rapid confront of this statistics with three other catalogs from literature.
A resume is visible in *Table 6.1*

DES: was performed on an area of ∼5000 square degrees to an AB photometric depth at 10σ of 20.1 mag in Y band. The magnitude of the Brightest Cluster Galaxy (BCG) is expressed in the catalog in i band, for which photometric depth in AB magnitude is of 22.5 mag. For a comparison with SHARKS results only DES clusters with BCG magnitudes < 22 mag were used. The total number of clusters in the DES catalog is 151244, with 138146 clusters with BCG magnitudes $<$ 22 mag. This means 27.6 clusters per square degree. However the catalog contains clusters in the redshift range $0.1 \leq z \leq 1.5$, while SHARKS catalog covers a redshift range $0.6 \leq z \leq 3$.

KiDS: The Kilo-Degree Survey cluster catalog covers an area of 114 square degrees containing 1858 galaxy clusters in a redshift range $0 \leq z \leq 0.7$. Clusters with BCG magnitudes < 22 mag are 1543, for an average of 10.7 clusters at the same depth of SHARKS catalog.

Planck high z: The catalog contains 2151 candidate sources at redshift $z \geq 1.5$ in an area covering 26% of the sky. Approximating this area to 158665 square degrees, we obtain an average of 0.2 sources per square degree.The depth limit of this survey is a flux density of 500 mJy at 545 GHz. This corresponds to a flux of 4.85e-7 mJy inside an aperture of 2″ which is equivalent to the aperture used to calculate the depth of SHARKS catalog. Converting to AB magnitudes we end up with a limiting magnitude of 24.6 mag. This would imply that the number of sources would be even more reduced if we account for SHARKS depth. So the number of objects per square degrees in the SHARKS cluster catalog is highly inconsistent with the one resulting from Planck high redshift catalog of extragalactic sources.

Survey name	Catalog coverage (squared degree)	Number of clusters per square degree
SHARKS	8	23
KiDS	114	11
DES	5000	27
Planck	158665	0.2

Table 6.1. Comparison of the number of clusters per square degree between SHARKS and other cluster catalogs from literature

6.1 Color of cluster members

In *Figure 6.1* color-magnitude diagrams are presented for some clusters identified in image 1000000074498. Each diagram shows the Y-Ks color index for all sources in the area covered by the cluster, including all foreground and background galaxies that are projected on the sky plane together with the actual cluster members identified by FoF. Cluster members are highlighted (in orange) in each plot.
It is possible to see how the typical color-magnitude diagram of a candidate cluster tends to resemble the color-magnitude diagram of the entire catalog of sources. This is due to foreground and background sources having different redshift and different spectra than the red sequence galaxies used to identify the candidate.
However the spread in color of sources in the cluster area is smaller than the one obtained for the entire catalog. This is expected if the dominant fraction of sources in the cluster area is constituted by cluster members, so red sequence galaxies with very similar redshifts. This means that the observed structure is dominated by galaxies that are physically close together and do not appear clustered by exclusively a projection effect.

6.2 Flux distribution

The total flux for each cluster is calculated considering at first the sum of the fluxes of all cluster members detected by the FoF algorithm. Then a second flux is calculated considering the sum of fluxes coming from all projected galaxies that resides in the area spanned by the cluster.
A simple counting function was used to confront the contribution of foreground and background sources with respect to cluster members. The result is plotted in *Figure 6.2*. As expected, the counting function that considers fluxes from all objects projected in the cluster area is peaked to larger values of flux. This is due to the contribution of galaxies along the line of sight and of galaxies that are part of the candidate cluster, but that are not in the galaxy red sequence, which fluxes sum up with the fluxes from the candidate cluster members. The peaks of the two distributions are at a distance of $\sim$0.4 dex.

6.3 Redshift distribution

The redshift of each cluster was calculated with an interpolation of the color-redshift relation. Average color of the cluster members were computed for each cluster and used in the interpolated relation to obtain the corresponding average redshift.
The error on redshift is calculated as half of the redshift bin used to select galaxies for the

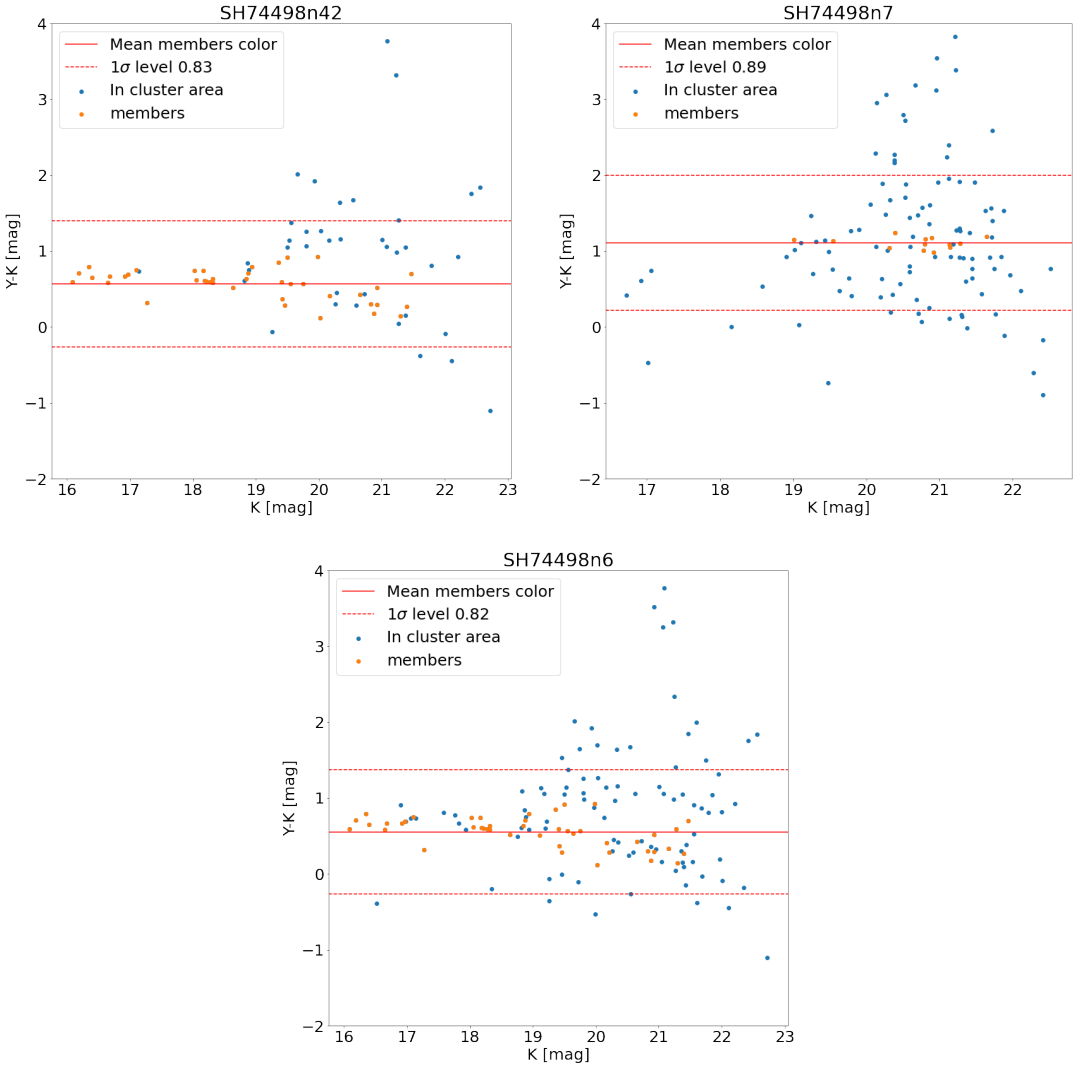


Figure 6.1. Color-magnitude diagram for the 3 calibration images shown in *Figure 5.5*. The order of candidates is maintained the same. Mean color is traced by the continuous line, while dashed lines show the level of dispersion at 1σ. Candidate SH74498n7, whilst being convincing at first glance, seems to have a low number of cluster members and many objects of very different colors projected on its surface. This would suggest that such a candidate is not real, rather a projection effect. However, the associated redshift of its members is z=1.4, meaning that a lot of contamination is possible by the large number of galaxies that can reside between us and the cluster. Moreover, it is possible that the low amount of red sequence galaxies found in the structures comes from Malmquist bias. This could justify the large dispersion in color.

application of the FoF. In this case it is $\Delta z = 0.3$, so the error is $z_{err} = 0.15$.

The distribution of the clusters with redshift is shown in *Figure 6.3*. As one could expect, the number of detected candidates gets lower with redshift. There seems to be no cluster identified for $z < 0.2$. In this region, in fact, we can find local clusters, usually characterized by very large angular sizes and by low surface number densities.

These characteristics makes these kind of structures difficult to observe within the limits of the FoF application due to the large average number density of objects in the survey.

The most distant candidate clusters in the catalog are found to be at redshift z∼3. This is consistent with what predicted by the survey itself, that aimed to reach a depth of z∼3 for a fair number of galaxies and then to go even further observing the most luminous and active objects.

We can see a gap on the distribution around z∼2, however this seems to change slightly with

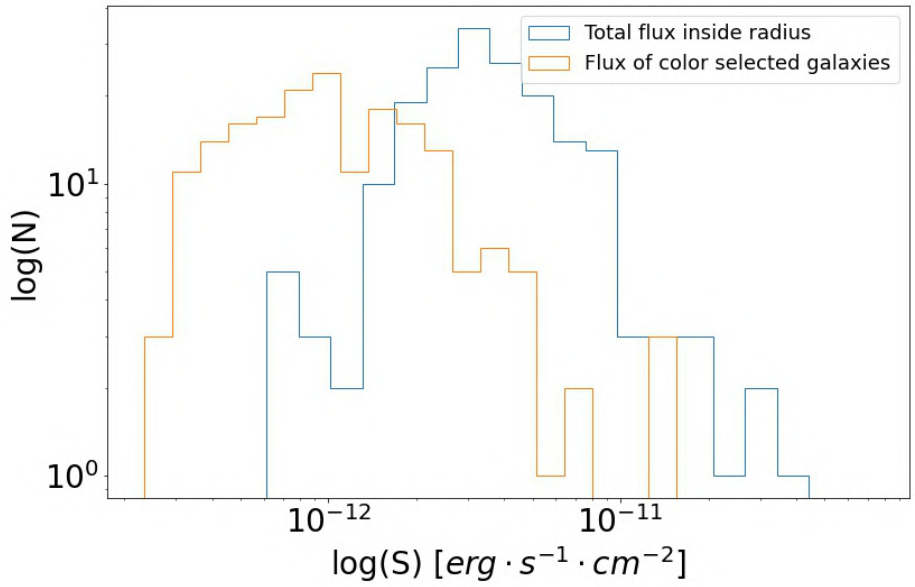


Figure 6.2. Normalized counting functions for SHARKS galaxy clusters in log-log space. In *cyan*, fluxes are calculated summing fluxes over all sources projected on the cluster area. In *orange* fluxes are calculated summing exclusively over the cluster members.

different values of Δz. For this reason such a gap is probably just a statistical effect. Increasing the area of the SHARKS cluster catalog could lead to discovering more sources at this redshift.

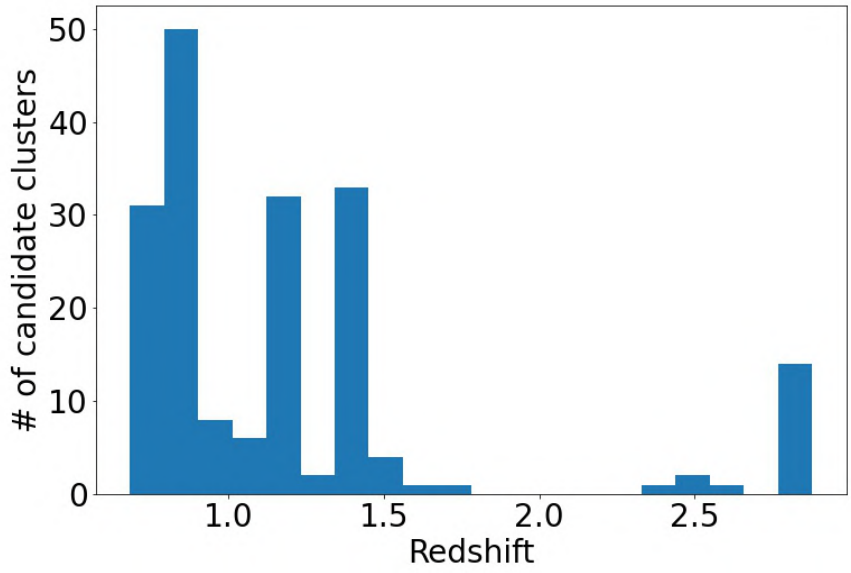


Figure 6.3. Distribution of the redshift values for SHARKS galaxy clusters. The number of objects decreases with redshift as expected. No candidates for z $<$ 0.2 are found, probably since clusters in such regions are much more extended and with lower surface number density.

6.4 Clusters luminosity

Having redshift and total flux for each cluster, luminosity was calculated using the relation

$$L = 4 \times \pi \times {D_l}^2 \times F$$

where L is the luminosity, ${D_l}^2$ the luminosity distance [1] of the cluster and F its total flux within the radius.
Distribution of cluster luminosity with redshift is shown in *Figure 6.5*. In the plot a visibly large fraction of the candidate clusters is found to be at redshift lower than z∼1. Some candidates are found also at much higher redshift, close to z∼3. These distant objects shows luminosity of at least one order of magnitude higher than candidate clusters at lower redshift. This is consistent with expectations since for a blind survey as SHARKS, that do not look for objects in a specific redshift interval, Malmquist bias [2] has a dominant effect.
An absence of candidate clusters is found in the redshift range $1.8 < z < 2.4$. This is probably just a statistical effect, to be explored in more detail extending the catalog to the entire SHARKS survey.

A pseudo-luminosity function was constructed expressing the number of candidate clusters observed for each luminosity bin. The differential counting function is not weighted by volume, as usually performed for a luminosity function. This is then just an indicative distribution, showing that the largest fraction of candidate clusters have a luminosity of the order of $10^{13} M_{\odot}$. This is again consistent with the typical values observed for galaxy clusters. A small fraction of objects shows luminosity of the order of $10^{14} M_{\odot}$, probably associated to the distant extremely bright galaxies.
I have calculated the error associated to each value of luminosity using the propagation with partial derivatives from the luminosity-flux relation. The error in luminosity distance was extracted with a Montecarlo method. For this I had to sample 1000 values of redshift from a normal distribution centered in the estimated value of the cluster redshift and having width $\sigma = 0.15$ (the error associated to each redshift measurement). For each sampled redshift I have calculated the corresponding luminosity distance, generating in turn a distribution. The error on luminosity distance is the standard deviation of such distribution of luminosity distances.
To better understand the influence of foreground and background galaxies on the light coming from the clusters, we plotted in *Figure* **??** the value

$$\Delta L = \frac{L_{area} - L_{memb}}{L_{area}}$$

at different redshift where L_{area} is the luminosity accounting for all sources in the cluster's area and L_{memb} is the luminosity accounting only for cluster members.
Distant clusters seem on average more affected by the presence of foreground sources than closer

[1] Cosmological distance can be defined in different ways depending on the proprieties we use as references. It depends mainly on the characteristics of the metric used to describe the geometry of space-time in the Universe. Luminosity distance is defined as the distance relating apparent and absolute magnitude of an object, accounting for the cosmological effects encoded in the metric. Another measure of distance can be obtained from the relation between real and apparent angular size of an extended object or Angular Diameter Distance.

[2] Observations at a cosmological level shows that the fraction of very luminosity objects increases with distance. This is not a propriety of the Universe, but an effect of limitations of the instrumental apparatus. The flux of an object decreases in fact with its distance squared. This means that to observe a distant source requires it to be very luminous. Fluxes from faint sources become instead undetectable even at low distances.

clusters, an effect due simply for the fact that foreground sources are in greater number than background ones, due to fluxes decreasing generally rapidly with distance.

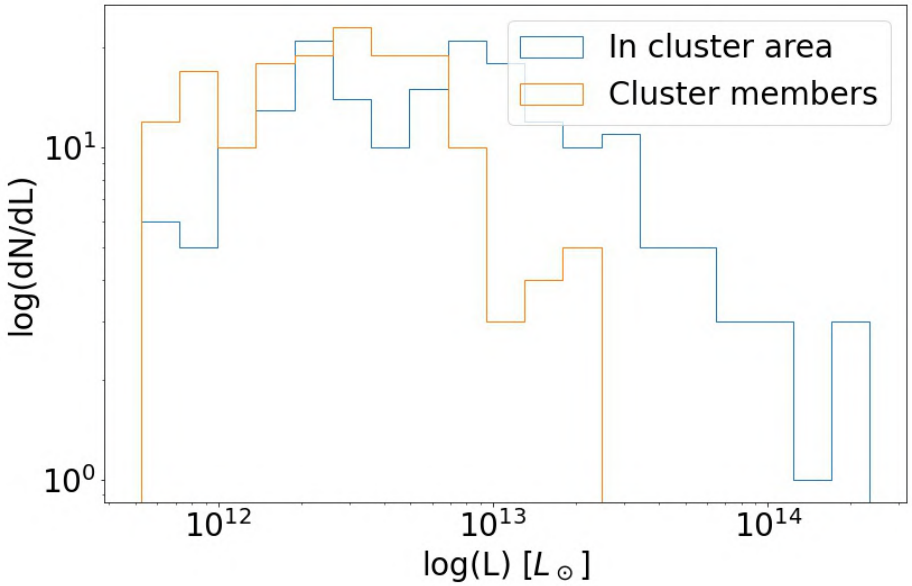


Figure 6.4. Counting function for SHARKS clusters using luminosity bins. There is no weighting for the volume of the survey occupied by each cluster.

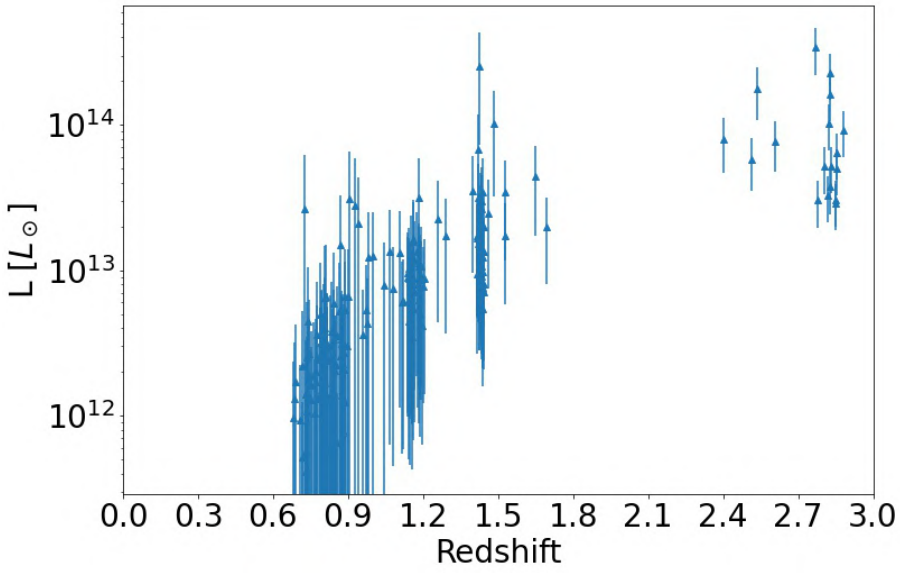


Figure 6.5. Distribution of candidate clusters' luminosity with redshift. Luminosity is in units of solar luminosity. It is calculated using the flux coming from all sources projected on the cluster area. Error on luminosity is calculated by propagation of the error in flux for each galaxy using partial derivatives. The most luminous objects are found to be at higher redshift. Less luminous objects have much larger error bars associated.

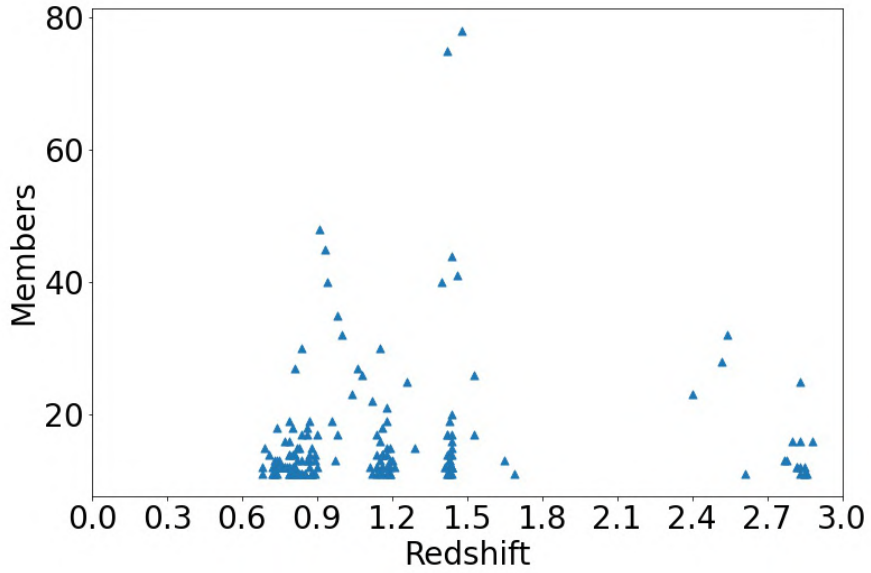


Figure 6.6. Distribution of the number of members for each candidate cluster with redshift. The most populous objects are at lower redshift. This is expected since at lower redshift more sources appear visible even with relatively small luminosity.

6.5 Existing clusters in literature

A search for known clusters in SHARKS area was performed, focused on the 4 contiguous fields in SGP. Four correspondences was found, two in the Abell (1989) catalog and two in Planck catalog of high redshift sources. The catalogs and their sources are presented in the following sections.

6.5.1 Matching the Abell 1989 catalog

Abell 2794: According to the Abell catalog of rich clusters (Abell, Corwin, and Olowin 1989) two massive structures are present in SHARKS catalog 1000000074498. Abell 2794 is part of the Pisces-Cetus supercluster, a collection of gravitationally bound clusters at z $\sim$ 0.06 (Porter and Raychaudhury 2005). There is no much information in literature about Abell 2794.
Three SHARKS candidate clusters are identified in the same region of the Abell structure: SH74498n12, SH74498n14, SH74498n56 with estimated redshift respectively at $z = 0.80 \pm 0.15$, $z = 1.16 \pm 0.15$ and $z = 0.68 \pm 0.15$. Images of the single clusters can be found in *Figure 6.8*. The redshift value of Abell 2794 is $z = 0.062$. SH74498n12 and SH74498n56 have a redshift consistent with the Abell cluster, while SH74498n56 seems to be at greater distance.
It is possible that these candidate clusters are not part of Abell 2794, but that they are instead projected onto it. The radius of Abell clusters are usually much larger than the ones found in SHARKS, mainly due to the fact that they are local structures occupying large patches of the sky plane.
Moreover, due to the vicinity of the structure, it is possible for the number density of a local cluster to be much lower than the one used to search structures in SHARKS. It is then probable that the candidates found in the area of Abell 2794 contains some of Abell's galaxies, but that they are not the main members used by the FoF algorithm to detect the structure. In *Figure 6.7* the area covered by Abell 2794 is shown together with the three SHARKS clusters.

Abell 2802: The second known cluster in the region is Abell 2802 at redshift z = 0.125.

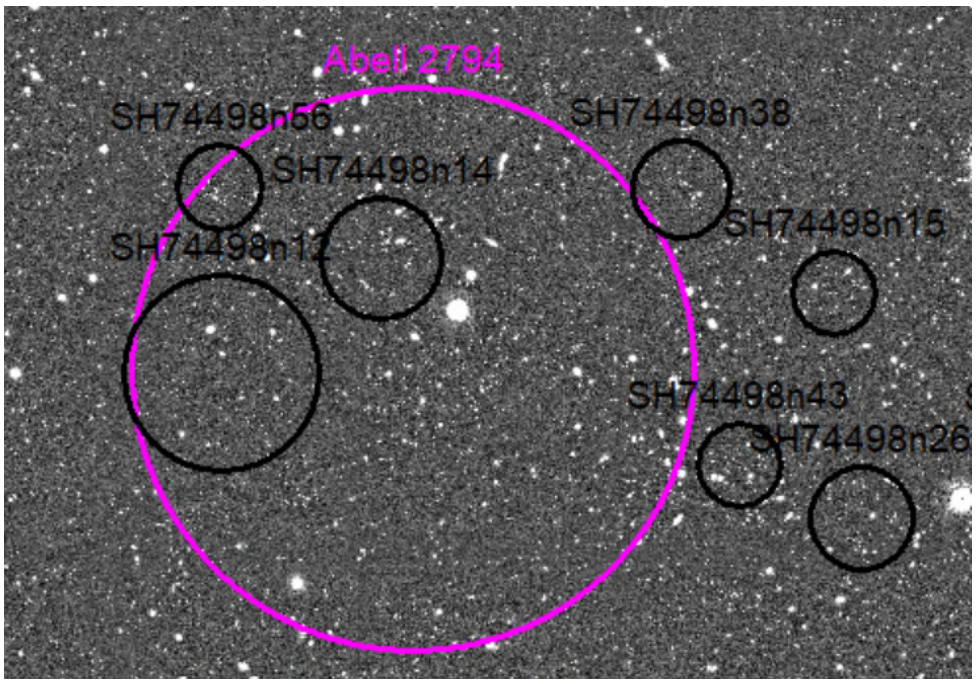


Figure 6.7. The region covered by Abell 2794 (magenta) together with SHARKS catalogs (black). Three SHARKS candidates apparently coincide with the Abell cluster. However their redshift is very different from the one of Abell 2794, indicating that they are probably not part of the cluster.

Basically no information is found in literature about this object if not in the Abell (1989) catalog itself.
At the position of Abell 2802 a structure is well recognizable even upon a simple visual inspection as shown in *Figure 6.9*. The structure presents itself almost as a filament, with what would seem some smaller substructures around the main elongated body.
SHARKS' counterpart for this local cluster is again constituted not by a single object, but by 3 different candidates. However the 3 SHARKS candidates have very similar redshift z $\sim$ 0.9 and so they are likely to represent sections of a single large structure.
Four smaller candidates are identified around the main body by the algorithm, their estimated redshift ranging between z $\sim$ 0.74 and z $\sim$ 1.44.
Considering the error in the redshift estimation as half the width of a redshift bin, it is possible for all seven candidates to be part of the same large system.
The redshift value estimated from the color-redshift relation is however far from the value found in literature.
Abell 2802, together with SHARKS' matching candidates, are shown in *Figure 6.10*

6.5.2 Matching the Planck high-redshift catalog

Two Planck's sources are situated in the SHARKS area spanned by the 4 contiguous mosaics in the SGP region. Each of them has counterparts in SHARKS cluster catalog as visible in *Figure 6.11*.
PHz G328.82-86.23: contains one SHARKS counterpart: SH74506n23, visible in *Figure 6.13*. The estimated angular radius for this counterpart is of 1.84 arcminutes, while the Planck source is contained in an area of radius 11.4 arcminutes. The estimated redshift for the SHARKS candidate is $z = 1.14 \pm 0.3$, while for the Planck source it is $z = 1.76$. Both sizes and redshift are then inconsistent.

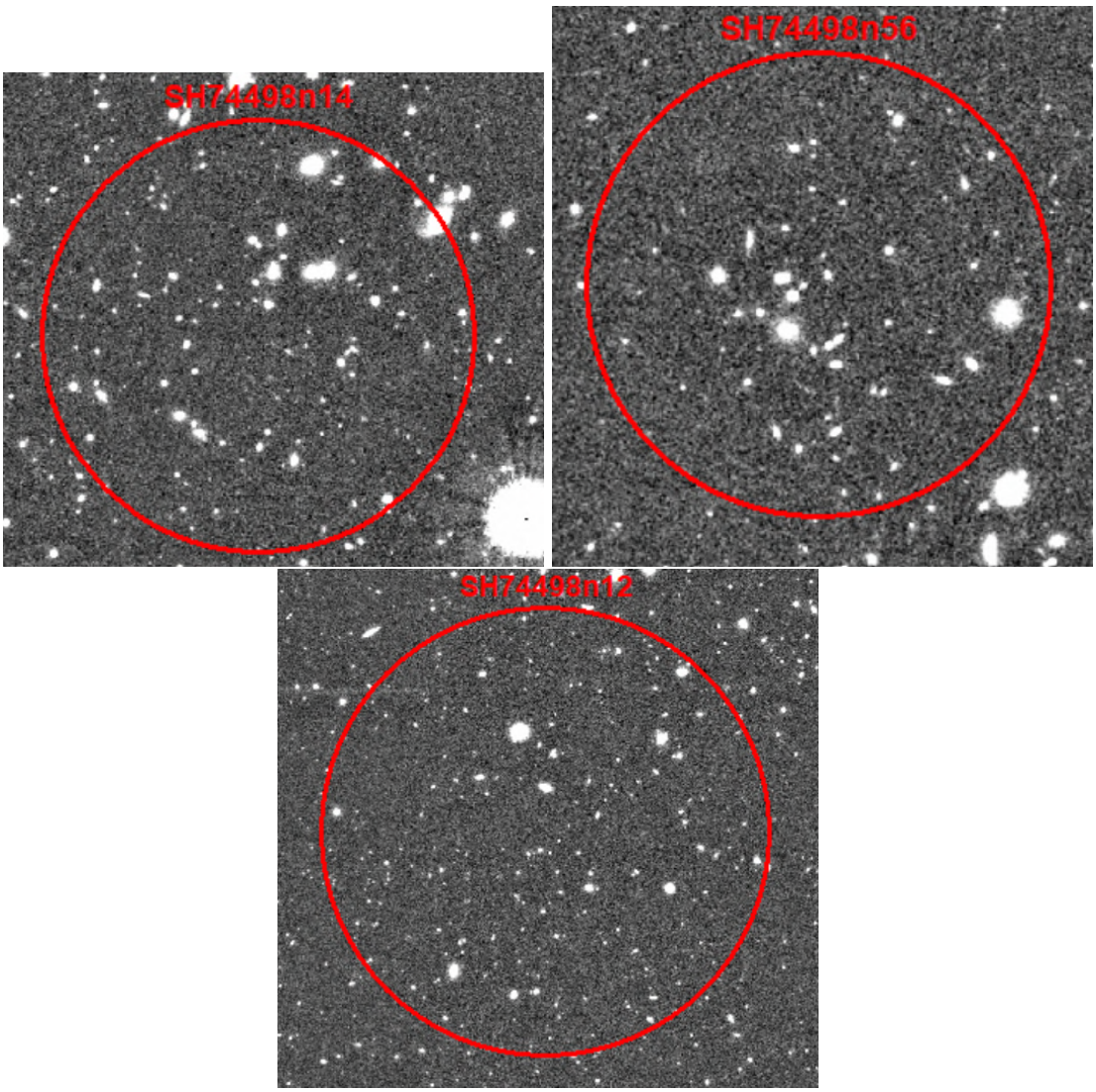


Figure 6.8. Closer view of the three SHARKS candidate clusters in the region of Abell 2794. *Top left:* SH74498n14, *Top right:* SH74498n56, *Bottom center:* SH74498n12.

PHz G333.47-86.90: Three SHARKS candidates are found in the area of this source: SH74506n13, SH74506n29, SH74506n32 with radii respectively of r=0.9, r=1.1, r=0.8 arcminutes and estimated redshift z=1.2, z=0.88, z=1.16. Planck source has instead a radius of 13 arcminutes and an estimated redshift of 1.91. No consistency is found between the two catalogs.
Despite the inconsistency of size and distances between Planck high-z sources and SHARKS candidate counterparts, it is still possible that this correspondance between SHARKS and Planck sources is real. According to Negrello et al. 2017, in fact, it is possible that some Planck high-z sources are the results of blending of the signal from different structures along the line of sight.

ID	**RA [deg]**	**DE [deg]**	**R ["]**	**z**	**Members**
SH74498n1	9.68334204432859	-31.60117778710629	99.0	0.91	48
SH74498n2	9.301089861924092	-31.800150314074326	123.0	1.43	19
SH74498n3	9.191133953588462	-31.474129488845524	76.0	1.18	19
SH74498n4	9.843251042962967	-30.779325713499254	76.0	1.26	25
SH74498n5	8.893666764978931	-31.28536807698479	107.0	0.81	27
SH74498n6	9.712197914128927	-31.608049013360453	100.0	0.93	45

SH74498n7	8.60594446564157	-31.353388115835525	92.0	1.43	13
SH74498n8	10.306217531947336	-31.867359659159835	113.0	2.54	32
SH74498n9	10.286434692194492	-31.64619603268091	93.0	0.79	19
SH74498n10	9.215376417788757	-31.747399888724825	94.0	0.87	19
SH74498n11	10.405879302126925	-31.734965873919045	71.0	0.86	18
SH74498n12	9.138497569859808	-30.956583707849948	126.0	0.8	18
SH74498n13	8.919079841662365	-31.440785029916512	120.0	2.83	16
SH74498n14	9.185505613592287	-31.01366662615061	77.0	1.16	18
SH74498n15	9.170905009032163	-31.17403205691693	53.0	0.84	17
SH74498n16	9.027962598475598	-31.530074949048025	82.0	2.88	16
SH74498n17	8.856935321459957	-31.191535985391074	56.0	2.83	12
SH74498n18	10.057266251940922	-31.306546666874812	87.0	1.15	16
SH74498n19	9.799254757412529	-31.15532941767881	65.0	1.44	12
SH74498n20	8.63620709978397	-31.592381512134995	63.0	1.19	15
SH74498n21	8.76983001749735	-31.48563868776972	52.0	0.89	14
SH74498n22	9.342053296083352	-31.633097592283846	79.0	0.87	14
SH74498n23	9.964025122666847	-31.451177179410077	60.0	0.71	14
SH74498n24	10.101654848250407	-31.803503815790013	74.0	1.14	14
SH74498n25	9.840161365341968	-31.686817784015187	103.0	0.74	13
SH74498n26	9.07759491238453	-31.18320278483572	66.0	1.16	14
SH74498n27	9.502106417679279	-31.23587450974239	69.0	1.16	12
SH74498n28	9.561919981862857	-30.9055483215611	59.0	0.82	13
SH74498n29	9.283397516413084	-31.50523266117635	62.0	0.89	13
SH74498n30	8.899014706149362	-31.501905934034948	33.0	0.86	13
SH74498n31	8.969043847182864	-30.949886572857213	54.0	0.68	12
SH74498n32	9.56672748337006	-31.650900570363547	89.0	1.44	12
SH74498n34	8.712415967505065	-31.99449595855042	6645.0	0.81	12
SH74498n35	10.40254296286004	-31.848670456849977	64.0	1.43	12
SH74498n36	9.413450028808116	-31.412559184813738	114.0	1.11	12
SH74498n37	9.1449857721783	-31.618976902908845	91.0	1.44	12
SH74498n38	9.04517337474033	-31.674089804618703	61.0	0.9	12
SH74498n39	9.213769085372755	-31.119864109149518	62.0	1.15	30
SH74498n40	9.862444992046552	-30.783007980768858	60.0	1.43	12
SH74498n41	8.835564523791838	-31.222183897742248	64.0	0.81	12
SH74498n42	8.900691315143128	-31.11185447351158	62.0	0.94	40
SH74498n43	9.704330699122087	-31.60737153522772	72.0	0.78	12
SH74498n44	9.09957451335635	-31.140077712647706	54.0	0.85	11
SH74498n45	10.096384883351018	-30.982857488136037	67.0	0.74	12
SH74498n46	8.87360135688136	-31.263758871682796	59.0	1.12	22
SH74498n47	9.866624598043211	-30.789916620209304	55.0	0.73	11
SH74498n48	10.023622205954094	-30.90173464299949	59.0	1.42	11
SH74498n49	9.654934568983382	-31.03506974447511	55.0	0.88	11
SH74498n50	9.739486706480767	-31.952888165680182	62.0	1.44	11
SH74498n51	8.702334950132089	-30.96961231940737	58.0	0.84	11

SH74498n52	9.096458354384295	-31.259714856955167	62.0	0.72	11
SH74498n53	9.965015102709568	-31.340699824207842	57.0	0.8	11
SH74498n54	9.768856591904594	-31.510000858695616	57.0	1.19	11
SH74498n55	9.034959641814671	-31.572089197752494	57.0	1.14	11
SH74498n56	8.93141570631257	-30.88482796199882	59.0	0.68	11
SH74498n57	9.214601814982835	-30.956432435869477	55.0	0.72	11
SH74498n58	9.396107099710056	-31.313888017272298	39.0	2.85	11
SH74498n59	9.00401811390259	-31.51122326331726	60.0	2.52	28
SH74498n60	10.321049550607626	-31.864605509346273	80.0	0.82	11
SH74501n1	10.477765470527505	-31.788329055391994	90.0	0.84	30
SH74501n2	10.309114953937202	-31.866254646766762	118.0	2.83	25
SH74501n3	11.486175179600956	-30.978878382152097	72.0	1.16	14
SH74501n4	10.95609052619918	-31.207022404517847	73.0	1.46	41
SH74501n5	10.521826114851114	-30.938792790117123	78.0	1.18	14
SH74501n6	10.945589295568814	-31.194720030797487	90.0	1.44	44
SH74501n7	10.660749558969504	-31.359773458857532	75.0	0.9	17
SH74501n8	10.956816782955942	-31.25021620795849	171.0	1.48	78
SH74501n9	10.47310912231048	-30.947662000218436	107.0	1.42	17
SH74501n10	10.47885039980211	-31.75900398139853	66.0	1.16	14
SH74501n11	11.220754547504232	-31.631488023369403	76.0	0.79	16
SH74501n12	10.404192861185464	-31.73654820655129	61.0	0.88	15
SH74501n13	10.368177529472137	-31.931485943443125	82.0	1.16	14
SH74501n14	11.363357599591604	-30.83450803910685	91.0	1.44	15
SH74501n15	11.74567129763095	-31.42686876964242	86.0	0.69	15
SH74501n16	10.652664233609185	-31.713379349504297	101.0	1.42	14
SH74501n17	10.933453759472586	-31.26267562664501	50.0	1.08	26
SH74501n18	11.914182549422467	-31.873019143544717	57.0	2.82	12
SH74501n19	11.01119242954799	-31.975744532269395	85.0	1.15	13
SH74501n20	10.60880779492684	-31.402408363011084	69.0	0.81	14
SH74501n21	12.026308640639154	-31.7357258833511	80.0	1.29	15
SH74501n22	11.812026260241515	-31.60661156677948	56.0	2.85	12
SH74501n23	10.287008657049658	-31.640903352549998	76.0	0.86	17
SH74501n24	10.922175983825698	-31.186571716980385	75.0	1.44	20
SH74501n25	10.920695337305293	-31.245568717863872	108.0	1.4	40
SH74501n26	11.815312877411987	-31.484562300941516	53.0	0.77	12
SH74501n27	11.498237102634096	-31.074463425355024	48.0	0.87	12
SH74501n28	10.954789290900672	-31.2187917917135	135.0	1.42	75
SH74501n29	11.711221510583695	-31.244509051101787	55.0	1.44	12
SH74501n30	11.253310354442565	-31.581266520472493	83.0	2.85	11
SH74501n31	10.497348417514557	-30.910219989848283	62.0	0.74	12
SH74501n32	11.752562916998976	-30.950568450825333	70.0	1.14	12
SH74501n33	11.102045108886415	-31.488012222971008	69.0	0.8	12
SH74501n34	12.060841969622105	-31.455495273442498	64.0	0.79	12
SH74501n35	10.878320023744845	-31.1008167957774	56.0	0.98	17

SH74501n36	10.862744438376023	-31.21031776246236	87.0	2.61	11
SH74501n37	10.449265934452153	-31.991289781015738	53.0	0.81	11
SH74501n38	10.612723159829523	-30.993131853177577	62.0	1.41	12
SH74501n39	12.076389934717245	-31.15905901313504	50.0	0.88	11
SH74501n40	11.67083147746646	-31.210225886608818	92.0	1.18	12
SH74501n41	10.739097096557531	-31.38265791554715	41.0	0.88	11
SH74501n42	11.060859333885723	-31.169693731200017	62.0	0.81	11
SH74501n43	10.30097260175695	-31.616373179475655	80.0	1.16	12
SH74501n44	10.473318596659983	-31.243101317448385	60.0	1.42	11
SH74501n45	11.2645129304687	-31.36276876234329	77.0	0.79	11
SH74501n46	10.856912859291533	-31.09876520867442	89.0	1.04	23
SH74501n47	10.706927857955385	-31.806048327421365	66.0	2.86	11
SH74501n48	11.478614564408327	-31.644958606992038	91.0	0.79	11
SH74501n49	10.286510550123198	-31.800372841051505	73.0	2.83	11
SH74501n50	11.634664866992674	-31.32750051155504	52.0	0.89	11
SH74501n51	12.011494718740122	-31.70552338038921	98.0	1.65	13
SH74501n52	10.717027041872074	-31.066013474139424	64.0	1.15	11
SH74501n53	11.959769916100205	-31.683734673552276	72.0	0.85	11
SH74501n54	10.979370525963025	-31.209418449092944	45.0	0.97	13
SH74506n2	11.444898124313337	-29.712965487270434	6852.0	0.98	35
SH74506n3	10.06801926870178	-30.56700623277325	99.0	0.82	15
SH74506n4	11.120403400999988	-30.768024982635634	95.0	0.79	14
SH74506n5	11.629142089731005	-30.118570963816246	68.0	1.0	32
SH74506n6	10.060135452194459	-30.55381010604584	93.0	0.79	14
SH74506n7	11.047448756735887	-30.930792108537617	67.0	0.96	19
SH74506n8	10.211920748522177	-30.54636335960102	50.0	0.72	12
SH74506n9	10.477601550695482	-30.071591520923008	85.0	1.44	17
SH74506n10	11.44717460657064	-30.011946060359975	98.0	0.79	12
SH74506n11	10.024570798685167	-30.902105187210523	60.0	0.76	12
SH74506n12	10.425165092976343	-30.90678784844895	64.0	1.43	14
SH74506n13	11.501197113401547	-30.06074231267098	54.0	1.2	13
SH74506n14	10.967279938423504	-29.972933866610575	83.0	2.8	16
SH74506n15	11.68696760484187	-30.45355078637725	92.0	0.73	12
SH74506n16	10.607882375321966	-29.86419098899793	44.0	0.76	12
SH74506n17	11.066874211475765	-30.857134039545752	80.0	1.44	14
SH74506n18	10.565423931495266	-30.54856305016042	106.0	1.06	27
SH74506n19	10.212528192621887	-30.533818496629962	123.0	0.77	12
SH74506n20	9.944254697436179	-30.57659432147601	80.0	0.73	11
SH74506n22	11.380799867891351	-30.60258769736205	70.0	1.19	12
SH74506n23	9.868233335982694	-30.07108542842912	8685.0	1.14	12
SH74506n24	11.469537283156464	-30.742159737475674	50.0	0.73	11
SH74506n25	10.988559830946386	-30.448551875492537	107.0	0.74	11
SH74506n26	10.228299126753896	-30.09569803584059	49.0	0.8	11
SH74506n27	10.892310129823635	-30.141346558089246	66.0	0.85	11

SH74506n28	10.156200928552622	-30.554210834592244	68.0	1.43	13
SH74506n29	10.129630367399013	-30.82414166788164	65.0	0.88	11
SH74506n30	11.36292257067358	-30.832654500831616	88.0	1.44	16
SH74506n31	11.090432525870494	-29.952533184142556	43.0	1.21	12
SH74506n32	11.58745606860821	-29.756704062630774	45.0	1.16	11
SH74506n33	10.28693772651366	-30.595733704606065	80.0	1.43	12
SH74506n34	10.905153737409282	-29.853904243185994	79.0	2.78	13
SH74506n36	10.54031709212658	-29.74163426947781	95.0	1.53	26
SH74506n37	11.615116118823735	-30.521861947928397	65.0	2.77	13
SH74506n38	11.72290833914451	-30.048871609496583	8700.0	0.83	11
SH74506n39	11.591561770402315	-29.769837826381274	97.0	2.82	12
SH74506n40	10.318820163388583	-30.190250354687517	116.0	1.43	11
SH74511n1	8.55515042069142	-30.625503496576982	115.0	1.18	21
SH74511n2	9.37652302340727	-30.93031032777609	84.0	0.74	18
SH74511n4	9.13774144454052	-29.701321164313587	6823.0	0.77	16
SH74511n5	8.504953771654673	-30.48114793747244	83.0	0.83	15
SH74511n6	8.716185141623122	-30.35596167530503	70.0	1.14	17
SH74511n7	10.05911152822121	-30.552718269523687	50.0	0.8	14
SH74511n8	9.26164890924333	-30.612501847077123	61.0	0.87	13
SH74511n9	8.756330951624486	-30.550274948003185	71.0	0.75	13
SH74511n10	9.729409836772358	-30.27079168578518	68.0	1.18	15
SH74511n11	9.517977195516487	-30.179317757799392	70.0	0.73	13
SH74511n13	9.609899526475592	-30.277859342303977	60.0	0.84	13
SH74511n14	8.970398392560107	-29.694883305452716	6840.0	1.42	17
SH74511n15	8.921342760264185	-30.666538165300302	62.0	1.53	17
SH74511n16	9.566253549981083	-30.137084192049414	66.0	0.75	12
SH74511n17	9.57707532269292	-29.909602711384494	79.0	1.43	13
SH74511n18	9.944555442778146	-30.5765975813111	80.0	1.19	13
SH74511n19	9.290536375319538	-30.022808201019927	81.0	1.16	12
SH74511n20	9.431035829362179	-29.888446326591822	63.0	0.8	11
SH74511n21	9.780765439429327	-30.228675663832394	59.0	1.19	11
SH74511n22	8.434924134771926	-30.085523927546237	55.0	1.41	12
SH74511n23	9.364172547063323	-29.949479557960796	69.0	1.18	11
SH74511n24	9.871381671126334	-30.78916054767362	76.0	0.8	11
SH74511n25	9.357700967241867	-29.923061208195747	43.0	0.79	11
SH74511n26	8.45898951759632	-30.77006467669773	52.0	1.43	12
SH74511n27	10.023622878367316	-30.901734763216123	59.0	2.4	23
SH74511n28	9.83722911935717	-29.85716535255061	67.0	1.44	12
SH74511n29	9.56837261159003	-29.938764148492062	114.0	0.88	11
SH74511n30	9.431933797530844	-30.071266309826047	62.0	0.84	11
SH74511n31	9.124572164946112	-29.8076799031192	36.0	1.12	11
SH74511n32	9.969394008571026	-30.86852246594852	47.0	1.44	12
SH74511n33	8.966785617525098	-30.82983847645523	72.0	1.69	11
SH74511n34	9.160690310794514	-30.289578416218884	60.0	1.43	11

SH74511n35	9.855271706851543	-30.908549279307874	75.0	0.88	11
SH74511n36	9.13755813692029	-30.024030177640896	94.0	0.84	11
SH74511n37	9.536053151550153	-30.116533055205498	66.0	0.83	11
SH74511n38	9.562537698531335	-30.904977140772576	45.0	1.19	11

Table 6.2: List of all 186 candidate clusters for the first 8 squared degrees of SHARKS cluster catalog. Clusters are found applying a Friends-of-Friends algorithm using a linking length of 0.1 and a minimum number of galaxies of 10. The algorithm was applied to redshift slices of the catalog, obtained from a color-redshift relation using the color Y-Ks (Y band from Dark Energy Survey and Ks from SHARKS). Redshift bins have a width $\Delta z = 0.3$.

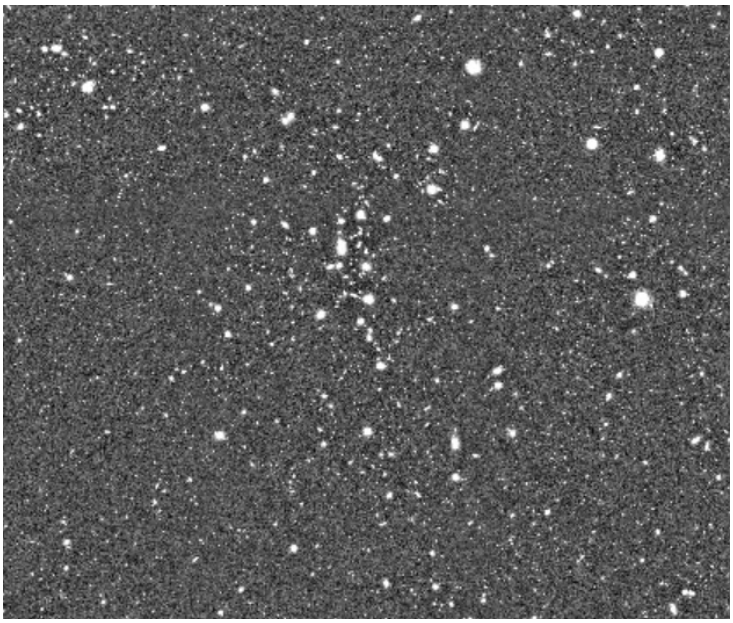


Figure 6.9. Close view of the region of Abell 2802. The image covers an area of approximately 200 square arcminutes. The main filamentary structure appears to have a length of 4 arcminutes.

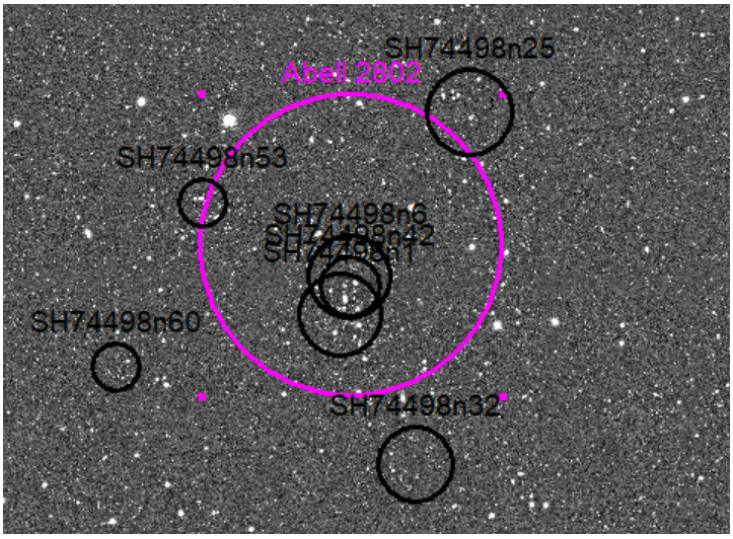


Figure 6.10. Abell 2802 (magenta) together with SHARKS candidate clusters in the same region (black). Three candidates are found along the filamentary region. They seem to have almost the same redshift, so to be part of the same structure. However SHARKS clusters have z $\sim$ 0.9 $\pm$ 0.15, while Abell 2802 has z = 0.125. This could mean that a systematic error in the calibration of the color-redshift relation is present.

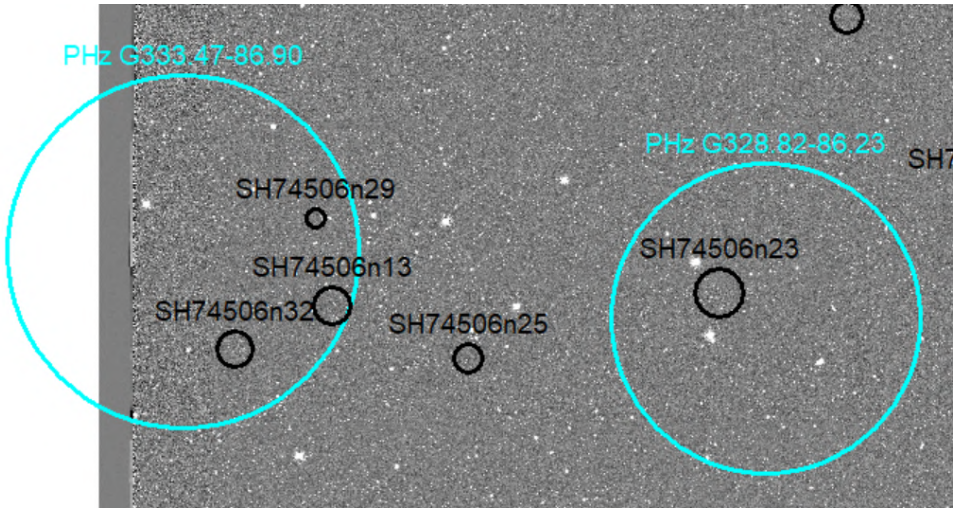


Figure 6.11. A view of the two Planck Cold Sub-millimiter sources together with SHARKS candidate clusters in the same region. Plank sources are huge, of the order of 10 arcminutes in radius, while SHARKS clusters are much smaller, of the order of 1-2 arcminutes in radius. This suggests that is not likely a connection between the 2 sources. Not enough information is then provided on the possibility for Planck's sources to be blended emissions from different clusters projected along the line of sight.

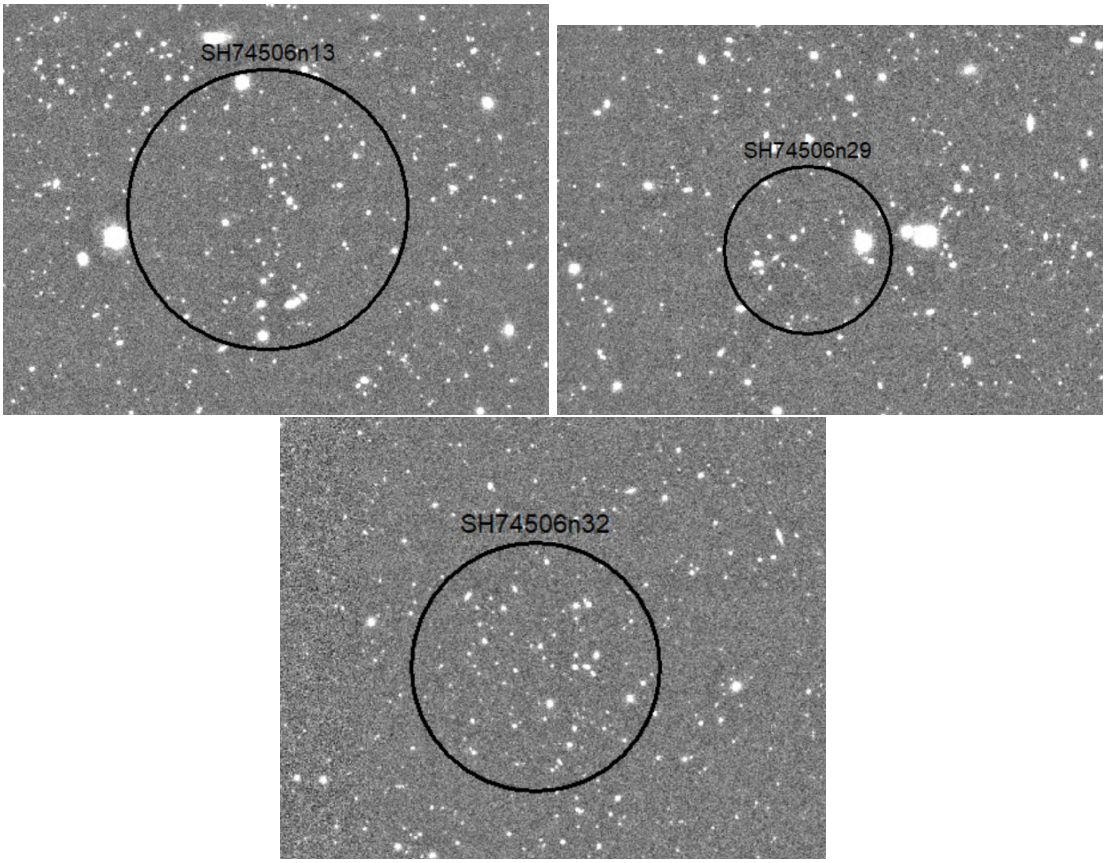


Figure 6.12. Closer view of the three SHARKS candidate clusters in the region of Planck source PHz G333.47-86.90. *Top left:* SH74506n13, *Top right:* SH74506n29, *Bottom center:* SH74506n32

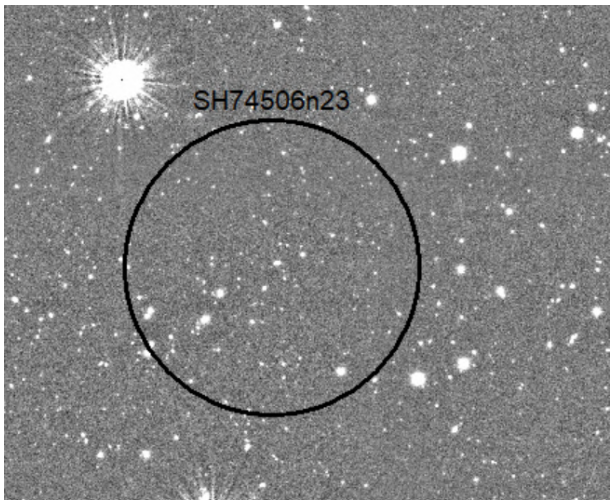


Figure 6.13. Candidate cluster SH74506n23 in the area of PHz G328.82-86.23

Chapter 7

Conclusions

7.1 Summary and comments

Here I resume the processes that led to the generation of SHARKS cluster catalog.
The catalog is constructed from public data realized by the ESO SHARKS survey. This is a 300 square degrees single band survey. It is performed by the 4 m telescope VISTA using the InfraRed camera VIRcam with a Ks filter centered at 2.2 μm. The point spread function of the telescope, corrected with assistance from an adaptive optical system, is of 0.5 arcsec.
The objective of the survey is to observe with enhanced resolution and in a complementary band a vast area of the Herschel-ATLAS survey. The depth reached by SHARKS at 5σ level is of $\sim$ 22.7 mag.
On January 2022 the first data release was published, containing 10 reduced images and corresponding catalogs for a total of $\sim$ 20 square degrees of survey area. The area is divided in 6 mosaics in the SGP region and 2 mosaics for GAMA-12 and GAMA-15 fields.
I cleaned dr1 catalog from spurious sources due to diffraction artifacts around saturated foreground stars. Sources contaminated by background flux had already been identified during the catalog reduction and were presented with a badness flag in the dr1 release. However fake sources generated by diffraction patterns had not been identified and removed efficiently. Such sources are generally caused by bright saturated stars generating multiple reflection and diffraction patterns in the instrumental apparatus. Since these patterns are regular,depending mainly on the instrumentation, linear relations between their characteristics can be constructed. I used five SHARKS images to calibrate the linear relations. Then I used the relations to identify automatically these patterns in all other dr1 images.
I calculated the completeness level for the catalog using the flux of each source measured in two different apertures of respectively 2″ and 8″.
Bright sources shown a tight relation between these two measures, while faint sources had a much higher dispersion, allowing for the establishment of a magnitude limit were the dispersion overcomes a certain threshold.
Completeness is then found to be of 22.1 mag on average.
I removed the foreground stellar component contaminating the galaxy catalog using Stellarity index and half-light radius parameters provided by dr1 data. This parameter comes from the SExtractor algorithm which was used to extract sources from SHARKS images and to construct the dr1 catalogs.
The stellarity index gives the probability of a source to be a star accounting for its shape on the

CCD. The half-light radius instead gives indications on the concentration and on the symmetry of a source. Stars tend to be distributed in tight sequences when plotting the stellarity index or half-light radius over the sources' magnitude. For this reason it is possible to identify and remove them. The process was estimated to be reliable up to a depth of 20.56 mag.
I have identified cluster candidates using a Friends-of-Friends algorithm. This is a grouping algorithm that groups galaxies according to their reciprocal distance. Galaxies that are closer together than a threshold distance are said to belong to the same group.
The propriety of belonging to a group is transitive, meaning that if a galaxy A groups with B and B groups with C, then A,B and C must be part of the same group. Groups that contain at least a given minimum number of galaxies are considered candidate clusters. The algorithm thus depends on two parameters called the linking length (b) and the minimum number of members of galaxies to form a cluster (mingal).
I tested the algorithm for different values of the two parameters on the entire dr1 catalog.
I produced simulations of the dr1 catalogs reassigning to each galaxy random positions in the survey area in order to establish a relation between the two parameters. The minimum number of galaxies needed to form a candidate was calculated in this way. The meaning of this number is to only find the clusters that are sufficiently populated so to have a low probability of being formed by chance from random fluctuations of the galaxy distribution. This approach produces a very small number of candidates characterized by small angular sizes (of the order of $\sim$1 arcmin) and large number of members (>25).
The definitive method chosen for the construction of the catalog was, however, to divide galaxies of the dr1 catalog in redshift bins. I then applied FoF algorithm to each bin using b=0.1 and mingal=10. Redshift was calculated using a color-redshift relation calibrated using the spectral energy distribution model of a typical old red sequence galaxy. Color was obtained using SHARKS Ks photometric band and Dark Energy Survey (DES) Y band. DES and SHARKS, however, overlap only in the SGP region. The region is covered in dr1 only by the 4 contiguous fields 1000000074498, 1000000074501, 1000000074506 and 1000000074511. For this reason the cluster catalog can cover only an area of $\sim$8 square degrees.
Redshift bins were chosen to be of width $\Delta z = 0.3$. Each redshift value is calibrated for red sequence galaxies. These galaxies are expected to be present in higher fraction in clusters relatively to the field, and so can be used as tracers for dense environments.
The resulting cluster catalog includes 186 candidate clusters distributed along an estimated redshift interval of $0.2 \leq z \leq 3$. The error on each estimated redshift is of half the width of the adopted redshift bins so $z_{err} = 0.15$.
Counting function performed on the cluster catalog shows a peak around the typical cluster luminosity $L = 10^{13} L_{\odot}$. The most distant structures, about $z \sim 3$ are also the most luminous ones, reaching values of $L = 10^{14} L_{\odot}$. The number of clusters per square degrees (N) is consistent by order of magnitude with statistics from KiDS and DES.

$$N_{SHARKS} = 23$$

$$N_{KiDS} = 11$$

$$N_{DES} = 27$$

. There is an overlap of SHARKS clusters with 2 Abell objects and 2 Planck high redshift submillimetric sources. However it is thought that the correspondence is only apparent, since the

overlapping sources have very different angular sizes and estimated redshifts.
The SHARKS cluster catalog (Lorenzon et al., in prep.) will be made publicly available upon acceptance of the related publication.

7.2 Future perspectives

The SHARKS catalog of galaxy clusters and proto-clusters was constructed in order to provide useful information for the study of environment-dependent galaxy evolution.
Galaxy evolution is essential to understand the processes leading to the formation of different stellar populations and to the metal enrichment of the inter-stellar medium, responsible, at last, for the formation of planetary systems and life itself.
It seems that the evolution of galaxies depends on the environment, with denser regions having faster evolution (Madau and Dickinson 2014, eg). This could be linked to the fact that in such environments it is more frequent and strong the interaction between galaxies themselves and the diffuse component of gas and dust (Chiang, Overzier, and Gebhardt 2014, eg).
It is then important to measure the amount of inter-stellar medium present in galaxies residing in dense regions, as clusters and proto-clusters to compare with less dense regions as the field. Moreover a large amount of information resides in the Star Formation Rate (SFR) for galaxies inside clusters. This is the measure of the amount of stars produced per year by a galaxy. It is possible to express the amount of stars produced per comoving volume of the Universe using the star formation rate density (SFRD). This quantity evolves with time and changes depending on the environment density.
The first important information that SHARKS cluster catalog is able to provide is a study of a precise mass profile for the clusters luminous matter. The majority of the luminous mass of a galaxy, in fact, comes from the low mass stellar component (Kroupa 2001, eg), with masses of the order M$\leq 0.1 M_{\odot}$. The energy distribution of photons from a star is similar to a Planck distribution, deviating from it due to absorption lines. The photon frequency corresponding to the peak of this distribution depends on the temperature of the star, and this is related to its mass. Low mass stars thus have a peak emission around 1.6μm which is well detected by SHARKS photometric Ks band. Furthermore, Ks band is much less affected by dust absorption than higher frequency bands, allowing it to detect the luminous mass component of galaxies even if they are immersed in thick layers of gas and dust, as expected from galaxies in proto-cluster environment. Once the total 300 squared degrees of the SHARKS survey will be published, it will be possible to extend the cluster catalog to such a large area. In this way the statistics on the number density and redshift distribution will be more complete.
The redshift value assigned to each cluster through the color-redshift relation is a rough estimate: better measurements are required, especially to increase the information on the distribution of cluster members.
Photometric redshift will be constructed using all different surveys performed in the GAMA fields and SGP region including GALEX deep for the u band, Dark Energy Survey for g,r,i,z,Y, SHARKS for the Ks band and H-ATLAS for IR region. Sub-millimeter and radio observations from the Australian Square Kilometer Array Pathfinder (ASKAP) and Low-Frequency Array (LOFAR) will be also available.
Spectroscopic surveys covering 100% of the catalog area are already planned as WEAVE, Deep-WAVES from ground and Euclid from space. They will provide precise redshift measurements for a large number of SHARKS sources contributing to the confirmation and characterization of the

structures in the SHARKS clusters catalog.
Using the fairly large area of the entire survey, together with more precise redshift information it will be possible to construct cluster luminosity functions and study their evolution with redshift up to z∼3. It will be possible also to calculate Star Formation Rates (SFR) for galaxies inside SHARKS clusters. This is the measure of the amount of stars produced per year by a galaxy. This can be normalized to the volume in star formation rate density (SFRD). This quantity expresses the amount of stars produced per comoving volume of the Universe and it is linked to the evolution process of galaxies. It is important to understand how SFRD evolves in time and how it depends on the matter density of the environment. Large and statistically consistent catalogs of clusters at high redshift as the complete SHARKS cluster catalog are needed to increase the knowledge on how the environment influences the star formation process and thus the evolution and mass accretion of galaxies.

Acknowledgements

- I want to thank my supervisor Professor Giulia Rodighiero for the patience, support and encouragement during the construction of this master thesis. I am also grateful to my co-supervisor Dr. Francesco Sinigaglia for the many fundamental advices and discussions, as well as for his huge help on the writing process of this work.

- A fundamental part of this project would not have been possible without the help of Professor Claudia Maraston, who provided the galaxy evolution models for the construction of a color-redshift relation.

- I thank also my opponent Professor Bianca Maria Poggianti for the precious suggestions on how to improve the quality of the project.

- Based on data products created from observations collected at the European Organization for Astronomical Research in the Southern Hemisphere under ESO programme 198.A-2006.

- For the creation of the data used in this work, the SHARKS team at the Instituto de Astrofísica de Canarias has been financial supported by the Spanish Ministry of Science, Innovation and Universities (MICIU) under grant AYA2017-84061-P, co-financed by FEDER (European Regional Development Funds), by the Spanish Space Research Program "Participation in the NISP instrument and preparation for the science of EUCLID" (ESP2017-84272-C2-1-R) and by the ACIISI, Consejería de Economía, Conocimiento y Empleo del Gobierno de Canarias and the European Regional Development Fund (ERDF) under grant with reference PROID2020010107.

- We thank the support of the Wide-Field Astronomy Unit for testing and parallelising the mosaic process and preparing the releases. The work of the Wide-Field Astronomy Unit is funded by the UK Science and Technology Facilities Council through grant ST/T002956/1.

- This research has made use of the services of the ESO Science Archive Facility

- This project used public archival data from the Dark Energy Survey (DES). Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana–Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.

- The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Enérgeticas, Medioambientales y Tecnológicas–Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois

at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the OzDES Membership Consortium, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

- Based in part on observations at Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

Bibliography

Abell, George O., Harold G. Corwin Jr., and Ronald P. Olowin (May 1989). "A catalog of rich clusters of galaxies". In: *The Astrophysical Journal Supplement Series* 70, p. 1. ISSN: 0067-0049, 1538-4365. DOI: 10.1086/191333. URL: http://adsabs.harvard.edu/doi/10.1086/191333 (visited on 08/24/2022).

Aghanim, N. et al. (Sept. 1, 2020). "Planck 2018 results - VI. Cosmological parameters". In: *Astronomy & Astrophysics* 641. Publisher: EDP Sciences, A6. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201833910. URL: https://www.aanda.org/articles/aa/abs/2020/09/aa33910-18/aa33910-18.html (visited on 08/07/2022).

Allen, Steven W., August E. Evrard, and Adam B. Mantz (2011). "Cosmological Parameters from Observations of Galaxy Clusters". In: *Annual Review of Astronomy and Astrophysics* 49.1. _eprint: https://doi.org/10.1146/annurev-astro-081710-102514, pp. 409–470. DOI: 10.1146/annurev-astro-081710-102514. URL: https://doi.org/10.1146/annurev-astro-081710-102514 (visited on 08/09/2022).

Bergh, Sidney van den (Sept. 1, 1999). "The local group of galaxies". In: *Astronomy and Astrophysics Review* 9.3, pp. 273–318. ISSN: 0935-4956, 1432-0754. DOI: 10.1007/s001590050019. URL: http://link.springer.com/10.1007/s001590050019 (visited on 07/29/2022).

Binggeli, B. and L.M. Cameron (Dec. 1991). "Dwarf galaxies in the Virgo cluster. I. The systematic photometric properties of early-type dwarfs." In: *Astronomy & Astrophysics* 252, p. 27.

Binggeli, Bruno, GA Tammann, and Allan Sandage (1987). "Studies of the Virgo cluster. VI-Morphological and kinematical structure of the Virgo cluster". In: *The Astronomical Journal* 94, pp. 251–277.

Bleem, L. E. et al. (Jan. 2015). "GALAXY CLUSTERS DISCOVERED VIA THE SUNYAEV-ZEL'DOVICH EFFECT IN THE 2500-SQUARE-DEGREE SPT-SZ SURVEY". In: *The Astrophysical Journal Supplement Series* 216.2, p. 27. DOI: 10.1088/0067-0049/216/2/27. URL: https://doi.org/10.1088/0067-0049/216/2/27.

Böhringer, H. et al. (Oct. 1, 2004). "The ROSAT-ESO Flux Limited X-ray (REFLEX) Galaxy cluster survey - V. The cluster catalogue". In: *Astronomy & Astrophysics* 425.1. Number: 1 Publisher: EDP Sciences, pp. 367–383. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361:20034484. URL: https://www.aanda.org/articles/aa/abs/2004/37/aa0484-03/aa0484-03.html (visited on 08/10/2022).

Bond, J. Richard, Lev Kofman, and Dmitry Pogosyan (Apr. 1996). "How filaments of galaxies are woven into the cosmic web". In: *Nature* 380.6575, pp. 603–606. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/380603a0. URL: http://www.nature.com/articles/380603a0 (visited on 07/24/2022).

Busca, N. G. et al. (Apr. 1, 2013). "Baryon acoustic oscillations in the Ly_α forest of BOSS quasars". In: *Astronomy & Astrophysics* 552. Publisher: EDP Sciences, A96. ISSN: 0004-6361, 1432-0746.

DOI: 10.1051/0004-6361/201220724. URL: https://www.aanda.org/articles/aa/abs/2013/04/aa20724-12/aa20724-12.html (visited on 08/07/2022).

Casey, Caitlin M. (June 8, 2016). "THE UBIQUITY OF COEVAL STARBURSTS IN MASSIVE GALAXY CLUSTER PROGENITORS". In: *The Astrophysical Journal* 824.1, p. 36. ISSN: 1538-4357. DOI: 10.3847/0004-637X/824/1/36. URL: https://iopscience.iop.org/article/10.3847/0004-637X/824/1/36 (visited on 07/20/2022).

Casey, Caitlin M., Desika Narayanan, and Asantha Cooray (Aug. 10, 2014). "Dusty star-forming galaxies at high redshift". In: *Physics Reports*. Dusty star-forming galaxies at high-redshift 541.2, pp. 45–161. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2014.02.009. URL: https://www.sciencedirect.com/science/article/pii/S0370157314000477 (visited on 08/09/2022).

Cautun, Marius et al. (July 11, 2014). "Evolution of the cosmic web". In: *Monthly Notices of the Royal Astronomical Society* 441.4, pp. 2923–2973. ISSN: 0035-8711. DOI: 10.1093/mnras/stu768. URL: https://doi.org/10.1093/mnras/stu768 (visited on 07/20/2022).

Chapman, S. C. et al. (Oct. 20, 2004). "A Population of Hot, Dusty Ultraluminous Galaxies at $z \approx 2$". In: *The Astrophysical Journal* 614.2. Publisher: IOP Publishing, p. 671. ISSN: 0004-637X. DOI: 10.1086/423833. URL: https://iopscience.iop.org/article/10.1086/423833/meta (visited on 08/09/2022).

Chiang, Yi-Kuan, Roderik Overzier, and Karl Gebhardt (Jan. 21, 2014). "DISCOVERY OF A LARGE NUMBER OF CANDIDATE PROTOCLUSTERS TRACED BY ∼15 Mpc-SCALE GALAXY OVERDENSITIES IN COSMOS". In: *The Astrophysical Journal* 782.1, p. L3. ISSN: 2041-8205, 2041-8213. DOI: 10.1088/2041-8205/782/1/L3. URL: https://iopscience.iop.org/article/10.1088/2041-8205/782/1/L3 (visited on 07/20/2022).

Copeland, Edmund J., M. Sami, and Shinji Tsujikawa (Nov. 2006). "Dynamics of dark energy". In: *International Journal of Modern Physics D* 15.11. Publisher: World Scientific Publishing Co., pp. 1753–1935. ISSN: 0218-2718. DOI: 10.1142/S021827180600942X. URL: https://www.worldscientific.com/doi/abs/10.1142/S021827180600942X (visited on 08/09/2022).

Copi, Craig J. et al. (Oct. 7, 2010). "Large-Angle Anomalies in the CMB". In: *Advances in Astronomy* 2010. Publisher: Hindawi, e847541. ISSN: 1687-7969. DOI: 10.1155/2010/847541. URL: https://www.hindawi.com/journals/aa/2010/847541/ (visited on 08/09/2022).

Dalton, G.B. et al. (June 2006). "The VISTA infrared camera". In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Ed. by Ian S. McLean and Masanori Iye. Vol. 6269. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 62690X, p. 62690X. DOI: 10.1117/12.670018.

Dicke, R. H. et al. (July 1, 1965). "Cosmic Black-Body Radiation." In: *The Astrophysical Journal* 142. ADS Bibcode: 1965ApJ...142..414D, pp. 414–419. ISSN: 0004-637X. DOI: 10.1086/148306. URL: https://ui.adsabs.harvard.edu/abs/1965ApJ...142..414D (visited on 08/09/2022).

Dressler, A. (Mar. 1, 1980). "Galaxy morphology in rich clusters: implications for the formation and evolution of galaxies." In: *The Astrophysical Journal* 236. ADS Bibcode: 1980ApJ...236..351D, pp. 351–365. ISSN: 0004-637X. DOI: 10.1086/157753. URL: https://ui.adsabs.harvard.edu/abs/1980ApJ...236..351D (visited on 08/10/2022).

Driver, S. P. et al. (Apr. 2011). "Galaxy and Mass Assembly (GAMA): survey diagnostics and core data release". In: *Monthly Notices of the Royal Astronomical Society* 413.2, pp. 971–995. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2010.18188.x. eprint: https://academic.oup.com/mnras/article-pdf/413/2/971/18586750/mnras0413-0971.pdf. URL: https://doi.org/10.1111/j.1365-2966.2010.18188.x.

Eisenstein, Daniel J., Hee-Jong Seo, and Martin White (Aug. 2007). "On the Robustness of the Acoustic Scale in the Low-Redshift Clustering of Matter". In: *some journal name* 664.2, pp. 660–674. DOI: 10.1086/518755. arXiv: astro-ph/0604361 [astro-ph].

Eisenstein, Daniel J., Idit Zehavi, et al. (Nov. 2005). "Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies". In: *some journal name* 633.2, pp. 560–574. DOI: 10.1086/466512. arXiv: astro-ph/0501171 [astro-ph].

Emerson, J., A. McPherson, and W. Sutherland (Dec. 2006). "Visible and Infrared Survey Telescope for Astronomy: Progress Report". In: *The Messenger* 126, pp. 41–42.

Farrens, S. et al. (Oct. 2011). "Friends-of-friends groups and clusters in the 2SLAQ catalogue". In: *Monthly Notices of the Royal Astronomical Society* 417.2, pp. 1402–1416. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2011.19356.x. eprint: https://academic.oup.com/mnras/article-pdf/417/2/1402/3073223/mnras0417-1402.pdf. URL: https://doi.org/10.1111/j.1365-2966.2011.19356.x.

Haider, M. et al. (Apr. 11, 2016). "Large-scale mass distribution in the Illustris simulation". In: *Monthly Notices of the Royal Astronomical Society* 457.3, pp. 3024–3035. ISSN: 0035-8711. DOI: 10.1093/mnras/stw077. URL: https://doi.org/10.1093/mnras/stw077 (visited on 08/06/2022).

Hand, Nick et al. (Sept. 2018). "nbodykit: An Open-source, Massively Parallel Toolkit for Large-scale Structure". In: *The Astronomical Journal* 156.4, p. 160. DOI: 10.3847/1538-3881/aadae0. URL: https://doi.org/10.3847/1538-3881/aadae0.

Hansen, Sarah M et al. (2005). "Measurement of galaxy cluster sizes, radial profiles, and luminosity functions from SDSS photometric data". In: *The Astrophysical Journal* 633.1, p. 122.

Jauzac, Mathilde et al. (Aug. 2021). "Galaxy cluster cores as seen with VLT/MUSE: New strong-lensing analyses of RX J2129.4 + 0009, MS 0451.6 - 0305, and MACS J2129.4 - 0741". In: *Monthly Notices of the Royal Astronomical Society* 508.1. Conference Name: Monthly Notices of the Royal Astronomical Society, pp. 1206–1226. ISSN: 1365-2966. DOI: 10.1093/mnras/stab2270.

Katz, Harley et al. (Aug. 21, 2019). "Probing cosmic dawn with emission lines: predicting infrared and nebular line emission for ALMA and JWST". In: *Monthly Notices of the Royal Astronomical Society* 487.4, pp. 5902–5921. ISSN: 0035-8711. DOI: 10.1093/mnras/stz1672. URL: https://doi.org/10.1093/mnras/stz1672 (visited on 08/08/2022).

Kowalski, M. and Rubin (Oct. 2008). "Improved Cosmological Constraints from New, Old, and Combined Supernova Data Sets". In: *some journal name* 686.2, pp. 749–778. DOI: 10.1086/589937. arXiv: 0804.4142 [astro-ph].

Kron, R. G. (June 1980). "Photometry of a complete sample of faint galaxies." In: *apjs* 43, pp. 305–325. DOI: 10.1086/190669.

Kroupa, Pavel (2001). "On the variation of the initial mass function". In: *Monthly Notices of the Royal Astronomical Society* 322.2, pp. 231–246.

Kuchner, Ulrike et al. (Apr. 2020). "Mapping and characterization of cosmic filaments in galaxy cluster outskirts: strategies and forecasts for observations from simulations". In: *Monthly Notices of the Royal Astronomical Society* 494.4, pp. 5473–5491. ISSN: 0035-8711. DOI: 10.1093/mnras/staa1083. eprint: https://academic.oup.com/mnras/article-pdf/494/4/5473/33197796/staa1083.pdf. URL: https://doi.org/10.1093/mnras/staa1083.

Libeskind, Noam I. et al. (Jan. 1, 2018). "Tracing the cosmic web". In: *Monthly Notices of the Royal Astronomical Society* 473.1, pp. 1195–1217. ISSN: 0035-8711. DOI: 10.1093/mnras/stx1976. URL: https://doi.org/10.1093/mnras/stx1976 (visited on 07/20/2022).

Madau, Piero and Mark Dickinson (2014). "Cosmic Star-Formation History". In: *Annual Review of Astronomy and Astrophysics* 52, pp. 415–486.

Mandolesi, N. et al. (Feb. 1986). "Large-scale homogeneity of the Universe measured by the microwave background". In: *Nature* 319.6056. Number: 6056 Publisher: Nature Publishing Group, pp. 751–753. ISSN: 1476-4687. DOI: 10.1038/319751a0. URL: https://www.nature.com/articles/319751a0 (visited on 08/06/2022).

Manning, Sinclaire M. et al. (Jan. 2022). "Characterization of Two 2 mm detected Optically Obscured Dusty Star-forming Galaxies". In: *The Astrophysical Journal* 925.1, p. 23. DOI: 10.3847/1538-4357/ac366a. URL: https://doi.org/10.3847/1538-4357/ac366a.

Maraston, Claudia et al. (Mar. 2009). "Modelling the colour evolution of luminous red galaxies - improvements with empirical stellar spectra". In: *Monthly notices of the royal astronomical society* 394.1, pp. L107–L111. DOI: 10.1111/j.1745-3933.2009.00621.x. arXiv: 0809.1867 [astro-ph].

Mei, Simona et al. (Jan. 2007). "The ACS Virgo Cluster Survey. XIII. SBF Distance Catalog and the Three-dimensional Structure of the Virgo Cluster". In: *The Astrophysical Journal* 655.1, pp. 144–162. DOI: 10.1086/509598. URL: https://doi.org/10.1086/509598.

Mercurio, A. et al. (Nov. 11, 2015a). "Shapley Supercluster Survey: construction of the photometric catalogues and i -band data release". In: *Monthly Notices of the Royal Astronomical Society* 453.4, pp. 3686–3699. ISSN: 0035-8711, 1365-2966. DOI: 10.1093/mnras/stv1905. URL: https://academic.oup.com/mnras/article-lookup/doi/10.1093/mnras/stv1905 (visited on 07/20/2022).

— (Nov. 11, 2015b). "Shapley Supercluster Survey: construction of the photometric catalogues and i -band data release". In: *Monthly Notices of the Royal Astronomical Society* 453.4, pp. 3686–3699. ISSN: 0035-8711, 1365-2966. DOI: 10.1093/mnras/stv1905. URL: https://academic.oup.com/mnras/article-lookup/doi/10.1093/mnras/stv1905 (visited on 07/23/2022).

Mercurio, Amata and VST-GAME Team (2018). "VST-GAME: Galaxy Assembly as a Function of Mass and Environment with VST". In.

Muldrew, Stuart I., Nina A. Hatch, and Elizabeth A. Cooke (Sept. 21, 2015). "What are proto-clusters? – Defining high-redshift galaxy clusters and protoclusters". In: *Monthly Notices of the Royal Astronomical Society* 452.3, pp. 2528–2539. ISSN: 0035-8711. DOI: 10.1093/mnras/stv1449. URL: https://doi.org/10.1093/mnras/stv1449 (visited on 07/20/2022).

Muzzin, Adam et al. (Oct. 2008). "The Evolution of Dusty Star Formation and Stellar Mass Assembly in Clusters: Results from the IRAC 3.6, 4.5, 5.8, and 8.0 μm Cluster Luminosity Functions". In: *apj* 686.2, pp. 966–994. DOI: 10.1086/591542. arXiv: 0807.0227 [astro-ph].

Navarro, Julio F., Carlos S. Frenk, and Simon D. M. White (Dec. 1997). "A Universal Density Profile from Hierarchical Clustering". In: *The Astrophysical Journal* 490.2. Publisher: American Astronomical Society, pp. 493–508. ISSN: 0004-637X. DOI: 10.1086/304888. URL: https://doi.org/10.1086/304888 (visited on 08/09/2022).

Negrello, M. et al. (Sept. 2017). "On the statistics of proto-cluster candidates detected in the Planck all-sky survey". In: *Monthly Notices of the Royal Astronomical Society* 470.2, pp. 2253–2261. ISSN: 0035-8711, 1365-2966. DOI: 10.1093/mnras/stx1367. URL: https://academic.oup.com/mnras/article-lookup/doi/10.1093/mnras/stx1367 (visited on 07/20/2022).

Ouchi, Masami et al. (Aug. 2004). "Subaru Deep Survey. V. A Census of Lyman Break Galaxies at $z \leq 4$ and 5 in the Subaru Deep Fields: Photometric Properties". In: *The Astrophysical Journal* 611.2, pp. 660–684. DOI: 10.1086/422207. URL: https://doi.org/10.1086/422207.

Paillassa, M., E. Bertin, and H. Bouy (Feb. 1, 2020). "MAXIMASK and MAXITRACK: Two new tools for identifying contaminants in astronomical images using convolutional neural networks". In: *Astronomy & Astrophysics* 634. Publisher: EDP Sciences, A48. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201936345. URL: https://www.aanda.org/articles/aa/abs/2020/02/aa36345-19/aa36345-19.html (visited on 08/31/2022).

Peterson, J. R. and A. C. Fabian (Apr. 1, 2006). "X-ray spectroscopy of cooling clusters". In: *Physics Reports* 427.1, pp. 1–39. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2005.12.007. URL: https://www.sciencedirect.com/science/article/pii/S0370157306000020 (visited on 08/10/2022).

Pilbratt, G. L. et al. (July 1, 2010). "Herschel Space Observatory - An ESA facility for far-infrared and submillimetre astronomy". In: *Astronomy & Astrophysics* 518. Publisher: EDP Sciences, p. L1. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201014759. URL: https://www.aanda.org/articles/aa/abs/2010/10/aa14759-10/aa14759-10.html (visited on 08/09/2022).

Porter, Scott C. and Somak Raychaudhury (Dec. 21, 2005). "The Pisces-Cetus supercluster: a remarkable filament of galaxies in the 2dF Galaxy Redshift and Sloan Digital Sky surveys". In: *Monthly Notices of the Royal Astronomical Society* 364.4, pp. 1387–1396. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2005.09688.x. URL: https://doi.org/10.1111/j.1365-2966.2005.09688.x (visited on 08/25/2022).

Radovich, Mario et al. (Sept. 24, 2020). "AMICO galaxy clusters in KiDS-DR3: galaxy population properties and their redshift dependence". In: *Monthly Notices of the Royal Astronomical Society* 498.3, pp. 4303–4315. ISSN: 0035-8711, 1365-2966. DOI: 10.1093/mnras/staa2705. arXiv: 2009.03563[astro-ph]. URL: http://arxiv.org/abs/2009.03563 (visited on 07/20/2022).

Ramella, M. et al. (Jan. 23, 2001). "Finding Galaxy Clusters using Voronoi Tessellations". In: *Astronomy and Astrophysics* 368, p. 776. DOI: 10.1051/0004-6361:20010071.

Saavedra Esquivel, Sergio et al. (2021). "SHARKS DR1 validation and the quest for Extremely Red Objects". In: *None*.

Sand, David J. et al. (Mar. 2004). "The Dark Matter Distribution in the Central Regions of Galaxy Clusters: Implications for Cold Dark Matter". In: *The Astrophysical Journal* 604.1. Publisher: American Astronomical Society, pp. 88–107. ISSN: 0004-637X. DOI: 10.1086/382146. URL: https://doi.org/10.1086/382146 (visited on 08/09/2022).

Santos, Joana S et al. (2015). "The reversal of the SF–density relation in a massive, X-ray-selected galaxy cluster at z= 1.58: results from Herschel". In: *Monthly Notices of the Royal Astronomical Society: Letters* 447.1, pp. L65–L69.

Scaramella, R. et al. (May 2014). "Euclid space mission: a cosmological challenge for the next 15 years". In: *Proceedings of the International Astronomical Union* 10 (S306). Publisher: Cambridge University Press, pp. 375–378. ISSN: 1743-9213, 1743-9221. DOI: 10.1017/S1743921314011089. URL: https://www.cambridge.org/core/journals/proceedings-of-the-international-astronomical-union/article/euclid-space-mission-a-cosmological-challenge-for-the-next-15-years/C1C2127FE02AC0900D9F75D08E6B7DD6 (visited on 08/08/2022).

Shectman, Stephen A. et al. (Oct. 1, 1996). "The Las Campanas Redshift Survey". In: *The Astrophysical Journal* 470. ADS Bibcode: 1996ApJ...470..172S, p. 172. ISSN: 0004-637X. DOI: 10.1086/177858. URL: https://ui.adsabs.harvard.edu/abs/1996ApJ...470..172S (visited on 08/07/2022).

Smith, Graham P. et al. (May 2005). "A Hubble Space Telescope lensing survey of X-ray luminous galaxy clusters - IV. Mass, structure and thermodynamics of cluster cores at z = 0.2". In:

Monthly Notices of the Royal Astronomical Society 359.2, pp. 417–446. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2005.08911.x. eprint: https://academic.oup.com/mnras/article-pdf/359/2/417/3093384/359-2-417.pdf. URL: https://doi.org/10.1111/j.1365-2966.2005.08911.x.

Strateva, Iskra et al. (Oct. 2001). "Color Separation of Galaxy Types in the Sloan Digital Sky Survey Imaging Data". In: *The Astronomical Journal* 122.4, pp. 1861–1874. DOI: 10.1086/323301. URL: https://doi.org/10.1086/323301.

Sutherland, Will et al. (Mar. 2015). "The Visible and Infrared Survey Telescope for Astronomy (VISTA): Design, technical overview, and performance". In: *Astronomy & Astrophysics* 575, A25, A25. DOI: 10.1051/0004-6361/201424973. arXiv: 1409.4780 [astro-ph.IM].

Valiante, E. et al. (July 2016). "The Herschel-ATLAS data release 1 - I. Maps, catalogues and number counts". In: *Monthly Notices of the Royal Astronomical Society* 462.3, pp. 3146–3179. ISSN: 0035-8711. DOI: 10.1093/mnras/stw1806. eprint: https://academic.oup.com/mnras/article-pdf/462/3/3146/8009213/stw1806.pdf. URL: https://doi.org/10.1093/mnras/stw1806.

Vieira, J. D. et al. (Mar. 11, 2013). *Dusty starburst galaxies in the early Universe as revealed by gravitational lensing*. DOI: 0.1038/nature12001. arXiv: 1303.2723[astro-ph]. URL: http://arxiv.org/abs/1303.2723 (visited on 07/20/2022).

Werner, M. W. et al. (Sept. 2004). "The Spitzer Space Telescope Mission". In: *The Astrophysical Journal Supplement Series* 154.1, pp. 1–9. DOI: 10.1086/422992. URL: https://doi.org/10.1086/422992.

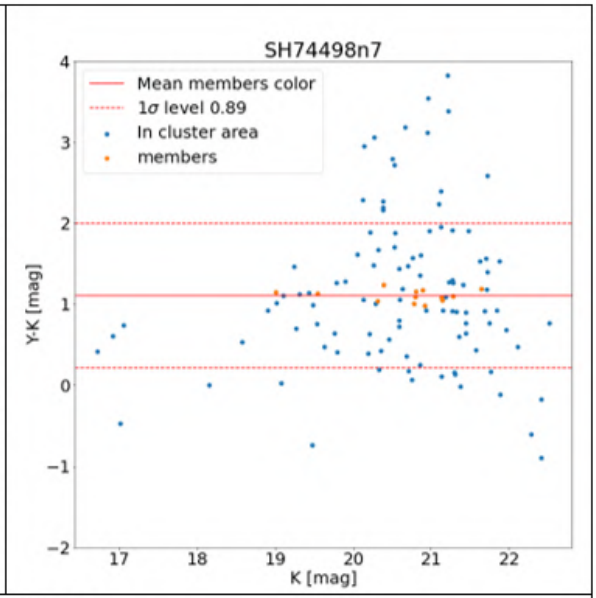
Wright, Edward L. et al. (Nov. 2010). "THE WIDE-FIELD INFRARED SURVEY EXPLORER (WISE): MISSION DESCRIPTION AND INITIAL ON-ORBIT PERFORMANCE". In: *The Astronomical Journal* 140.6, pp. 1868–1881. DOI: 10.1088/0004-6256/140/6/1868. URL: https://doi.org/10.1088/0004-6256/140/6/1868.

Zavala, J. A. et al. (Dec. 2019). "On the Gas Content, Star Formation Efficiency, and Environmental Quenching of Massive Galaxies in Protoclusters at z ≈ 2.0-2.5". In: *The Astrophysical Journal* 887.2. Publisher: American Astronomical Society, p. 183. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ab5302. URL: https://doi.org/10.3847/1538-4357/ab5302 (visited on 08/09/2022).

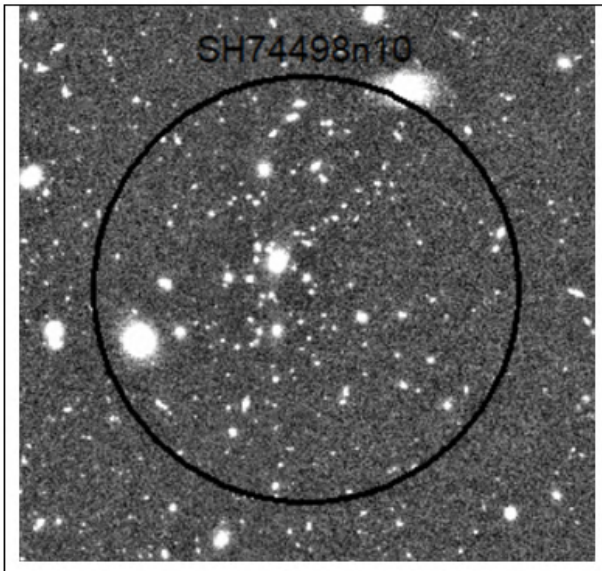
Appendix A

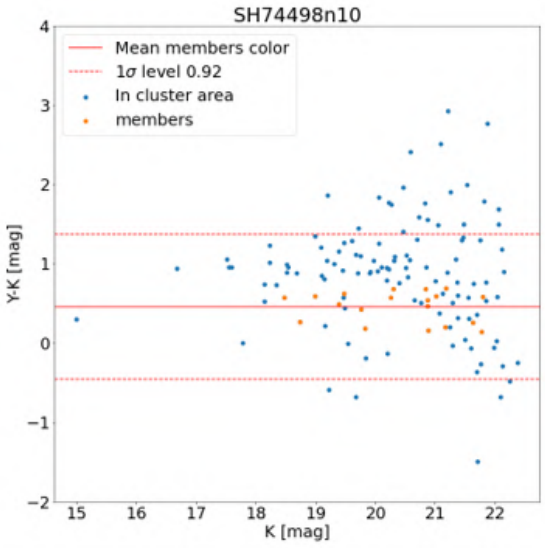
Candidates identity chart

Here I present a series of 14 charts describing some SHARKS candidate clusters. For each chart, on the left image it is shown the candidate cluster, while on the right image it is shown the Y-Ks color magnitude diagram of cluster members (in orange) and of all the galaxies projected on the candidate area (in blue). The continuous red line represents the average color of the cluster members. The dashed lines indicate the 1σ dispersion level of the color of all galaxies in the cluster area from the average color of the cluster members. The table contains ID, position in equatorial coordinates (degrees), redshift estimate, number of cluster members, total flux of the cluster members and estimated luminosity of the cluster.

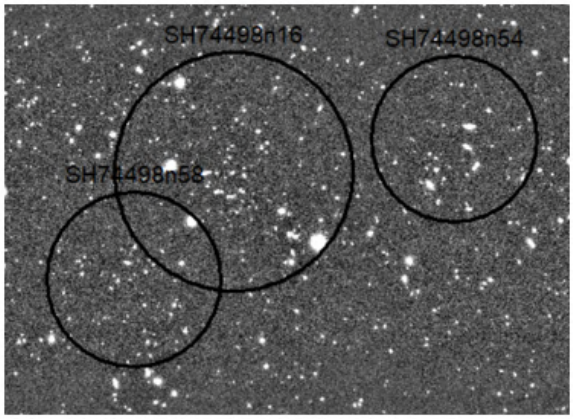


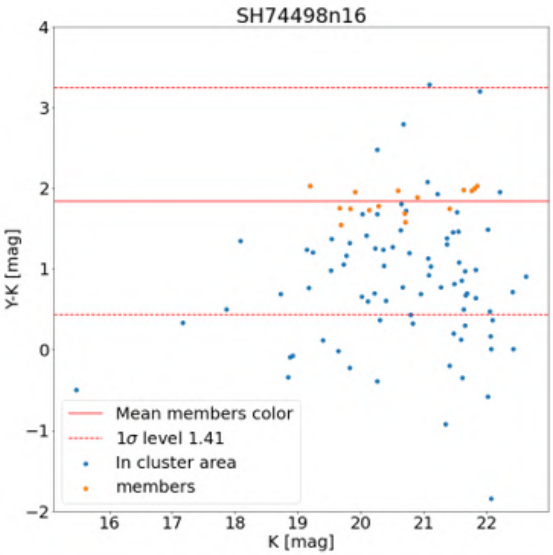
SH74498n7	
RA (deg): 8.60594446564157	**DEC (deg):** -31.353388115835525
Z (±0.15)	1.43
R (arcsec)	92.0
MEMBERS	13
FLUX (erg·s^{-1}· cm^{-2})	4.44e-13 ± 8.41e-15
LUMINOSITY (L_{sun})	1.55e12 ± 1.10e12



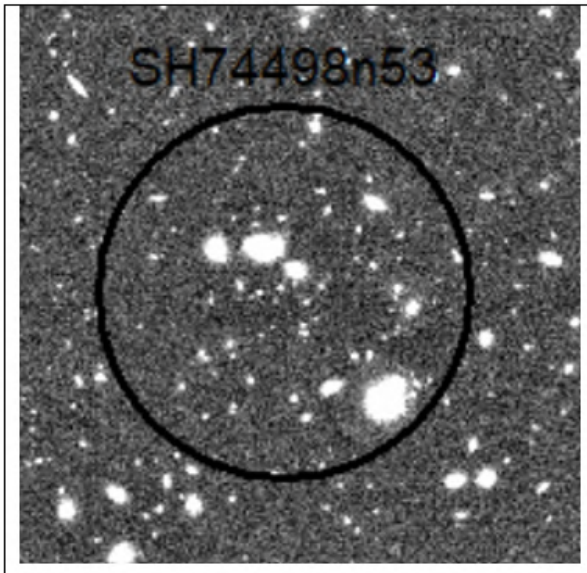


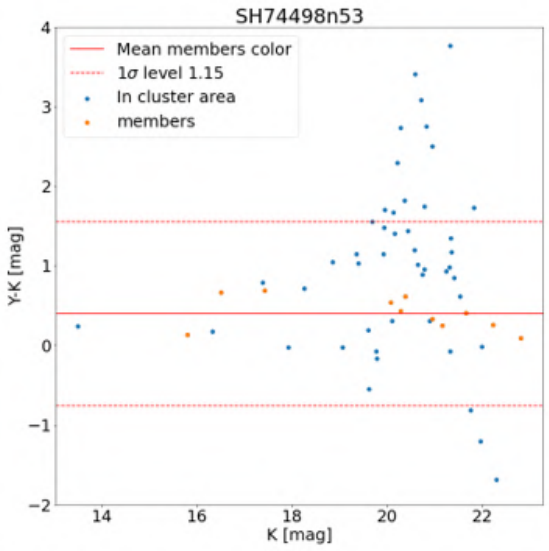
SH74498n10	
RA (deg): 9.215376417788757	**DEC (deg):** -31.747399888724825
Z (±0.15)	0.87
R (arcsec)	94.0
MEMBERS	19
FLUX (erg·s^{-1}· cm^{-2})	1.02e-12 ± 9.88e-15
LUMINOSITY (L_{sun})	1.04e12 ± 1.21e12



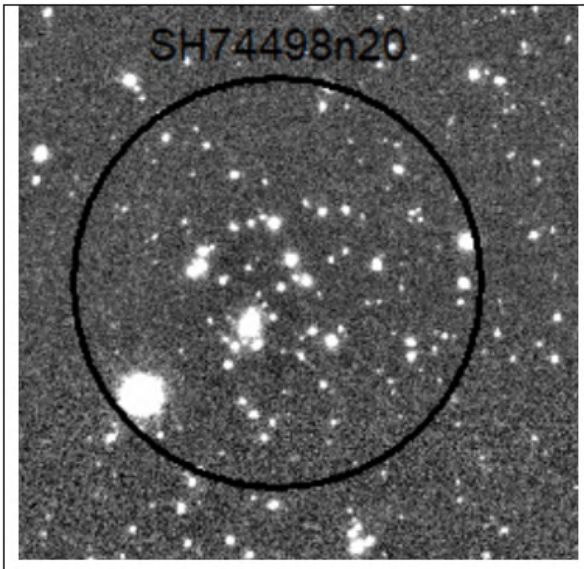


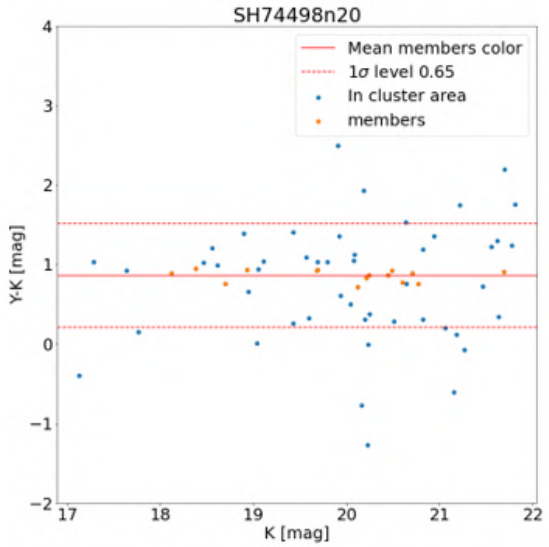
SH74498n16	
RA (deg): 9.027962598475598	**DEC (deg):** -31.530074949048025
Z (±0.15)	2.88
R (arcsec)	82.0
MEMBERS	16
FLUX (erg·s^{-1}· cm^{-2})	5.84e-13 ± 8.68e-15
LUMINOSITY (L_{sun})	1.12e13 ± 3.87e12



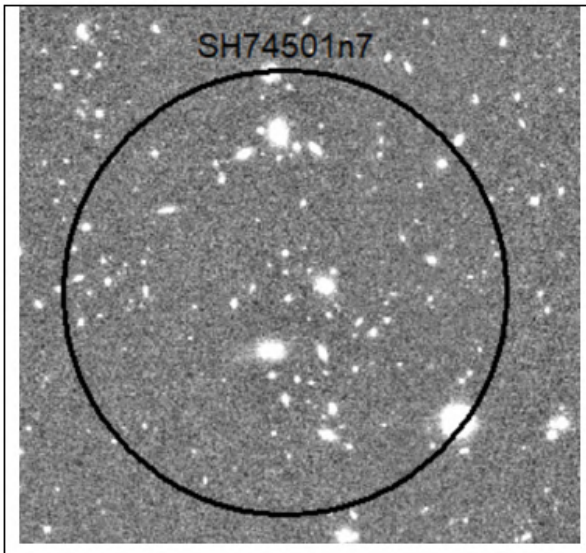
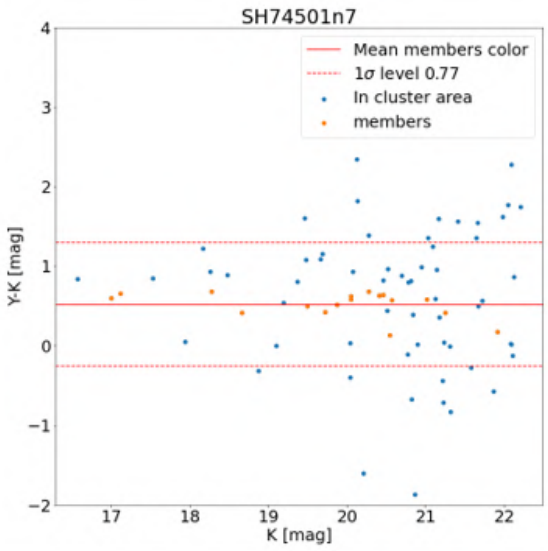


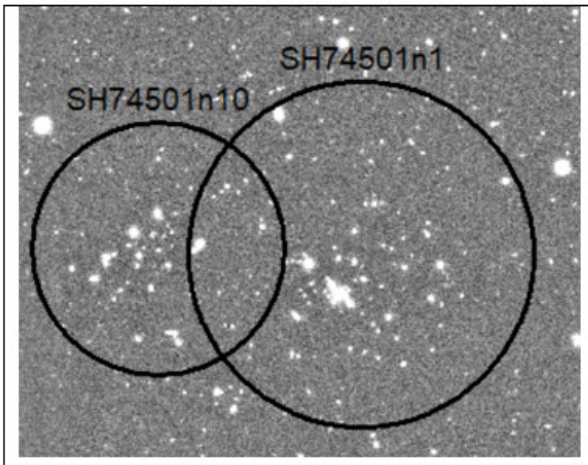
SH74498n53	
RA (deg): 9.965015102709568	**DEC (deg):** -31.340699824207842
Z (±0.15)	0.8
R (arcsec)	57.0
MEMBERS	11
FLUX (erg·s^{-1}· cm^{-2})	4.27e-12 ± 1.79e-14
LUMINOSITY (L_{sun})	3.55e12 ± 4.47e12

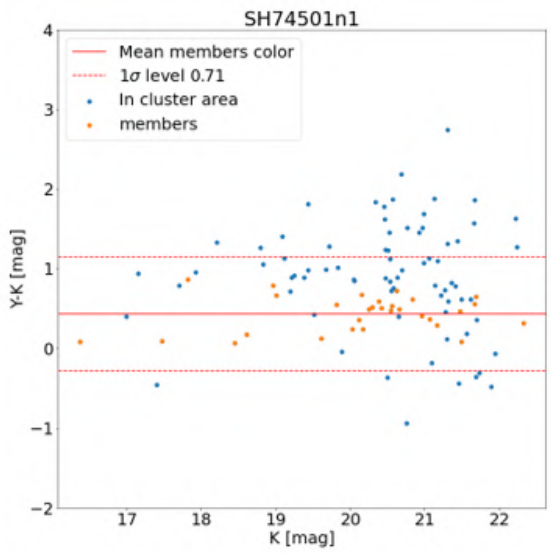




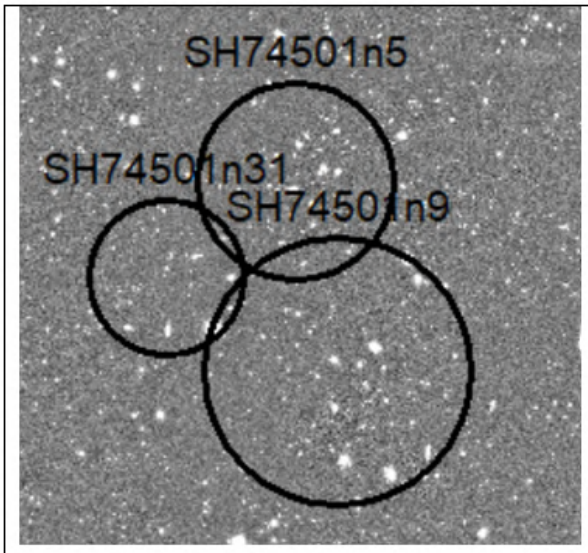
SH74498n20	
RA (deg): 8.63620709978397	**DEC (deg):** -31.592381512134995
Z (±0.15)	1.19
R (arcsec)	63.0
MEMBERS	15
FLUX (erg·s^{-1}· cm^{-2})	1.21e-12 ± 8.56e-15
LUMINOSITY (L_{sun})	8.24e11 ± 1.02e12

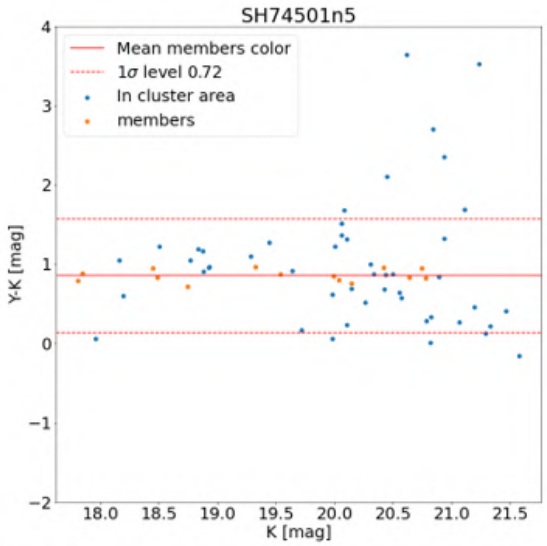
<table>
<tr>
<td> 
</td>
<td> 
</td>
</tr>
<tr><td colspan="2">SH74501n17</td></tr>
<tr><td>RA (deg): 10.660749558969504</td><td>DEC (deg): -31.359773458857532</td></tr>
<tr><td>Z (±0.15)</td><td>0.90</td></tr>
<tr><td>R (arcsec)</td><td>75.0</td></tr>
<tr><td>MEMBERS</td><td>17</td></tr>
<tr><td>FLUX (erg·s^{-1}· cm^{-2})</td><td>2.4e-12 ± 1.55e-14</td></tr>
<tr><td>LUMINOSITY (L_{sun})</td><td>4.20e12 ± 3.94e12</td></tr>
</table>



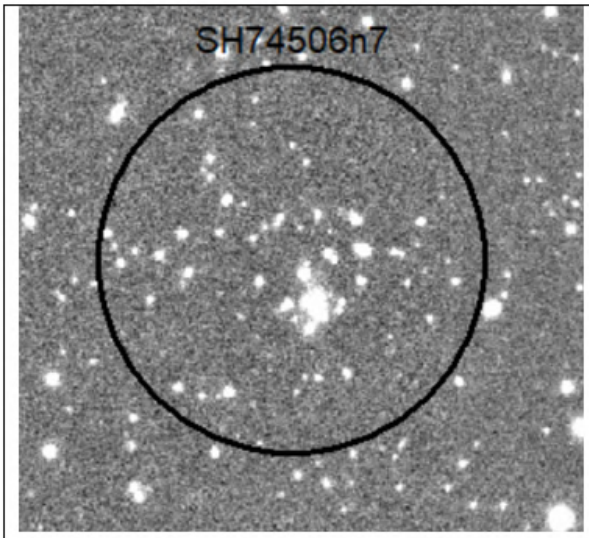
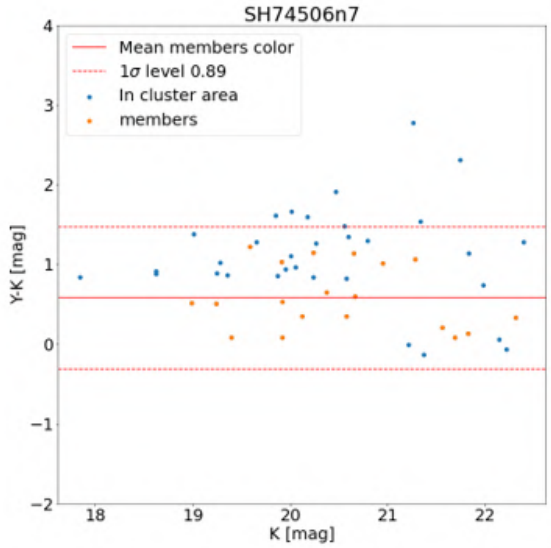


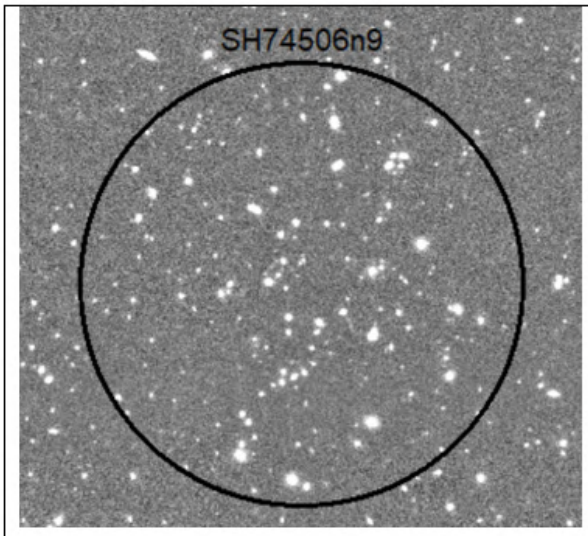
SH74501n1	
RA (deg): 10.477765470527505	**DEC (deg):** -31.788329055391994
Z (±0.15)	0.84
R (arcsec)	90.0
MEMBERS	30
FLUX (erg·s^{-1}· cm^{-2})	3.6e-12 ± 1.67e-14
LUMINOSITY (L_{sun})	3.38e12 ± 4.05e12

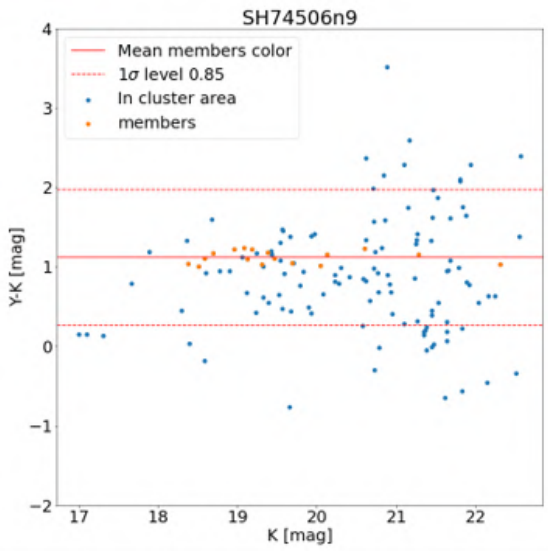




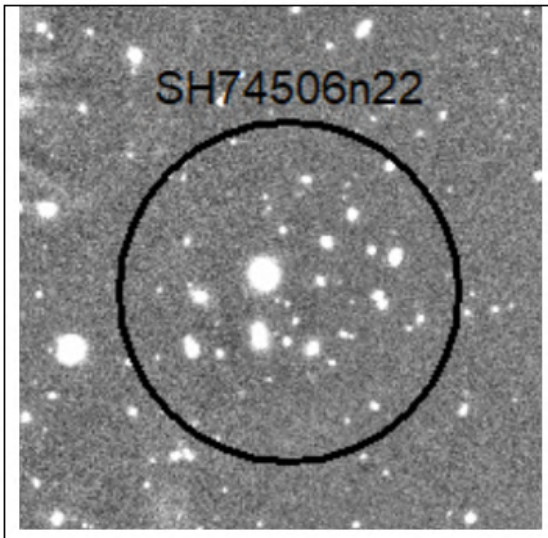
SH74501n5	
RA (deg): 10.521826114851114	**DEC (deg):** -30.938792790117123
Z (±0.15)	1.18
R (arcsec)	78.0
MEMBERS	14
FLUX (erg·s^{-1}· cm^{-2})	1.71e-12 ± 1.15e-14
LUMINOSITY (L_{sun})	3.71e12 ± 3.19e12

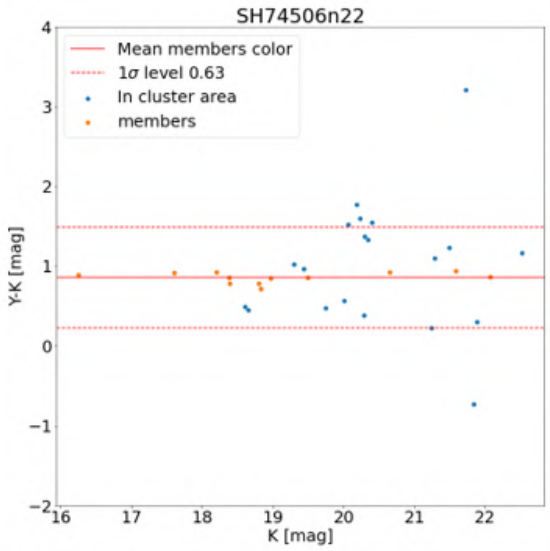
<table>
<tr><td> 
</td><td> 
</td></tr>
<tr><td colspan="2">SH74506n7</td></tr>
<tr><td>RA (deg): 11.047448756735887</td><td>DEC (deg): -30.930792108537617</td></tr>
<tr><td>Z (±0.15)</td><td>0.96</td></tr>
<tr><td>R (arcsec)</td><td>67.0</td></tr>
<tr><td>MEMBERS</td><td>19</td></tr>
<tr><td>FLUX (erg·s^{-1}· cm^{-2})</td><td>8.19e-13 ± 1.06e-14</td></tr>
<tr><td>LUMINOSITY (L_{sun})</td><td>1.07e12 ± 1.13e12</td></tr>
</table>



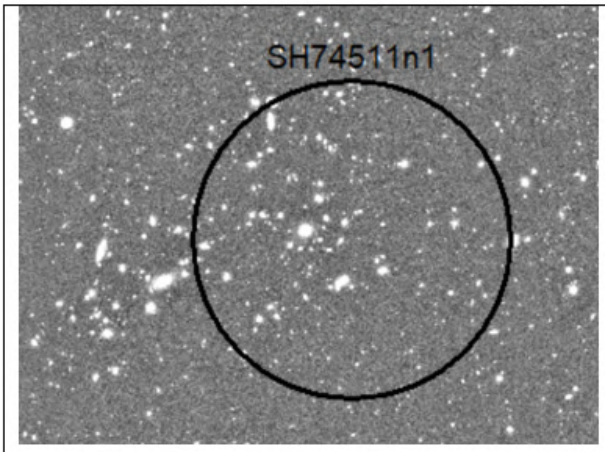


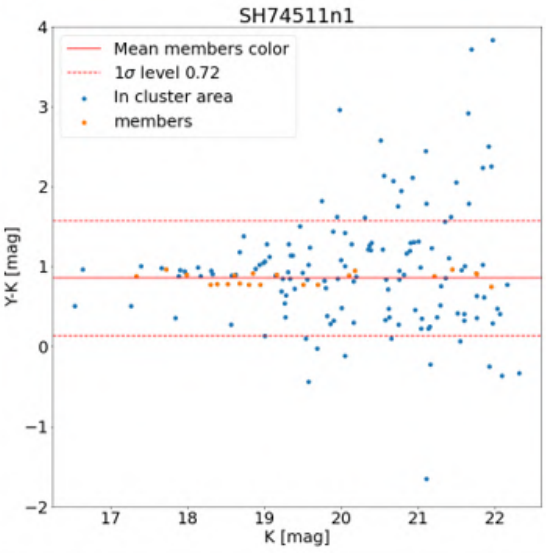
SH74506n9	
RA (deg): 10.477601550695482	**DEC (deg):** -30.071591520923008
Z (±0.15)	1.44
R (arcsec)	85.0
MEMBERS	17
FLUX (erg·s^{-1}· cm^{-2})	1.69e-12 ± 1.16e-14
LUMINOSITY (L_{sun})	5.99e12 ± 4.22e12



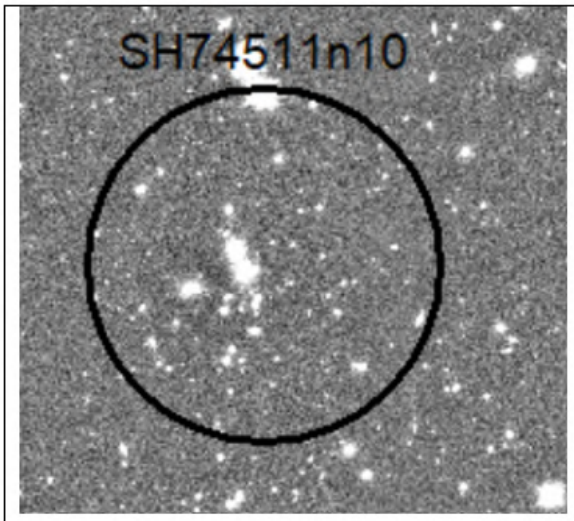


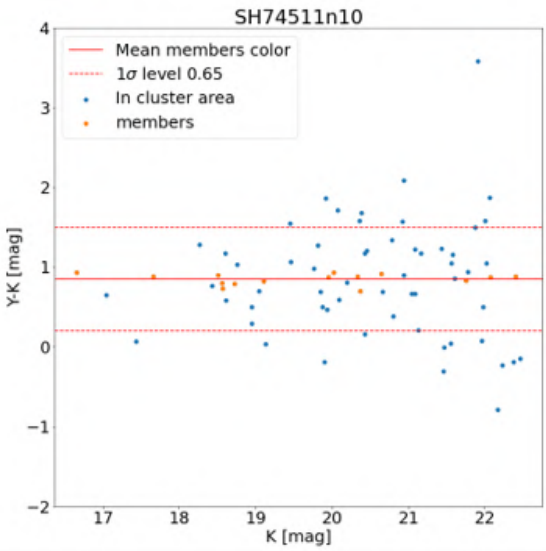
SH74506n22	
RA (deg): 11.380799867891351	**DEC (deg):** -30.60258769736205
Z (±0.15)	1.19
R (arcsec)	70.0
MEMBERS	12
FLUX (erg·s^{-1}· cm^{-2})	3.24e-12 ± 1.46e-14
LUMINOSITY (L_{sun})	7.18e12 ± 6.12e12



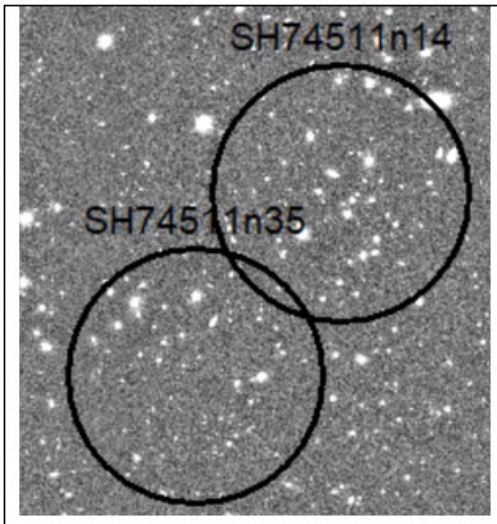
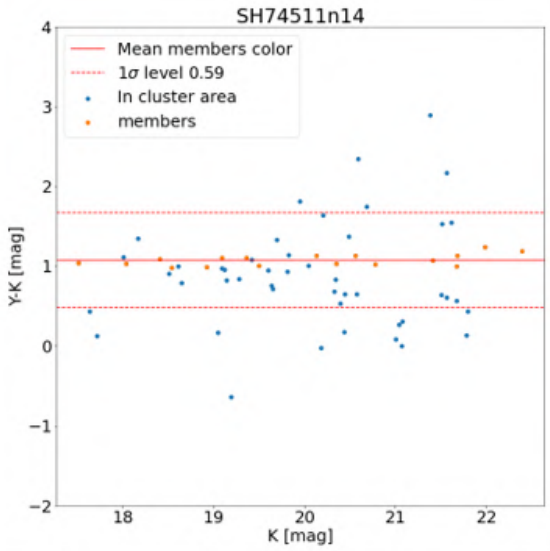


SH74511n1	
RA (deg): 8.55515042069142	**DEC (deg):** -30.625503496576982
Z (±0.15)	1.18
R (arcsec)	115.0
MEMBERS	21
FLUX (erg·s^{-1}· cm^{-2})	3.09e-12 ± 1.15e-14
LUMINOSITY (L_{sun})	6.71e12 ± 5.77e12





SH74511n10	
RA (deg): 9.517977195516487	**DEC (deg):** -30.179317757799392
Z (±0.15)	0.73
R (arcsec)	70.0
MEMBERS	13
FLUX (erg·s^{-1}· cm^{-2})	6.28e-13 ± 7.78e-15
LUMINOSITY (L_{sun})	5.55e12 ± 4.77e12

<table>
<tr>
<td> 
</td>
<td> 
</td>
</tr>
<tr><td colspan="2">SH74511n14</td></tr>
<tr><td>RA (deg): 8.970398392560107</td><td>DEC (deg): -29.694883305452716</td></tr>
<tr><td>Z (±0.15)</td><td>1.42</td></tr>
<tr><td>R (arcsec)</td><td>66.0</td></tr>
<tr><td>MEMBERS</td><td>17</td></tr>
<tr><td>FLUX (erg·s^{-1}· cm^{-2})</td><td>1.79e-12 ± 1.17e-14</td></tr>
<tr><td>LUMINOSITY (L_{sun})</td><td>6.14e12 ± 4.38e12</td></tr>
</table>