UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI FISICA E ASTRONOMIA
"GALILEO GALILEI"

CORSO DI DOTTORATO DI RICERCA IN ASTRONOMIA
Ciclo XXXV

THE GALAXY DENSITY FIELD AS A TOOL TO DETERMINE COSMOLOGICAL PARAMETERS AND CHARACTERIZE GALAXY EVOLUTION

Tesi redatta con il contributo della Fondazione Cariparo

Ph.D. School Coordinator:	Ch.mo Prof. Giovanni Carraro
Supervisor:	Ch.mo Prof. Giulia Rodighiero
Co-supervisor:	Ch.mo Prof. Amata Mercurio
	Ch.mo Prof. Benedetta Vulcani
Ph.D. Candidate:	Gilberto Nicolas Estrada Martinez

Grazie

Our world is now understood to be a world where something really happens; the whole story of the world need not have been written down in the first quantum like a song on the disc of a phonograph. The whole matter of the world must be present at the beginning, but the story it has to tell may be written step by step.

The beginning of the world from the point of view of quantum theory

G. Lemaître, Nature, 127, page 706 (1931).

Contents

Abstract

In the current cosmological model, the large-scale structures of the Universe are the result of matter's evolution under gravity through cosmic time. Galaxy clusters ($\sim$ 1 Mpc), filaments of the cosmic web ($\sim$ 10 Mpc), and in general the distribution of galaxies across large volumes ($\sim$ 100 Mpc) are some of the tracers of the large-scale structure. Galaxies are the fundamental observable of large-scale structures, and the preferential scenario where the baryonic matter condenses and becomes luminous by the process of star formation.

In this thesis, I have explored the relationship between galaxies and the large-scale structure of the Universe. The key tool that I have exploited is the galaxy density field, used as a tracer of the local environment which influences the galaxy assembly and evolution, and as a tracer of the clustering of matter across large scales.

Observational studies have widely demonstrated that a galaxy's physical properties are strongly affected by the surrounding environment. On one side, gas inflows provide galaxies with new fuel for star formation. On the other side, the high temperatures and densities of the medium are expected to induce quenching in the star formation. Observations of large structures, in particular filaments at the cluster outskirts ($r>2r_{200}$), are currently limited to the low redshift Universe. Deep and wide photometric data, better if combined with spectroscopic redshifts, are required to explore several scenarios on galaxy evolution at intermediate redshift ($z \sim 0.4$).

To have a deeper understanding of the process of galaxy formation and evolution on large-scale structures, I have worked with data of the Galaxy Assembly as a function of Mass and Environment with the VLT Survey Telescope (VST-GAME) survey. The project is aimed at gathering deep ($r<24.4$ mag) and wide (approx. 20x20Mpc^2) observations at optical (u, g, r, i, VST) wavelengths for six massive galaxy clusters at $0.2<z<0.6$, complemented with near-infrared data (Y, J, Ks, VISTA, ESO public survey GCAV). The aim is to investigate galaxy evolution in a wide range of stellar masses and environmental conditions.

The first part of this thesis is therefore focused on the multiband catalog construction for the cluster MACS J0416.1-2403 (z=0.397), from the source extraction up to photometric redshift computation,

using optical and near-infrared bands. In particular, I developed a procedure to mask the OmegaCAM fields from artifacts generated by bright stars. This procedure can be applied to other deep VST and VISTA surveys. In the second part, I define the cluster membership, and the sequence of red galaxies, and compute the galaxy density field around MACSJ0416, up to more than 10 Mpc of the cluster core.

Studying galaxy colors according to their local environment I found that the $g-r$ colors show bimodal behaviors in all the environments, but the peak of the distribution of red galaxies shifts toward redder colors with increasing density. Additionally, I also found three overdense regions in the cluster outskirts at r$\sim$5r$_{200}$. The galaxies situated within these structures exhibit average densities and luminosities akin to those found within the cluster core. The similarity in color between galaxies on these substructures and those in the cluster core provides compelling evidence for pre-processing phenomena occurring within these substructures.

The second part of this thesis is focused on cosmological scales, using the VIMOS Public Extragalactic Redshift Survey (VIPERS), over the redshift range $0.6 < z < 1$. The analysis is optimized to extract the cosmological parameters while fully accounting for the complex survey geometry. I employ a Gibbs sampling algorithm to iteratively draw samples of the galaxy density field. Despite the high number of degrees of freedom, the samples converge to the joint posterior distribution and give self-consistent constraints on the model parameters. I validate the approach using VIPERS mock galaxy catalogs. I find that the precision of the results matches those of the traditional analyses applied to the VIPERS data that use more constrained models. As a result, I obtained constraints on the values of key cosmological parameters such as the matter density, the baryon fraction, and the amplitude of matter fluctuations.

Overall, this work has contributed to a better understanding of the relationship between galaxies and their environment on large scales and has provided valuable insights into the structure and evolution of the Universe.

Sommario

Nel modello cosmologico attuale, le strutture a grande scala dell'Universo sono il risultato dell'evoluzione della materia sotto l'influenza della gravità nel corso del tempo cosmico. Gli ammassi di galassie ($\sim$ 1 Mpc), i filamenti della rete cosmica ($\sim$ 10 Mpc) e in generale la distribuzione delle galassie su grandi volumi ($\sim$ 100 Mpc) sono alcuni dei traccianti della struttura a grande scala. Le galassie sono l'osservabile fondamentale di queste strutture ed il luogo preferenziale in cui la materia barionica si condensa e diventa luminosa attraverso il processo di formazione stellare.

In questa tesi, ho esplorato il rapporto tra le galassie e la struttura a grande scala dell'Universo. Lo strumento chiave a questo scopo è il campo di densità di galassie, utilizzato sia come indicatore dell'ambiente locale delle galassie, sia come indicatore dell'aggregazione di materia su grandi scale.

Le osservazioni hanno ampiamente dimostrato che le proprietà fisiche delle galassie sono fortemente influenzate dall'ambiente circostante. Da un lato, gli afflussi di gas forniscono alle galassie nuovo carburante per la formazione stellare. Dall'altro lato, le alte temperature e turbolenze del mezzo intergalattico sopprimono la formazione stellare. Le osservazioni di grandi strutture, in particolare i filamenti nei pressi dei cluster ($r>2r_{200}$), sono attualmente limitate all'Universo a basso redshift.

Per avere una comprensione più profonda del processo di evoluzione delle galassie sulla struttura a grande scala, ho lavorato con i dati dell'*Galaxy Assembly as a function of Mass and Environment with VST* (VST-GAME). Il progetto ha lo scopo di raccogliere osservazioni profonde ($r<24.4$ mag) e su grandi aree (circa 20x20Mpc^2) alle lunghezze d'onda nello spettro dell'ottico (u, g, r, i, VST) per sei cluster di galassie a redshift $0{,}2<z<0{,}6$. Le osservazioni somo completate con dati a lunghezze d'onda nel vicino-infrarosso (Y, J, Ks, VISTA, survey pubblicoa ESO GCAV).

La prima parte di questa tesi si concentra quindi sulla costruzione del catalogo multibanda dell'ammasso MACS J0416.1-2403 ($z=0.397$), dall'estrazione delle sorgenti fino al calcolo del redshift fotometrico, utilizzando bande nell'ottico e vicino-infrarosso. In particolare, ho sviluppato una procedura per mascherare i campi OmegaCAM da

artefatti generati da stelle luminose. Questa procedura può essere applicata ad altri survey profondi VST e VISTA. Nella seconda parte, definisco l'appartenenza al cluster, la sequenza di galassie rosse e calcolo il campo di densità di galassie intorno a MACSJ0416, fino a oltre 10 Mpc del nucleo dell'ammasso.

Studiando i colori delle galassie in base al loro ambiente locale, ho scoperto che i colori $g - r$ mostrano comportamenti bimodali in tutti gli ambienti, ma il picco della distribuzione delle galassie rosse si sposta verso colori più rossi con l'aumento della densità. Inoltre, ho anche trovato tre regioni sovra-dense negli estremi del cluster a $r \sim 5r_{200}$. Le galassie situate all'interno di queste strutture presentano densità e luminosità medie simili a quelle riscontrate all'interno del nucleo del cluster. La similarità nel colore tra le galassie presenti in queste sottostrutture e quelle del nucleo del cluster fornisce una forte evidenza di fenomeni di pre-elaborazione che si verificano all'interno di tali sottostrutture.

La seconda parte di questa tesi si concentra sulla scala cosmologica, utilizzando il *VIMOS Public Extragalactic Redshift Survey* (VIPERS), nell'intervallo di redshift $0.6 < z < 1$. L'analisi è ottimizzata per estrarre i parametri cosmologici tenendo conto pienamente della geometria complessa della survey. Ho utilizzato un algoritmo del tipo Gibbs sampler per generare campioni del campo di densità di galassie. Nonostante l'elevato numero di gradi di libertà, i campioni convergono ad una distribuzione congiunta e forniscono vincoli coerenti sui parametri del modello. Ho validato l'approccio utilizzando i cataloghi simulati della survey. Riscontro che la precisione dei risultati corrisponde a quella delle analisi tradizionali applicate ai dati VIPERS che utilizzano modelli più vincolanti. Come risultato, ho ottenuto vincoli sui valori dei principali parametri cosmologici come la densità di materia, la frazione di barioni e l'ampiezza delle fluttuazioni di materia.

In generale, questo lavoro ha contribuito a una maggiore comprensione del rapporto tra le galassie e il loro ambiente su grandi scale e ha fornito preziose intuizioni sulla struttura e l'evoluzione dell'Universo.

Thesis outline

Introduction: Brief description of the formation of the large-scale structures according to the cosmological model, and presentation of the current picture of galaxy evolution according to the environment.

The first part of the thesis is focused on *galaxy evolution*:

- **Chapter 2** introduces the Galaxy Assembly as a Function of Mass and Environment with VST (VST-GAME) survey and in particular my work on the photometric assessment that led to the publication of the first multi-band catalog using survey data for the cluster MACSJ0416 (first part of Estrada et al. 2022b).
- **Chapter 3** presents the cluster member selection to compute the density field around MACSJ0416 and to study environmental effects in shaping the properties of galaxies at large clustercentric distances (second part of Estrada et al. 2022b).

The second part of the thesis is focused on *cosmology*:

- **Chapter 4** presents the VIMOS extra galactic galaxy survey (VIPERS, Guzzo et al. 2014), and explains my work to unmask the VIPERS survey from the induced effects of the selection function to bring joint probabilities on the parameters of the ΛCDM model Estrada et al. (2022a).

Summary and conclusions: The summary of my work and the future projects that can be done starting from the main chapters.

Convention

In this thesis, I assume the ΛCDM flat model using the Planck Collaboration et al. (2020) parameters, in particular: $A_s = 2.0968 \times 10^{-9}$ and $n_s = 0.9652$, $Y_{He} = 0.2454$. In the first part of the thesis it is assumed $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$ and $h = 0.7$. In the second part Ω_M and Ω_b are free parameters of the model, and $h = 0.6732$. The magnitudes are given in the AB photometric system and all the distance measurements are in comoving distances.

Introduction

Chapter 1

A structured Universe

The formation and evolution of galaxies are central topics in modern astrophysics, as they are the fundamental observables of the large-scale structure of the Universe. In this Ph.D. thesis, I investigate the role of the environment in shaping the properties of galaxies and the large-scale structure of the Universe. To do so, I use the density field, a measure of the distribution of matter in the Universe, as a tool to understand the influences on galaxy evolution.

The density field is a crucial quantity in the study of the large-scale structure of the Universe, as it encodes information about the distribution and clustering of matter. By analyzing the density field, is possible to investigate the role of various physical processes, such as gravitational collapse, gas cooling, and feedback from supernovae and active galactic nuclei, in shaping the properties of galaxies and the large-scale structure.

Through this thesis, I aim to gain a deeper understanding of the complex interplay between galaxies and their environments, and how this interplay shapes the large-scale structure of the Universe. By studying the observability of the large-scale structure through the lens of galaxy evolution, I hope to shed light on the fundamental processes that govern the formation and evolution of galaxies and the Universe as a whole.

1.1 The cosmological model and Formation of structures

This section contains a general introduction to the ΛCDM model and its parametrization. Additionally, it introduces the theory that describes the evolution of the cosmological density field, from a homogeneous state up to the structured state that is present in the late Universe.

1.1.1 General introduction to the ΛCDM model

The behavior of matter from small to large scale is determined by the dominant force in the Universe, gravity. The accepted theory of gravitation is General Relativity (GR),

formulated by Albert Einstein in 1916 (Einstein 1916). GR is based on two postulates: (1) our place in the Universe is not distinguished from other locations, *homogeneity*, and (2) the distribution of matter around us is isotropic, at least on large scales, *isotropy*. Both conditions are known as the *cosmological principle*.

The Einstein equations for GR govern how the metric of space-time responds to the energy-momentum of the matter:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G T_{\mu\nu}, \tag{1.1}$$

where $R_{\mu\nu}$ is the Ricci tensor, R is the Ricci scalar, $g_{\mu\nu}$ is the metric, G is the gravitational constant and $T_{\mu\nu}$ is the Energy-Momentum tensor. The left-hand side of the eq. 1.1 describes the curvature of space-time while the right-hand side is the energy-momentum content of the matter. For describing the Universe using the Einstein equation, a metric that describes the distance between two points in a generic space-time is needed. Models assuming a Universe that has an origin, and therefore not static but expanding or contracting, and following the cosmological principle are represented by the Friedmann–Lemaître–Robertson–Walker (FLRW) metric. According to the FLRW models, the Universe used to be smaller and hotter in the past and cools down as an effect of the expansion. Assuming that the known laws of physics were also valid in the past, we can trace back the Universe up to its primordial state, a very hot and very dense state, the so-called Big Bang. The FLRW metric is given by:

$$ds^2 = -dt^2 + \frac{a^2(t)}{c^2}\Big[\frac{dr^2}{1-kr^2} + r^2 d\Omega^2\Big], \tag{1.2}$$

where ds^2 is the four-distance between two points, dt is the time interval, $a(t)$ is the time-dependent scale factor and k is the curvature of the Universe. dr represents a distance coordinate and $d\Omega$ is the metric of a sphere. The curvature could be $k \in (-1, 0, 1)$, where the case $k = -1$ corresponds to constant negative curvature in space-time and is called a *open Universe*; $k = 0$ corresponds to no curvature and is called *flat Universe*; $k = 1$ corresponds to positive curvature and is called *closed Universe*.

Combining the FLRW metrics and the Einstein equation, we obtain the Friedmann equations, relating the scale factor to the energy-momentum of the Universe:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \tag{1.3}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) \tag{1.4}$$

In the precedent equations, p and ρ are the total pressure and the density of the Universe, composed of several species: radiation, matter, or vacuum. Those variables come from the energy-momentum tensor on the Einstein equation. The rate of expansion of the Universe is characterized by the Hubble parameter:

$$H = \frac{\dot{a}}{a}, \tag{1.5}$$

and its value at the present cosmological epoch is the Hubble constant, H_0. For practical characterization H_0 is parametrized as:

$$H_0 = 100\,h\,[km/sec/Mpc], \tag{1.6}$$

where h is an adimensional positive number that encodes the present rate of expansion of the Universe. The density parameter is defined as:

$$\Omega = \frac{8\pi G}{3H^2}\rho. \tag{1.7}$$

The density parameter can be decomposed into different species ρ_i according to the dominant components of the Universe and for each species can be specified an equation of state $p_i = p_i(\rho_i)$.

The conservation of energy of a fluid in comoving coordinates can be expressed as $\bigtriangledown_\mu T_0^\mu = 0$. A generic form of an equation of state is $p = w\rho$. Putting the generic equation of state into the energy conservation, we can obtain an equation that relates the evolution of the density content of the Universe with the main source of energy:

$$\frac{\dot{\rho}}{\rho} = -3(1+w)\frac{\dot{a}}{a}. \tag{1.8}$$

This expression sets that the evolution of the scale factor $a(t)$ depends on the main energy source during a cosmic epoch. For simplicity, the density can be decomposed into three main sources: *radiation* (relativistic matter), *matter* (collision-less, non-relativistic particles), and *vacuum* (perfect fluid with negative pressure effects). The generic equation of state $p = w\rho$ returns $w = \{1/3, 0, -1\}$ for radiation, matter, and vacuum, respectively. So the behavior of the density parameter goes as:

$$\rho_r \propto a^{-4} \qquad \rho_m \propto a^{-3} \qquad \rho_\Lambda \propto a^0$$

It is hence possible to derive the temporal evolution of the scale factor $a(t)$

$$a_r \propto t^{1/2} \qquad a_m \propto t^{2/3} \qquad a_\Lambda = \exp Ht.$$

Considering the main energy sources (radiation, matter, and vacuum), the Friedmann equation 1.3 can be expressed in the form:

$$H^2(t) = H_0^2\left[\frac{\Omega_r}{a^4(t)} + \frac{\Omega_m}{a^3(t)} + \Omega_\Lambda + \frac{\Omega_k}{a^2(t)}\right], \tag{1.9}$$

where Ω_r, Ω_m, and Ω_Λ are the dimensionless density parameter for radiation, matter, and dark energy, and Ω_k is the density associated with the curvature. These equations are the fundamental basis for modern cosmology.

The evolution history of the ΛCDM Universe can be divided into three main epochs according to the behavior of the scale factor, which is driven by the dominant component of the energy-momentum density: radiation, matter, and dark energy-dominated epochs.

The radiation epoch can be defined from the very first moments of the Universe, until nearly 47000 years after the Big Bang. During this epoch the main contribution to the energy density was in the form of radiation -photons and neutrinos- and so the scale factor evolved as $a(t) \propto t^{1/2}$. The successive -matter-dominated- epoch starts 47 kyr after the big bang and until 9.8 Gyr. During this epoch, the energy density of matter exceeded both the energy density of radiation and the vacuum energy density, so the scale factor evolves as $a(t) \propto t^{2/3}$. In this epoch, called the dark ages of the Universe, the decoupling of matter and radiation occurs, subsequently the reionization happens and starts galaxy formation. Finally, in the epoch of dark energy domination, the decelerated metric expansion of the Universe turns into an accelerated expansion due to the vacuum energy density. During this epoch, the scale factor goes as $a(t) = \exp Ht$. As long as this vacuum energy contribution is still not explained in terms of fundamental fields, it is called the dark energy era.

The ΛCDM model of cosmology is characterized by six independent parameters: the baryon density ($\Omega_b h^2$), the dark matter density ($\Omega_{DM} h^2$), the Hubble parameter (H_0), the scalar spectral index (n_s), the curvature fluctuation amplitude (A_s), and the reionization optical depth (τ). These parameters are constrained by a variety of observations, including the cosmic microwave background (Planck Collaboration et al. 2020), large-scale structure (Tegmark et al. 2004), and distant Type Ia supernovae (Perlmutter et al. 1999). The mass scale of neutrinos can also be included as a seventh parameter. Typical values for these parameters are: $\Omega_b h^2 \sim 0.022$, $\Omega_{DM} h^2 \sim 0.12$, $H_0 \sim 70$ km/s/Mpc, $n_s \sim 0.97$, $A_s \sim 2.1 \times 10^{-9}$, and $\tau \sim 0.07$.

1.1.2 Homogeneous or not homogeneous?

During the first 380000 years after the Big Bang, the temperature of the Universe was high enough to allow electrons and protons to exist as a highly ionized plasma. The plasma was opaque to electromagnetic radiation due to Thomson scattering by free electrons, which limited the distance a photon could travel before being scattered again. As the Universe cooled to approximately 3000 K, the plasma combined with neutral atoms, and photons were able to propagate freely. This process, known as recombination, made the Universe transparent. The photons that escaped during recombination constitute the Cosmic Microwave Background (CMB) radiation, which has cooled as a result of cosmic expansion. The CMB was first predicted in 1946 by George Gamow and was later discovered by Arno Penzias and Robert Wilson in 1965 (Penzias & Wilson 1965).

The most recent and precise measurement of the CMB has been made by the Planck mission (Planck Collaboration et al. 2020). The result follows the spectra of a black body at a temperature of $T_0 = 2.728 \pm 0.0004$K. The CMB radiation is observed reaching us with almost the same intensity in every direction, that is, nearly perfectly isotropic, suggesting that the Universe is isotropic on large angular scales. This thesis is focused on galaxies and the structures described by them, so in this context, the most remarkable fact about CMB is that it presents very small, but immensely important, temperature fluctuations with relative amplitude of the order of $\Delta T/T \sim 10^{-5}$, which will lead to the gravitational

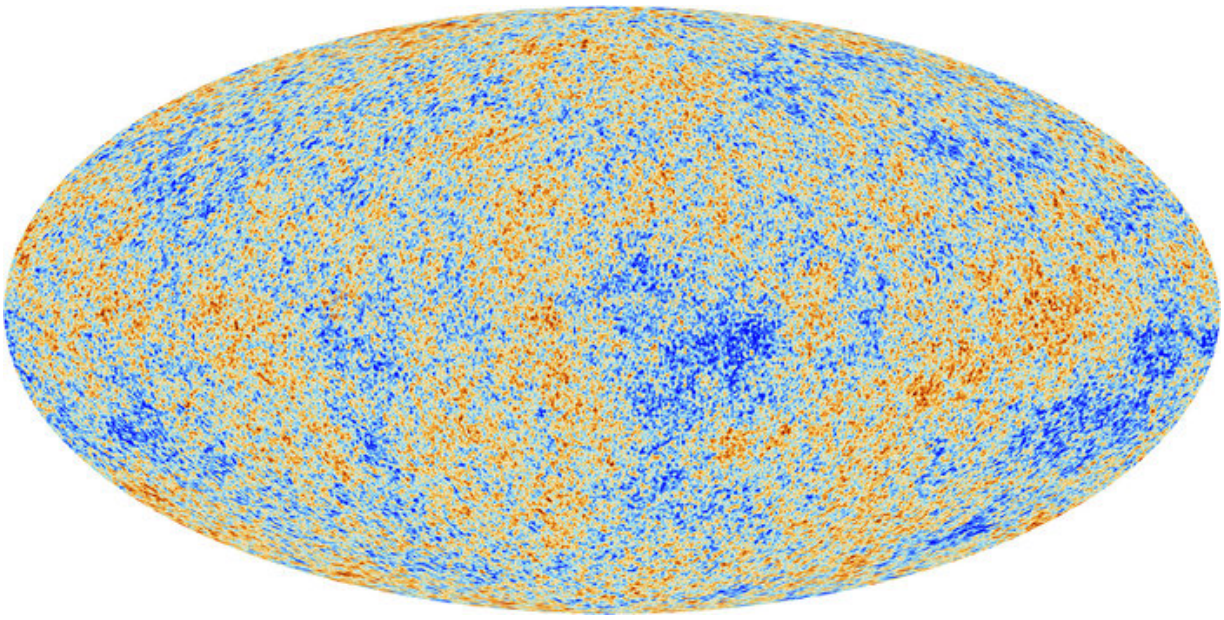


Figure 1.1: A map of the whole sky in galactic coordinates as observed by the Planck mission at millimeter wavelengths (Planck Collaboration et al. 2020). The emissions due to Galactic dust and synchrotron radiation have been subtracted from this map.

collapse of matter and consequent structure formation. Figure 1.1 shows the whole sky map of temperature obtained by Planck.

Observations show that the local Universe is not spatially homogeneous or isotropic, and instead exhibits a hierarchical aggregation of components, ranging from planetary systems around stars to galaxies in groups and clusters. Galaxy surveys have revealed a complex structure of voids, walls, filaments, and nodes in the evolved Universe, known as the Cosmic web (Bond et al. 1996). Despite this complexity, the majority of the total volume is in lower-density regions. An insight into the distribution of galaxies on a large scale is presented in Fig. 1.2. The reconciliation between the theoretical assumption of the cosmological principle and direct observations of structures is possible when considering large enough spatial scales. Indeed, averaging over scales of $\sim$100 Mpc, the structured Universe agrees with the cosmological principle.

The presence of structures in the evolved Universe is a highly interesting field of study. From a theoretical point of view, it is necessary to conciliate the effect of gravity, acting on the aggregation of matter, with the effect of the accelerated expansion of space, which tends to separate matter from itself. From the point of view of observational cosmology, there are two main observables: the initial state from the CMB, and the evolved state, which is given by the clustering of galaxies. It is expected that large-scale structures are an evolution of the low-amplitude anisotropies in the spectrum of the CMB, nevertheless, it is not trivial to connect those two states.

In this scenario, cosmological simulations have been a key instrument in advancing our understanding of the structure and galaxy formation in the Universe (Springel et al.

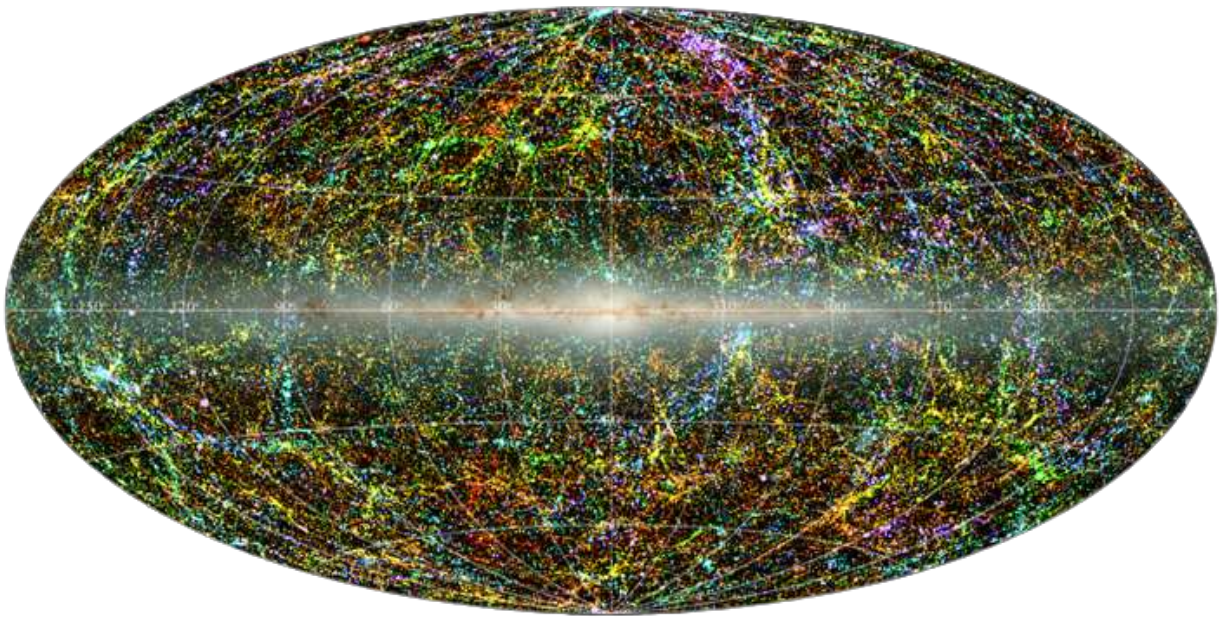


Figure 1.2: A map of the whole sky in galactic coordinates as observed by the 2MASS survey (Jarrett 2004)

2005; Vogelsberger et al. 2014). These simulations follow the nonlinear evolution of galaxies, modeling a variety of physical processes, such as the interplay between gravity, gas pressure, and radiation in an expanding medium, over an enormous range of time and length, scales (Keller et al. 2014; Dubois et al. 2016). A better understanding of the relevant physical processes arrives with improved numerical methods, and increased computing power (Springel 2010; Teyssier et al. 2010; Knollmann & Knebe 2009; Schaye et al. 2015; Kereš et al. 2005; Vogelsberger et al. 2014). Modern simulations model dark matter, dark energy, and ordinary matter in an expanding space-time starting from well-defined initial conditions. The modeling of ordinary matter is most challenging due to the large array of physical processes affecting this component (Springel et al. 2005). Cosmological simulations have also proven useful to study alternative cosmological models and their impact on the galaxy population (Vogelsberger et al. 2020). Figure 1.3, obtained from the TNG simulation as described in Springel & Hernquist (2003), shows the reproduction of a Universe's volume in three different moments of its evolution, with time going from left to right, putting in evidence the collapse of the matter, forming its large-scale structure.

1.1.3 Evolution of cosmological perturbations

The formation of structures is a key element in the evolution of the Universe because it connects the action of several puzzling elements, in particular the gravitational collapse of dark matter, the cycle of baryonic matter, and the expansion of the Universe, including its accelerating phase governed by dark energy. A full consideration of the problem requires a deep knowledge of General relativity, as well as quantum field theory, which exceeds

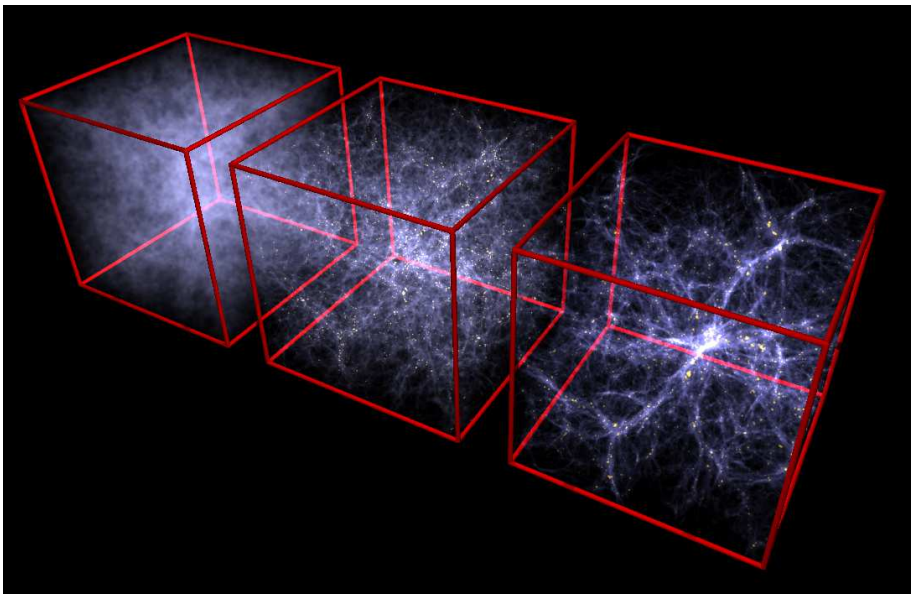


Figure 1.3: Snapshot from a cosmological simulation, described in Springel & Hernquist (2003), showing the behavior of matter on three moments: the initial state, *left*, is homogeneous and non-structured, while the final state, *right*, is an evolved state where the cosmic web, a large-scale structure, is visible.

the aim of this thesis. Nevertheless, I bring a more simple, still useful description, of the growth of perturbations in the standard Big Bang picture.

The velocity distribution of a fluid with density ρ and pressure p is given by $\mathbf{v}$. Starting from the equations of continuity, the equation of motion, and the equation for the gravitational potential, it is possible to analyze the behavior of perturbations, which corresponds to the small anisotropies of the CMB. Considering adiabatic perturbations, in particular, the density perturbation $\Delta = \delta\rho/\rho_0$, assuming an initial state homogeneous and isotropic so that $\nabla p_0 = 0$ and $\nabla\rho_0 = 0$, considering a background expanding uniformly and using comoving coordinates, it is possible to write the non-relativistic differential equation for gravitational instability, which is given by

$$\frac{d^2\Delta}{dt^2} + 2\left(\frac{\dot{a}}{a}\right)\frac{d\Delta}{dt} = \Delta(4\pi G\rho_0 - k^2c_s^2), \tag{1.10}$$

where c_2^2 is the adiabatic sound speed and $\mathbf{k}$ is the wavevector in comoving coordinates.

If we consider a static medium in Equation 1.10, with $\dot{a} = 0$, we can solve for waves of the form $\Delta = \Delta_0 \exp i(\mathbf{kr} - \omega t)$, which gives us the dispersion relation:

$$\omega^2 = c_s^2k^2 - 4\pi G\rho. \tag{1.11}$$

This equation opens two scenarios:

1. $c_s^2k^2 > 4\pi G\rho$, the right-hand side is positive and the perturbations are oscillatory, that is, they are sound waves in which the pressure gradient is sufficient to provide support for the region.

2. $c_s^2 k^2 < 4\pi G\rho$, the right-hand side of the dispersion relation is negative, corresponding to unstable models, which leads to the collapse of the field and the structure formation.

Eq 1.11 also defines the *Jeans' wavelength*, which is a characteristic length scale that determines the stability of density perturbations in a self-gravitating fluid. It is defined as:

$$\lambda_J = \frac{2\pi}{k_J} = c_s\big(\frac{\pi}{G\rho}\big)^{1/2}. \tag{1.12}$$

Jean's length determines the critical length scale below which density perturbations will be stable, resisting collapse under their self-gravity due to the internal pressure gradient (scenario 1 above). Perturbations with wavelengths larger than Jean's wavelength will be unstable and will tend to collapse, leading to the formation of structures in the Universe (scenario 2 above).

The evolution of inhomogeneities in the early Universe can provide insight into the nature of the initial conditions and the physical processes at play during the radiation- and matter-dominated epochs. To fully understand the behavior of these inhomogeneities, it is important to consider the scale of the perturbations and compare them with the particle horizon at a given epoch. The particle horizon is the maximum distance over which two points in the Universe can have been in causal contact since the Big Bang. Understanding how inhomogeneities evolve within the context of the particle horizon can provide insight into the initial conditions of the Universe and the physical processes at play during different epochs.

Super-horizon perturbations are density fluctuations that were larger than the horizon size at the time they were generated. These fluctuations are not influenced by cosmic expansion and are therefore preserved throughout the evolution of the Universe. In this regime, space curvature plays a special role in the development of density inhomogeneities. In the radiation-dominated era, the density contrast in Fourier space Δ_k is proportional to the square of the scale factor:

$$\Delta(k,t) \propto a^2(t), \tag{1.13}$$

while during the matter-dominated era, the contrast is proportional to the scale factor:

$$\Delta(k,t) \propto a(t). \tag{1.14}$$

Sub-horizon fluctuations were smaller than the horizon size at the time of their generation and are therefore influenced by the expansion of the Universe. In the regime of sub-horizon perturbations, the study of overdensities can be done through Eq. 1.10. The evolution of a density perturbation is influenced by the epoch in which it enters the horizon and reaches Jean's scale. The horizon scale λ_H increases with time, so larger scale fluctuations enter the horizon at later times. The moment in which radiation density equals matter density t_{eq} is the determinant for the evolution of density fluctuations. The

horizon at that event $\lambda_H(t_{eq})$ fixes a value for discriminating fluctuations. Fluctuations on scales smaller than $\lambda_H(t_{eq})$ enter the horizon during the radiation-dominated era, when the expansion is $H^2 \propto a^{-4}$, while fluctuations larger than $\lambda_H(t_{eq})$ enter during the matter-dominated era when $H^2 \propto a^{-3}$, as a consequence of eq. 1.9. When the Universe becomes dark-energy dominated, $H^2 \propto const$, and perturbations stop growing.

Galaxies are good tracers of the underlying dark matter density field because the formation and evolution of galaxies is tied to the distribution of dark matter in the Universe (White & Rees 1978; Davis et al. 1985; Mo et al. 1998; Blanton et al. 1999). Dark matter provides the gravitational scaffolding around which baryonic matter can collapse and form structures. Thus, the distribution of dark matter in the Universe is expected to be closely related to the distribution of galaxies (Peebles & Yu 1970; Peebles 1982; Frenk 1995).

1.2 Galaxies as tracers of Large-scale structures

The analysis of the previous section provides a simple description of the cosmological model and the theory of structure formation under the action of gravity. Now is the time to go one step forward by introducing observations of the evolved Universe. Those observations are done by measuring the position of galaxies across large areas of the sky.

From its definition, Dark Matter is not easy to observe, nevertheless its presence is a crucial point of the current cosmological model. Dark matter is supposed to overwhelm the baryonic matter, but, while the former is not directly observable, the baryonic matter is intrinsically observable and in addition, is a tracer for the dark matter distribution across the Universe at large scales. So, for simplicity purposes, it is convenient to think of galaxies as the building blocks of the Universe and tracers of its large-scale structure.

1.2.1 The two-point correlation function

The distribution of galaxies across large scales, the galaxy clustering, is determined by the content of matter of the Universe and its expansion history. Appropriate statistical methods are needed for describing the clustering properties on a wide range of scales. The simplest and more meaningful physical approach is the spatial two-point correlation function $\xi(r)$ which describes the number of galaxies in the volume element dV at distance r from any galaxy in the form:

$$dN(r) = N_0[1 + \xi(r)]dV, \tag{1.15}$$

Where N_0 is a suitable average background number density of galaxies. $\xi(r)$ describes the excess number of galaxies at distance r from any given galaxy. The function $\xi(r)$ can also be written in terms of the probability of finding pairs of galaxies separated by distance r:

$$dN_{pair} = N_0^2[1 + \xi(r)]dV_1dV_2 \tag{1.16}$$

The two-point correlation function can be directly related to the density contrast $\Delta = \delta\rho/\rho$, which was introduced in the previous section. Writing $\rho = \rho_0[1+\Delta(\mathbf{x})]$, where $\Delta(\mathbf{x})]$ is a generic function of space, the pairwise numbers of galaxies separated by distance $\mathbf{r}$ is

$$dN_{pair}(\mathbf{r}) = \rho_0^2[1 + \Delta(\mathbf{x})][1 + \Delta(\mathbf{x} + \mathbf{r})]dV_1dV_2. \tag{1.17}$$

Taking averages over a large number of volume elements, the mean value of Δ is zero by definition, and therefore the pairwise number of galaxies becomes

$$dN_{pair}(r) = \rho_0^2[1 + \langle\Delta(\mathbf{x})\Delta(\mathbf{x} + \mathbf{r})\rangle]dV_1dV_2. \tag{1.18}$$

This correlation shows explicitly the relation between the density contrast $\Delta(r)$ on different scales r and the two-point correlation function for galaxies:

$$\xi(r) = \langle\Delta(\mathbf{x})\Delta(\mathbf{x} + \mathbf{r})\rangle. \tag{1.19}$$

The description given by the two-point correlation function contains the assumption that in taking the averages, the distribution of galaxies is spherically symmetric around any randomly selected galaxy, which is nearly acceptable on large physical scales. If $\xi(r) = 0$ there is no clustering of galaxies. The excess or defect of clustering will be noticed by $\xi(r) > 0$ or $\xi(r) < 0$, respectively, with the presence of clusters or voids.

1.2.2 The power spectrum

The power spectrum is a statistical measure of the distribution of matter in the Universe on different scales. It is typically calculated by Fourier transforming the three-dimensional matter density field, and measuring the amplitude of the resulting Fourier modes as a function of their wavenumber. The convention for the Fourier transform pair for $\Delta(\mathbf{r})$ is:

$$\Delta(\mathbf{r}) = \frac{V}{(2\pi)^3} \int \Delta_k e^{-i\mathbf{k}\cdot\mathbf{r}} d^3k; \tag{1.20}$$

$$\Delta_\mathbf{k} = \frac{1}{V} \int \Delta(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} d^3x, \tag{1.21}$$

where V is the volume within which $\Delta(\mathbf{r})$ is defined. By using Parseval's theorem it is possible to relate equations 1.20 and 1.21. Using a spherical symmetry, averaging, and using the Fourier transform, the equation 1.19 for the two-point correlation function $\xi(r)$ becomes:

$$\xi(r) = \frac{V}{(2\pi)^3} \int |\Delta_k|^2 e^{i\mathbf{k}\cdot\mathbf{r}} d^3k. \tag{1.22}$$

The power spectrum is defined as

$$P(k) = |\Delta_k|^2. \tag{1.23}$$

Evolution of the Power Spectrum

The power spectrum of the cosmic microwave background (CMB) is a measure of the fluctuation in the temperature of the CMB on different scales. These temperature fluctuations are thought to be the result of primordial density fluctuations that were present in the early Universe and were amplified by the process of inflation (Guth 1981; Linde 1982). The primordial power spectrum can be described as a power law with two parameters: the amplitude A_s and the index n_s. This can be expressed as:

$$P_{prim}(k) \propto A_s k^{n_s}. \tag{1.24}$$

If we assume that the initial fluctuations of the gravitational potential are scale-invariant, the index n_s is equal to one, resulting in the Harrison-Zel'dovich power spectrum:

$$P_{prim}(k) \propto A_s k. \tag{1.25}$$

The power spectrum at the present epoch is composed of two regimes separated by a peak at the frequency $k = k_{eq}$, which is the mode associated with the horizon at equality

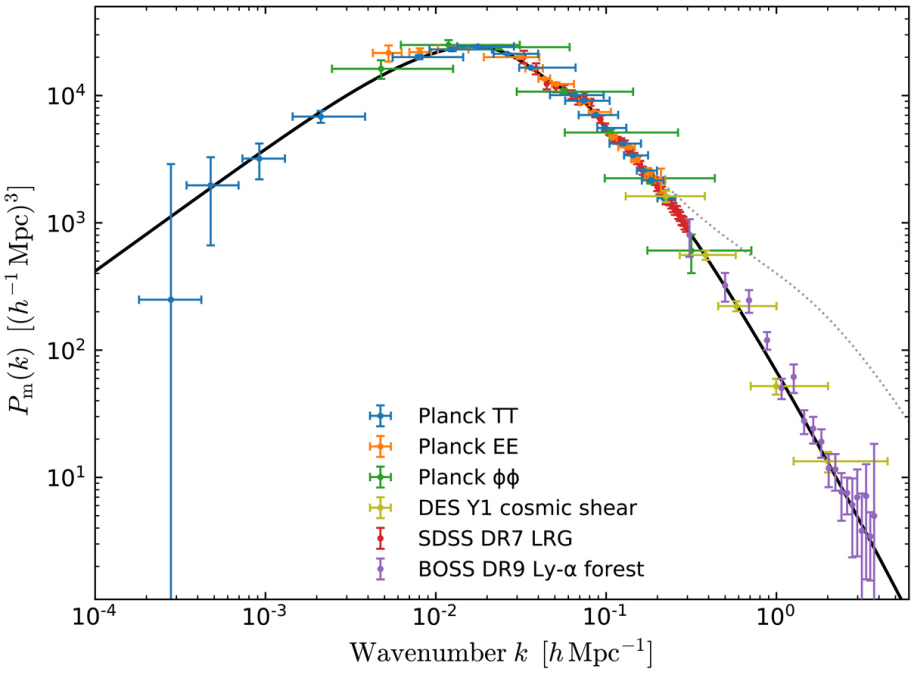


Figure 1.4: Linear theory matter power spectrum at present time, inferred from different cosmological probes. The dotted line shows the impact of non-linear clustering at the present epoch (Planck Collaboration et al. 2020).

$k_{eq} = 2\pi/\lambda_H(t_{eq})$. The shape of the power spectrum in this regime, $k < k_{eq}$, is the same as the Harrison-Zel'Dovich spectrum, $P(k) \propto k^{n_s}$. The power spectrum for $k > k_{eq}$ can be described as $P(k) \propto k^{n_s-4}$.

To describe the evolution of the matter power spectrum, it is possible to define a transfer function $T(k)$ which connects the primordial spectrum with the observed one:

$$P(k,z) = T^2(k)D^2(z)P_{prim}(k), \tag{1.26}$$

where $P(k,z)$ is the matter power spectrum measured at a given epoch (parametrized by z), $P_{prim}(k)$ is the primordial power spectrum, and $D(z)$ is the growth factor that describes the growth of fluctuations at different modes k. The transfer function has an asymptotic behavior in the ΛCDM model:

$$T(k) = \begin{cases} 1 & k \ll k_{eq} \\ k^{-2} & k \gg k_{eq} \end{cases} \tag{1.27}$$

One of the most complete measurements of the matter power spectrum at the present epoch, inferred from different cosmological probes, is presented in Fig. 1.4. The broad agreement of the model (black line) with such a disparate compilation of data, spanning 14 Gyr in time and three decades in scale, is an impressive testament to the explanatory power of ΛCDM.

1.2.3 Cosmological parameters measured from the Power spectrum

Galaxy Bias

So far, it has been assumed that galaxies trace the distribution of dark matter. The more general case will consider the *galaxy bias* as the offset between the galaxy density field and the dark matter density field. These densest haloes of dark matter enhance galaxy formation.

Different mechanisms can describe the biasing or anti-biasing effects on the local Universe. One possible mechanism that explains this bias is by considering that inherent to the description of the power spectrum of perturbations is the assumption that fluctuations are Gaussian with variance $\langle\Delta^2\rangle$. Thus, the probability of encountering a density contrast Δ at some point in space is proportional to $exp(-\Delta^2/\langle\Delta^2\rangle)$. So it is possible to argue that galaxies are more likely to form in the highest peaks of the density distribution. Thus, if structures are only formed if the density contrast exceeds a certain value Δ_{crit}, galaxy formation would be biased towards the highest-density perturbations over the mean background density. This could account for the fact that clusters of galaxies are much more strongly clustered than galaxies in general.

The *galaxy bias* definition that I follow is given by:

$$P_{gal}(k) = b^2 P_{DM}(k), \tag{1.28}$$

where $P_{gal}(k)$ is the power spectrum of the baryonic luminous matter density field and $P_{DM}(k)$ is the power spectrum of the dark matter density field.

So the galaxy power spectrum, measured directly from the position of galaxies, is a biased measure of the underlying dark matter power spectrum, predicted by the theory. In this work the galaxy bias is considered independent of scale and galaxy type; as a consequence, one of the most natural improvements of this thesis would be the study of different models of galaxy bias according to the scale or luminosity.

Redshift Space Distortions

The redshift is directly related to the distances of the galaxies according to Hubble's law. This is done by measuring the relative velocity between the galaxies and the Earth, which is by definition along the line of sight. But galaxies in general have peculiar velocities in all spatial directions, so measuring only radial velocities implies a distortion of the structure.

The Redshift Space Distortions (RSD) are at the origin of two effects, visible on different scales. The first effect is visible at small scales, the 'Fingers-of-God' (FoG): on redshift surveys, the spatial distribution of galaxies appears squashed and distorted when their positions are plotted as a function of their redshift rather than as a function of their distance. The second effect is associated with the fact that large-scale density perturbations induce potential motions. The mass flows from low-density regions onto high-density regions, and as a result, galaxies are expected to be observed 'falling into' large-scale density perturbations, and so the projected velocity component along the line of sight differs from that associated with its cosmological redshift Kaiser (1987).

The RSD can change the shape and amplitude of the power spectrum but is important to be careful of the scales on which the effects are present: on large scales, coherent velocities streaming from under densities to overdensities introduce a quadrupole anisotropy in the measured power spectrum. This in itself does not change the shape of the spherically-averaged power spectrum, rather, the amplitude is boosted by a scale-independent factor. On small scales, the FoG produces a damping of modes through the radial smearing due to virialized motions. This is hard to treat exactly because of the small-scale velocity correlations. A way to treat this problem is to consider the small-scale velocity field as an incoherent Gaussian scatter with 1D RMS dispersion σ. This approximation is reasonable because the observed pairwise velocity dispersion is a very slow function of separation.

The corrections due to RSD to the power spectrum can be represented by a model like:

$$P_{mod}(k) = P_{DM}(k)[b + f\mu^2]^2 e^{-k^2\mu^2\sigma_v^2}, \tag{1.29}$$

where $P_{mod}(k)$ is the redshift-space power spectrum; $P_{DM}(k)$ is the real-space power spectrum -the power spectrum due to the underlying dark matter density field-; b the galaxy bias, f the logarithmic growth rate of structure (see below); μ is the cosine of the angle between the wavevector k and the line-of-sight; σ_v is the velocity dispersion of galaxies due to the fingers-of-God effect and is in units of $h^{-1}Mpc$; the conversion to velocity units is $H(z)/(1+z) \approx 84.85 hMpc^{-1}\, kms^{-1}$, which is nearly constant over the range of interest. This expression is obtained by averaging the anisotropic effects by integrating over a uniform distribution of μ. The term $f\mu^2$ *-kaiser factor-* is related to the large-scale mass flows towards high-density regions, while the exponential term is due to the fingers-of-God effects on small-scale velocities. The RSD effects at different scales are anisotropic in k space, so they interfere before averaging; for this reason, the power spectrum described in equation 1.29 is not averaged on μ but calculated directly on each point of the k grid.

Growth rate of structures

In the analysis of the RSD, the logarithmic growth rate of the structure appears because the Universe expansion affects the process of structure formation. This factor can be defined as:

$$f(z) = \frac{d\ln\Delta}{d\ln a}, \tag{1.30}$$

where Δ is the amplitude of a growing mode and a is the cosmic scale factor (Lahav et al. 1991). The growth rate determines the amplitude of peculiar velocity flows and RSDs as a consequence of equations 1.10 and 1.29. The logarithmic growth rate of the structure can be expressed as:

$$f \approx \Omega_M^{\gamma}(a) \tag{1.31}$$

where $\Omega_M^{\gamma}(a)$ is the matter density at a given epoch and $\gamma = 0.55$ for standard gravity (Linder 2005). The growth rate, parametrized as $f\sigma_8$, is a measure of the amount by which perturbations in the density of the universe grow over time. It is defined as the variance in the counts of galaxies within a sphere of radius $8, h^{-1}Mpc$ (a scale chosen because the two-point correlation function for galaxies has roughly unit amplitude, $\xi \sim 1$,

at the present epoch) multiplied by a factor f that describes the rate of growth (Peebles 1980; Linder 2005; Wang & Steinhardt 1998).

1.2.4 Galaxy redshift surveys

In the previous sections, I introduced the theoretical framework for the formation of large-scale structures and how those structures, through the power spectrum, can be related to parameters of the cosmological model. Those predictions, nevertheless, require observational validation. In this context, galaxy redshift surveys provide a means of mapping the distribution of galaxies in the Universe and, by measuring the redshift of each galaxy, allow for the determination of the expansion history of the Universe.

The observed distribution of galaxies can be used to compute the galaxy power spectrum, which measures the clustering of galaxies on large scales, and is the main tool to compare the observations with theoretical predictions. With the observed power spectrum is possible to evaluate the accuracy of the model and constrain the values of key cosmological parameters such as the matter density, the dark energy density, and the amplitude of matter fluctuations (Percival et al. 2001; Springel et al. 2005; Laureijs et al. 2011; Alcock & Paczynski 1979; Efstathiou et al. 1990; Peacock et al. 2001). In this way, galaxy redshift surveys play a crucial role in our understanding of the structure and evolution of the Universe. In addition, redshift surveys can be used to study the evolution of galaxies and their relationship to the cosmic environment (e.g. (Dressler 1980; Balogh et al. 2004; Peng et al. 2010; Wetzel et al. 2012; Davidzon et al. 2016)).

Redshift surveys involve spectroscopic observations of selected galaxies, identified through photometry of a particular area of the sky. These observations are used to measure the redshifts of the galaxies by analyzing the wavelengths of their prominent spectral lines and comparing them to laboratory wavelengths (Peebles 1982). The selection of galaxies for the survey may be based on their brightness (e.g. Loveday et al. 1992) and, in some cases, their color (Baldry et al. 2004). The goal of the survey is to gain insight into the properties and distribution of the galaxy population (Fukugita et al. 1998), and to test theoretical models of the large-scale structure of the Universe (Peebles 1982; Springel et al. 2005).

The CfA Redshift Survey was the first systematic redshift survey, containing around 2400 galaxies (Huchra et al. 1983). The 2dFGRS survey (Colless et al. 2003) was notable for its use of the 2-degree Field instrument on the Anglo-Australian Telescope, which allowed for a wide-field survey covering a large portion of the sky. The DES survey (Abbott et al. 2018) utilized the Dark Energy Camera to survey a large area of the sky in multiple filters, allowing for the study of dark energy and the expansion of the Universe. The GAMA survey (Driver et al. 2011) focused on the study of galaxy evolution and the role of the environment on galaxy properties. The SDSS survey (York et al. 2000) used a dedicated 2.5-meter telescope to survey a large portion of the sky in multiple filters, leading to numerous discoveries in the fields of galaxy evolution and large-scale structure. The 2MASS redshift survey (Huchra et al. 2012) used near-infrared observations to study the properties and distribution of galaxies in the local Universe. The upcoming generation of instruments and telescopes, such as Vera C. Rubin (Ivezić et al. 2019), Euclid (Laureijs

et al. 2011), Roman Space Telescope (Spergel et al. 2015), and DESI (Levi et al. 2013), are expected to further advance our understanding of the distribution of galaxies and the underlying dark matter distribution.

Galaxy redshift surveys are sensitive to the velocity field projected along the line of sight due to the effect of redshift-space distortions (Kaiser 1987). The anisotropic signal encodes information about the logarithmic growth rate of structure and provides a key test of the gravity model on cosmological scales (Guzzo et al. 2008; Percival & White 2009).

The galaxy field from a redshift survey, over large scales, can be assumed to be a Gaussian random field. The power spectrum estimate for such a field has a formal error due to sample variance of $\sigma_P^2 = 2/N_{\rm modes}P(k)$ (Tegmark 1997). However, even in the ideal case of a Gaussian field, this limit cannot be achieved due to the loss of information inherent in the survey selection process, namely the sampling process and the window due to the finite survey volume. For a Poisson noise model, the Feldman, Kaiser Peacock (FKP) (Feldman et al. 1994) estimator minimizes the loss of information due to discrete sampling. The estimator can be further optimized by incorporating galaxy weights that account for the relationship between the amplitude of the power spectrum and galaxy properties (Montero-Dorta et al. 2020; Pearson et al. 2016; Cai et al. 2011; Percival et al. 2004). Fig. 1.5 presents the distribution of several redshift surveys, and simulations, showing the cosmic web of galaxies in the local Universe.

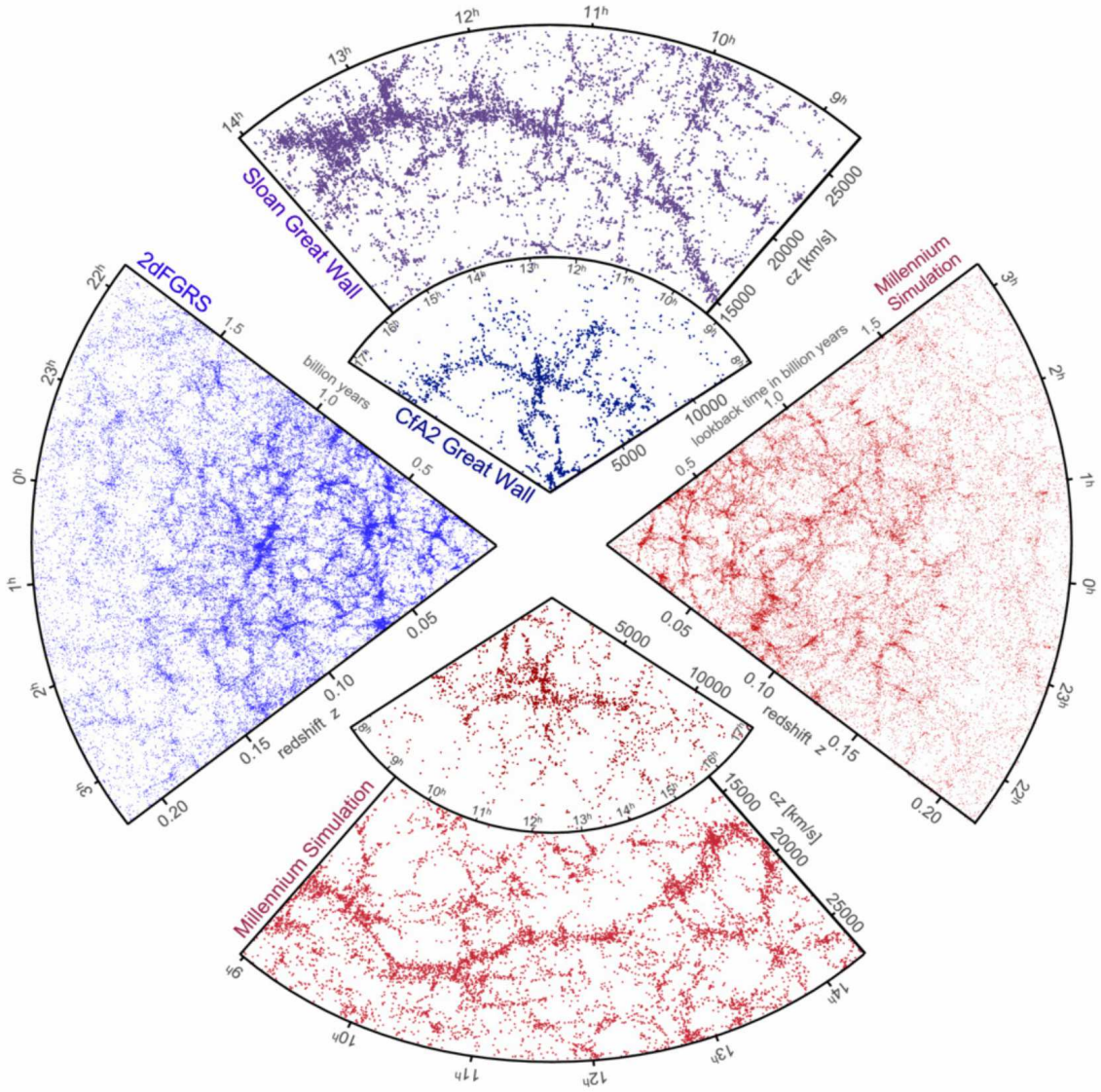


Figure 1.5: The figure shows the galaxy distribution obtained from various spectroscopic redshift surveys and mock catalogs constructed from cosmological simulations. The top slice shows the CfA2 'Great Wall', with the Coma cluster at the center. A small section of the SDSS is also shown, containing the larger 'Sloan Great Wall', which is one of the largest observed structures in the Universe. The cone on the left is one-half of the 2dFGRS. The bottom and right show mock galaxy surveys constructed using semi-analytic techniques to simulate the formation and evolution of galaxies within the evolving dark matter distribution of the 'Millennium' simulation, selected with matching survey geometries and magnitude limits. Image credits: Springel et al. (2006).

1.3 Galaxies into the Large-scale structure

The previous sections were focused on the cosmological scenario. This section connects large-scale structures with the galaxy formation process and the properties of galaxies inhabiting those structures.

The classical model of galaxy formation is the so-called "*monolithic collapse*" which proposes that galaxies formed from the monolithic gravitational collapse of a single, massive over-dense region of primordial gas very early in the lifetime of the Universe, without gaining significant mass thereafter (Eggen et al. 1962). However, modern evidence rules out this model. The current picture of galaxy formation is the so-called "*hierarchical*" model, which states that haloes assemble hierarchically, such that smaller haloes merge to form larger and more massive haloes in dense environments (Mo & White 1996; Sheth & Tormen 2002).

According to the hierarchical model, the early universe was filled with small, dense clouds of gas and dark matter that collapsed and formed the first generation of galaxies. These galaxies then grew and evolved through the merger and accretion of smaller structures, including dark matter haloes, dwarf galaxies, and gas clouds. The hierarchical model predicts that larger structures, such as galaxy clusters, form through the merger and accretion of smaller structures, including individual galaxies and galaxy groups. This model is supported by several pieces of evidence, such as the resemblance between the observations and the simulated distribution of matter in the Universe, observations of galaxy mergers, and accretion. Figure 1.6 presents a detailed description of the hierarchical model.

1.3.1 Galaxy components

To better understand the process of galaxy formation and evolution, it is mandatory to introduce galaxies as extended objects with complex morphology and to do a phenomenological description of galaxies, starting from their components in terms of matter and their more prominent properties.

The fundamental ingredient of a galaxy is its dark matter halo, which determines its total mass, velocity, and gravitational potential. Nevertheless, a magnificent structure of baryonic matter is developed inside the halo, including plenty of stars, gas, and dust, embedded into magnetic fields and an ocean of particles. The brief description of the main baryonic components of a galaxy is done according to Renzini (2006); Longhetti & Saracco (2009); Boselli & Gavazzi (2006) and references therein.

Dark Matter

Dark matter is a postulated form of matter that is believed to exist based on its gravitational effects on visible matter, radiation, and the structure of the universe. It does not interact with electromagnetic radiation in the same way as baryonic matter, making it challenging to be directly detected.

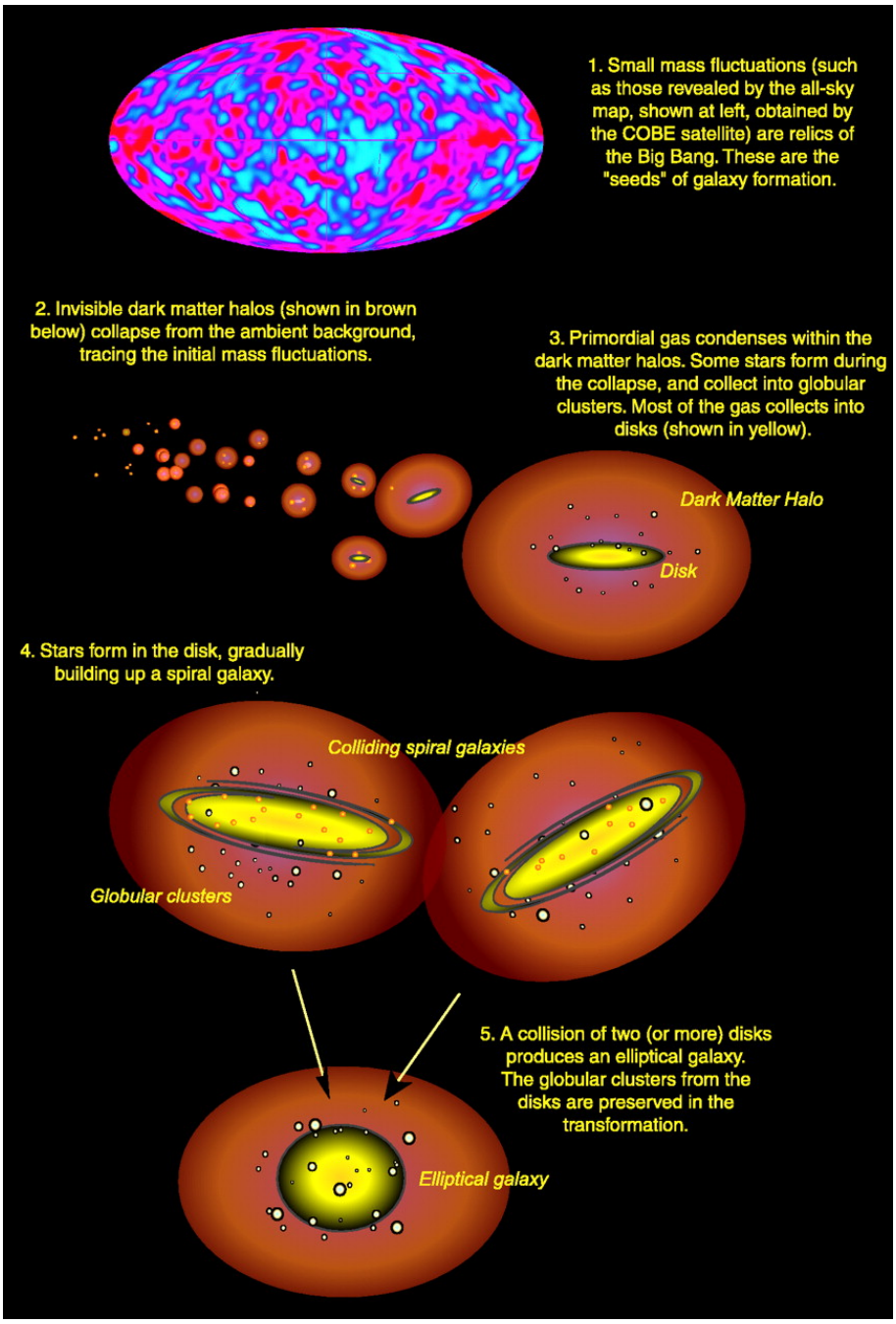


Figure 1.6: Galaxy formation process according to the hierarchical picture (Abraham & van den Bergh 2001)

Within a galaxy, dark matter is thought to be a vital component of the overall structure and dynamics. It is posited to comprise the majority of the mass in a galaxy and to exert a gravitational influence that helps to hold the galaxy together and shape its structure. Without dark matter, it is hypothesized that the baryonic matter in a galaxy would lack the mass necessary to generate the gravitational forces required to sustain its structure, and the galaxy would potentially disperse.

The precise role of dark matter within a galaxy remains poorly understood and this is an active area of research in astrophysics. Many theories have been proposed to explain its properties and behavior, and ongoing efforts are being made to gain a greater understanding of this enigmatic form of matter.

Stars

In a galaxy, stars are found throughout the disk, halo, and bulge components. The disk is characterized by a flattened, rotating structure, where most of the ongoing star formation takes place. The stars in the disk, for a spiral galaxy, follow nearly circular orbits around the galaxy's center. The halo is a roughly spherical, diffuse region that surrounds the disk, is dominated by old, metal-poor stars, and it often contains little gas or dust. The bulge component of a galaxy is a dense, roughly spherical structure located at the center of the galaxy. The bulge is formed through a combination of early star formation and mergers of smaller galaxies, so the stars in the bulge can have a variety of orbits. Stars can be classified based on their mass, age, and chemical composition. Thus, the properties of a galaxy, including its structure and evolution, are usually determined by the properties and distribution of its constituent stars.

Stars form from the collapse and contraction of dense clouds of gas and dust inside the galaxy. The process of star formation is complex and involves a variety of physical processes, including radiative cooling, gravitational collapse, and angular momentum transport. Stars play a key role in the evolution of galaxies through their production of heavy elements, their feedback, and their influence on the gravitational potential of the galaxy. The total mass of stars in a galaxy, known as the stellar mass, is an important parameter that can be used to study the galaxy's evolution and compare it to other galaxies.

Interstellar Medium

The interstellar medium (ISM) includes the gas and dust that fill the space between the stars in a galaxy. It is a complex and dynamic medium that plays a central role in the formation and evolution of galaxies and is composed of a mix of different phases, including atomic, molecular, and ionized gas, mainly hydrogen but also helium and metals, as well as dust particles. The properties of the ISM can vary significantly from one region to another within a galaxy.

The process of star formation occurs when dense clouds of gas and dust in the ISM collapse and contract under the influence of gravity. The ISM can be affected by the

feedback from stars, such as supernovae explosions, that inject energy and matter into the ISM. In more detail, the ISM is composed mainly of:

- **Atomic Hydrogen**: Atomic Hydrogen, H I, is the principal component of the interstellar medium (ISM) in early-type galaxies (see Morphology below). The H I gas distribution extends beyond the optical disk up to ~ 2 times the optical diameter, with a relatively flat radial distribution that sometimes shows a central dip. The distribution of the H I within a galaxy is influenced by various factors such as its mass, morphology, kinematics, and dark matter halo properties. It is thus reasonable to expect that the H I content and distribution in galaxies within clusters would differ from that in isolated galaxies. On average, cluster galaxies have a lower atomic gas content than their isolated counterparts.

- **Ionised Hydrogen**: Ionized hydrogen, HII, is formed when the temperature and radiation field in a region of the interstellar medium are sufficient to ionize the hydrogen atoms, breaking the bonds between the electrons and the protons and creating free electrons and ions. In the ISM, ionized hydrogen is found in regions of high temperatures, such as around hot, young stars or in the shock fronts of supernova remnants. Ionized hydrogen is typically associated with regions of ongoing or recent star formation, and it is a key tracer of the star formation activity in a galaxy. It can be detected through the Hα emission line.

- **Molecular Hydrogen**: Molecular hydrogen, H_2, is typically found in regions of the ISM with high density and low temperatures, such as in molecular clouds, which are dense clouds of gas and dust. Molecular hydrogen represents only the $\sim 15\%$ of the total gas reservoir in normal, late-type galaxies, and it is the ISM component that takes a direct part in the process of star formation. Molecular hydrogen is formed through the collision and combination of atomic hydrogen atoms, and it is an important coolant in the ISM, as it can radiate away energy through its rotational and vibrational transitions. The ratio of molecular hydrogen to atomic hydrogen can also be used to study the star formation activity in a galaxy. The H I gas has to condense inside molecular clouds before collapsing to form stars. It is thus expected that any possible external perturbation induced by the environment on the molecular gas can have important consequences on the star formation activity, and thus on the evolution of late-type galaxies. Cluster galaxies have an important molecular gas content, and the interpretation is that molecular gas, being more centrally peaked deep inside the potential well of the galaxy than atomic gas, cannot be easily removed by any stripping mechanism. Molecular gas removal is more effective in low-mass galaxies with shallow potential wells.

- **Helium**: Helium (He) is the second most abundant element in the universe, after hydrogen. Helium is a chemical element that belongs to the noble gases, a group of elements known for their low reactivity and tendency to remain in their elemental form. This property of helium makes it difficult for it to form compounds with other elements. However, observing helium in the ISM can be challenging due to

several factors, including its low abundance relative to hydrogen, the presence of other elements and molecules that can emit radiation at similar wavelengths, and its high stability, which makes it more difficult to dissociate an electron and generate an emission line.

- **Metal Content**: The metal content refers to the abundance of elements heavier than helium in the galaxy. These elements, such as carbon, oxygen, and iron, are commonly referred to as "metals" in astrophysics. The metal content is typically determined by the presence of absorption or emission lines of specific metal ions in the spectra of the galaxy. The strength of these lines is related to the abundance of the corresponding element, and the total metal content of the galaxy can be calculated by adding up the abundance of the different elements. The metal content can be affected by the star formation history of the galaxy, the initial mass function (IMF) of the stars, and the enrichment of the ISM through supernova explosions and other processes. Galaxies with a higher metal content tend to have a long history of star formation, as heavier elements are produced in the cores of stars and are released into the ISM through supernova explosions.

- **Dust**: Dust grains, small solid particles, made up mainly of silicates and carbonaceous materials, are present in the ISM. Dust is typically associated with regions of the ISM that have high densities and low temperatures, such as molecular clouds. Dust absorbs, scatters, and emits radiation at various wavelengths, mainly at infrared wavelengths, making it a valuable tracer of physical processes, like star formation. In fact, dust plays an important role in the process of star formation since hydrogen atoms condense on dust grains to form H_2 molecules. While most of the flux emitted by dust in normal galaxies is in the 60–200 micron domain (relatively warm $[T_{dust} > 30K]$, big grains), the bulk of the dust mass is colder ($[T_{dust} < 15K)$, and it radiates in the submillimeter domain at 1100 mm. At mid-IR wavelengths (15–60 microns), the emission is dominated by very small grains, while in the near-IR (5–15 microns), it mostly comes from unidentified infrared bands probably associated with planar molecules, called PAHs (polycyclic aromatic hydrocarbons; Desert et al. 1990). Dust is generally associated with the gaseous component: the dust-to-gas column density ratio is relatively well correlated to the metallicity of the ISM, and thus it varies with galaxy morphology and/or luminosity or, within a galaxy, follows the metallicity gradient.

Finally, the various components of a galaxy can be observed at different wavelengths. The stars can be observed at optical and near-infrared wavelengths. The cold molecular gas (H_2) can be detected at millimeter wavelengths. Dust can be observed at far-infrared wavelengths, while atomic hydrogen (HI) can be detected at radio wavelengths. These different wavelengths correspond to different energy levels and allow the study of the physical properties and distribution of the various components in a galaxy. A schematic view of a galaxy through different wavelengths is shown in Fig. 1.7.

Figure 1.7: Description of the different components of a galaxy as a function of the observed wavelength. Image from the multiwavelength astronomy program. [2]

1.3.2 Galaxy physical properties

Physical properties are intrinsic to each galaxy, but the values of these properties can depend on the tracer used to measure them, because the different observable components of a galaxy, i.e. stars, dust, and gas, have different properties that can affect the measured values of the physical properties.

Empirical relations between physical properties are often used to study the correlation between these properties and gain insight into the formation and evolution of galaxies.

Morphology

The first classification of galaxies was made by Hubble, considering their morphological shape (Hubble 1926). This classification divides galaxies into four classes and can be summarized in a diagram, shown in Fig. 1.8. Going from the left to the right, the morphological classes introduced are:

- **Ellipticals**: they appear as smooth and structureless. They vary in shape from round to fairly highly elongated in form. According to their ellipticity, they are divided into subtypes: E_1, E_2, E_3, ..., E_6, where the integer is the closest to $10(1 - b/a)$, where b/a is the apparent axial ratio.

- **Spirals**: they consist of a thin disk with spiral arm structures and a central bulge. They are divided into two branches, *barred spirals* and *normal spirals*, depending on

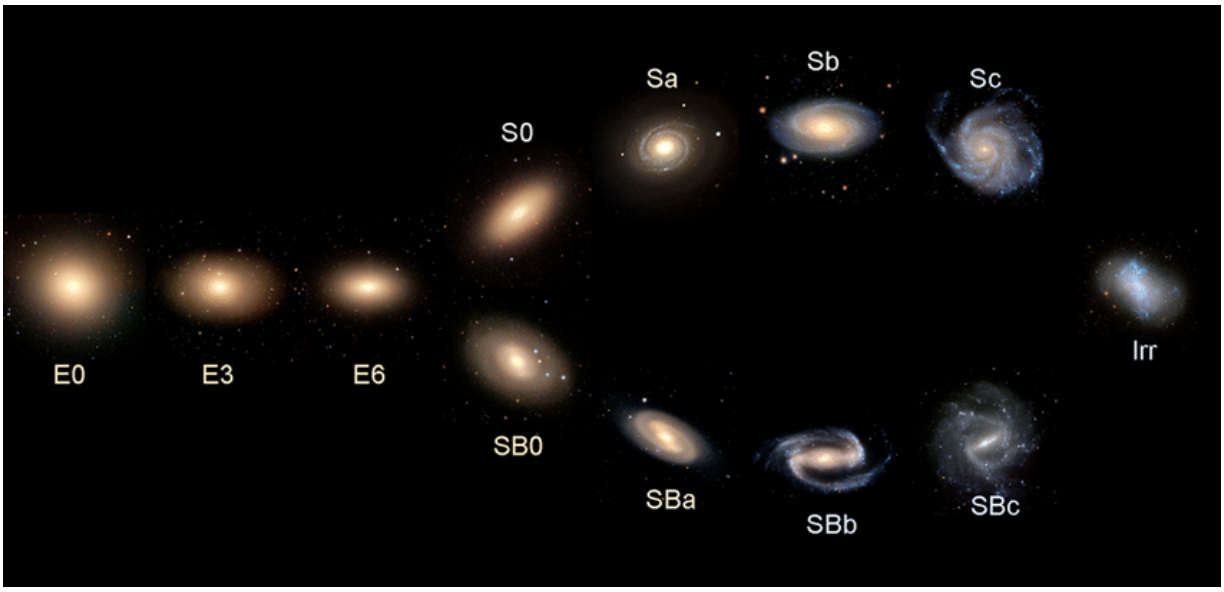


Figure 1.8: Schematic representation of the Hubble classification of galaxies (Cui et al. 2014). Galaxies are grouped as ellipticals (E), spirals (S and SB if a bar is present), irregulars (Irr), and lenticulars (S0s).

whether or not they show a bar. In each of these subclasses, a sequence is defined that is ordered according to the brightness ratio of the bulge and disk (a, b, c).

- **Lenticulars** or *S0s*: they were thought to be in transition between ellipticals and spirals. Like the ellipticals, S0s have a smooth light distribution with no spiral arms. Like the spirals, they show a thin disk and a bulge. Depending on whether or not they show a bar, they are subdivided into $S0$ and $SB0$.

- **Irregulars**: they have neither a dominating bulge nor a rotationally symmetric disk and lack any symmetry.

Elliptical galaxies tend to be redder in color than spiral galaxies, which are bluer in color. This is because elliptical galaxies have a larger population of old, cool stars, and spiral galaxies have a larger population of young, hot stars. Ellipticals and S0s together are called early-type galaxies, while spirals and irregulars are late-type galaxies.

Color

The color of a galaxy is a measure of the overall spectral energy distribution of the light emitted by the galaxy and is computed as the difference between the flux at different wavelengths. The color of a galaxy can be affected by a variety of factors, mainly the types of stars present, the amount of dust and gas in the galaxy, and the age of the galaxy.

Different types of stars emit light at different wavelengths, and the overall mix of star types can affect the color of a galaxy. For example, galaxies with a large number of hot,

young stars tend to have a bluish color, while galaxies with a larger fraction of cool, red stars tend to appear redder.

Dust and gas in a galaxy can also absorb and scatter light, which can affect the observed color of the galaxy. Galaxies with a lot of dust and gas tend to appear redder, while galaxies with less dust and gas tend to appear bluer.

The age of a galaxy can also play a role in its color. Younger galaxies tend to have a higher proportion of hot, blue stars, while older galaxies tend to have more cool, red stars. This can result in younger galaxies appearing bluer and older galaxies appearing redder.

Redder galaxies tend to have higher metallicities, as they contain more heavy elements that are produced by star formation and enrichment processes. Bluer galaxies tend to have higher star formation rates, as they contain more young, hot stars and more gas and dust, which are the raw materials needed for star formation.

Luminosity

Luminosity is a measure of the total amount of energy emitted by a galaxy in a given time period. Galaxies span a wide range of luminosity, from $\sim 10^3$ to $\sim 10^{12} L\odot$. There are several factors that are correlated with luminosity in a galaxy, including the mass and size of the galaxy, the number and properties of the stars it contains, the presence of active galactic nuclei (AGN), and the content of gas and dust within the galaxy. Gas and dust can absorb and re-radiate light, which can contribute to the overall luminosity of the galaxy, diminishing luminosity in the ultraviolet and optical wavelengths and increasing it in the infrared domain.

More massive galaxies tend to be more luminous, as they contain more stars. The mass, age, and chemical composition of the stars can also affect the luminosity of the galaxy, e.g. a galaxy with a large population of young, hot stars is likely to be more luminous than a galaxy with a large population of old, cool stars. Additionally, AGN can contribute significantly to the overall luminosity of the galaxy.

The brightness is a measure of the amount of energy received by an observer from an object per unit of time. The relation between luminosity and brightness depends on the distance between the galaxy and the earth and is also sensible to the observed wavelength. A galaxy with a given brightness can be either a low-mass near galaxy or a high-mass far galaxy.

Stellar masses

The stellar mass accounts for the mass of all the stars contained within a galaxy. The age of a galaxy can also affect its stellar mass, as older galaxies tend to have fewer young, massive stars and more low-mass, red dwarfs. The stellar mass can be derived from empirical relationships like the mass-to-light ratio.

The mass-to-light ratio (M/L) measures the amount of mass in a galaxy relative to the amount of light it emits. It is typically expressed in units of solar masses per solar luminosity and quantifies the relative contributions of different types of matter to the mass budget of the galaxy, as well as the distribution of the matter within the galaxy.

The mass of a galaxy can be estimated from the gravitational forces acting on the galaxy, e.g. through the rotation curves. The light emitted by the galaxy can be measured from observations of the galaxy's spectra or photometry.

The M/L of a galaxy can vary significantly depending on the type and distribution of the matter within the galaxy, e.g. galaxies containing a large amount of dark matter, which is non-luminous and therefore does not contribute to the overall brightness of the galaxy, will have a higher M/L ratio than galaxies with less dark matter. Similarly, the distribution of stars within a galaxy can also affect the M/L ratio, as galaxies with more concentrated or centrally located stars will be more luminous than galaxies with more diffuse or evenly distributed stars. In general, there is a correlation between the M/L and morphological type: late-type galaxies (i.e. spirals, irregulars) seem to have lower M/L, than early types (i.e. ellipticals and S0s).

Size

The size of a galaxy is a measure of its overall dimensions, and it can be difficult to define accurately because galaxies do not have a well-defined edge.

To measure the size of a galaxy, one way is to determine its diameter, which is the distance across the galaxy from one side to the other. This can be challenging to do accurately because the edges of galaxies can be diffuse or fuzzy, and depend on the component that is being observed, usually the gas and component of a galaxy extend far away than the stellar component. Additionally, the shape of the galaxy may not be symmetrical, and the edge of the galaxy may not be well-defined.

Star-formation Rate

The Star-formation Rate (SFR) is a measure of the rate at which new stars are being formed within the galaxy, typically expressed in units of solar masses per year, $M_\odot\ yr^{-1}$.

A large number of simulations have explained the star formation at low scales like the collapse of cold gas. The baryonic gas contained in the dark matter halos in which galaxies reside undergoes cooling processes on shorter timescales than the age of its parent halo. As a consequence, the gas is not supported by pressure anymore and starts streaming toward the halo center, increasing its density. In the case of efficient cooling mechanisms, when the density of the gas overcomes the density of the dark matter in the halo center, the gas becomes gravitationally unstable and collapses. This cooling process can lead to the formation of cold and dense clouds (molecular clouds) in which star formation takes place. High-resolution observations in the Milky Way and other nearby galaxies demonstrate that the star formation process occurs within such dense molecular clouds (Blitz 1993; Williams et al. 2000), and in particular, several pieces of observational evidence support the connection between the overall rate of star formation of a galaxy and its capacity of generating dense molecular clouds. Furthermore, observations of the carbon monoxide (CO) line, the main observational probe of molecular clouds, in starburst galaxies show that they are associated with the massive presence of molecular gas ($10^8 - 10^{10} M_\odot$) confined on small regions with dimensions typically lower than 2 kpc. Because relevant spatial and

temporal timescales of the formation of single stars are short compared with the formation and evolution of galaxies, the star formation process can be studied in terms of statistical relations between the star formation rate and the global properties of the gas involved.

The SFR can be determined through several observational techniques, involving tracers at different wavelengths. Since the observational capabilities do not allow to observe of the birth of a star except in the Milky Way or other few nearby galaxies, the SFR is considered as a global parameter of each galaxy, or lately thanks to integral field spectrographs, as a parameter of a galaxy region. Given those observational indicators of star formation are all sensitive to the presence of massive stars, the extrapolation to the total SFR, which includes also the contribution of low-mass stars, requires the assumption of an initial mass function (IMF)

The techniques developed for measuring the SFR evaluate the birth rate of massive stars (M $\geq 10M + \odot$), whose effective temperature during the main sequence phase is higher than $\sim$ 20000 K. Thus, the star formation process is associated with an intense UV radiation field. The energetic photons of this radiation ionize the hydrogen and other elements in the progenitor cloud, and further recombination produces strong emission lines, such as the Balmer (in the visual domain) and Lyman (UV) sequence lines. The UV radiation field is partially absorbed by the dust, which is abundant within the clouds where the star formation process takes place, and that radiates the energy absorbed at mid- and far-IR wavelengths. Finally, massive stars at the end of their evolutionary phase can explode as supernovae: relativistic electrons are accelerated as a consequence of the explosion and emit synchrotron radiation at radio frequencies. In the case of binary systems, the accretion of SN remnants which are compact objects (neutron stars, black holes) originates X-ray radiation.

Furthermore, observations of the specific SFR, that is the SFR per unit mass, in galaxies of the different stellar masses showed that the specific SFR of low-mass galaxies is on average higher than that of high-mass ones. This is related to the downsizing effect: more massive galaxies formed a high fraction of their stellar mass in the older epochs and on short timescales, while less massive galaxies have longer formation and evolution timescales (Cowie et al. 1996; Gavazzi et al. 2006; De Lucia et al. 2007; Sánchez-Blázquez et al. 2009).

1.3.3 Galaxy large-scale environment

Galaxies do not randomly populate the Universe. Their spatial distribution is a consequence of the physical formation processes and the tendency of structures to aggregate in overdense regions, as explained in previous sections. The galaxy environment refers to the physical conditions and properties of the region surrounding the galaxy, including the distribution and properties of other galaxies, the distribution and properties of the intergalactic medium (IGM), and the presence of any larger-scale structures such as galaxy clusters or cosmic filaments.

The characterization and full understanding of the galaxy environment has been a complex process, including misunderstandings due to several definitions of environmental

tracers, challenging observations, and comparing phenomena over different scales.

The first known large structures made up of galaxies were galaxy clusters, discovered in the first half of the last century (Coma cluster Zwicky (1933, 1937), Virgo cluster Smith (1936)). Initially, clusters were considered isolated structures. With the advent of galaxy surveys across larger sky areas, a larger structure connecting galaxy clusters was observed (de Lapparent et al. 1986; Klypin & Shandarin 1993; Pogosyan et al. 1996; Bond et al. 1996). In the current picture, clusters are massive and populous galaxy agglomerates embedded into a filamentary network, the cosmic web.

The cosmic web is the largest structure in the Universe and is ultimately responsible for determining the environment of a galaxy. It was first observed by de Lapparent et al. (1986) using galaxies in the local Universe, and later theorized (Klypin & Shandarin 1993; Pogosyan et al. 1996; Bond et al. 1996). It is composed of under-dense void regions bounded by sheet-like walls embedded in a web-like filamentary network branching on high-density dark matter haloes, that spans over a wide range of scales larger than the Megaparsec (Aragón-Calvo et al. 2007; Cautun et al. 2013; van de Weygaert 2016; Libeskind et al. 2018).

The most prominent and defining features of the cosmic web are the filaments. Filaments appear to be the highways of the Universe, the transport channels along which mass and galaxies get into the higher-density cluster regions (van Haarlem & van de Weygaert 1993; Knebe et al. 2004), and which define the connecting structures between higher-density complexes (Bond et al. 1996; Colberg et al. 2005; van de Weygaert & Bond 2008; Aragón-Calvo et al. 2010). Although this picture is well established by cosmological simulations, it has been difficult to demonstrate observationally due to the low galaxy density and gas temperature of filaments.

By contrast, the tenuous sheetlike membranes are considerably more difficult to find in the spatial mass distribution traced by galaxies. Their low surface density makes them far less visible than the surrounding filaments, while they are populated by galaxies with a considerably lower luminosity (Cautun et al. 2014). Along with filaments, the large void regions represent the most prominent aspect of the Megaparsec scale Universe. These are enormous regions with sizes in the range of 20 - 50 h^{-1} Mpc that are practically devoid of any galaxy, usually roundish in shape and occupying the major fraction of space in the Universe (van de Weygaert 2016).

Galaxy clusters

Galaxy clusters are the largest gravitationally bound structures in the Universe and are embedded at the intersections of filaments of the cosmic web. Their high surface density and the elevated number of bright member galaxies make them identifiable at huge distances and particularly suitable as cosmological tracers. In the hierarchical structure formation model, galaxy clusters grow and evolve through a series of mergers and accretion from the surrounding filaments in their outer envelope (Kauffmann et al. 1999; Benson et al. 2001; Springel et al. 2005; De Lucia et al. 2006).

Clusters of galaxies have typically masses greater than $10^{14} M_{\odot}$ and a size ranging from 1 to 5 Mpc. Galaxy clusters consist of hundreds to thousands of galaxies, orbiting the

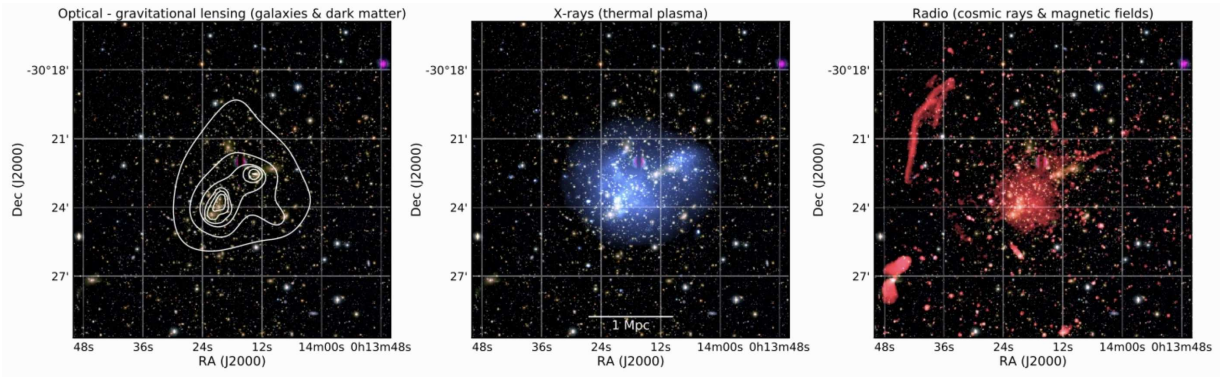


Figure 1.9: The galaxy cluster Abell 2744 as observed in multiple wavelengths. *Left*: optical image of the cluster from the Subaru telescope in the BRz band, with white contours indicating the mass surface density derived from weak lensing measurements. *Middle*: X-ray emission, in blue, from the ICM as observed by the Chandra telescope in the 0.5-2.0 keV energy range. *Right*: radio image of the cluster at 1-4 GHz, in red, taken with the Very Large Array (VLA) telescope, which traces cosmic rays and magnetic fields. Image credit: van Weeren et al. (2019).

cluster with a velocity that is determined by the gravitational potential of the cluster. The typical velocity dispersion of a galaxy cluster member is $\sigma > 400 km/s$. Galaxy clusters contain additionally large amounts of hot, ionized gas, known as the intracluster medium (ICM). Finally, the largest amount of mass within a cluster is given by its dark matter component. The distribution of matter in a cluster can be inferred, among others, through gravitational lensing techniques, the velocity dispersion of cluster galaxies, or the temperature of the ICM.

According to their dynamics, clusters can be classified as virialized or evolving. Virialized galaxy clusters have reached a state of dynamic equilibrium, where individual galaxies and other matter within the cluster are in a state of dynamic balance, moving around within the cluster but not escaping or collapsing. Evolving galaxy clusters are still in the process of forming and undergoing changes, including the accretion of new matter, the merging of smaller clusters or groups into the larger cluster, and the disruption of galaxies within the cluster due to the gravitational forces at work. Evolving galaxy clusters are typically less dynamically relaxed than virialized clusters, and may show the presence of substructures within the cluster or the presence of large-scale gas shocks.

Galaxy clusters present also intracluster light (ICL), which is the diffuse light emitted by stars that are not gravitationally bound to a particular galaxy. The ICL is typically found in the regions between galaxies within the cluster, but it is normally more concentrated around the brightest cluster galaxy (BCG). The dominant ICL production occurs in tandem with the formation and growth of the BCG and/or through the accretion of preprocessed stray stars. The amount of ICL in a cluster can be used as a measure of the past interaction and merging history of the cluster, as the stripping of stars from galaxies is thought to be more common in dynamically active environments (Montes 2022; Joo &

Jee 2023).

Besides clusters, other significant large-scale structures traced by galaxies are galaxy groups. Galaxy groups typically have masses in the range $10^{13} M\odot < M < 10^{14}\odot$ and velocity dispersions $\sigma < 400 km/s$. The fraction of galaxies located in groups increases with cosmic time, up to the Local Universe values (Huchra & Geller 1982; Tully 1987; Eke et al. 2004). Today, over 50% galaxies are in groups, span a wide range in local density, and show properties that range from cluster-like to field-like (Zabludoff & Mulchaey 1998).

Observing galaxy clusters is a challenging task due to the diverse nature of their baryonic constituent components, i.e. galaxies and ICM, which can be observed across a range of wavelengths on the electromagnetic spectrum:

- The galaxy component of a cluster, in particular the baryonic luminous matter traced by its stars, is usually observed at optical and near-infrared wavelengths (Girardi et al. 1998; Biviano et al. 2002, 2006; Koester et al. 2007; Haines et al. 2009; Annunziatella et al. 2016, 2017). Clusters appear as over-densities of galaxies over the field (Binggeli 1982; Postman & Geller 1984). To detect a galaxy cluster using optical or near-infrared data, it is necessary to compute the surface or volume density, according to the photometry or spectroscopy available, and to look for the densest regions of the field. However, the detection of clusters via spectroscopic or photometric surveys is affected by contamination due to projection effects (Postman et al. 1996; Koester et al. 2007; Annunziatella et al. 2017).

- The hot, diffuse gas present in the ICM can be observed through its X-ray emission. The dark matter halos have trapped hot plasma where the deflection of free electrons caused by the interaction with the electric field of protons produces the so-called bremsstrahlung radiation emitting at X-ray frequencies Sarazin (1988). The X-ray luminosity of the hot gas increases with the total mass of the cluster, with typical X-ray luminosities in a wide range of values from 10^{43} to 10^{45} erg/s Rosati et al. (2002). This makes clusters the most luminous X-ray sources in the Universe, except for quasars (Brunetti & Jones 2014). The advantage of relying on this kind of observations for identifying structures in the Universe is the possibility of tracing the potential well in which galaxy processes related to the environment occur, and directly measuring X-ray-related quantities such as temperatures and luminosities which can be used as proxies for the mass of the halo (Vikhlinin et al. 2006).

- Galaxy clusters can be observed in the microwave domain thanks to the Sunyaev-Zel'dovich (SZ) effect (Sunyaev & Zeldovich 1970, 1972). The SZ effect is a phenomenon that occurs when the CMB radiation is scattered by hot gas in the ICM. As the CMB photons pass through the ICM, they can interact with the high-energy electrons in the gas, resulting in a characteristic shift in the energy of the CMB photons. This effect can be observed as a change in the temperature of the CMB radiation at different frequencies. By measuring this temperature shift, it is possible to infer the presence and properties of the ICM in a galaxy cluster (Carlstrom et al. 2002).

An example of a galaxy cluster observed through different wavelengths is presented in Fig. 1.9.

The core of a galaxy cluster can be identified by its virial radius. The virial radius is the region within which the mean density of the cluster exceeds the critical density of the Universe. The critical density depends on the mass-energy content of the Universe and can be calculated using the equation $\rho_c = \frac{3H^2}{8\pi G}$, where H is the Hubble constant and G is the gravitational constant. To define the virial radius, it is commonly used the r_{200} parameter. r_{200} is the radius at which the average density of the cluster is 200 times the critical density of the Universe. Within the viral radius, the cluster is gravitationally bound and will continue to grow over time through the accretion of matter from its surroundings. The total mass of the cluster, and therefore its gravitational influence on the surrounding cosmic web, is usually determined as the mass contained within the virial radius (Evrard et al. 1996; Eke et al. 1998).

Outskirts of galaxy clusters

The outskirts of a galaxy cluster are defined as the regions beyond the virial radius, where the transition from the dense cluster environment to the lower-density field takes place (Balogh et al. 2004; Cortese et al. 2006; Haines et al. 2015; Pimbblet et al. 2013; Paccagnella et al. 2016; Wetzel et al. 2015). The accretion of matter from the filaments of the cosmic web into a galaxy cluster can be observed and characterized on the outskirts of the cluster, as shown in Figure 1.10.

The study of the galaxy population in the outskirts of a galaxy cluster provides valuable insights into the processes of galaxy evolution and the role of the cluster environment in shaping the properties of galaxies, as well as information about the structure and formation of the cluster itself: the accretion flows of matter towards the cluster core leave characteristic marks, especially in the cluster outskirts, giving rise to caustics in the dark matter density profile (Mansfield et al. 2017; Diemer et al. 2017), internal bulk and turbulent gas motions (Lau et al. 2009; Vazza et al. 2009; Battaglia et al. 2012), non-equilibrium electrons in the ICM (Rudd & Nagai 2009), and an inhomogeneous gas density distribution (Nagai & Lau 2011; Roncarelli et al. 2013).

Despite the presence of these signatures, the complete accretion physics taking place in the cluster outskirts still needs to be explored (Dekel et al. 2009; Danovich et al. 2012; Welker et al. 2020; Walker et al. 2019). This is primarily because complete observations in the cluster outskirts are challenging, and comparisons with simulations are required to understand the underlying physics. To date, most studies have focused on the relatively dense central regions of galaxy clusters, which represent only the tip of the iceberg of the ICM, corresponding to the inner $\sim 10\%$ in terms of volume.

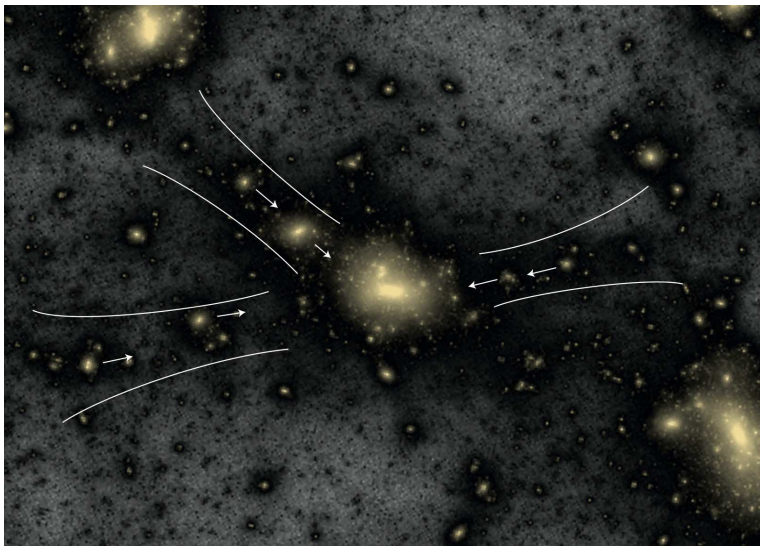


Figure 1.10: The growth of galaxy clusters. Galaxy clusters grow by the accretion of galaxies from their anisotropic surroundings. Infalling galaxies (motions indicated with white arrows) inject angular momentum and induce tides on the galaxy cluster, which may cause it to align with the surrounding matter distribution, and a similar effect may apply to the central galaxy. Image from Sifón (2017).

1.4 Environment as a driver of galaxy evolution

Previously, I introduced the connection between the large-scale structure with the properties of galaxies. This section focuses on the formation of new stars as a way to understand the "life" of a galaxy and the process of star formation quenching as a key factor in the evolution of a galaxy.

Galaxy clusters, being the largest gravitationally bound structures in the Universe, provide a unique opportunity to study the evolution and properties of galaxies in a high-density environment. In recent years, there has been increasing interest in studying the outskirts of galaxy clusters, as these regions provide a transition between the dense cluster environment and the lower-density field, and are thought to be an important site for galaxy evolution.

One key aspect of galaxy evolution that has garnered attention is the process of star formation quenching, or the mechanisms that act on individual galaxies to reduce their star formation activity. Understanding this process is crucial for understanding how galaxies evolve and how they are affected by their environment. By studying the galaxy population in the outskirts of galaxy clusters, it is possible to gain insight into the role of the cluster environment in shaping the properties of galaxies, as well as the processes of galaxy evolution.

1.4.1 Environment definition

The environmental dependence of galaxy properties is well established as it will be explained below. Nevertheless, 'galaxy environment' is a very general concept that has been used in a variety of ways. Its definition, what it measures, and how it is measured, can vary from author to author (Muldrew et al. 2012). Nevertheless, the galaxy environment can be divided into two broad categories. The *global* or *large-scale environment*, refers to the larger-scale structure where the galaxy is located, in particular is related to the mass of its host halo. The *local environment* refers to the immediate surroundings of the galaxy, including the other galaxies and gas that are close to it in space. This can include other galaxies in the same group or cluster, as well as any intergalactic gas or dark matter in the region.

In general, the global environment is measured by estimating the mass of the host halo, (e.g. through the velocity dispersion of galaxies within a cluster, or using the temperature of the ICM), or by measuring the galaxy density, in projected or redshift-space, using fixed aperture methods, e.g. counting the number of galaxies within cells of a fixed size. The local environment is best measured with nearest neighbor methods to determine the local density of glaxies (Kovač et al. 2010; Cucciati et al. 2010; Muldrew et al. 2012; Vulcani et al. 2012; Darvish et al. 2015; Fasano et al. 2015).

1.4.2 Introduction to the environmental influence over galaxies

The pioneering studies of Zwicky (1951) and Abell (1958) introduced the study of galaxy populations inside clusters. A starting point for the study of the influence of environment over galaxies was the discovery of the Butcher-Oemler effect. The Butcher-Oemler effect refers to the observation that galaxy clusters at $z \sim 0.5$ and $z \sim 0.3$ have a higher fraction of blue, actively star-forming galaxies than their low-redshift counterparts Oemler (1974); Butcher & Oemler (1978, 1984). Subsequent studies have found evidence for the Butcher-Oemler effect in clusters at a range of redshifts (Ellingson et al. 2001; Demarco et al. 2007; Haines et al. 2009). Another milestone in the study of environment in relation with galaxies is the morphology density relation. In high-density regions of galaxy clusters, the fraction of early-type (elliptical and lenticular) galaxies is higher, while in low-density regions, the fraction of late-type (spiral and irregular) galaxies is higher. This relationship was first reported by Dressler (1980) in a study of galaxy clusters in the Coma and Abell 1367 clusters. The morphology density relation has since been observed in many other galaxy clusters and is thought to be related to the role of the cluster environment in shaping the properties of galaxies (Postman & Geller 1984; Dressler et al. 1997).

Considering larger scales, instead, the advent of large spectroscopic galaxy surveys, such as the Sloan Digital Sky Survey (SDSS; York et al. 2000) and the two-degree Field Galaxy Redshift Survey (2dFGRS; Colless et al. 2001), has greatly enhanced the ability to study the connection between galaxies and their environments. Using large redshift surveys, Balogh et al. (2004) found a smooth transition from high-density regions populated with red galaxies towards blue-dominated low-density regions. Blanton et al. (2005) showed that the typical rest-frame color, luminosity, and morphology of nearby galax-

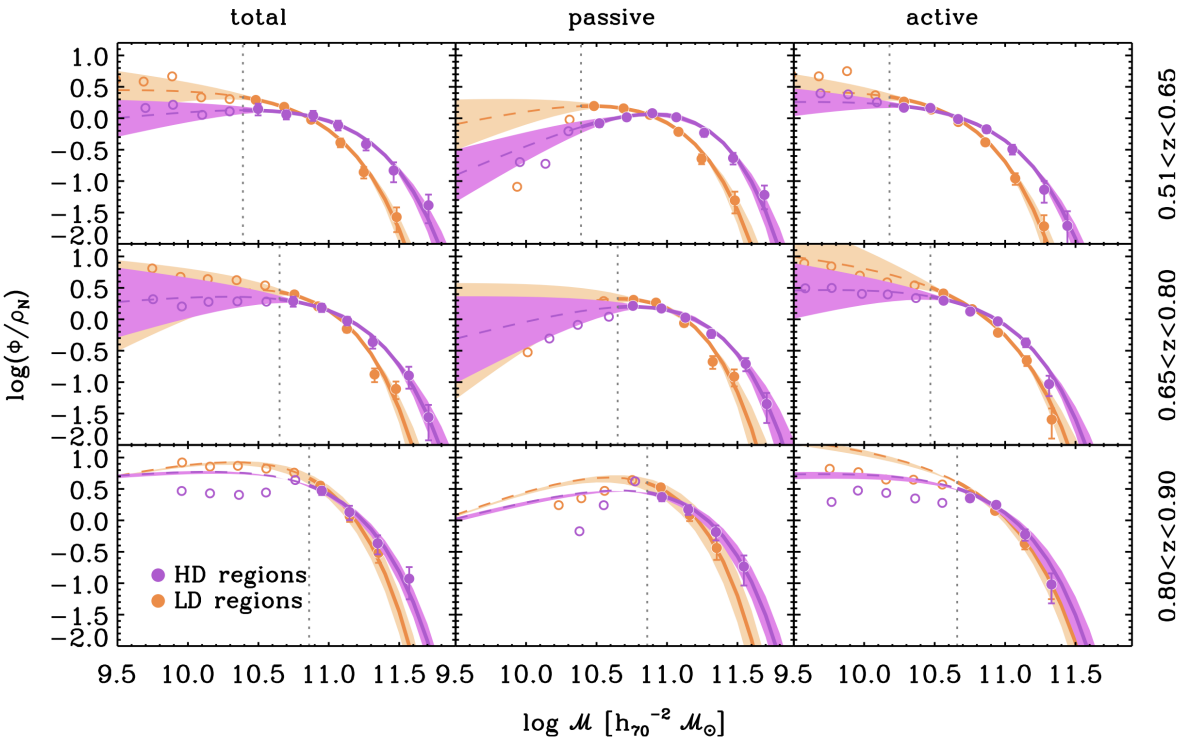


Figure 1.11: Stellar mass functions of galaxies at low density (orange) and high density (violet) in three different redshift bins from the VIPERS survey. Right panels show the GSMFs of active galaxies, while central panels refer to passive ones. The GSMFs of the whole sample in the same z-bins are shown on the left (Davidzon et al. 2016).

ies are highly correlated with the local galaxy density on $\sim 1\ h^1$ Mpc scales. Cucciati et al. (2006), studying the VVDS survey, showed that the color-density relation shows a dramatic change as a function of cosmic time: at the lowest redshifts the fraction of the red(/blue) galaxies increasing(/decreasing) as a function of density, at previous epochs (up to z$\sim$1.5) blue and red galaxies seem to be mostly insensitive to the surrounding environment, with a nearly flat distribution of the fraction of the bluest and reddest objects over the whole over-density range. The absence of the color-density relation at the highest redshift bins investigated implies that quenching of star formation was more efficient in high density regions. Davidzon et al. (2016) conducted a study on the evolution of the galaxy stellar mass function (GSMF), using a the VIPERS survey, finding that the GSMF evolves differently in high and low density regions: high density environments in general contain more massive galaxies, and the number of passive galaxies is larger on high density regions, as shown in Fig. 1.11.

1.4.3 Color-magnitude diagram

An important tool for represent statistically galaxy properties is the color-magnitude diagram, which is a graphical representation of a sample of galaxies, plotted according to

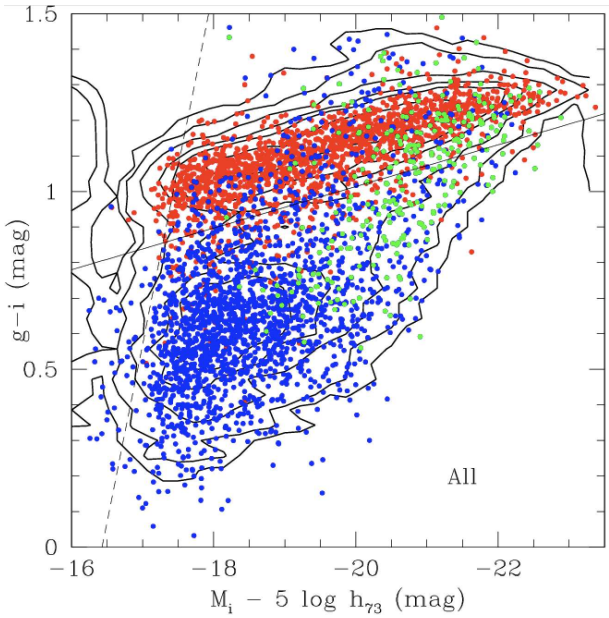


Figure 1.12: g i color versus i-band absolute magnitude relation of all galaxies in the Coma cluster coded according to Hubble type: red = early-type galaxies (dE-E-S0-S0a); blue = disk galaxies (Sbc-Im-BCD); green = bulge galaxies (Sa-Sb). Contours of equal density are given. Image credits: Gavazzi et al. (2010)

their colors and absolute magnitudes. The color of a galaxy is measured as the difference between the magnitude in two photometric bands. The color-magnitude diagram provides information about ages, metallicities, and star formation histories of galaxies, e.g., young, actively star-forming galaxies tend to have blue colors, while older, quiescent galaxies tend to have red colors. The position of a galaxy on the color-magnitude diagram can also reveal its mass, as more massive galaxies tend to be more luminous (Blanton et al. 2005; Baldry et al. 2006; Gavazzi et al. 2010)). An example of the color-magnitude diagram for the galaxies in the Coma cluster is presented in Fig 1.12.

1.4.4 Major drivers of galaxy evolution

From an observational perspective, to address outstanding questions on the evolution of galaxies, homogeneous and sizable galaxy samples are needed, covering a wide redshift range, and with reliable measurements of galaxy rest-frame colors, luminosities and masses. The evolutionary history of isolated galaxies, by definition free as much as possible from environmental influences, can then be used as the benchmark for exploring the presence of environment-dependent forces taking place in group and cluster galaxies.

Classical discussions contrast a scenario in which the fate of a galaxy is determined primarily by physical processes coming into play after the galaxy has become part of a group or a cluster (“nurture”), to one in which the observed environmental trends are established before these events and primarily determined by internal physical processes

("nature"). Among a galaxy's internal properties, a galaxy's stellar mass is considered the main contributor, as galaxies are characterized by a wide range of total stellar masses whose role in determining galaxy properties is crucial. For example, Kauffmann et al. (2003) found that color, specific star formation rate, and internal structure are strongly correlated with galaxy stellar mass. Among external processes, the major actors are considered to be the environment in which a galaxy is found during its lifetime and mergers.

This concept is represented intuitively in the left panel Fig. 1.13 (Figure 15 from Peng et al. (2010)), in the context of the build-up of the passive population. Galaxy stellar mass leads the quenching of star formation at all cosmic epochs above a certain stellar mass (which slowly increases with time), while the low-mass regime is dominated by environmental quenching in the recent history of the Universe ($z<0.5$) and by merging processes in the past. Focusing on galaxies in the local Universe, the diagram is observationally motivated by the trends represented in the right panel of Figure 1.5. In the figure, which analyzes a sample of low-redshift galaxies in the SDSS (York et al. 2000), the variation in the fraction of red galaxies is represented as a function of environment, parametrized as local overdensity of galaxies, and stellar mass. It emerges that galaxies become redder in any given environment when increasing their stellar mass, and in turn, the red fraction increases when galaxies reside in denser environments at fixed stellar mass. These results are supported by many studies in the literature which found that environment seems to be more relevant for lower mass galaxies, at least as far as quenching is concerned: galaxies in denser environments tend to be redder than galaxies in less dense environments (Haines et al. 2007; Pasquali et al. 2010; Peng et al. 2010; McGee et al. 2011; Sobral et al. 2011; Muzzin et al. 2012; Smith et al. 2012; Wetzel et al. 2012; La Barbera et al. 2014; Lin et al. 2014; Vulcani et al. 2015; Guglielmo et al. 2015). In contrast, on average, more massive galaxies have formed their stars and completed their star formation activity at higher redshift than less massive galaxies, regardless of the environment.

However, this dichotomy is simplistic because stellar mass and environment are interrelated. The parameterization of the latter is often connected to the gravitational mass of the hosting halo, which is also physically coupled to galaxy stellar mass. Therefore, in a scenario of hierarchical accretion, it is expected that most massive galaxies show a correlation with overdensities (Kauffmann et al. 2004; Abbas & Sheth 2005; Scodeggio et al. 2009). For this reason, it is misleading to contrast stellar mass and environment as two separate aspects of galaxy evolution (see the discussion in De Lucia et al. (2012)).

In the hierarchical model, dark matter collapses into haloes in a bottom-up fashion: small systems form first and subsequently merge to form progressively larger systems. As the structure grows, galaxies join more and more massive systems, therefore experiencing a variety of environments during their lifetime. In this context, the nature-nurture debate appears to be ill-posed, as these two elements of galaxy evolution are inevitably and heavily interconnected (De Lucia et al. 2012).

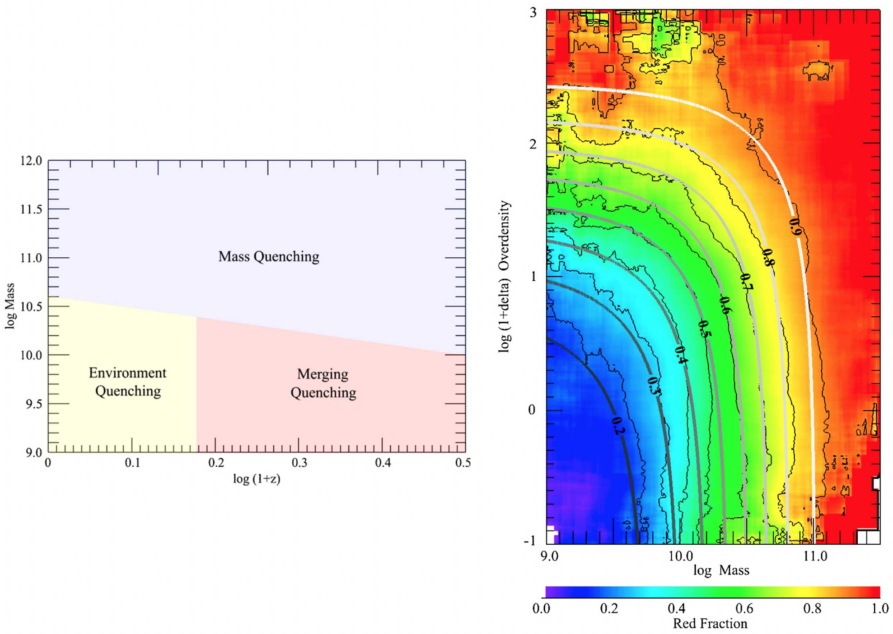


Figure 1.13: Left panel: Diagram showing the dominant mechanism for the quenching of galaxies as a function of mass and redshift in typical environments (Peng et al. 2010). According to the authors, merging and environmental quenching reflect both the underlying merger of dark matter halos. Right panel: Fraction of red galaxies in the SDSS sample studied in Peng et al. (2010) as a function of stellar mass and environment.

1.4.5 Physical process acting in dense environments

Several physical processes can affect the evolution of galaxies. These processes can be divided into those related to the interaction with the ICM and those related to the galaxy-galaxy interactions. The description of these phenomena is described following (De Lucia et al. 2007, 2010; Boselli & Gavazzi 2006; Cortese et al. 2021):

Interaction with the ICM:

Galaxies interacting with the hot ionized gas of the ICM can experience a plethora of physical mechanisms, usually leading to some form of gas removal or consumption. This gas removal entails a direct decrease or inhibition of future star formation, leading to galaxy quenching. As a consequence, galaxies in clusters have properties that differ from those in the field (Dressler 1980; Poggianti et al. 1999; Bai et al. 2009).

- **Gas starvation or strangulation**: Current theories of galaxy formation assume that, when a galaxy is accreted onto a larger structure, the gas supply can no longer be replenished by cooling that is suppressed by the removal of the hot gas halo associated with the infalling galaxy. This process is usually referred to as 'strangulation' or 'starvation' (Larson et al. 1980; McCarthy et al. 2008). It is expected that strangulation affects the star formation of cluster galaxies on relatively long timescales, and therefore causes a slow decline in the star formation activity. Nevertheless, in semi-analytic models, this process is usually associated with strong supernovae feedback and is assumed to be instantaneous. As a consequence, galaxies that fall onto a larger system consume their cold gas rapidly, moving onto the red sequence on very short time scales. Numerical simulations have recently pointed out that the stripping of the hot halo associated with infalling galaxies should not happen instantaneously.

- **Ram pressure stripping**: Galaxies, when entering into a dense ICM suffer a strong ram-pressure stripping that can sweep cold gas out of the stellar disc (Gunn & Gott 1972). Depending on the binding energy of the gas in the galaxy, the ICM will either blow through the galaxy removing some of the diffuse ISM or will be forced to flow around the galaxy. Ram-pressure stripping is expected to be more important at the center of massive systems because of the large relative velocities and higher densities of the ICM (Brüggen & De Lucia 2008). Virtually all cluster galaxies suffered weaker episodes of ram pressure Vulcani et al. (2022), suggesting that this physical process might have a significant role in shaping the observed properties of the entire cluster galaxy population. Figure 1.14 shows the ESO 137-001 galaxy undergoing ram-pressure stripping.

- **Thermal evaporation**: Another mechanism to slow down the star formation rate of galaxies is the thermal evaporation of the cold gas in disc galaxies via heat conduction from the surrounding hot ICM (Cowie & Songaila 1977). If the ICM temperature is high compared to the galaxy velocity dispersion, at the interface between the hot ICM and the cold ISM, the temperature of the ISM rises rapidly and the gas

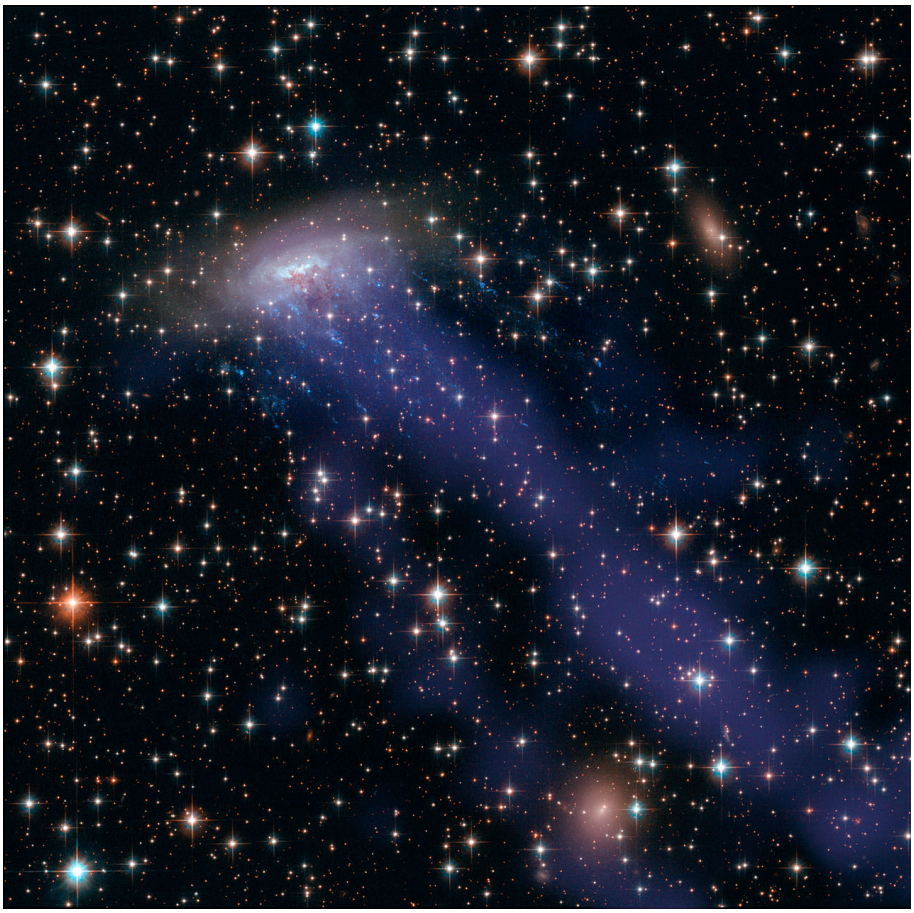


Figure 1.14: The spiral galaxy ESO 137-001 entering the Norma cluster. This image combines NASA/ESA Hubble Space Telescope observations with data from the Chandra X-ray Observatory. As well as the electric blue ram pressure stripping streaks seen emanating from the galaxy, a giant gas stream can be seen extending towards the bottom of the frame, only visible in the X-ray part of the spectrum. Credit: NASA, ESA, CXC [4]

evaporates. The mass loss rate is proportional to the dimension of the galaxy and the density and temperature of the ICM. A typical galaxy (with a radius of 15 kpc and 5×10^9M of atomic gas) can be completely stripped on timescales of the order of some 10^7 - 10^8 yr. (Cowie & Songaila 1977; Bureau & Carignan 2002)

Galaxy - galaxy interactions:

Galaxy-galaxy interactions refer to the gravitational interactions between two or more galaxies. These interactions can lead to a variety of phenomena, including the merging of galaxies, the formation of tidal tails and bridges, and the triggering of star formation or quenching in the affected galaxies (e.g., Toomre & Toomre (1972); Kennicutt (1998); Scudder et al. (2012).

- **Tidal interactions among Galaxies or Galaxy mergers**: Galaxy mergers are the process by which two or more galaxies merge to form a single, larger galaxy. Numerical simulations have shown that close interactions can lead to a strong internal dynamical response driving the formation of spiral arms and, in some cases, of strong bar modes (Mihos 2004; Toomre & Toomre 1972; Barnes & Hernquist 1996; Cox et al. 2008). Sufficiently close encounters can destroy the disc, leaving a kinematically hot remnant with photometric and structural properties that resemble those of elliptical galaxies. Mergers, act on gas, dust, and stars, as well as on dark matter, with an efficiency depending on the gravitational bounding of the various components. Mergers are, however, more efficient in the infalling group environment than in massive clusters because of the large velocity dispersion of massive clusters. Mergers are an important factor throughout the lifetime of a galaxy cluster: at early times when the cluster is first collapsing, and still at later times in the outskirts of the cluster, as it gathers galaxies which will infall into the core. Mergers drive gas towards the center of galaxies, where it can trigger a burst of star formation and fuel the central black hole, the feedback from which can heat the remaining gas and eventually quench star formation (Mihos & Hernquist 1996; Wild et al. 2007; Schawinski et al. 2009; Gnedin 2003; Villalobos et al. 2014).

- **Harassment**: Galaxy harassment is a mechanism that strips a galaxy of part of its mass and drives a morphological transformation as a consequence of repeated fast encounters, coupled with the effects of the global tidal field of the cluster (Spitzer & Baade 1951; Farouki & Shapiro 1981; Moore et al. 1996, 1999). A detailed study on the impact of harassment on the properties of galaxies was conducted by Moore et al. (1996), which concluded that harassment has the potential to change any internal property of a galaxy within a cluster, including the gas distribution and content, the orbital distribution of stars, and the overall shape. It is expected to have an important role in the formation of dwarf ellipticals, in the transformation of gas-rich spirals into gas-poor lenticular galaxies, or in the destruction of low surface brightness galaxies in clusters, but is less able to explain the evolution of luminous cluster galaxies. In addition, at the typical velocity dispersions of bound groups, galaxy-galaxy harassment (i.e. frequent galaxy encounters) leads to the morphological transformation of disc galaxies. Indeed, it results in the loss of the galaxy's gaseous component, partly ablated and partly falling into the center, entailing a dramatic conversion of discs into spheroidals. (e.g., Spitzer & Baade 1951; Moore et al. 1996, 1999)

- **Tidal interactions with the cluster potential**: Not only tidal interactions between galaxies can induce some relevant perturbation, but also tidal interactions between galaxies and the whole cluster potential, induce gas inflow, bar formation, nuclear, and perhaps disc star formation. Models by Fujita (1998); Valluri (1993); Henriksen & Byrd (1996) showed that tidal compression of galactic gas via interaction with the cluster potential can accelerate molecular clouds of disc galaxies falling towards the core, increasing the star formation rate. The efficiency of this process

depends on the cluster-to-galaxy mass ratio, the distance of the galaxy from the center, and the galaxy radius. The first galaxies to experience this kind of interaction are large galaxies passing within a few hundred kpcs of the cluster center. The net result could be a first increase in the nuclear activity of cluster galaxies and eventually a decrease of the total gas consumed through star formation events, while a complete removal of the gas due directly to the interaction is unlikely. At group scales, the combinations of tides and ram pressure stripping efficiently remove the gas from spirals, quenching the star formation in galactic discs, while triggering it in the arms and at the leading edge of the gaseous disc, with the net result of a suppression of star formation on timescales of several Gyr (Moore et al. 1996).

1.4.6 Pre-processing in groups and filaments

Preprocessing refers to the processes that affect a galaxy before it becomes a member of a larger structure, such as a galaxy cluster. Preprocessing can occur in a variety of environments, affecting the properties of a galaxy as it moves through the cosmic filaments and approaches a galaxy cluster, shaping its structure and properties and preparing it for its future evolution within the cluster. Preprocessing can affect the star formation activity, morphology, and metallicity of a galaxy, and it can have a significant impact on the galaxy's evolution. (e.g. (Moore et al. 1996; Muzzin et al. 2008; Wetzel et al. 2012; Haines et al. 2013).

It is essential to notice that many of the physical processes outlined above acting on dense cluster environments are relevant also in the preprocessing phase. Additionally, a large fraction of the preprocessing effects takes place in galaxy groups. In numerical simulations, massive clusters have accreted up to 50% and 45% of their stellar mass and galaxies, respectively, through galaxy groups (McGee et al. 2009). Hence, being fundamental building blocks of both mass and galaxy populations in clusters, galaxy groups could also play a significant role in shaping the evolution of cluster galaxies. Studying preprocessing effects, therefore, is related to the study of groups and galaxy overdensities in the outskirts of clusters.

As the properties of the ICM vary from the cluster core to the outskirts (Nagai & Lau 2011; Ichikawa et al. 2013; Lau et al. 2015; Biffi et al. 2018; Mirakhor & Walker 2021), galaxies in the external regions are expected to have a different degree of interaction with the ICM, and the efficiency of the aforementioned physical processes could be different. In the cluster envelopes, environmental effects could also accelerate the consumption or remove the gas reservoir before galaxies enter a cluster (e.g., Zabludoff & Mulchaey 1998; Mihos 2004; Fujita 2004). Hence the cluster envelopes are a key region infalling galaxies cross even before they reach the cluster core. To fully characterize the effect of the cluster environment on galaxies, it is mandatory to have a thorough characterization of the population of galaxies in the outskirts of clusters. An additional motivation to study this region at intermediate and high redshift is because pre-processing mechanisms do not work in the same way in the local Universe as they worked in the past (van der Burg et al. 2020).

At redshift $z < 0.3$, observational studies have found that the properties of galaxies such as star formation, gas content, and color are indeed affected by the cluster environment at large clustercentric distances, up to at most ∼3 virial radii (e.g., Solanes et al. 2002; Lewis et al. 2002; Gómez et al. 2003; Verdugo et al. 2008; Park & Hwang 2009; Braglia et al. 2009; von der Linden et al. 2010; Dressler et al. 2013; Haines et al. 2015; Rhee et al. 2017; Paccagnella et al. 2016). In particular, spiral galaxies with low star formation rates are found in the outskirts of clusters (Couch et al. 1998; Dressler et al. 1999) and the fraction of blue/star-forming galaxies in the outskirts is intermediate between the field and core values (Wetzel et al. 2013; Haines et al. 2015; Guglielmo et al. 2018; Bianconi et al. 2018; Just et al. 2019). Exploiting cosmological simulations, Bahe et al. (2013) recovered a deficit of star-forming galaxies in the infalling region of clusters, while Zinger et al. (2018) found that the hot gas is removed from galaxy halos, and this process is more effective between 1 and 3 virial radii. They argue that the removal of gas from the galaxy's halo sets the stage for the quenching of the star formation by starvation over 2-3 Gyr before it enters the cluster.

The studies mentioned above have analyzed the infall region considering all possible directions, while other works consider the fall through filaments only. Galaxies are indeed preferentially accreted into clusters through filaments (e.g., Colberg et al. 1999; Ebeling et al. 2004; Castignani et al. 2022), and to a lesser extent from other directions. Both Martínez et al. (2016) and Salerno et al. (2020) at z<0.15 exploiting SDSS and WINGS clusters and Salerno et al. (2019) at $0.43 < z < 0.89$ using VIPERS data compared the properties of galaxies falling into clusters along filaments or from other directions (isotropic infalling), finding an enhancement of quenching for galaxies in filaments.

Until now, a few studies have focused on cluster outskirts up to a large distance from the core with deep data at intermediate redshift (Lubin et al. 2009; Schirmer et al. 2011; Lemaux et al. 2012; Verdugo et al. 2012; Lu et al. 2012; Just et al. 2019; Sarron et al. 2019; Lemaux et al. 2019; Tomczak et al. 2019). The emerging picture is that cluster outskirts play a major role in cluster evolution as early as $z \sim 1.4$ (Lemaux et al. 2019; van der Burg et al. 2020). However, a deep systematic mapping of the cluster envelopes out to very large clustercentric distances, $r \sim 5r_{200}$, is still missing beyond the local universe.

1.5 This thesis

In this thesis, I aim to study the galaxy distribution across large scales, using the galaxy density field as a tool to simultaneously probe both the evolution of galaxies and the ΛCDM cosmological model. Therefore, the thesis is organized into two main sections, each corresponding to a different cosmic scale. At cluster scales, around 10 Mpc h^{-1}, I study the role of the environment in shaping the properties and evolution of galaxies. At larger scales, around 100 Mpc h^{-1}, I study the overall structure and evolution of the Universe from a cosmological perspective.

The different targets and approaches followed in this thesis are connected materially by the fluctuations of the density field, most commonly referred to as the density field, which is generally expressed as $\delta = \rho/\bar{\rho} - 1$, where ρ is the local density of galaxies and $\bar{\rho}$ is the mean local density of galaxies at a given redshift. I aim to demonstrate the potential of using the density field as a bridge between astrophysics and cosmology since it is a single observable that encodes information on both scales due to its sensitivity to the large-scale structure of the Universe determined by the ΛCDM model, as well as its ability to reveal the underlying physical processes that govern the formation and evolution of galaxies.

In the first part of this thesis, I introduce the Galaxy Assembly as a Function of Mass and Environment with VST (VST-GAME) survey, to study the galaxy density field and characterize the galaxy evolution scenario. The work corresponding to this section is presented in Estrada et al. (2022b)

The first part of **Chapter 2**, covers the VST-GAME observational program, including a description of the VST telescope and its photometric instrument, OmegaCAM, as well as the complementary NIR photometry from the VISTA telescope provided by the GCAV survey. The second part of **Chapter 2**, focuses on the galaxy cluster MACS J0416.1-2403 and provides a detailed description of the photometric analysis of the VST and VISTA images, as well as the process of source extraction and catalog construction, including photometric redshifts. In **Chapter 3**, I use the density field around M0416 to classify the environment of galaxies according to their local density, analyzing a large galaxy sample ranging from the cluster center to the outskirts at epochs when the galaxy population is still rapidly evolving Poggianti et al. 2006; Desai et al. 2007. This allows me to study environmental effects in shaping the properties of galaxies at large clustercentric distances.

The section focused on cosmology of this thesis is a work carried over spectroscopic data from the VIMOS extra galactic galaxy survey (VIPERS) presented in Guzzo et al. (2014). The work corresponding to this section is presented in Estrada et al. (2022a).

Chapter 4 presents the VIPERS survey, introduces the Bayesian code that I developed to unmask the VIPERS survey, and summarizes the main results. The final aim of my work was the reconstruction of the galaxy density field to bring joint probabilities on the parameters of the ΛCDM model.

The third part of this thesis, *final remarks*, contains not only the summary and conclusions of my work but also an analysis of the future projects that can be done starting from the main chapters.

The density field around a galaxy cluster

Chapter 2

The Galaxy Assembly as a function of Mass and Environment: VST-GAME

The Galaxy Assembly as a function of Mass and Environment with VST (VST-GAME, P.I. Amata Mercurio) is a large observing program aimed at studying the evolution of galaxies in the outskirts of galaxy clusters at intermediate redshift (z$\sim$0.4). The survey uses photometric data from the OmegaCAM instrument at the VLT Survey Telescope (VST) to observe a wide range of galaxies in the intermediate Universe. VST-GAME makes use of near-infrared observations from the VISTA Public Survey program Galaxy Clusters At Vircam (G-CAV, P.I.: M. Nonino), and spectroscopic redshift measurements provided by the CLASH-VLT project (P.I.: P Rosati). The data collected by the VST-GAME survey will be used to investigate the physical processes that drive the growth and evolution of galaxies in these dense environments and to understand better how they are affected by their surroundings.

The VST-GAME survey (Mercurio & VST-GAME Team 2018; Estrada et al. 2022b)is one of the first projects to focus on the large outskirts of 6 galaxy clusters (r>2r$_2$00) at intermediate redshift (z$\sim$0.4), using deep photometry ($r \sim 24.4$ mag). Outskirts of galaxy clusters are an essential but under-explored area for the growth and evolution of galaxies.

The data collected by the VST-GAME survey will be used to answer a range of important questions about galaxy evolution in the outskirts of galaxy clusters, e.g. the survey will investigate the role of environmental effects in shaping the growth and evolution of galaxies. The results of the VST-GAME survey will provide new insights into the physical processes that drive the evolution of galaxies, will advance our understanding of the role of galaxy clusters in the evolution of the Universe, and will open the door to the study of filaments of the cosmic web on an epoch still under-explored.

Reduced images from the VST optical bands and near-infrared VISTA bands are the starting point of this thesis. The image reduction was implemented by Mario Nonino and the OATS.

2.1 Data description

2.1.1 VST telescope and OmegaCAM instrument

The VLT Survey Telescope (VST) is a telescope with a 2.6-meter aperture that is situated at the European Southern Observatory's Paranal Observatory, located at an altitude of 2635 meters in Chile. VST is designed for surveying the sky in visible light and is equipped with a wide-field imager, OmegaCAM, which has a field of view of 1 degree and a resolution of 268 megapixels. This combination of the telescope's wide field of view and the high resolution of OmegaCAM allows for the capture of detailed images of large portions of the sky in a single exposure.

The VST is focused on the optical range of wavelengths of 0.3 to 0.7 microns. However, while the largest telescopes, such as the VLT, can only study a small region of the sky at any point, the VST is designed to rapidly and thoroughly photograph large areas of the sky. With a field of view totaling 1 deg^2, the VST supports the VLT's observations by detecting and pre-characterizing sources that the VLT Unit Telescopes can then further study.

The VST telescope is designed using a Ritchey-Chretien configuration, which consists of two mirrors: a primary mirror (M1) with a diameter of 2.61 meters and a smaller secondary mirror (M2) with a diameter of 93.8 cm. This design allows the telescope to achieve a wide field of view and high image quality. The VST is mounted on a stable alt-azimuth platform, allowing it to track astronomical objects as they move through the sky. Additionally, the telescope is equipped with an advanced active optics system, which automatically adjusts the shape of the primary mirror to maintain a clear image across the field of view.

OmegaCAM is a CCD camera located at the Cassegrain focus of the VST telescope. It consists of a mosaic of 32 CCD detectors, resulting in a total of 16,000 x 16,000 pixels, allowing for the capture of high-resolution images of the sky. In addition, OmegaCAM is equipped with a set of filters that enable observations of the sky across a range of wavelengths, from ultraviolet to near-infrared.

The VST telescope and OmegaCAM instrument work together to provide a powerful tool for surveying the sky. The camera's advanced technology, the quality of the telescope's optics, and the exceptional observing conditions at Paranal enable the VST to capture detailed images of large areas of the sky in a single exposure. This allows astronomers to study a wide variety of astronomical phenomena, such as distant galaxies and nearby star clusters. Figure 2.1 shows the VST telescope.

2.1.2 VST-GAME survey

The VST-GAME survey gathers observations for six massive galaxy clusters at $0.2 \lesssim z \lesssim 0.6$ (Abell 2744, MACSJ0416-2403, Abell S1063, MACSJ0553.4-3342, PLCK G287.0+32.9, RXC J1514.9-1523) to investigate galaxy evolution down to 10^9 $M_\odot$ in stellar mass, in a wide and largely unexplored range of local densities. The uniqueness of this dataset is given by the wide field coverage ($\sim$20$\times$20 Mpc^2 at z=0.4), combined with long exposures

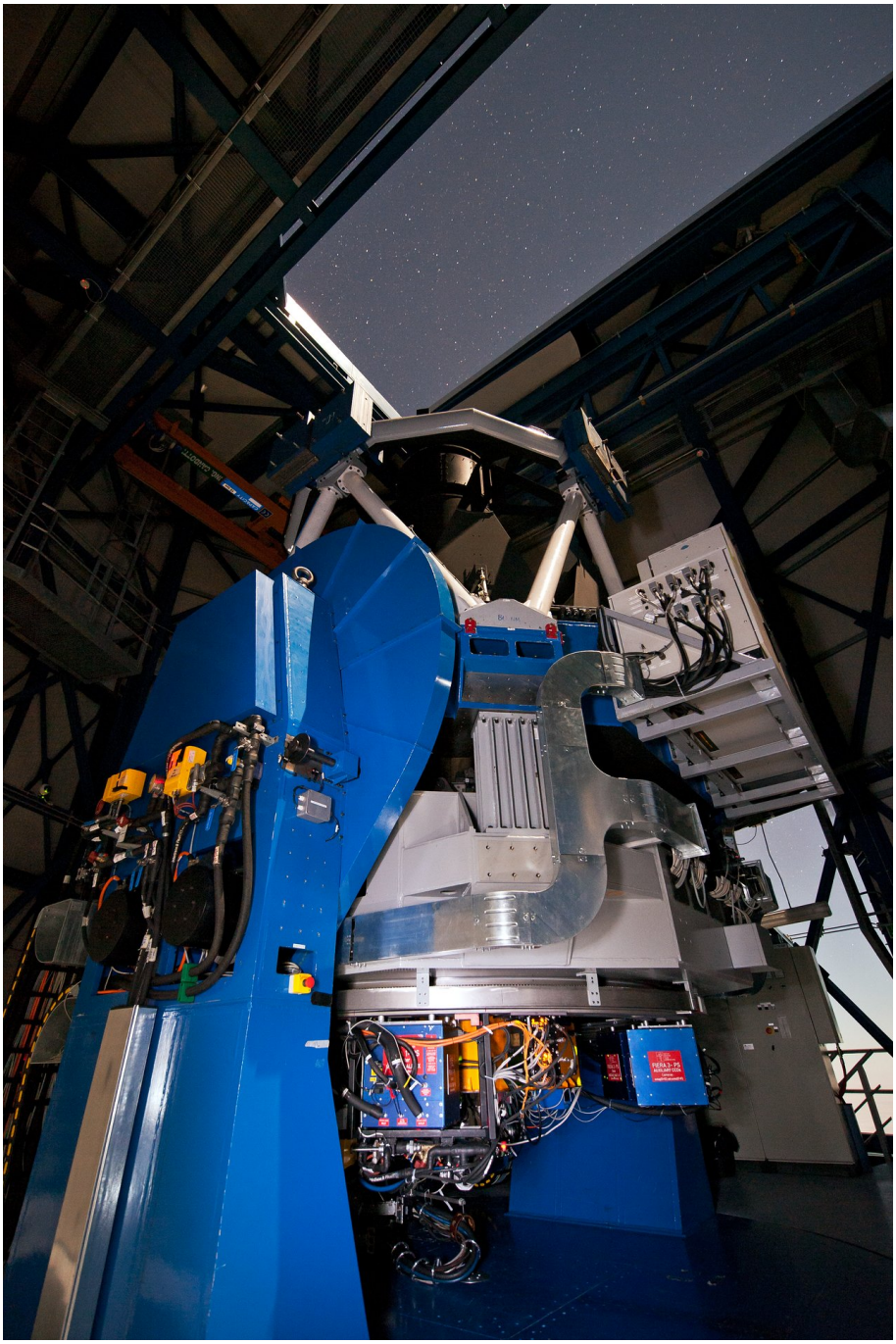


Figure 2.1: The VLT Survey Telescope (VST) at Cerro Paranal. The VST is a state-of-the-art 2.6-meter telescope equipped with OmegaCAM, a 268-megapixel CCD camera with a field of view of 1 deg^2. It will survey the visible-light sky. Credit:ESO, G.Hüdepohl

(between 28 to 60 ks by band to reach the dwarf galaxy regime), to determine the relative importance of different cluster assembly processes in driving the evolution of galaxies as a function of mass and environment. The depth of the VST-GAME survey is characterized by the completeness magnitude limit of 24.4 in the r band, as described in Section 2.2. A comparison of the depth of other VST surveys, such as the VST ATLAS survey with a median 5σ magnitude limit of 22.6 in the r band and the KIDS survey with a constant median 5σ magnitude limit of 25.2 in the r band, illustrates the depth of the VST-GAME survey. It should be noted that the median 5σ magnitude limit is deeper than the completeness limit.

The VST-GAME survey started in ESO period P99 and finished in P110. It is carried out using 300h of the Italian INAF Guaranteed Time Observations (GTO) with Omega-CAM.

Image reduction, alignment, and co-adding were made by the OATS team, using the standard procedure which includes overscan, bias, and flat corrections. After these steps, for each image, a weight map is created before an extraction step. The catalogs are then fed to Scamp (Bertin 2006) using GAIA DR2 as a reference. The astrometric solutions are then used with Swarp (Bertin et al. 2002) to create the final, global stacked images. A catalog of point sources is matched to PanStarr (Chambers et al. 2016; Flewelling et al. 2020) to create and then apply an illumination correction map.

2.1.3 VISTA Telescope and GCAV data

The Visible and Infrared Survey Telescope for Astronomy (VISTA) is a state-of-the-art ESO telescope located at the Paranal Observatory in Chile. It is designed to observe the sky in both visible and near-infrared wavelengths, and is equipped with a large aperture mirror that is 4.1 meters in diameter. The telescope is specifically designed to carry out large-scale surveys of the sky, and its unique combination of sensitivity and field of view make it ideal for studying a wide range of astronomical phenomena.

The VISTA Infrared Camera (VIRCAM) instrument is mounted on the VISTA telescope and is the main imaging instrument for the telescope. It is a 67-megapixel camera that operates in the near-infrared range, and is equipped with 16 individual detector arrays and available broad band filters at Z, Y, J, H, Ks and narrow band filters at 0.98, 0.99, and 1.18 micron. The instrument has a field of view of 1.65 square degrees, which is much larger than most other infrared cameras currently in use, and allows for a rapid survey of large areas of the sky.

The VISTA Public Survey program Galaxy Clusters At Vircam (G-CAV) is a second-generation ESO public VISTA survey (program 198.A-2008, P.I. M. Nonino) that started in P98 and was completed in P108. It observed 20 massive galaxy clusters covering ∼30 deg^2 in Y, J, and Ks bands.

Reduction of the images has been made by the OATS team (M. Nonino private comm). Raw images have been first corrected for nonlinearity, and then darks and flats have been applied. First-pass sky subtraction is then performed on an OB basis. For each image, a weight image, which also takes into account pixels flagged in darks and/or flats, is created.

Thus similar to optical data, the sources extracted are used to obtain an astrometric solution. A first stack, band per band, is then generated. *SExtractor* (Bertin & Arnouts 1996) is then run to create a segmentation map, which is then dilated and converted into a 0-objects/1-sky image. This global mask is mapped back to each chip of each reduced image, using the precise astrometric solution, to create an image per image mask, which is then used in the second run of sky estimation employing the classical sliding window approach. The resulting sky-subtracted images are then used to perform the final stacks under analysis.

All data are publicly accessible via the ESO Science Portal[1] or the Science Archive Programmatic and Tools Access[2].

2.1.4 MACSJ 0416

This thesis is focused on the galaxy cluster MACS J0416.1-2403 (hereafter M0416; Ebeling et al. 2001) as a first step towards studying the other clusters in the VST-GAME survey. M0416 was observed in P100 (0100.A-0570), P102 (0102.A-0603) and P104 (0104.A-0531) in four optical VST bands, i.e. u, g, r and i (see Table 2.2). By studying this particular cluster in detail, this thesis aims to gain a better understanding of the processes that drive the evolution of galaxies in a cluster environment, especially on the larger outskirts. The findings from this study will provide valuable insights that can be applied to the other clusters in the VST-GAME survey.

The galaxy cluster M0416, located in the constellation Eridanus, is one of the most studied galaxy clusters in the Universe. It is located at redshift $z = 0.397$ (Balestra et al. 2016), which corresponds to a distance of approximately 4.3 billion light-years away from Earth. This galaxy cluster has long been known to be extremely massive (M_{200}=0.88±0.13 $\times 10^{15} M_{\odot}$). It has also been observed to be extremely X-ray luminous ($L_X \sim 10^{45}$ erg s^{-1}). The galaxy cluster is named after the MACS (Massive Cluster Survey) project, which is a large-scale survey of galaxy clusters using the Hubble Space Telescope.

M0416 was first observed by HST as part of the CLASH survey (Postman et al. 2012). Then the cluster was observed as part of the Hubble Frontier Fields (HFF) initiative (Lotz et al. 2017), obtaining deep images (5σ point-source detection limit of ∼29 AB-mag). M0416 has been also observed in the second semester of 2022, by the James Webb Space Telescope (JWST) as part of the CAnadian NIRISS Unbiased Cluster Survey (CANUCS) (Willott et al. 2022).

M0416 was also observed as part of the ESO Large Programme "Dark Matter Mass Distributions of Hubble Treasury Clusters and the Foundations of ΛCDM Structure Formation Models" (CLASH-VLT; Rosati et al. 2014) with the VIsible Multi-Object Spectrograph (VIMOS) at the ESO/VLT. Using these data, Balestra et al. (2016) confirmed an overall complex dynamical state of this cluster. Early works identified M0416 as a merger, given its unrelaxed X-ray morphology and the separation (∼ 200 kpc in projection) of the

[1] http://archive.eso.org/scienceportal/home?data_collection=GCAV&publ_date=2020-12-07h

[2] http://archive.eso.org/programmatic/#TAP

two BCGs (Mann & Ebeling 2012). Balestra et al. (2016) showed the presence of two main subclusters and supported the hypothesis that they are being observed in a pre-collisional phase, in agreement with the findings from radio and deep X-ray data of Ogrean et al. (2015), which provides an opportunity to study the dynamics of galaxy cluster mergers and cluster formation. Finally, Olave-Rojas et al. (2018) using photometric and spectroscopic data, determined dynamically the presence of substructures up to $r<2r_{200}$ around M0416.

Observations of M0416 have revealed the presence of multiple images of background galaxies around the center of the cluster, as a result of gravitational lensing. Gravitational lensing occurs when the massive gravity of the galaxy cluster bends and amplifies the light from distant galaxies, creating a distorted image of the background universe, and revealing valuable information about the cluster itself. Using CLASH HST data, Zitrin et al. (2013) performed the first strong lensing analysis, discovering an elongated projected mass distribution in the cluster core, typical of merging clusters. Jauzac et al. (2015) detected two main central mass concentrations, and two possible secondary ones to the SW and NE, both at ~ 2 from the cluster center. This model was improved using HFF images and including the mass component associated with the hot gas (Bonamigo et al. 2017, 2018), and the kinematic measurements of a large sample of clusters galaxies (Bergamini et al. 2019). A state-of-the-art strong lensing model of M0416 is presented in Bergamini et al. (2022), obtained utilizing 237 spectroscopically confirmed multiple images, which is the largest sample of secure multiply lensed sources utilized to date, including also stellar kinematics information of 64 cluster galaxies and the hot-gas mass distribution of the cluster determined from Chandra X-ray observations (Bergamini et al. 2022; Mestric et al. 2022; Vanzella et al. 2019, 2021).

Bonamigo et al. (2017), from a combined analysis of X-ray and gravitational lensing, measured a projected gas-to-total mass fraction of approximately 10% at 350 kpc from the cluster center and showed that the dark matter over total mass fraction is almost constant, out to more than 350 kpc. Moreover, Annunziatella et al. (2017) showed that there is no significant offset between the cluster stellar and dark-matter components in the core of the cluster.

Following Balestra et al. (2016), the coordinates of the cluster center coincide with the NE-BCG at RA 04 : 16 : 09.14 and DEC -24 : 04 : 03.1. For the virial region, in this thesis, I will consider a region of radius $r_{200} = 1.82 \pm 0.11$ Mpc, as estimated from the weak lensing by Umetsu et al. (2014). The M0416 cluster core is shown in Fig. 2.2 using the HFF multiband photometry.

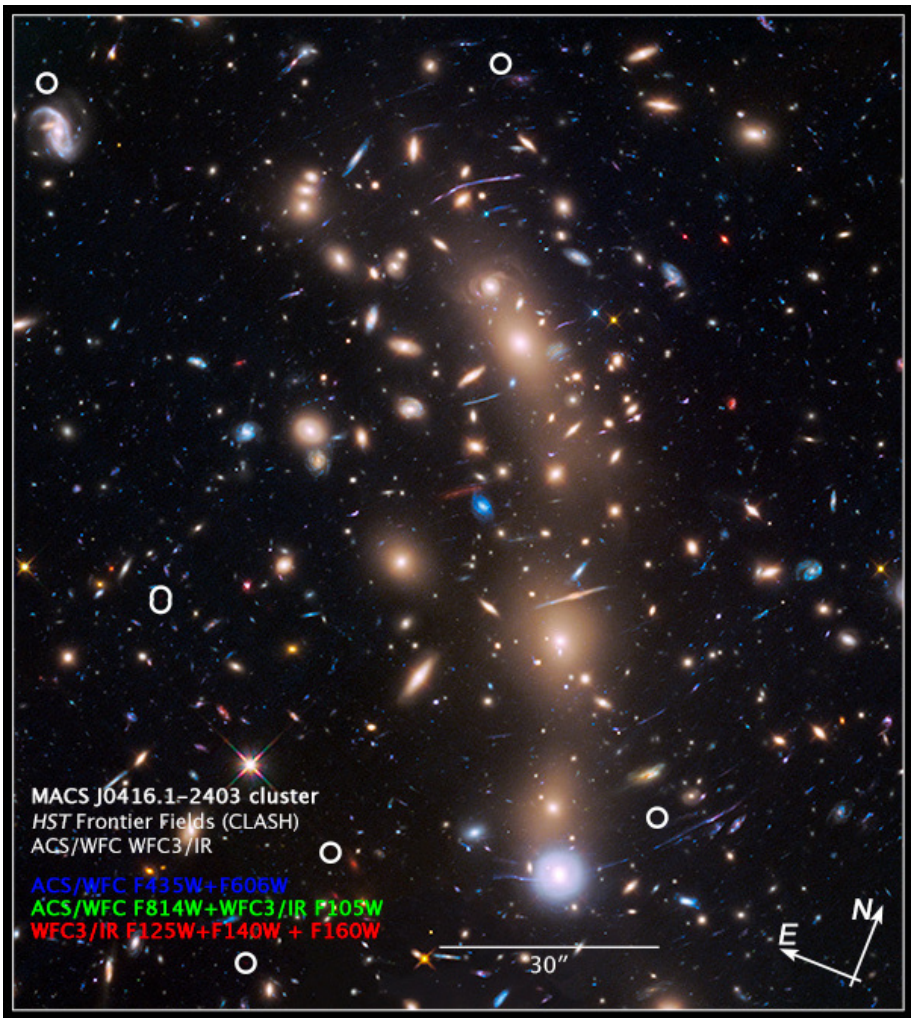


Figure 2.2: M0416 using HFF data. This image is a composite of separate exposures acquired by the WFC3/IR and ACS/WFC instruments (Lotz et al. 2017). Credits image: https://hubblesite.org [4]

2.2 VST-GAME: Photometry assessment and catalog construction

The photometry assessment described in this chapter led to the publication of a multi-band catalog for M0416, the first peer-reviewed publication of the VST-GAME program (Estrada et al. 2022b). The starting point is the reduced image of the cluster in the seven available bands, and the final multi-band catalog encloses the data synthetically and cleanly, adding information about the quality of the measurements and their distance.

Here I present a summary of the process followed to build the catalog; every step will be described in detail in this chapter. The first step on which I worked is source extraction, which was tuned for detecting both bright galaxies and also a large number of faint sources, paying particular attention to avoiding splitting large bright galaxies that show subclumps and deblending close sources in crowded fields. To have a reliable sample of galaxies, I implemented a robust separation between extended and point-like sources (hereafter the star/galaxy separation) up to the completeness limit magnitude. Then, I developed a procedure for regions with spurious photometry. Indeed, the internal optics of OmegaCAM at the VST often produces spikes, haloes and ghosts (that in the context of this thesis are called *spurious regions*) near bright stars. The photometry of sources detected inside these haloes could be affected by larger uncertainties, so I flagged these sources through the use of masks. Finally, using the magnitudes obtained in the previous step, I worked on the computation of photometric redshifts using Spectral Energy Distribution (SED) fitting techniques. The results of all those steps are contained in the multi-band catalog presented at the end of this chapter.

As a summary of the main part of the work carried out during this thesis, Fig. 2.3 shows the whole VST field centered on M0416. The image is the composition of the g, r, and i bands.

2.2.1 Source detection with SExtractor

The sources are extracted using the software *SExtractor* (Bertin & Arnouts 1996), which is a software tool that is used to detect and measure the properties, e.g. brightness, size, or shape, of astronomical objects in photometrical images.

To detect objects in an image, *SExtractor* first divides the image into a grid of pixels, and it identifies groups of connected pixels that have a brightness higher than a chosen average. These groups of pixels are then tested to see if they meet certain criteria, such as having a minimum size and a minimum contrast with the surrounding pixels, to determine if they are likely to be astronomical objects. If an object passes these tests, it is considered to be a candidate for further analysis.

Once candidate objects have been identified, *SExtractor* measures their properties using a variety of techniques, e.g., it can measure the flux, or brightness, of an object by summing the light from all of the pixels within its boundary. It can also measure the size, shape, and orientation of an object by fitting a model to the pixels within its boundary. In addition, *SExtractor* can measure the color of an object by comparing its brightness

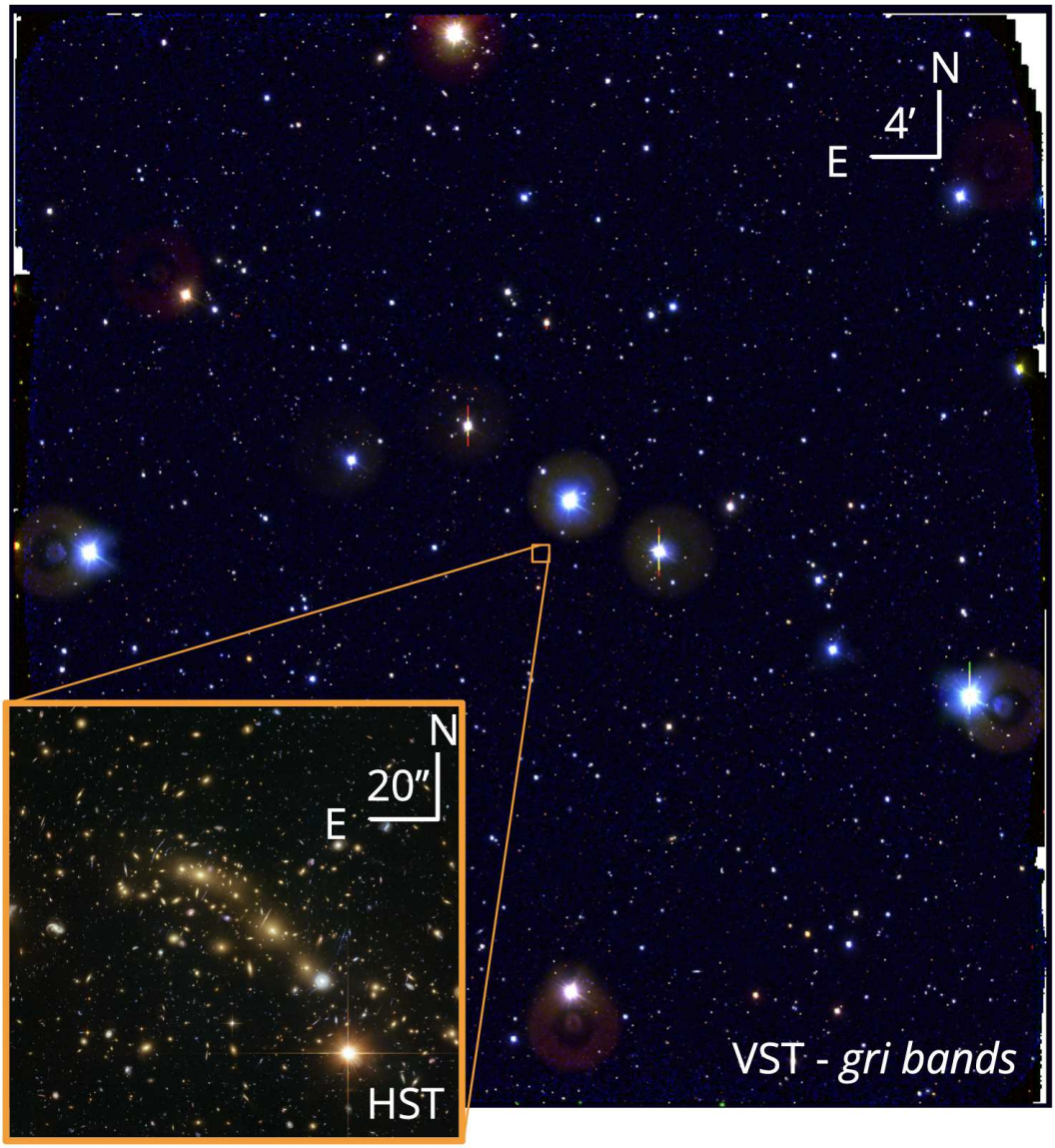


Figure 2.3: Complete view of the VST FoV for M0416 using g, r and i bands. The image size is 1 deg^2 and is centered on the cluster. The orange box shows as a comparison the HFF phtometry of the cluster, the same presented in Fig. 2.2

in different filters, and it can use this information to classify the object based on its type (e.g. star, galaxy, quasar).

Once *SExtractor* has detected and measured the properties of objects in an image, it outputs the results in a catalog file that can be used for further analysis. The catalog file contains information about the position, brightness, size, and other properties of each object, as well as an estimate of the uncertainty in these measurements. This information can be used to study the properties of the objects in the image and to compare them to other objects in the sky.

PSFEx (Point Spread Function Extractor Bertin 2011) is a software tool that is used to perform fitting photometry on astronomical images. The Point Spread Function (PSF) is a mathematical representation of the way that light from a point source is spread out, or "smeared", by the telescope and instrument as it passes through the atmosphere and optics. By modeling the PSF of an image, PSFex can determine the shape and size of the PSF, as well as its variation across the image.

Once the PSF has been modeled, *SExtractor* uses the *PSFex* models as input to carry out the PSF-corrected model fitting photometry for all sources in the image. This involves fitting a model of the PSF to the light from each object and using the fit to determine the object's brightness and other properties. The advantage of using PSF photometry is that *SExtractor* can take into account the distortions of the PSF measured by *PSFEx* along the field, which is particularly relevant in a large FoV. A specific validation of the PSF-corrected photometry is presented in Annunziatella et al. (2013).

To efficiently set the *SExtractor* parameters, it is necessary to keep in mind that the survey aims to trace a large population of galaxies (from the brightest elliptical galaxies to the dwarf regimes) in a wide range of cluster environments (from the crowded center to the outskirts). In this context, the source extraction should be tuned and optimized for detecting both bright galaxies and also a large number of faint sources but avoiding fake defections, in particular in the halos around bright stars, and also for deblending close sources in crowded fields, avoiding splitting large bright galaxies that show subclumps.

SExtractor can be used in *single-* or in *dual-mode*. In *single-mode*, both the detection and the flux measurements are computed over a single image. In dual mode, the detection is done in a reference image, while the photometry is measured in a second image. As a first step, I extracted the catalogs in each band (g, r, i, Y, J, Ks), except for the u band, using the *single-mode*. For the u band I extracted the sources directly in *dual-mode* using the r band, which has the best seeing, as detection image. This choice was required to improve both the extraction of galaxies characterized by the presence of sub-clumps, which often were split into sub-components due to the nature of UV star-forming emission and the detection of faint sources in this band. Finally, I used the r band as detection image, in *dual-mode*, to measure the photometry in g, i, Y, J, and Ks bands of sources detected in the r band up to the completeness limit, but not extracted in *single-mode* in the other bands, mainly because they are faint and, therefore, below the corresponding completeness limit (see below).

I adopted the same setup of the *SExtractor* parameters for VST images as described in Mercurio et al. (2015, 2021) for a similar dataset. I made several tests on *SExtrac-*

Table 2.1: Main input parameters into the *SExtractor* configuration file. This setup is established for the aims of our survey, i.e. to optimize the number of individual sources, especially faint, detected in all the photometric bands. `DETECT_MINAREA` and `BACK_SIZE` are expressed in pixels. The other quantities are dimensionless and defined on *SExtractor* documentation. I measured 9 `PHOT_APERTURES`, 7 are expressed in arcsec (first row) and 2 as a function of the measured FWHM (second row)

Parameters	Values
DETECT_MINAREA	5
DETECT_THRESH	1.5 σ
ANALYSIS_THRESH	1.5 σ
DEBLEND_NTHRESH	32
DEBLEND_MINCONT	0.001
BACK_SIZE	64
BACKFILTER_SIZE	3
PHOT_APERTURES	1.5, 2, 3, 4, 16, 30, 45
	3*FWHM, 8*FWHM

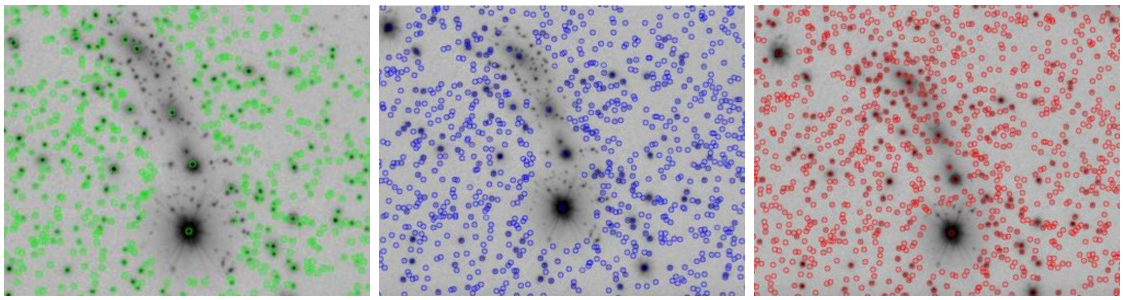


Figure 2.4: M0416 cluster core observed in the VST r band. In green (left) are shown the detections of *SExtractor* using a `BACK_SIZE` of 256 pixels, in blue (middle) using a `BACK_SIZE` of 128 pixels, and in red (right) are shown the detections of *SExtractor* using a `BACK_SIZE` of 64 pixels.

tor parameters to optimize the detections for our purposes, in particular, the `BACK_SIZE` was found to be directly linked with the total number of cluster members detected on each band. The `BACK_SIZE` parameter determines the mesh size employed to model the background around each image pixel. The `BACKFILTER_SIZE` determines the size of the median filter to suppress possible local overestimations due to bright stars. Dealing with a large field implies that there are several kinds of objects, from point sources to diffuse galaxies, so the background level is not uniform across the field. If `BACK_SIZE` is too small, the background estimation is affected by the presence of objects and random noise. Most importantly, part of the flux of the most extended objects can be absorbed in the background map. If the mesh size is too large, it cannot reproduce the small-scale variations of the background. Some tests done with several values of `BACK_SIZE` are shown in Fig. 2.4. As a result of those tests, I fixed the `BACK_SIZE` to 64 pixels and `BACKFILTER_SIZE` to 3 to detect the larger number of sources.

The threshold parameters (`DETECT_THRESHOLD` and `ANALYSIS_THRESHOLD`) are chosen to maximize the number of detected sources, while simultaneously keeping the number of spurious detection to a minimum. The value for the threshold is chosen according to the procedure detailed in Fig. 4 of Mercurio et al. (2015). A summary of the key *SExtractor* parameters and the aperture diameters used in this work is given in Table 2.1.

The *SExtractor* configuration described above is therefore optimized for this thesis and can be considered as the reference configuration. In addition, I run again *SExtractor* in single mode on the r band using a higher value for the threshold parameters, to recover objects in the most crowded regions of the field. This so-called 'hot-mode' configuration in r band was used as a reference for a second source extraction in the dual mode in all bands. There is a larger probability of fake detections on this 'hot-mode' configuration. From the sources identified in this 'hot mode', only those previously identified spectroscopically and missing in the reference photometric extraction were added to the final catalog. The number of objects recovered with this test is nearly 200, few in comparison with the whole catalog ($\sim$72000) but significant in comparison with the spectroscopic sample ($\sim$4000).

Completeness limit

The completeness limit refers to the faintest magnitude at which a source can be detected and measured with a high level of confidence. It is determined by the sensitivity of the instrument and the level of background noise present in the observations. A photometric survey with a high completeness limit can detect and measure faint sources, while a survey with a low completeness limit may only be able to detect and measure the brightest sources.

The completeness magnitude limit of the extracted catalogs is determined by studying the dispersion of source magnitudes within apertures of 1.5 and 8 arcsecond. A cartesian plot is used to visualize the relationship between these magnitudes, where the vertical axis of the plot represents the magnitude within the 8-arcsecond diameter aperture, while the horizontal axis represents the magnitude within the 1.5-arcsecond diameter aperture. A linear regression was performed on the data, and the one-sigma dispersion was measured. The point at which the distribution of the data began to deviate significantly from this linear relation was then identified, and a vertical line is potted at that detection limit.

Table 2.2: VST-GAME and VISTA-GCAV photometry

Band	Exp. Time (ks)	FWHMs (arcsec)	Completeness (mag)
u	38.5	1.01	23.6
g	60.0	0.92	24.8
r	40.5	0.67	24.4
i	38.0	0.62	23.6
Y	28.8	0.77	24.4
J	23.0	0.81	22.9
Ks	25.2	0.77	22.2

The completeness limit was defined then as the horizontal line that intersects the vertical line and the one sigma level of the fitted linear regression (Garilli et al. 1999; Mercurio et al. 2015). The procedure to determine the completeness level is shown in Fig. 2.5 for the VST optical bands and Fig. 2.6 for the VISTA NIR bands. In the last column of Table 2.2 the completeness limiting magnitude for each band is reported, using the single-mode extractions. It is computed as a single-mode extraction for the u band just for computing its completeness limit. In Fig. 2.7 is shown the magnitude distribution of the sources detected in *single-mode* in each band.

Considering sources within the magnitude limit proper of each band, the band catalogs include 14591 detections in the u band, 78871 in the g band, 84964 in the r band, 64547 in the i band, 103862 in the Y band, 109724 in the J band, and 98275 in the Ks band. The larger number of sources in the NIR bands is motivated by the larger Field of View of the VISTA telescope.

A summary of total exposure times, FWHMs of the images, and completeness limits analyzed in this thesis are reported in Tab. 2.2.

2.2.2 Star/Galaxy separation

The procedure described in the previous section returns a catalog of sources for each band. *SExtractor* is an optimal tool for the detection of luminous sources, nevertheless, given the depth of our photometry, there is the risk of misidentifying small and far-away galaxies, so it is convenient to calibrate a more accurate star/galaxy separation process.

Several methods can be used to distinguish galaxies from stars using only deep photometric data. One approach is to compare the size and shape of the objects. Stars are point-like in appearance with angular sizes of a few arcseconds or less, while galaxies typically exhibit extended shapes and sizes ranging from a few arcseconds to several arcminutes, depending on their type and distance. By analyzing the relation between the luminosity profile (half-light radius) and the apparent magnitude of the sources, a stellar

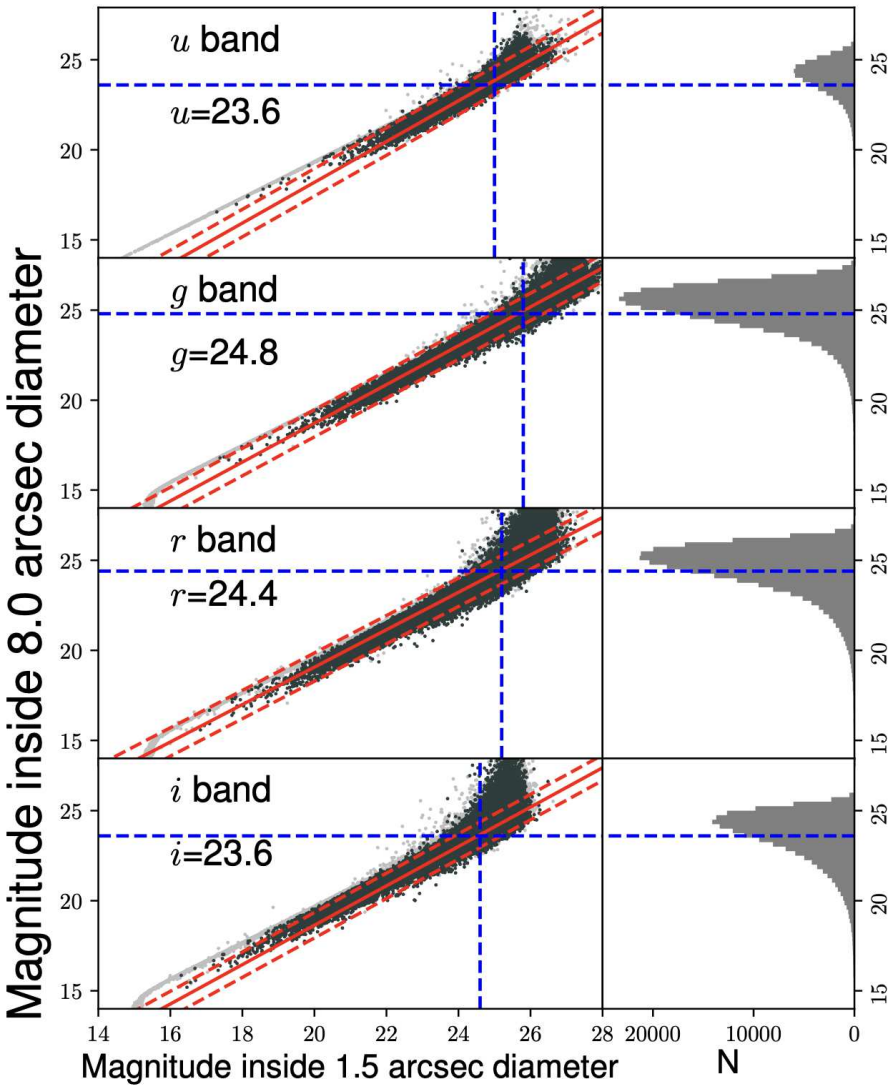


Figure 2.5: Distribution of the *SExtractor* magnitude inside an 8.0 arcsec diameter as a function of the magnitude inside a detection aperture of 1.5 arcsec diameter for u, g, r and i bands. The horizontal and vertical blue dashed lines indicate completeness and the detection limits, respectively. The red continuous lines are the linear relation between the magnitude within the 8.0 arcsec aperture diameter and the magnitude within the detection aperture, minus/plus 1σ (red dashed lines). Image credits: A. Mercurio.

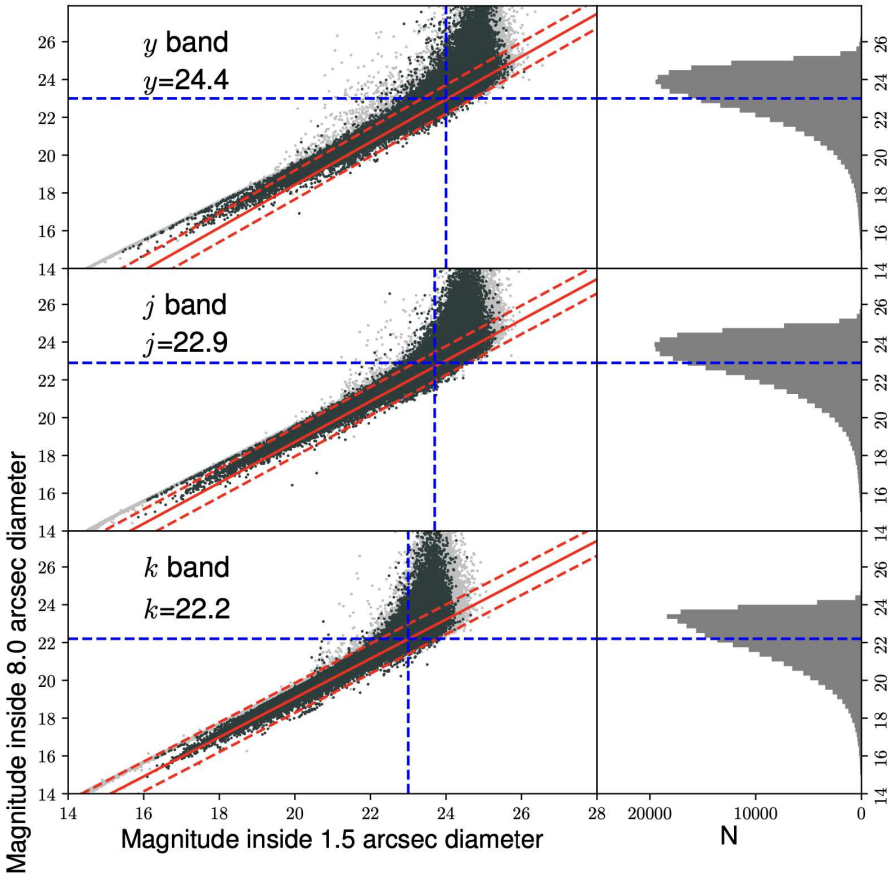


Figure 2.6: Distribution of the *SExtractor* magnitude inside an 8.0 arcsec diameter as a function of the magnitude inside a detection aperture of 1.5 arcsec diameter for Y, J, and Ks bands. The horizontal and vertical blue dashed lines indicate completeness and the detection limits, respectively. The red continuous lines are the linear relation between the magnitude within the 8.0 arcsec aperture diameter and the magnitude within the detection aperture, minus/plus 1σ (red dashed lines). Image credits: A. Mercurio

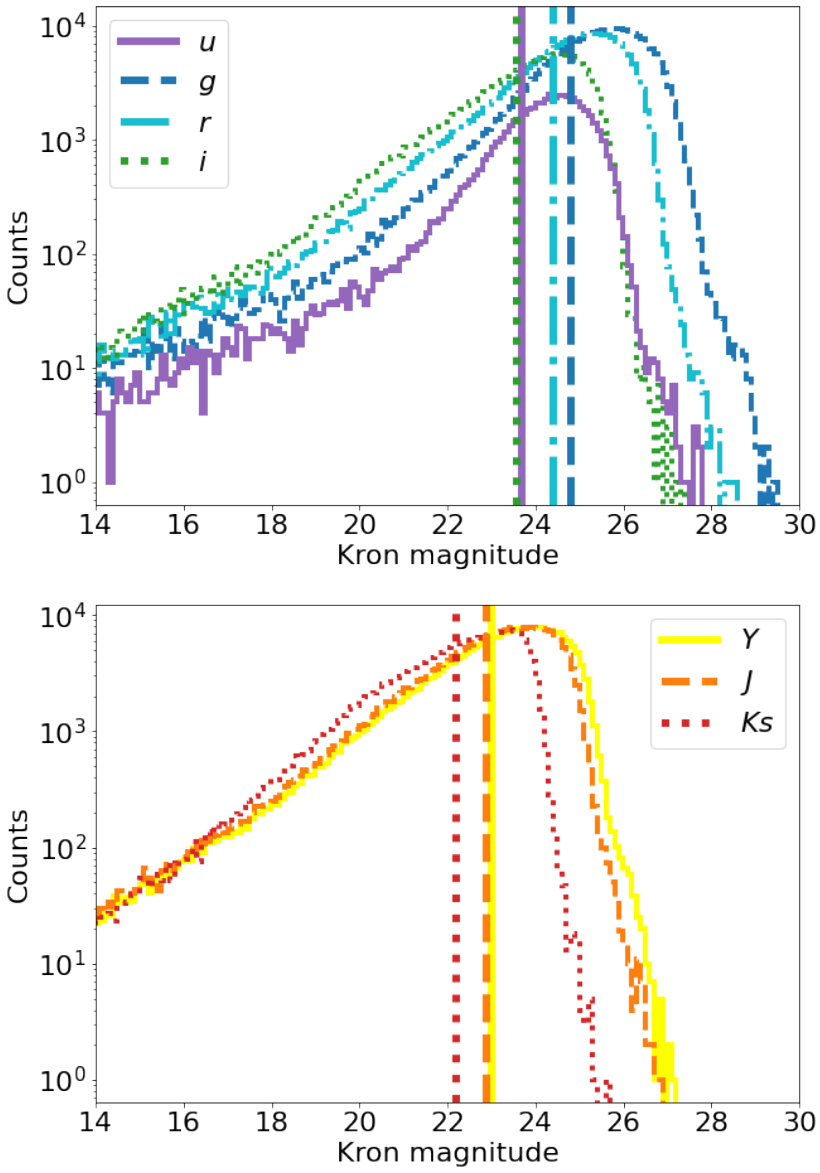


Figure 2.7: Magnitude distribution of sources extracted in each band in the *SExtractor* single-mode. The vertical lines represent the completeness limits for each band (See Table 2.2).

locus and a galaxy locus can be identified.

Another way to distinguish between stars and galaxies is to examine their spectral energy distributions (SED) or colors. A stellar SED is based on the spectrum of a single star, which is determined by its physical properties such as its surface temperature, gravity, chemical composition, and interstellar extinction. A galaxy SED is typically more complex, with contributions from many different stellar populations (the number and types of stars in the galaxy), star formation history (how fast and when stars were formed in the galaxy), dust and gas content, and non-thermal processes (such as synchrotron radiation from relativistic electrons in magnetic fields). Color-color or color-magnitude diagrams can be used to separate stars from galaxies based on their position in the color space.

Separation through luminosity profiles

To separate point-like and extended sources I adopt a procedure similar to that described in Annunziatella et al. (2013), using the following parameters: (i) the stellarity index; (ii) the half-light radius; (iii) the new *SExtractor* classifier `SPREAD_MODEL`; and (iv) the peak of the surface brightness above the background (μmax). The procedure is done for each band separately.

Preliminary, the catalogs are cleaned, removing spurious detections with incomplete isophotal values or corrupted memory overflow that occurred during deblending or extraction. I also remove detections with the `FLUX_RADIUS_50` below 1 pixel, which are likely warm pixels or residuals from the cosmic ray rejection. Each step of the classification aims at going deeper in magnitude using one *SExtractor* parameter and follows the order of the four panels in Fig. 2.8, where black points refer to all sources, orange points are detections brighter than the saturation limit $r = 16.1$ mag (see below), and the gray shaded region indicates those that are outside the completeness limit of the band:

1. The stellarity index (`CLASS_STAR`): *SExtractor* identifies objects with `CLASS_STAR`=0 as galaxies and those with `CLASS_STAR`=1 as stars, so traditionally, the star galaxy separation has been done exclusively using this parameter. However, the top-left panel of Fig. 2.8 shows the presence of multiple sequences in the `CLASS_STAR` - Kron magnitude (Kron 1980) plane for r band sources, suggesting that a unique value of the `CLASS_STAR` parameter is not enough in our case. Taking advantage of the spectroscopic data, I found that sources with `CLASS_STAR` > 0.8 are both stars and galaxies. Then I decide to consider as stars sources under the saturation limit $r < 16.1$ mag with a stellarity index greater than 0.5 and sources from all the magnitudes, $r \geq 16.1$ mag with a stellarity index greater or equal to 0.99 which are surely stars. I use other *SExtractor* parameters to further separate stars and galaxies.

2. The half light radius (`FLUX_RADIUS_50`): this is the radius in pixels containing half of the galaxy's light and can be used as a direct measure of source concentration. The top-right panel shows the `FLUX_RADIUS_50` as a function of the Kron magnitude for r band sources. Since I assume that stars are point-like sources, the stellar

locus is therefore the sequence visible in the left and lower part of the plot. This sequence remains distinguishable from the extended sources only down to approximately r∼22.5 mag. Considering that the 95% completeness limit of the r band is 24.4, I have to use another parameter to separate stars from galaxies at magnitudes fainter than 22.5.

3. The spread classifier (`SPREAD_MODEL`): this parameter considers the difference between the model of the source and the model of the local point-spread function obtained with *PSFex* (Desai et al. 2012). By construction, it is close to zero for point sources and positive for extended sources. In the bottom-left panel, the sequence of stars is visible, having `SPREAD_MODEL`∼0 as a function of the Kron magnitude. This parameter allows us to extend the classification to fainter magnitudes up to the completeness limit r=24.4 mag.

4. The peak of the surface brightness above the background (`MU_MAX`): this value represents the peak in the surface brightness given in magnitudes per square arcsecond. This parameter is used as a test for the star/galaxy separation computed in previous steps, and for the magnitude to separate saturated from non-saturated stars. In the bottom-right panel `MU_MAX` is shown as a function of the Kron magnitude. The horizontal loci of saturated stars (orange) is visible, having an almost constant value of `MU_MAX` down to $r = 16.1$ mag. The sequence of stars previously identified is shown in red and galaxies are shown in blue. The fact that stars identified in the previous steps follow a clear pattern in this panel validates our procedure.

Testing the star/galaxy separation with colors

As a further validation of the star/galaxy classification, I plot the optical vs. NIR color-color diagram. According to Baldry et al. (2010) and Jarvis et al. (2013), the locus of stars is defined by a sequence in the region with $J - Ks < 0.2$, and the locus of galaxies in the upper left cloud of the diagram. As shown in fig. 2.9, I found a very good agreement between the defined star/galaxy loci and the star/galaxy separation in the r band up to its completeness limit. I also found that 14.8 % of the sources classified as stars (`NSFLAG_r`>0) have $J - Ks > 0.2$, which is in agreement with the spectroscopic sample. 17.3% of the spectroscopically confirmed stars are located in the region with $J - Ks > 0.2$. This also suggests that the optical-NIR color-color plot is a useful tool to separate galaxies from stars, but with residual contamination of stars of $\sim 15\%$.

The star/galaxy separation procedure identifies 805 saturated stars, 7510 nonsaturated stars, and 66323 galaxies within the completeness limit of the r band.

2.2.3 Detecting and masking regions with spurious photometry

Spurious photometry refers to the detection and measurement of false or misleading light intensity values in an image. Those regions are generated by the combination of the optical

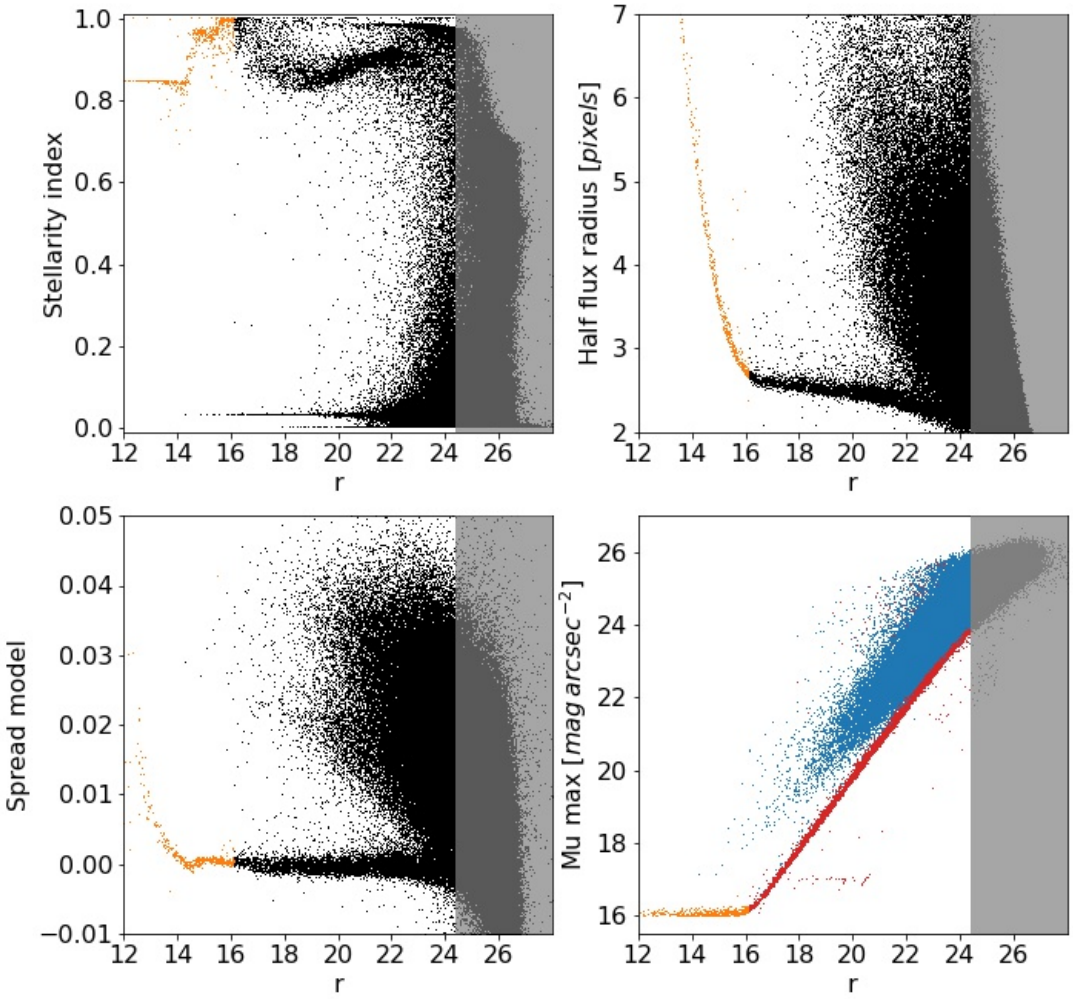


Figure 2.8: Star/galaxy separation. Diagram of the *SExtractor* `CLASS_STAR` (*Top-left panel*), `HALF_FLUX_RADIUS_50` (*Top-right panel*), and `SPREAD_MODEL` (*Bottom-left panel*), as a function of Kron magnitude for sources detected in the r band (black points), with the loci of saturated stars colored in orange. The gray shaded area corresponds to the area beyond the completeness limit of the r band. In the *bottom-right panel* I plot the *SExtractor* `MU_MAX` parameter as a function of the Kron magnitude for sources detected in the r band, color-coded according to their classification: galaxies in blue, saturated stars in orange, unsaturated stars in red.

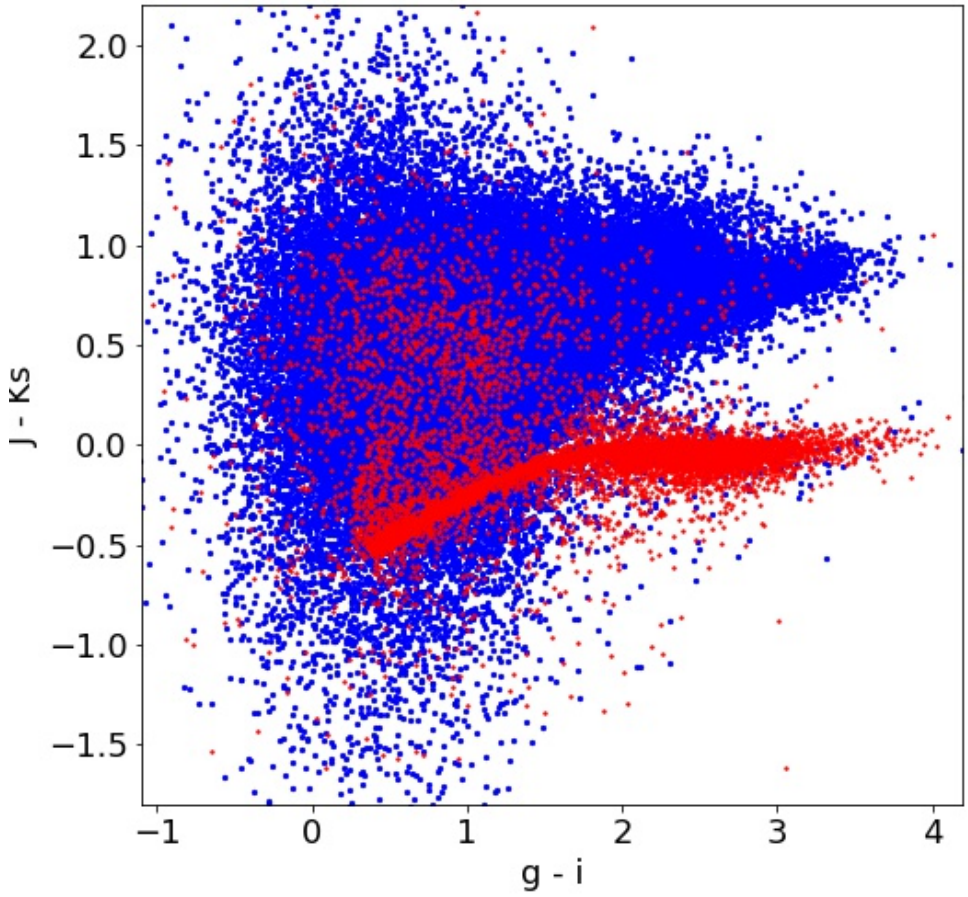


Figure 2.9: $J - Ks$ vs $g - i$ color-color diagram, using aperture magnitudes within 3 arcsec diameter to visualise the star/galaxy separation. Here stars are in red and galaxies in blue. To assess the classification precision, the dataset was bifurcated into two distinct groups, one with and the other without spectroscopy. The quantification of impurity was performed by calculating the proportion of stars (red points) within the $J - Ks > 0.2$ sequence of galaxies (blue region). The contamination of stars on the galaxy region is therefore 14.8 % for the sample without spectroscopy and 17.3 % for the spectroscopic one.

design of the instrument used to capture the image and the presence of bright objects in the field of view producing false signals.

The presence of bright stars in the observed images can produce complex regions with enhanced surface brightness due to multiple reflections of the light in the internal optics of OmegaCAM. A visual inspection of the sources lying in and around these regions suggests that these sources might have their flux measurements affected by these irregular patterns, it is, therefore, necessary to properly identify them.

There are mainly three kinds of impurities due to bright stars inside the field: spikes, halos, and ghosts, an example of which is shown in Fig. 2.10. Spikes are narrow bright lines coming out from the star in the vertical direction (see the green rectangle in Fig. 2.10). Halos are circular regions with higher luminosity centered on the star (see the red circle in Fig. 2.10). Ghosts are composed of a central region with low surface brightness surrounded by an outer corona with magnified surface brightness characterized by an irregular pattern. These features are not centered on the star that generated them, but outside in a radial direction with respect to the center of the OmegaCAM field (see the orange circle in Fig. 2.10).

I developed a tool to mask haloes and ghosts using the position and magnitude of the stars that generate them. Spikes are almost always contained inside haloes and involve no fake sources, so there is no need to mask them. I start with the visual inspection of the r band image and verify that indeed all the haloes and ghosts are generated by saturated stars. I compare with the GAIA catalog (Gaia Collaboration et al. 2018) and find that stars with a G magnitude brighter than 12 generate the strongest haloes and ghosts. Stars with $12 < G < 13$ still generate small spurious regions with a low noise level and affect only a few detected objects. Using the DS9 regions tool, I create circular masks around stars with $G \lesssim 13$, where the photometry is affected by haloes and ghosts. Based on the size and distribution of some of these visually selected masks, shown in the left panel of Fig. 2.11, I develop a general procedure to automatically generate circular masks around bright stars.

Haloes are always centered on the star that generates them, and their radius, r_{halo} in pixels, depends linearly on the mag of the star: $r_{halo} = b_1 \times mag + b_2$, where $b_1 = -150.29$ and $b_2 = 1821.34$, according to the handmade haloes drawn for the r band. This calibration has been made individually for all of the available VST and VISTA bands. Table 2.3 presents for all the bands the coefficients to determine the halo radius (b_1 and b_2).

The location of the ghosts depends on the position of the parent star in the OmegaCAM or VIRCAM field. They are aligned in the radial direction, towards the external regions, from the center of the field with the generating star. The distance from the star to the center of the ghost (d_{SG} in pixels) depends linearly on the distance between the center of the field and the star (d_{CS}): $d_{SG} = c_1 \times d_{CS} + c_2$, where $d_{center-star}$ is measured in pixels, and the coefficients for the r band are $c_1 = 0.07716$ and $c_2 = 38.16$. The radius of the ghost depends on the GAIA G magnitude of the star: 1000 pixels for $G \leq 8.7$, 900 pixels for $8.7 < G \leq 9.7$, 850 pixels for $9.7 < G \leq 10.2$, and 200 pixels for $10.2 < G \leq 13$. Following this procedure, I find that for stars near the center of the field, haloes and ghosts are usually overlapped, while for stars in the outskirts, the two spurious regions are separate.

The size of haloes and ghosts depends on the band, being larger in the r, g, and i bands. The u band presents the smallest spurious regions of the VST OmegaCAM fields. On the other hand, VISTA fields with VIRCAM present a similar behavior but the sizes of haloes and ghosts are smaller because near infrared light is less affected by scattering and diffraction than visible light. Near infrared light has longer wavelengths than visible light, which makes it less prone to scattering and diffraction by dust particles, air molecules, and optical elements. Table 2.4 presents for all the bands the coefficients to determine the ghosts position (c_1 and c_2) and the ghost size as a function of the magnitude of the star in the GAIA G band.

The final mask for the r band is shown in the right panel of Fig. 2.11. Additionally, Fig. 2.12 and Fig. 2.13 show the mask for the VST and VISTA bands. The mask covers 5.29%, 12.29%, 12.93% and 9.85% of the 1 deg^2 VST FoV in the u, g, r and i bands respectively. The mask covers 2.20%, 2.84% and 1.73% of the 1.5 deg^2 VISTA FoV in the Y, J and Ks bands, respectively. Regardless of the independent masking of every single band, in the matched catalog I consider the r band to be the reference masking image for all the objects in the field.

Additionally, due to the *SExtractor* setup adopted for this work, I identified a subset of sources that exhibit anomalous patterns corresponding to the spurious regions within the mask, e.g. a sequence of sources arranged in concentric circles with the halo or ghost. These sources are likely false positives and constitute about 0.4% of the single-band catalogs within the completeness limit. I manually eliminated these sources from the single-band catalogs. In contrast, the multi-band catalog (see Sec. 2.2.5) does not suffer from this problem because false positives are unlikely to occur simultaneously across seven bands due to the different geometries of the spurious regions in each band. Therefore, false positives are effectively filtered out by the band matching procedure.

After the masking procedure on the multi-band sample (see Sec. 2.2.5), I find that 7169 sources (9.6%) are inside the mask (`MASK_FLAG`=1), while 66945 sources (90.4%) are outside the mask (`MASK_FLAG`=0).

The total area covered by the masking procedure for the VST bands is 5.29 % for the u-band, 12.29 % for the g-band, 12.93 % for the r-band, and 9.85 % for the i-band. For the VISTA photometry, the area covered by the mask is 2.20 % on the Y-band, 2.84 % on the J-band, and 1.73 % on the Ks-band.

The masking procedure that I developed for this thesis could be applied to new surveys done with VST and VISTA telescopes, as shown by Lorenzon et al. (in prep.), where the 30 deg^2 of the Southern H-Atlas Regions Ks-band (SHARKS) ESO survey have been masked with the relations described above.

Table 2.3: Coefficients b_1 and b_2 for all bands to determine the radius of the halo, in pixels, generated by stars brighter than 13 in the GAIA G band

Band	b_1	b_2
u	-108.78	1338.92
g	-130.76	1573.16
r	-150.29	1821.34
i	-116.58	1389.50
Y	-50.59	694.38
J	-71.82	920.87
Ks	-48.12	648.73

Table 2.4: Coefficients c_1 and c_2 to determine the distance in pixels from the generating star to the ghost in all bands. The radius of the ghost is related to the GAIA G band magnitude in the intervals: $G \leq 8.7$, $8.7 < G \leq 9.7$, $9.7 < G \leq 10.2$, and $10.2 < G \leq 13$

Band	c_1	c_2	ghost radius
u	0.09295	-89.92	900, 400, 400, 200
g	0.07729	29.27	1000, 900, 850, 200
r	0.07716	38.16	1000, 900, 850, 200
i	0.07145	79.50	950, 900, 850, 200
Y	0.01651	38.70	400, 400, 400, 300
J	0.00459	14.88	600, 500, 400, 300
Ks	0.00470	-2.57	400, 300, 300, 300

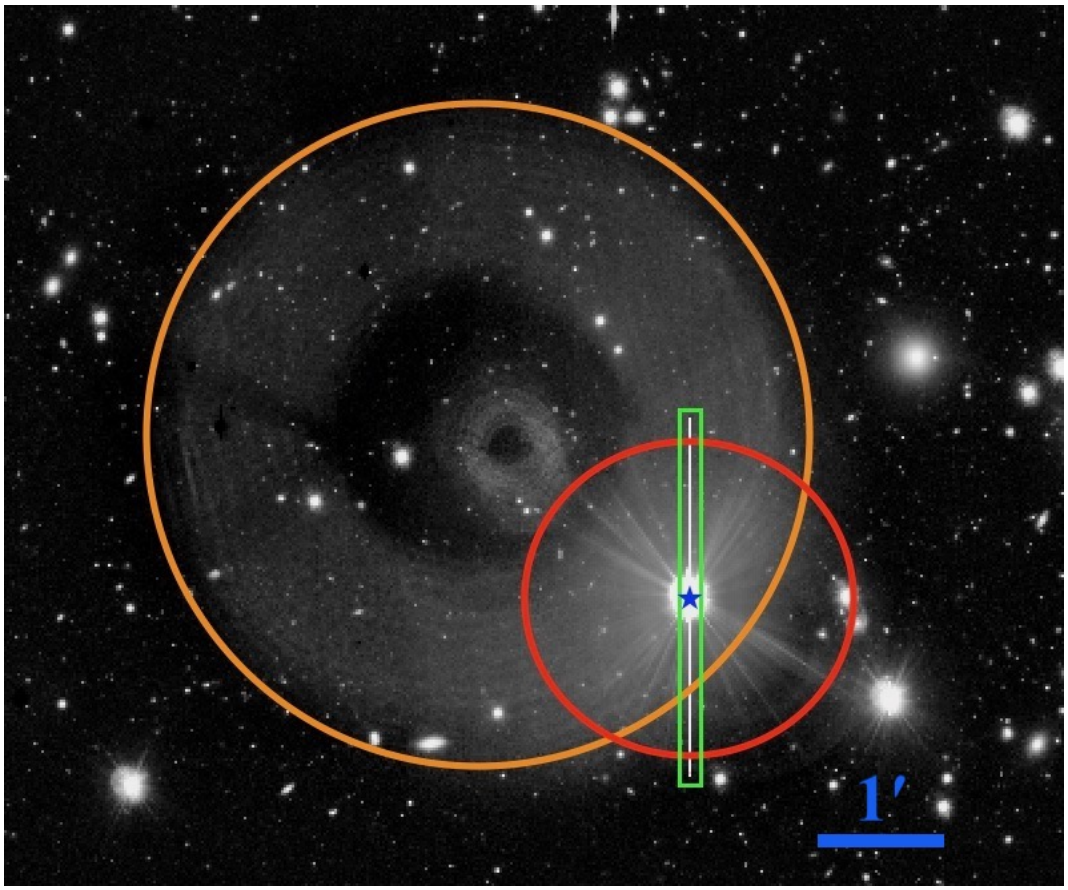


Figure 2.10: Example of the three kinds of regions with enhanced surface brightness generated by a star (blue) in the r band image: the vertical spike is shown in green, the halo centered on the star in red, and the ghost aligned in the radial direction with respect to the center of the OmegaCAM field in orange.

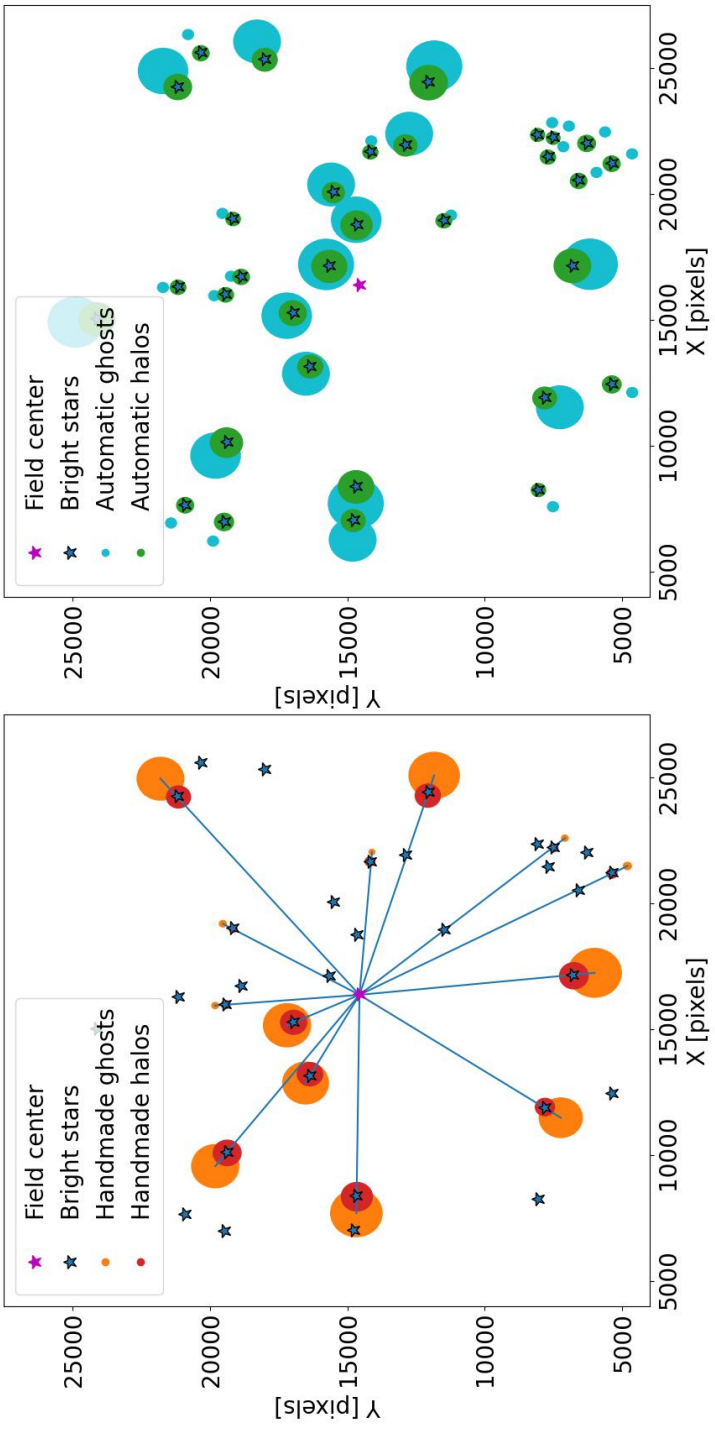


Figure 2.11: *Left Panel*: star haloes and ghosts manually drawn around bright stars in the r band. The magenta star is the central pixel of the OmegaCAM field, the blue stars in the plot are all the stars brighter than 13 in the GAIA G band, the red circles are the handmade halos centered into the corresponding star, and the orange circles are the handmade ghosts placed in the radial direction with respect to the field center. Handmade haloes and ghosts are not drawn for all the bright stars. Blue lines connect the center of the ghosts with the center of the field. Using the sample of handmade haloes and ghosts (position, diameter, and magnitude of the star) I build a mask for all the bright stars. *Right Panel*: mask for all the stars brighter than 13 in the GAIA G band. The procedure to generate the mask is calibrated from the regions visually identified in the left panel and applied to all the bright stars in the field.

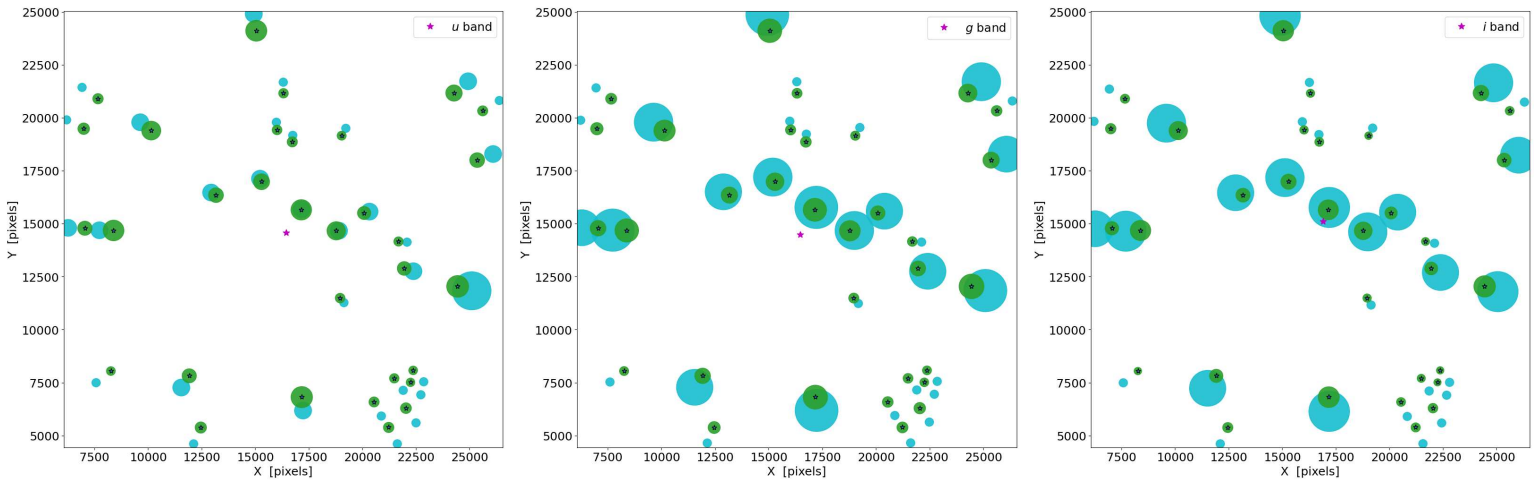


Figure 2.12: Mask generated for the u, g and i VST bands. The fraction of the area covered by those masks is 5.29 %, 12.29 %, and 9.85 % respectively. Color code follows right panel of Fig. 2.11.

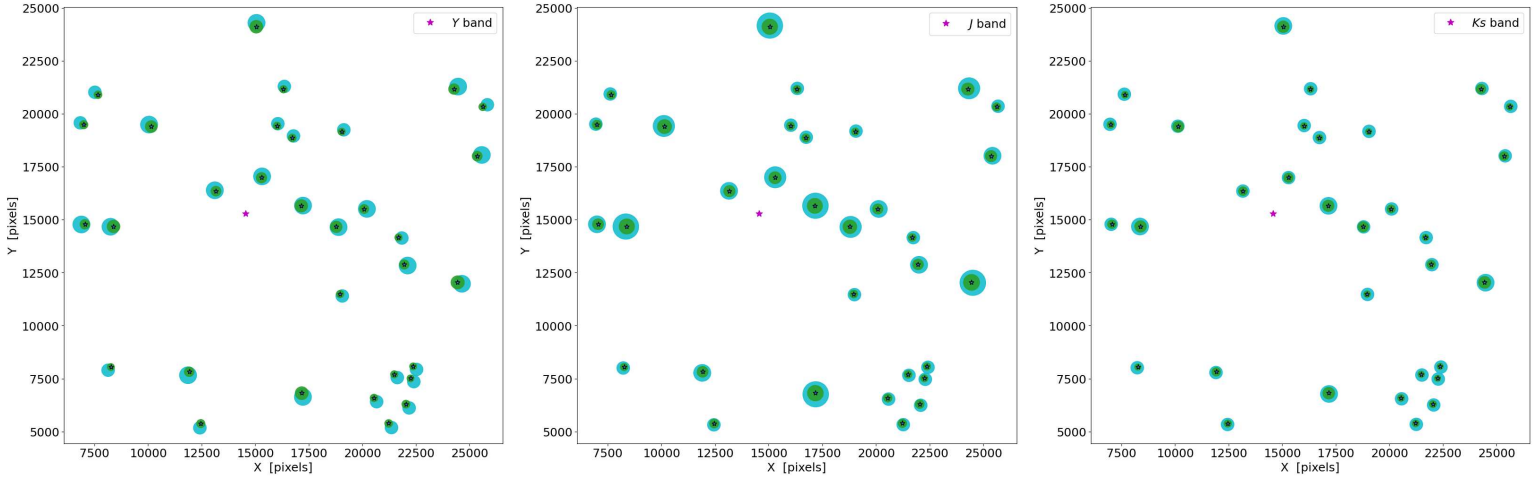


Figure 2.13: Mask generated for the Y, J and Ks VISTA bands. The fraction of the area covered by those masks is 2.20 %, 2.84 %, and 1.73 % respectively. Color code follows right panel of Fig. 2.11

2.2.4 Redshift computation

Due to the expansion of the Universe, the light emitted by a galaxy experiments a shift toward the red wavelengths. This redshift can be used to calculate the distance to the galaxy, based on the known properties of the Universe.

The redshift of a galaxy can be measured mainly through spectroscopic or photometric techniques.

Spectroscopic redshift, is a method for determining the redshift of a galaxy based on the spectrum of its light. It is possible to determine the redshift by looking for specific features on the galaxy spectrum, such as absorption or emission lines of certain elements or molecules. These features are shifted to longer wavelengths by the expansion of the universe, and the amount of shift can be used to calculate the redshift. Spectroscopic redshift is the most precise method to compute the distance to a galaxy, but it requires spectroscopic observations, which are more time-consuming and expensive to obtain.

Photometric redshift, on the other hand, is a method for estimating the redshift of a galaxy based on the comparison of the measured flux on several photometric bands to theoretical models or templates that predict the expected luminosity at different redshifts. This method is less expensive from a point of view of observational time, but it is less precise than spectroscopic redshift and can have larger uncertainties in the resulting redshift estimates.

The quality and depth of the VST and VISTA photometry, allow us to assign accurate photometric redshifts to the sources detected in the VST field centered on the M0416 cluster. In this section, I describe the iterative procedure adopted to tune the photometric redshift determination.

To calibrate the photometric redshifts, I take advantage of the spectroscopic data from the CLASH-VLT survey (Rosati et al. 2014), which collected 4386 source spectra in a field of view (FoV) of $\sim$25$\times$25 arcmin^2 around the cluster center up to $\sim 2r_{200}$, leading to the spectroscopic confirmation of $\sim$800 cluster members (Balestra et al. 2016). In this work, galaxies falling within the spectroscopic redshift range of $0.397 \leq z \leq 0.410$ are classified as cluster members.

Spectral Energy Distribution

Photometric redshifts are determined through a Spectral Energy Distribution (SED) fitting approach, using the code *LePhare* (Arnouts et al. 1999; Ilbert et al. 2006).

A spectral energy distribution (SED) is the function of the energy emitted by an object as a function of different wavelengths. The SED fitting consists in fitting a model to the observed flux of an object to infer either its redshift or its physical properties, e.g. mass, age, star formation rate, and metallicity. For each galaxy, there will be as many points to fit as the number of observed bands, so in general having more bands to fit, especially covering a larger wavelength range, is the key factor to improve the performance of the fitting procedure. In Fig. 2.14 are represented the transmission coefficients that are given to *LePhare* to perform the fitting. In my fitting procedure, I employ the COSMOS spectral energy distribution templates as introduced by Ilbert et al. (2009). These templates are

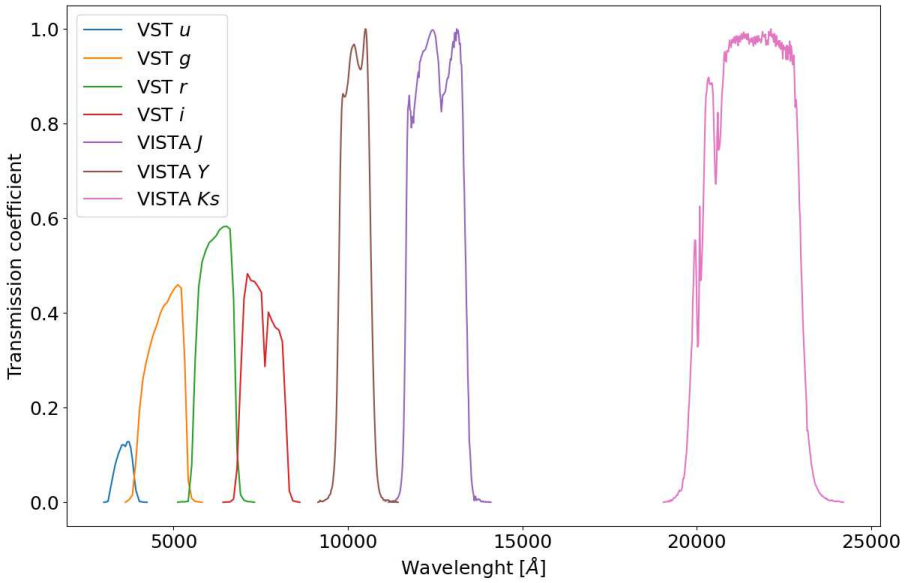


Figure 2.14: Set of filters of the VST OmegaCAM and VISTA VIRCAM used to compute the photometric redshifts.

derived from the works of Bruzual & Charlot (2003) and Polletta et al. (2007).

Spectroscopic sample to calibrate photo-z

To calibrate the photometric measurements and quantify their accuracy, I benefit from the rich spectroscopic sample available for M0416 (Balestra et al. 2016; Caminha et al. 2017; Vanzella et al. 2019), covering the area at r$<$2 r_{200}. I match the spectroscopic sample with the multi-band catalog, identifying 4123 objects with both VST-GAME photometry and CLASH-VLT spectroscopy. I calculate zero point corrections for each photometric band using the `AUTO_ADAPT` parameter in the *LePhare* algorithm. This parameter minimizes the fitting procedure by using spectroscopic redshifts as a reference and estimates the required zero point corrections. The `CONTEXT` parameter, an integer value utilized to determine the filter combination for fitting each source, is calculated by taking into account the number of available bands for each source. This parameter is employed as input for *LePhare* to exclude any erroneous values that may be present for a given band. This is motivated by the nearly 1% of the objects in the final catalog with a missing aperture magnitude mainly in the u and Ks bands.

The spatial distribution of spectroscopic redshifts in comparison with our photometric field, selecting the cluster redshift interval, is shown in Fig. 2.15.

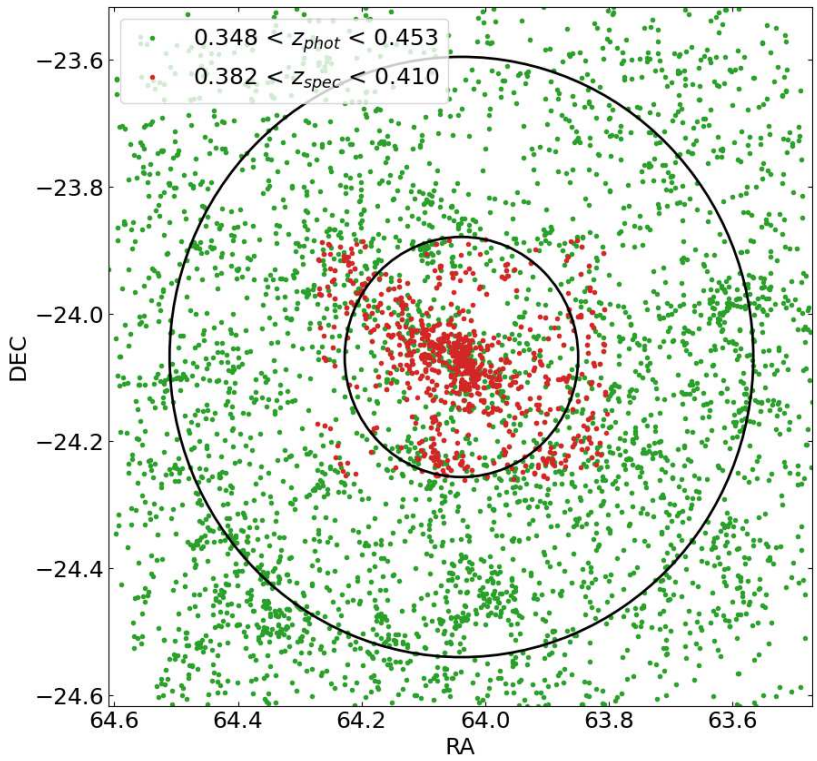


Figure 2.15: Distribution across the VST field of spectroscopic (red) and photometric (green) cluster members. The inner circle indicates $2r_{200} = 3.64\,Mpc$ and the outer circle indicates $5r_{200} = 9.1\,Mpc$.

Photo-z statistical parameters

To quantify the goodness of the photometric redshifts, I follow Euclid Collaboration et al. (2020) to define the quantity:

$$\Delta z = (z_{phot} - z_{spec})/(1 + z_{spec}). \tag{2.1}$$

Objects with $|\Delta z| > 0.15$ are regarded as outliers, and their abundance indicates the accuracy of the fitting. After excluding the outliers, I compute the following photo-z statistical estimators (Razim et al. 2021; Hong et al. 2022):

- the mean of $|\Delta z|$, *bias*;
- the standard deviation of $|\Delta z|$, σ;
- the standard deviation of the normalized median absolute deviation, $\sigma_{NMAD} = 1.4826 \times median(|\Delta z|)$;
- the fraction of outliers, η.

Optimization of photometric redshifts on M0416

Optimizing the accuracy of photometric redshifts has been a challenging task throughout this thesis. While not all of the tests can be described in detail due to space constraints, this section outlines the process of experimentation and improvement that led to the final result.

In the first iteration, I analyze the whole multi-band catalog by utilizing all available bands (u, g, r, i, Y, J, Ks), considering all the sources within the completeness limit $r < 24.4$, and evaluate the performance of the fitting procedure using various magnitudes as inputs. Since I am using PSF fitting photometry, the aperture magnitudes already take into account the differences in seeing among the bands, so the fraction of the source's flux is the same for a given aperture in all the bands.

The initial series of tests were made using aperture magnitudes and the errors on magnitudes as given by *SExtractor*. The tests were computed by tuning *LePhare* parameters. One of the best results obtained with this configuration is shown in the top panel of Fig. 2.16, where is reported the distribution of photometric versus spectroscopic redshifts, and can be described by the following statistics: $bias = 0.0446$, $\sigma = 0.0357$, $\sigma_{NMAD} = 0.0516$, $\eta = 17.07\%$.

A subsequent series of tests were made with three aperture magnitudes inside 1.5, 4.0, and 8.0 arcsec diameter. I find that the most accurate results are obtained using the smallest aperture magnitude in our catalog, 1.5 arcsec. This fact confirms the results of Jouvel et al. (2014), who found that the optimal *LePhare*'s performance is obtained with small aperture magnitudes. A run of *LePhare* computed with the whole multi-band catalog, using the 1.5 arcsec aperture magnitude is shown in the bottom-left panel of Fig. 2.16 and can be described by the following statistics: $bias = 0.0380$, $\sigma = 0.0318$, $\sigma_{NMAD} = 0.0431$, $\eta = 8.88\%$.

Table 2.5: Coefficients taken from Mercurio et al. (2015) to compute the multiplicative factor $R = a_1 \times exp[a_2 \times (mag - 20)] + a_3$ to increase nominal errors given by *SExtractor*. The i band coefficients are used for the NIR bands.

Band	a_1	a_2	a_3
u	1.26	-0.985	1.353
g	18.55	-0.762	1.144
r	9.39	-0.839	1.397
i	3.67	-0.858	1.467

A possible way to improve photometric redshifts is to adopt a larger error on photometric magnitudes than the one given by *SExtractor*. So, I multiply the nominal error given by *SExtractor* by the factors given in Table 6 of Mercurio et al. (2015): $R = a_1 \times \exp[a_2 \times (\mathrm{mag} - 20)] + a_3$, where mag is the magnitude of the source and the a_i coefficients are calibrated for each band. For the VISTA magnitudes, I use the same coefficients as for the i band. After increasing errors on magnitudes, I obtain an improvement in the accuracy of photo-z, as shown in the bottom right panel of Fig. 2.16. On this run, it was used the whole photometric sample and the 1.5 arcsec aperture magnitude can be described with $bias = 0.0358$, $\sigma = 0.0296$, $\sigma_{NMAD} = 0.0418$ and $\eta = 6.06\%$.

The final step of the optimization process is to focus on the fitting procedure and prioritize the analysis of galaxies within the cluster's redshift range. So, I only included galaxies with `NSFLAG_r`=0 in the analysis. Additionally, I required that spectroscopically confirmed galaxies must have redshifts within the range $0 < z_{spec} < 1$, to avoid issues with the under-sampled galaxies at high redshift. These selection criteria allowed us to effectively analyze cluster member galaxies and mitigate the impact of potential biases in the data. With these constraints, I obtain a final photometric sample of 66211 galaxies and a spectroscopic sample of 3840 galaxies. The final SED fitting accuracy is shown in Fig. 2.17, where I compare the photo-z measurements with spectroscopic redshifts. The statistics of the run are $bias = 0.0361$, $\sigma = 0.0298$, $\sigma_{NMAD} = 0.0411$, $\eta = 3.75\%$, confirming that those are the best photo-zs for the sample. The results of this run are those present in the released catalog.

One last attempt to improve the *LePhare* performance was using the (`Mag_Model`), which models the flux of each source according to a bulk/disk model and accounts additionally for PSF effects. A run using the sample with increased flux errors and focused on the $0 < z < 1$ range returns even less accurate results, with $bias = 0.0478$, $\sigma = 0.0368$, $\sigma_{NMAD} = 0.0569$, $\eta = 16.72\%$, so this was not an improvement at all. The reason for the bias is that `Mag_Model` adjusts the magnitude of an object that has already undergone PSF correction.

Finally, I also check the impact of a bad signal-to-noise (S/N) ratio on the NIR bands. By excluding NIR sources with a S/N of less than 5 in either the Y, J, or Ks band, a refinement in the accuracy of photometric redshifts was achieved, reaching a $bias = 0.0354$, $\sigma = 0.0293$, $\sigma_{NMAD} = 0.0406$ and $\eta = 3.16\%$. In the following analysis, I use only galaxies

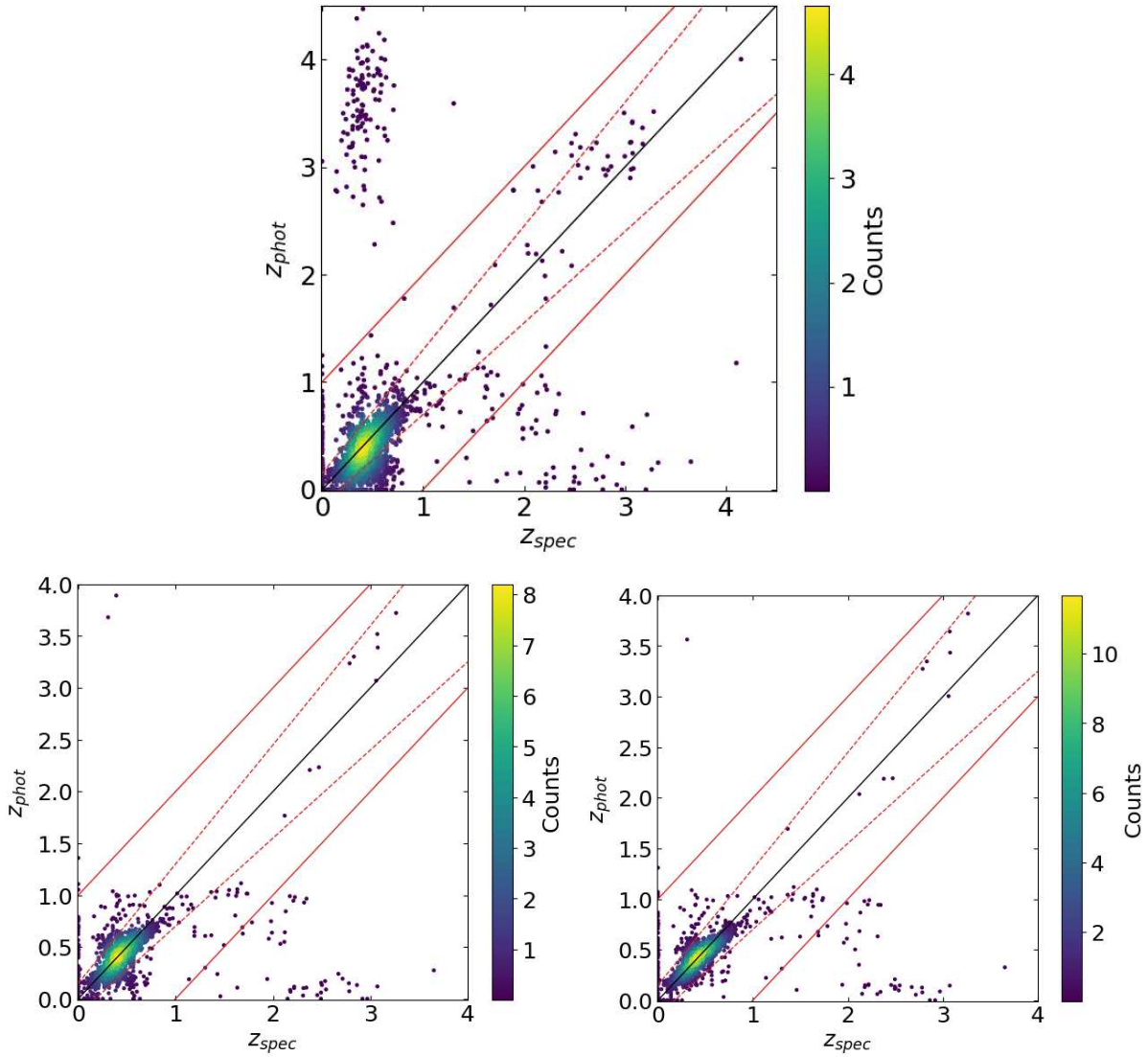


Figure 2.16: Photometric redshift as a function of spectroscopic redshift for some preliminary test, before optimizing for the $0 < z < 1$ interval. The internal red dotted lines indicate outliers, $|\Delta z| > 0.15$, and red continuos lines catastrophic outliers, $|z_{phot} - z_{spec}| > 1$. ***Top center***: Photo-z computes using aperture magnitude and before increasing the *SExtractor* errors. The statistics of this run are: $bias = 0.0446$, $\sigma = 0.0357$, $\sigma_{NMAD} = 0.0516$, $\eta = 17.07\%$. ***Bottom left***: Photo-z computed using the aperture magnitude in 1.5 arcsec and before increasing the *SExtractor* errors. The statistics of this run are: $bias = 0.0380$, $\sigma = 0.0318$, $\sigma_{NMAD} = 0.0431$, $\eta = 8.88\%$. ***Bottom right***: Photo-z computed using the aperture magnitude in 1.5 arcsec and having increased the *SExtractor* errors. The statistics of this run are: $bias = 0.0358$, $\sigma = 0.0296$, $\sigma_{NMAD} = 0.0418$, $\eta = 6.06\%$

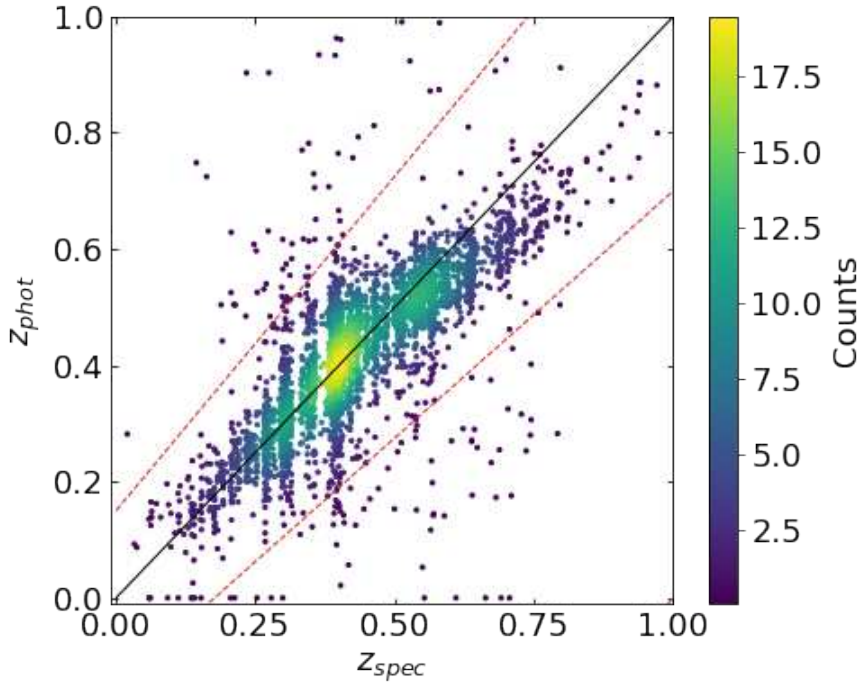


Figure 2.17: Photometric redshift as a function of the spectroscopic redshift optimized at the cluster interval. This SED fitting is calibrated with a sample of galaxies which excludes stars (`NSFLAG_r`>0 or $z_{spec} = 0$) and galaxies with $z_{spec} > 1$. The fraction of outliers is 3.75% and the σ_{NMAD} is 0.0411.

with $S/N \geq 5$ in the NIR bands (N_{gal} =49779) to enhance the robustness of the scientific results. However, in the multiband catalog (Estrada et al. 2022b), I compute photometric redshifts for all the galaxies, without considering any cut in signal-to-noise.

Analysis of the photo-z final sample

Fig. 2.19 shows the histogram of the whole sample of photometric redshift found in this thesis (purple) in comparison with the whole sample of spectroscopic redshifts (yellow). This histogram evidences the numerical abundance of photometric measurements in comparison with spectroscopy. A particular feature in the photometric histogram is the presence of two peaks redshift values different than the cluster core. Those peaks indicate a larger number of sources, but not an excess of overdense structures. In Fig. 2.20 is shown a test done using the density field (as will be described in the next chapter), to verify that the abundance of galaxies with photometric redshift z$\sim$0.6 or z$\sim$1 does not comport an increment of dense structures across the field.

Photo-z using machine learning

In recent years, the use of machine learning (ML) techniques for estimating photometric redshifts has become increasingly popular, due to the ability of these methods to accurately predict redshifts for large, complex datasets. The computation of photometric redshifts using machine learning techniques involves training a model on a dataset of galaxies with known spectroscopic redshifts. The model is then used to predict the redshift of a galaxy based on its observed magnitudes. The model is then trained on this dataset using a process known as cross-validation, in which the data is divided into a training set and a validation set. The model is trained on the training set and then tested on the validation set to assess its performance. The performance of the model can be improved by fine-tuning the parameters of the algorithm or by adding additional features to the input dataset (e.g., galaxy shape or size). Several different machine learning algorithms can be used for this purpose, including artificial neural networks, decision trees, and support vector machines (e.g., Wang et al. 2008; Way et al. 2009; Carrasco Kind & Brunner 2013; Brescia & Cavuoti 2014; Hoyle 2016.

Several studies have demonstrated the effectiveness of machine learning for photometric redshift estimation (e.g., Budavari 2009; Hoyle 2016; Carrasco Kind & Brunner 2014. However, it is important to carefully select and pre-process the training data, as well as to evaluate the performance of the model on independent test samples, to obtain reliable results.

During the process of optimization of the photometric redshifts, I tested PhotoRApToR tool, presented by Brescia & Cavuoti (2014) to obtain photometric redshifts for the whole M0416 sample. PhotoRApToR is composed of a Multi-Layer Perceptron (MLP; Rosenblatt 1961) neural network, which is among the most used feed-forward neural networks in a large variety of scientific and social contexts. The MLP is trained by a learning rule based on the Quasi-Newton Algorithm (QNA), which uses a variable metric method for finding

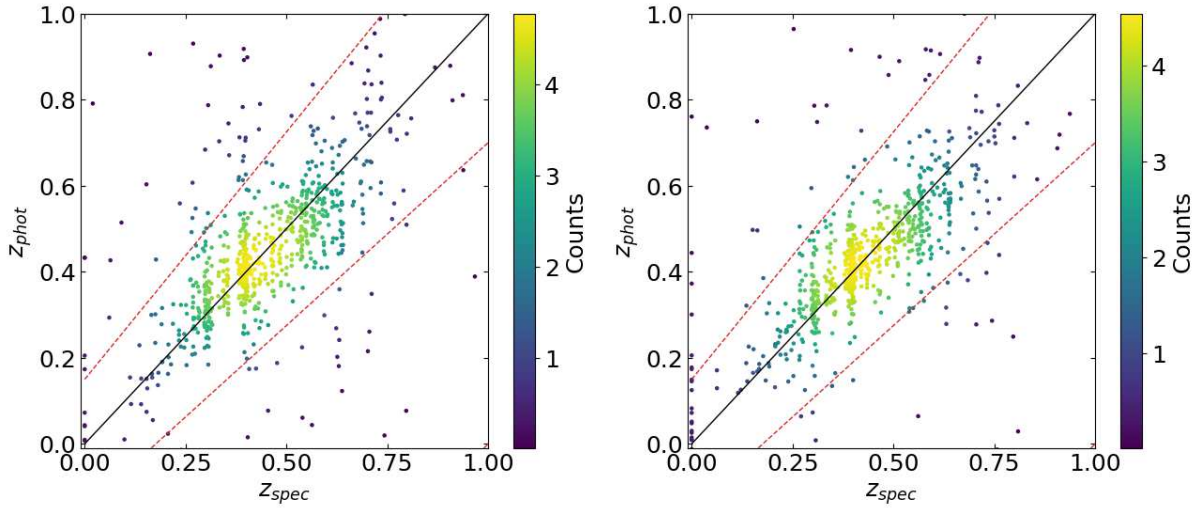


Figure 2.18: Results of the photo-z's computed with the MLPQNA algorithm (Cavuoti et al. 2015). With respect to previous plots, here I used only 20% of all the objects with spectroscopic redshift. *Left*: Results using Kron magnitude. $bias = 0.0424$, $\sigma = 0.0344$, $\sigma_{NMAD} = 0.0513$, $\eta = 16.00\%$ *Right*: Results using aperture magnitude. $bias = 0.0377$, $\sigma = 0.0337$, $\sigma_{NMAD} = 0.0415$, $\eta = 14.17\%$.

local maxima and minima of functions (Davidon 1991). The complete architecture of the MLPQNA is described on Cavuoti et al. (2015).

To validate the model, I randomly split the sample of $\sim$ 4000 spectroscopic galaxies into a training set (80%) and a test set (20%). This ensures the independence of the results. However, the resulting photometric redshift estimates were not of sufficient quality. This can be motivated by the small and biased nature of the training sample, only 3200 galaxies. This sample did not adequately cover the full range of parameter space, i.e. a large fraction of the training is composed of cluster members at z$\sim$0.4, leading to a bias towards the cluster redshift. Figure 2.18 presents the results of the two most successful tests using the ML approach among all the tests performed, using kron and 1.5 arcsec aperture magnitudes as input. According to the statistics shown in the figure caption, the performance of the MLPQNA algorithm is discretely accurate, but not sufficient to compare with the results obtained through SED fitting techniques.

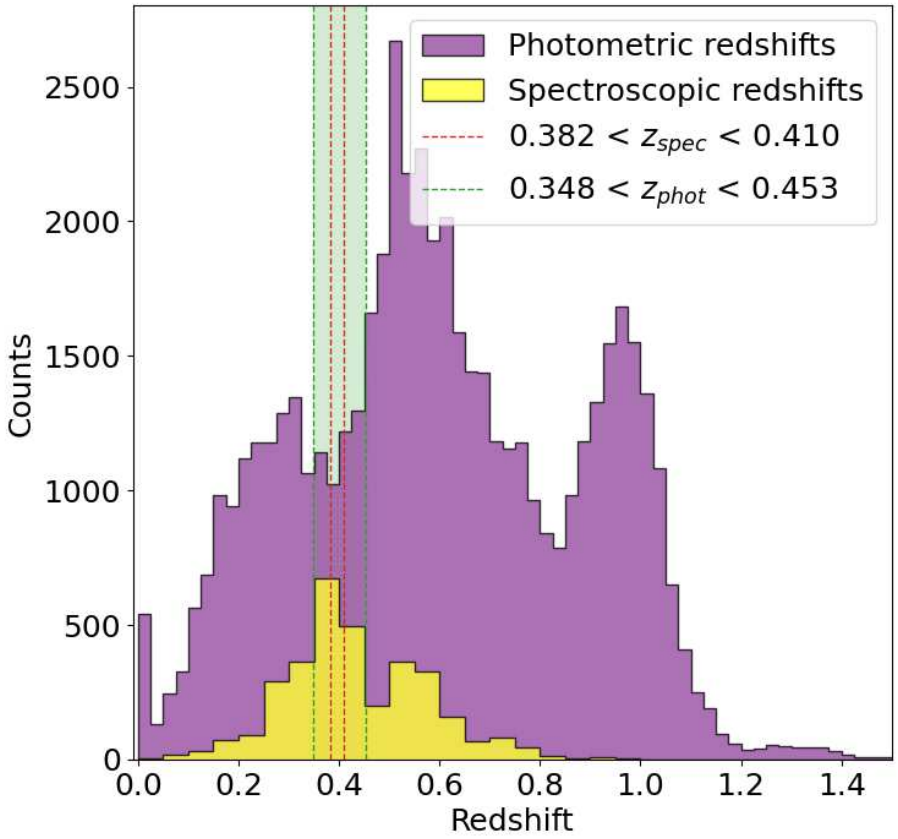
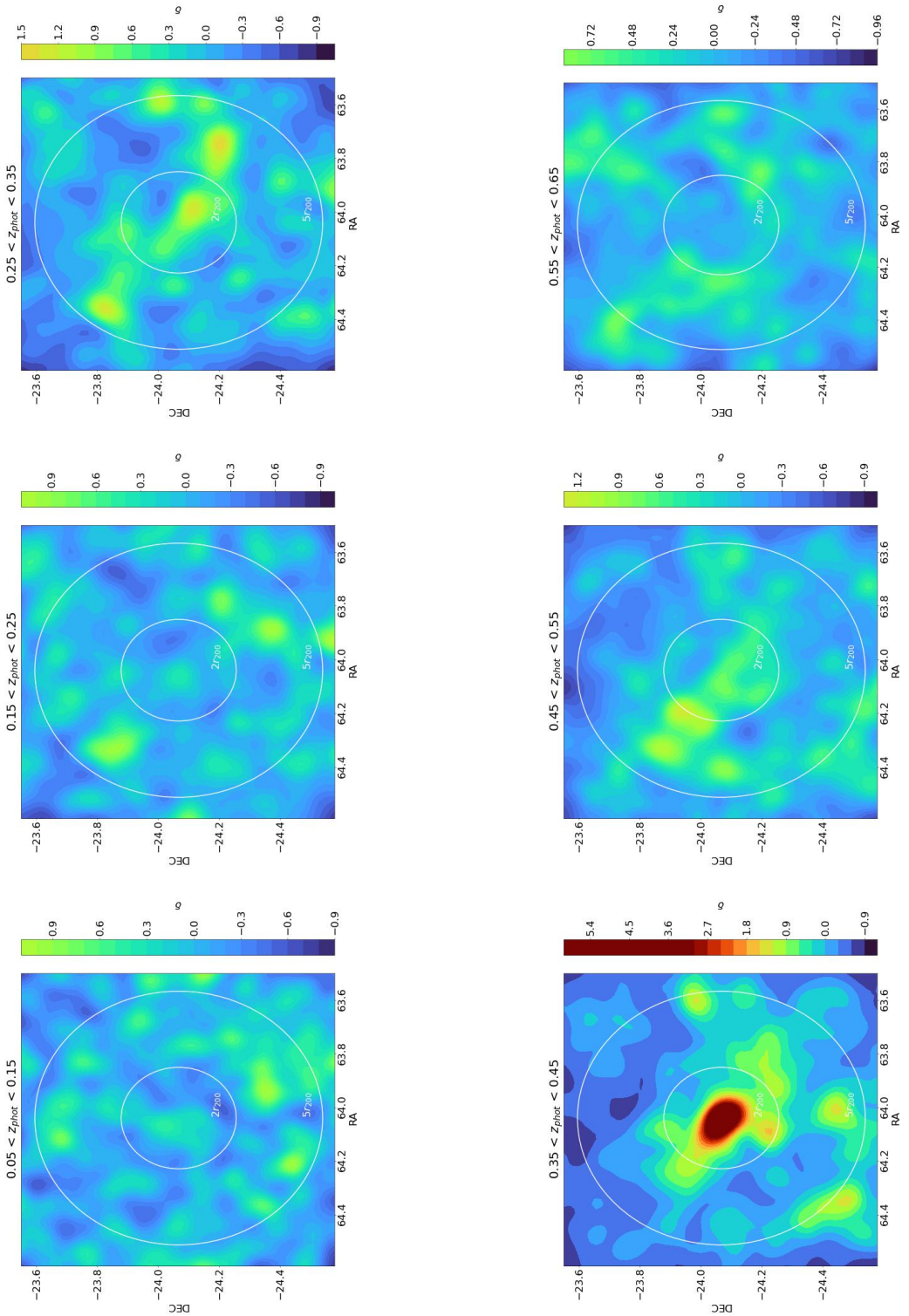


Figure 2.19: Histogram with all the redshifts in the $0 < z < 1.5$ interval. In magenta are shown objects with photometric redshifts across the whole field, and in yellow objects with spectroscopy ($r < 2R_{200}$). The green area indicates the photometric cluster interval while the red lines indicate the spectroscopic cluster interval.


0.05 < z_{phot} < 0.15
0.15 < z_{phot} < 0.25
0.25 < z_{phot} < 0.35
0.35 < z_{phot} < 0.45
0.45 < z_{phot} < 0.55
0.55 < z_{phot} < 0.65
$2r_{200}$
$5r_{200}$
RA
DEC
δ

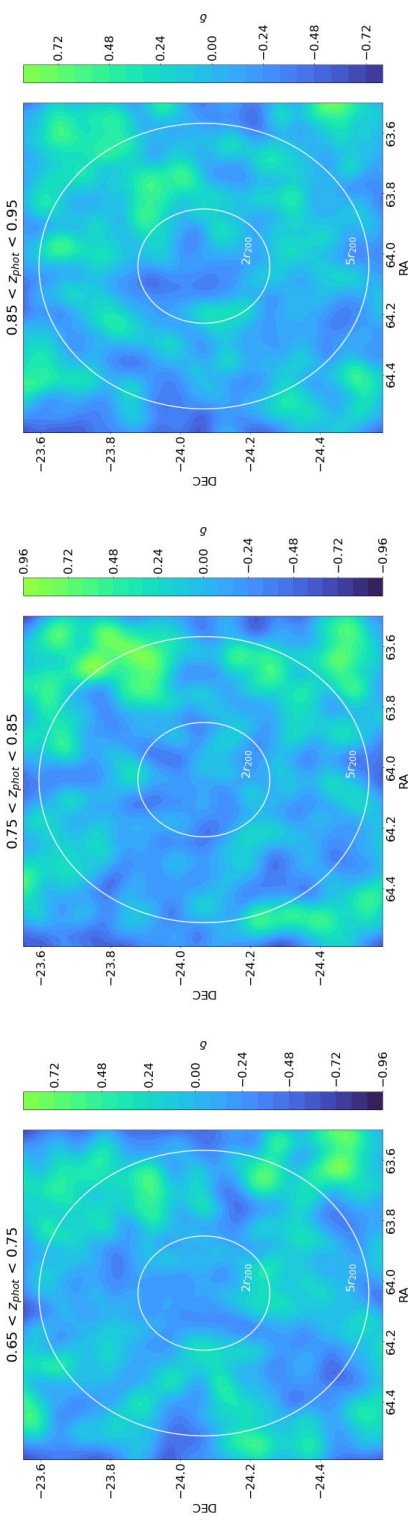


Figure 2.20: Density field over the VST field of view, computed as described in section 3.3. This series of plots scans the whole redshift interval $0.5 < z < 0.95$ in slices of z∼0.1. this figure is proof to demonstrate that despite the abundance of photometric redshifts at z∼0.1 or z∼1 (see Fig 3.1), the only considerable overdensity coincides with the cluster redshift.

2.2.5 Band merging and final setup of the multi-band catalog

In Estrada et al. (2022b), we publish the multi-band catalog for the cluster M0416 using VST and VISTA photometry. This catalog is obtained by matching all the sources in the r band within the completeness limit $r < 24.4$, using the Starlink Tables Infrastructure Library Tool Set `STILTS` (Taylor 2006). Detections in the u, g, i, Y, J and Ks bands within a 1 arcsec distance from r band sources are identified as a match. Magnitudes in the $ugiYJKs$ bands can be fainter than their corresponding completeness limit since I do not cut these catalogs before the match. In the catalog, I report a unique primary key (`GAME_ID`), which univocally identifies sources. It is composed of a string containing 19 characters, where four of them are "GAME" and fourteen for the digits hhmmsss-ddmmsss barycentre coordinate, that I report in degrees (`RA`, `DEC`). I also report the parameters of the ellipse that describes the shape of the objects: the semi-major and semi-minor axes (`A` and `B`), and position angle (`THETA`), together with the half flux radius (`R_50`) and the Kron radius (`R_Kron`) measured in the r band. Among all the measured magnitudes, I report three aperture magnitudes inside 1.5, 3, and 4 arcsec diameter (`AP_15, AP_30, AP_40`), the Kron magnitude (`Mag_Kron`) and the model magnitude obtained from the sum of the spheroid and disc components of the fitting (`Mag_Model`), as well as the PSF magnitude (`Mag_PSF`). Magnitudes are corrected for the galactic extinction, according to Schlafly & Finkbeiner (2011), and I also report in the catalog the adopted corrections (`A_u, A_g, A_r, A_i, A_Y, A_J, A_Ks`).

Finally, I provide the stellarity index obtained for the r band from *SExtractor* (`SI`) and two additional flags: the star/galaxy flag (`NSFLAG_r`) and the mask flag (`MASK_FLAG_r`). `NSFLAG_r` flag is set to 0 for extended objects and greater than zero for point-like sources identified in the r band, according to the star/galaxy separation described in Sect. 2.2.2. `MASK_FLAG_r` is set to one for objects inside the r band mask and zero otherwise, according to the identification of star haloes described in Sect. 2.2.3. I also report the measured photometric redshift (`photo_`z), as described in the previous section.

The multi-band catalog contains 74114 sources. Table 2.6 provides some statistics on the number of objects, and Table 2.7 contains a general view of the first rows of the multi-band catalog.

Table 2.6: Number of objects in the multi-band catalog, including classification according to star/galaxy separation, masking, and the availability of spectroscopic redshift.

Total number of objects	74 114
Galaxies (`NSFLAG_r`=0)	66 323
Stars (0 <`NSFLAG_r`< 9)	7 510
Saturated stars (`NSFLAG_r`≥ 9)	805
Non masked objects (`MASK_FLAG_r`=0)	66 945
Masked objects (`MASK_FLAG_r`=1)	7 169
Objects without *zspec*	69 991
Objects with *zspec* (cluster core)	4 123

Table 2.7: Extract of the first lines of the catalog. All the quantities with the subindex _u in this table are present for all the bands in the place of the [...] column. The other quantities are reported from the r band, and the photo-z includes information in all bands.

GAME_ID	RA [deg]	DEC [deg]	A [pixel]	B [pixel]	THETA [deg]	R_50 [pixel]	R_KRON [pixel]
GAME0414225-2436203	63.593922	-24.605635	28.11	23.52	-52.19	22.95	3.50
GAME0414248-2436296	63.603519	-24.608214	10.91	7.38	-84.86	7.80	3.50
GAME0416207-2436258	64.086120	-24.607161	16.08	8.37	82.97	12.16	3.50
GAME0414229-2437001	63.595250	-24.616692	4.04	2.49	89.69	4.42	3.53
GAME0415287-2437023	63.869606	-24.617305	2.58	2.14	-31.40	2.78	3.51

GAME_ID	AP_15_u [mag]	APERR_15_u [mag]	AP_30_u [mag]	APERR_30_u [mag]	AP_40_u [mag]	APERR_40_u [mag]	Mag_Kron_u [mag]	MagERR_Kron_u [mag]
GAME0414225-2436203	21.63	0.06	20.26	0.06	19.72	0.06	17.86	0.11
GAME0414248-2436296	15.20	0.06	14.60	0.07	14.48	0.07	14.33	0.11
GAME0416207-2436258	22.39	0.07	21.15	0.07	20.75	0.07	19.60	0.09
GAME0414229-2437001	25.04	0.47	23.79	0.32	23.79	0.41	23.75	0.44
GAME0415287-2437023	25.55	0.75	24.46	0.63	23.93	0.52	24.05	0.49

GAME_ID	Mag_Model_u [mag]	MagERR_Model_u [mag]	Mag_PSF_u [mag]	MagERR_PSF_u [mag]	A_u [mag]	[...] -	SI -	NSFLAG -	MASK_FLAG -	photo_z -
GAME0414225-2436203	17.86	0.12	21.46	0.06	0.188	[...]	0.03	0.0	0.0	0.17
GAME0414248-2436296	14.31	0.16	14.89	0.07	0.192	[...]	0.75	9.0	0.0	-99.00
GAME0416207-2436258	19.60	0.07	22.29	0.07	0.193	[...]	0.03	0.0	0.0	0.30
GAME0414229-2437001	23.75	0.65	24.67	0.36	0.186	[...]	0.03	0.0	0.0	0.18
GAME0415287-2437023	21.60	4.23	25.05	0.49	0.185	[...]	0.96	0.0	0.0	0.32

Chapter 3

Results on the density field and galaxy colors according to environment

Galaxies form and evolve in a variety of environments, depending on the initial conditions of the halo and gas from which they are formed, as well as the surrounding structures and conditions in the Universe.

The physical environment in which a galaxy forms and evolves is a complex and multifaceted concept (Muldrew et al. 2012), that can be characterized by the density of the surrounding matter, including other galaxies, dark matter, and gas. By comparing these measurements with theoretical simulations (**?**Vogelsberger et al. 2014, 2020), it is possible to make predictions about the environment in which the galaxy formed and evolved.

In this chapter, I will introduce the environment classification done using that catalog and the study of galaxy properties in the core and outskirts of the cluster. Sec. 3.1 introduces the definition of cluster members using our photometry. In Sec 3.2 is identified the red sequence galaxies, and in Sec. 3.3 is presented the characterization of the density field. These two ingredients allow us to study the variation of galaxy colors with the environment (Sec. 3.4.1 and 3.4.2).

3.1 Cluster membership

The determination of cluster membership in this study was not based on velocity dispersion as the photometric data used was not suitable for such analysis. Instead, cluster membership was defined as all galaxies within a specific redshift interval across the field of view. The selection of this interval was based on the criteria of completeness and purity, without directly taking into account the uncertainties associated with photometric redshift. However, the balance between completeness and purity, as measured against spectroscopic data, considers the low precision of the photo-z estimation. The completeness is defined as the ratio between the number of galaxies identified as members (both spectroscopic and photometric) and the number of spectroscopic members $C_M = N_{pm \cap zm}/N_{zm}$. The purity is defined as the ratio between the number of photometric members that are also confirmed by spectroscopy and the number of photometric members that have z_{spec},

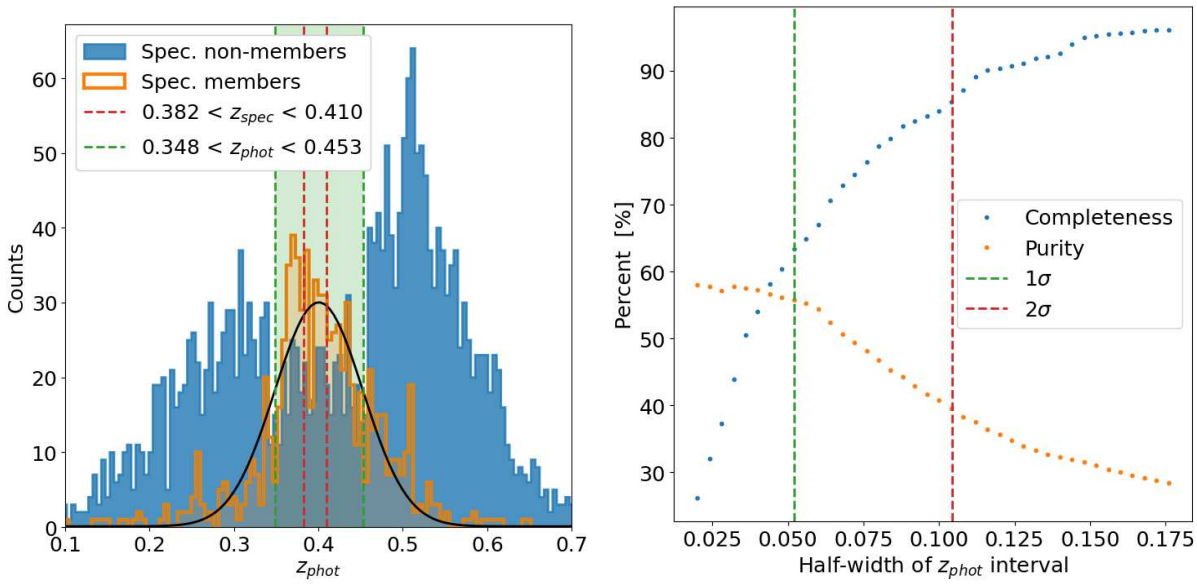


Figure 3.1: *Left*: Histogram showing the distribution of photometric redshift for objects with spectroscopic information. In orange are shown the spectroscopic cluster members, described by a Gaussian curve (black line). The photometric cluster interval is shown in green and corresponds to the one-sigma dispersion of the Gaussian curve $\sigma = 0.052$. In blue are shown the objects with a spectroscopic redshift outside the cluster interval. *Right*: Values of completeness and purity as a function of the half-width of the z_{phot} interval. Vertical lines represent the dispersion of the Gaussian represented on the left panel.

$P = N_{pm\cap zm}/N_{pm\cap z}$. Enlarging the redshift interval increases the completeness of cluster members' samples but at the expense of purity.

The left plot of Figure 3.1 shows the histogram of photometric redshifts for the objects in our catalog with spectroscopic information and with a $S/N > 5$ in the NIR bands. In orange are shown the photo-z of the 801 spectroscopic cluster members within the interval $0.382 < z_{spec} < 0.410$ (red dotted lines in the figure, according to Fig. 5 of Balestra et al. 2016), which corresponds to a rest-frame velocity of ± 3000 km s^{-1}. In blue are shown the photo-z of 2494 objects with a spectroscopic redshift outside the cluster interval.

To define the photometric cluster membership, I fit a Gaussian curve over the photo-z histogram of spectroscopic cluster members (black line in the left panel of Fig. 3.1) and study the behavior of completeness and purity as a function of σ-levels of that Gaussian. In the right panel of Fig. 3.1 the values of completeness and purity as a function of the half-width of the z_{phot} interval are shown. The best choice for the cluster membership interval is therefore the one sigma level of the Gaussian [1] plotted on the left panel of Fig. 3.1, which leads us to $0.348 < z_{phot} < 0.453$ (green shaded area in the left panel of Fig. 3.1).

The number of photometric cluster members is 4067 galaxies across the whole VST field. The completeness of the sample of cluster members on the interval is 63.3 % and the purity is 55.8 %.

3.2 Sequence of red galaxies

Since the population of red galaxies dominates the densest regions of large-scale structures, they are efficient tracers of the cosmic web. In this section, I present the analysis of the color-magnitude diagram, to select red galaxies across the field and to identify overdensities.

The color-magnitude plane of all members is shown in Fig. 3.2. I obtain the red sequence relation using spectroscopically confirmed cluster members ($0.382 < z_{spec} < 0.410$) in the $g - r$ observed color within a 3 arcsec -diameter aperture, and the r band Kron magnitude. The sequence of red galaxies obtained in this section and shown as the red shaded area in Fig. 3.2, will be plotted as a reference in all the color-magnitude plots of the next sections.

Exploiting the software presented in sec. 3.2 of Cappellari et al. (2013), I first fit a linear relation over the whole sample, to split it in two. Then, a subsample of galaxies with $g - r > -0.0985 \times (r - 21.99) + 1.386$, the upper part of the diagram, is chosen to determine the red sequence relation: $g - r = -0.0662 \times (r - 21.86) + 1.6631$, with a standard deviation of $\sigma = 0.0813$. Finally, I define as red sequence galaxies all photo-z-based members contained within 2.6 times the standard deviation of this sequence. It is expected that $\sim$99% of red galaxies are included within this confidence level. The black stars in Fig. 3.2 are the spectroscopic cluster members while the grey and red points are photometric cluster members over the whole field.

[1] I understood as one sigma level the interval: $(\mu - \sigma, \mu + \sigma)$.

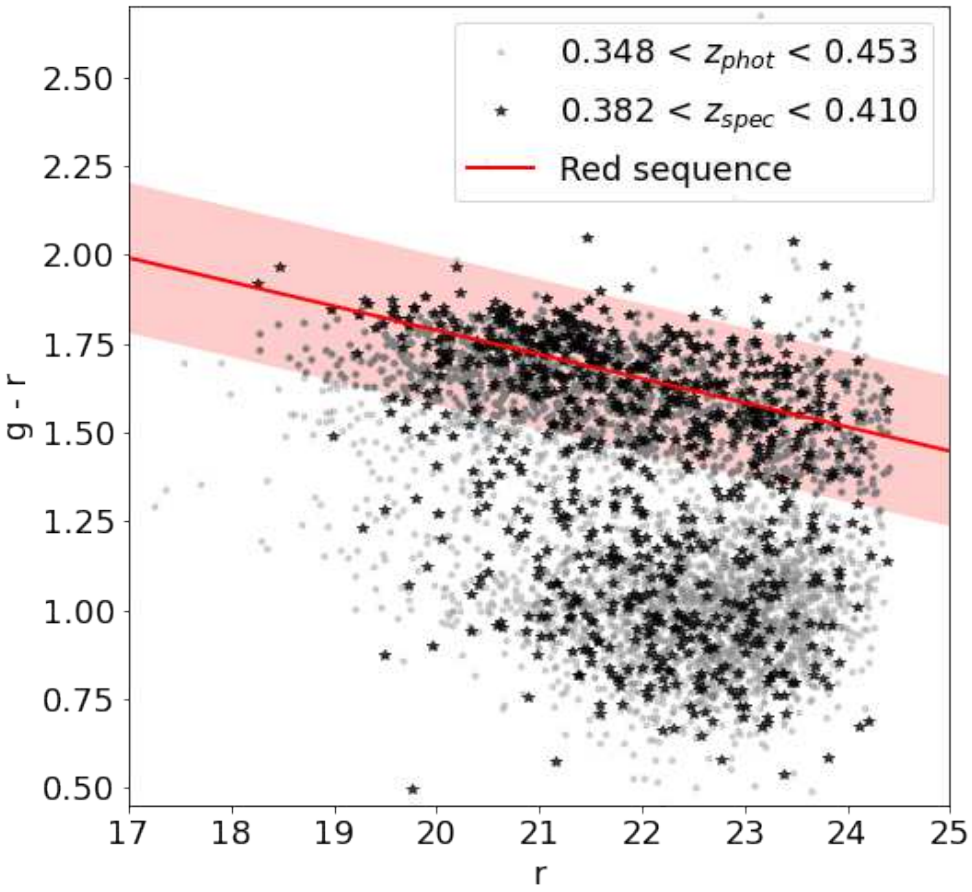


Figure 3.2: Color magnitude diagram for MACS0416 cluster members using VST photometry. Color is determined using g and r band aperture magnitudes of 3.0 arcsec. The x-axis shows the Kron magnitude for the r band. The sequence of red galaxies is determined with spectroscopic cluster members (black stars in the figure), first splitting the sample in two and then fitting a linear relation in the upper part of the plane. The red line determines the best fit for red galaxies and the red shadow area represents the 2.6σ region around the fit. Grey points are galaxies with photometric redshift within the cluster interval.

Considering the sample of red galaxies in the cluster redshift interval of $0.348 < z_{phot} < 0.453$, it is obtained a completeness of 100.0% and a purity of 71.1% with respect to the sample of spectroscopic cluster members within the red sequence. This improvement confirms that red galaxies have a more accurate photometric redshift than the whole sample and that can be used as an efficient tracer of dense structures at the studied redshift.

In the following, I will use the red-sequence galaxies to compare galaxy populations across environmental density (Sec. 3.4.1) and to identify the most significant overdense regions in the cluster outskirts (Sec.3.4.2).

3.3 Density field using cluster members

A way to determine the different environments affecting galaxies and their properties is to compute the galaxy density field (e.g., Gómez et al. 2003; Kovač et al. 2010; Muldrew et al. 2012; Annunziatella et al. 2014; Granett et al. 2015; Annunziatella et al. 2016; Malavasi et al. 2017; Gargiulo et al. 2019; Ata et al. 2021).

I derive the projected density field using the galaxies lying in the photometric redshift interval of $0.348 < z_{phot} < 0.453$. To calculate the density field, I utilized photometric redshifts for galaxies with confirmed spectroscopic redshifts, as my research is primarily concerned with the outer regions of the cluster where spectroscopic measurements are lacking. Initially, I set a rectangular grid using square cells of 1 arcmin a side, corresponding to a comoving distance of ∼300 kpc at the cluster redshift. On this grid, I use the nearest grid point (NGP) scheme. The NGP consists of a binning of the field in rectangular cells and counting the number of galaxies into a given cell, n_i. The output of this step is shown in the top-left panel of Fig. 3.3.

Then, I compute the fluctuations over the mean value for each cell, $\delta_i^* = n_i/\bar{n} - 1$, where $\bar{n} = 0.92$ is the mean number of galaxies per cell across the whole field at the cluster redshift. This allows me to obtain an adimensional density measurement, which is geometrically equal to the density field in the previous step but describes better the fluctuations around a mean value. This step is shown in the top-right panel of Fig. 3.3.

Additionally, δ_i^* is convolved with a Gaussian kernel to connect the local density with adjacent cells, generating a stronger environment tracer, which takes into account not only the cell density but also the density of nearby cells. The standard deviation of the Gaussian kernel corresponds to two cells of the grid. This step is shown in the bottom panel of Fig. 3.3. This smoothed density field on a particular cell (for simplicity δ) will be the environment tracer, and its value is assigned to each galaxy inside the cell.

Complementary, I compute the density field using only red galaxies δ_{rs} following the same procedure as described before ($\bar{n}_{rs} = 0.31$).

In general, the numerical values of the density field are dependent on the method used to compute it, and, in our case, on the cell size and width of the Gaussian kernel. Cosmological studies like Monaco & Efstathiou (1999); Kitaura et al. (2009); Schirmer et al. (2011); Granett et al. (2015); Ata et al. (2021) use a grid size between 1 and 5 Mpc h^{-1} and Gaussian kernels up to 10 Mpc to better identify structures over large regions of the sky, while studies like Annunziatella et al. (2014); Malavasi et al. (2017) that were focused on galaxy evolution, on scales from 300 kpc up to 2 Mpc to reconstruct the local environment. In our case, the cell size and width of the Gaussian kernel are chosen to find an optimal trade-off between large-scale structures and small-scale number count fluctuations.

In Fig. 3.5 is shown the galaxy density field for M0416 above the r band completeness magnitude limit $r < 24.4$, an more graphical elaborated -but physically identical- version of the bottom panel of Fig. 3.3. Fig. 3.5 is the density field δ computed with all the galaxies in the redshift interval, while Fig. 3.4 is the density field computed using only red galaxies δ_{rs}. The former is used in Sec. 3.4.1 to study populations of galaxies across

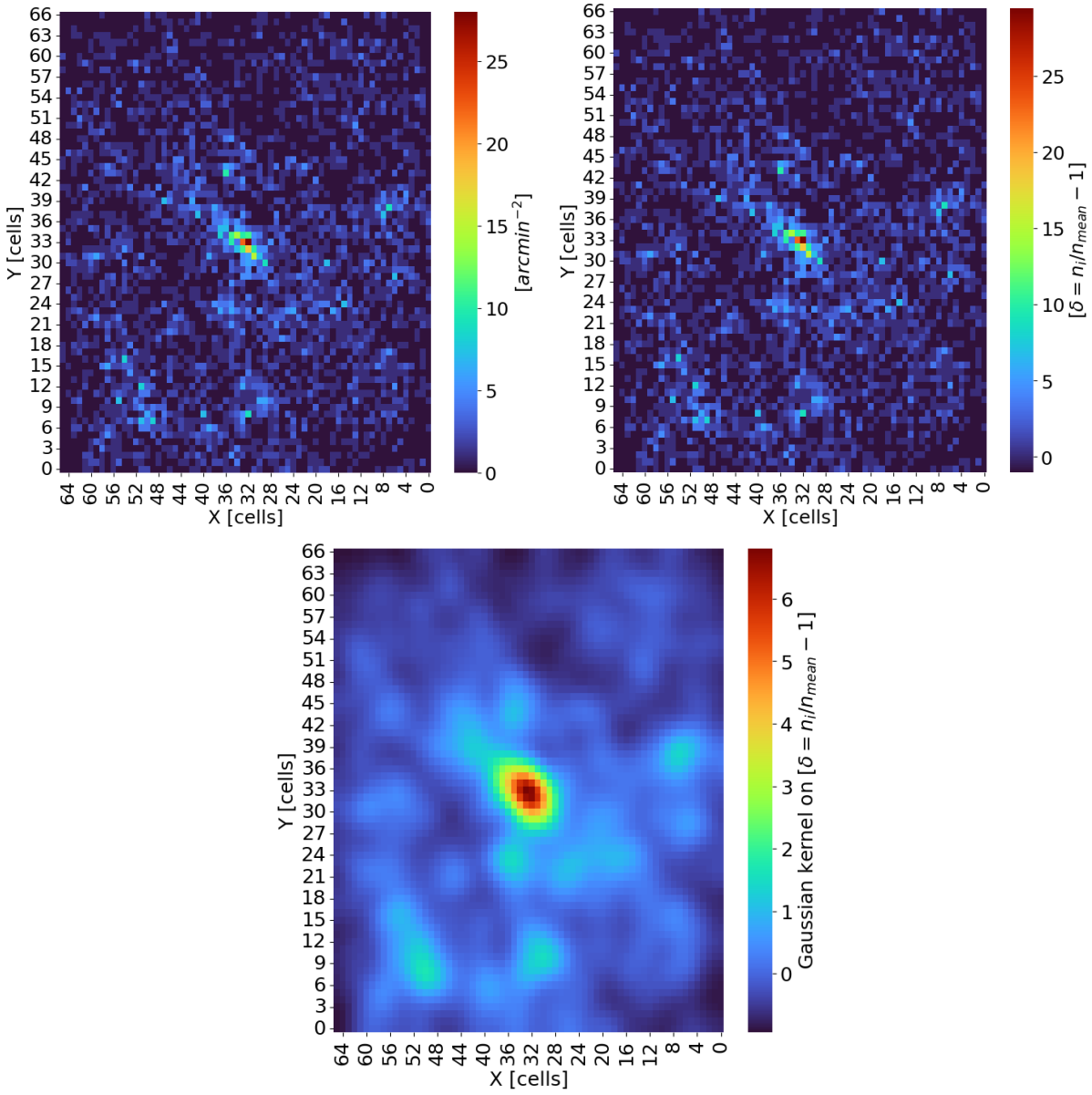


Figure 3.3: The three steps followed in the construction of the density field. *Top left*: Histogram on cubic cells of ∼1 arcmin. *Top right*: Fluctuations around the mean value. *Bottom*: Convolution with the Gaussian kernel.

several levels of environmental density, while the latter is used in Sec. 3.4.2 to detect overdense regions in the cluster outskirts at $r>2r_{200}$. As expected, in both figures, 3.5 and 3.4, panels it appears evident that the highest density values correspond to the central region of the cluster.

Additionally, I used the galaxy sample to estimate the cumulative projected number of cluster members and the differential number of density profiles. I computed the profiles as a function of the projected radius in units of r_{200} and rescaled them by the number of members, n_0, found within the radius $r/r_{200} = 0.15$ to follow the convention implemented by Angora et al. (2020). In Fig. 3.6 I show the cumulative projected number and the differential projected number density profiles of cluster members after applying such renormalization where the shaded areas correspond to 68% confidence levels. The slope of the cumulative profile reveals the presence of a structure, denser in the core, less dense but still structured up to the periphery of the field. The obtained profile follows the same distribution found by Angora et al. (2020) for the four clusters Abell S1063 (z=0.347), M0416, MACSJ1206.2-0847 (z=0.438), and MACS J1149+2223 (z=0.544), which also reproduces the results found by Bonamigo et al. (2018) and Caminha et al. (2019) using strong lensing modeling.

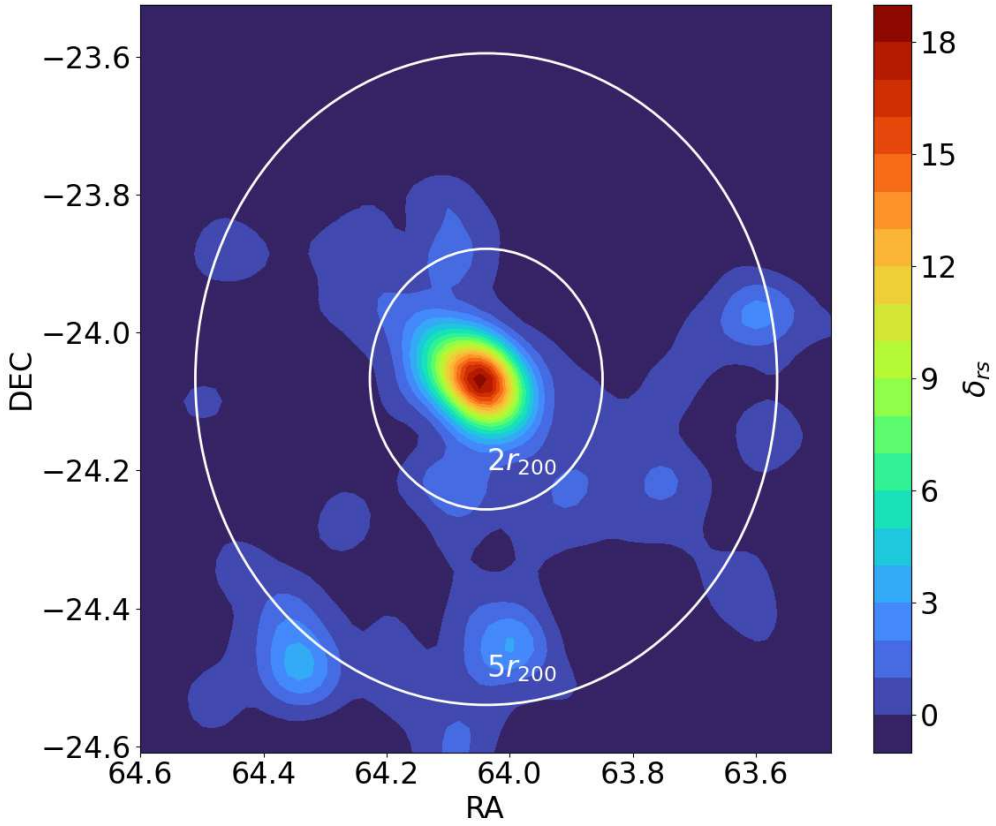


Figure 3.4: MACS0416 density field computed using red galaxies in the $0.348 < z_{phot} < 0.453$ redshift interval. The inner circle indicates $2\mathrm{r}_{200} = 3.64\,Mpc$ and the outer circle indicates $5\mathrm{r}_{200} = 9.1\,Mpc$. Both circles are centered into the NE-BCG which coincides with the center of the cluster (Balestra et al. 2016).

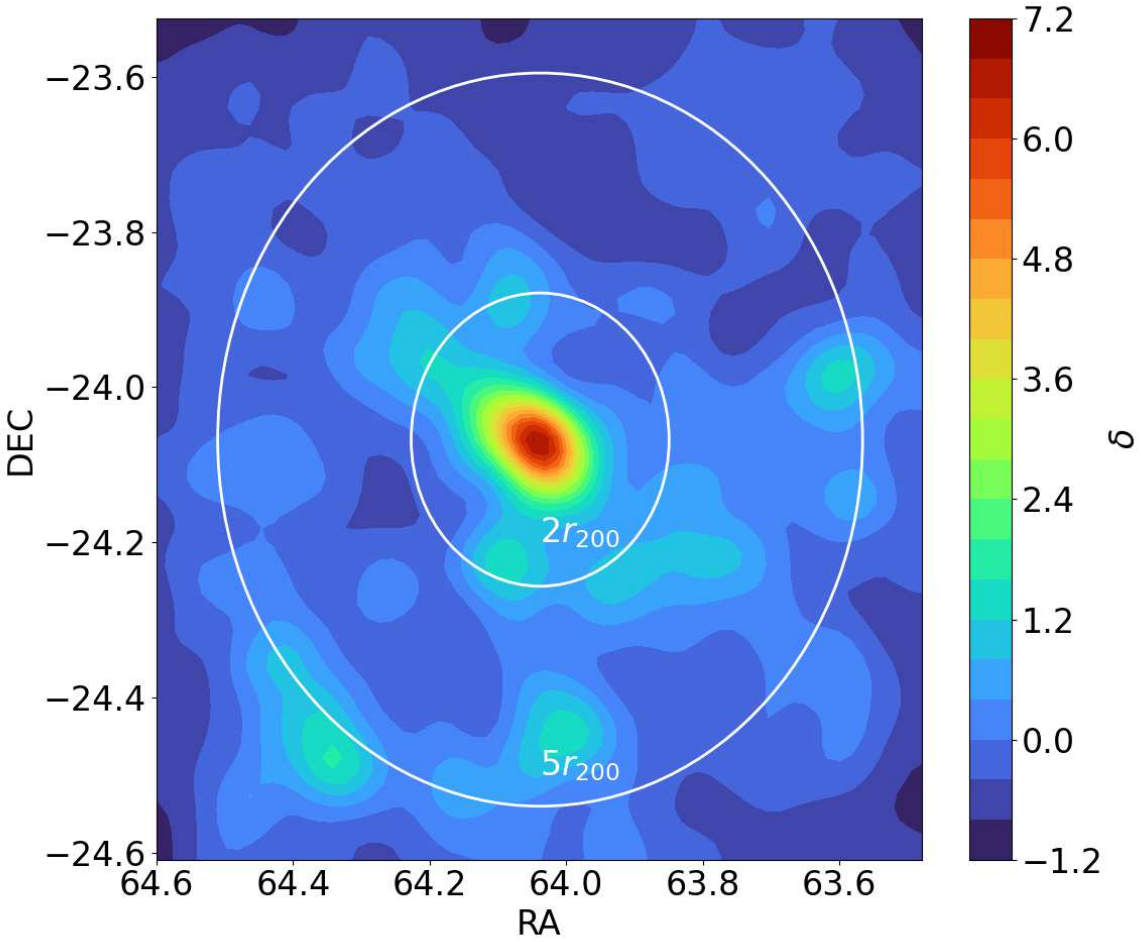


Figure 3.5: MACS0416 density field computed using all the VST-GAME galaxies in the $0.348 < z_{phot} < 0.453$ redshift interval. The inner circle indicates $2r_{200} = 3.64\,Mpc$ and the outer circle indicates $5r_{200} = 9.1\,Mpc$. Both circles are centered into the NE-BCG which coincides with the center of the cluster (Balestra et al. 2016)

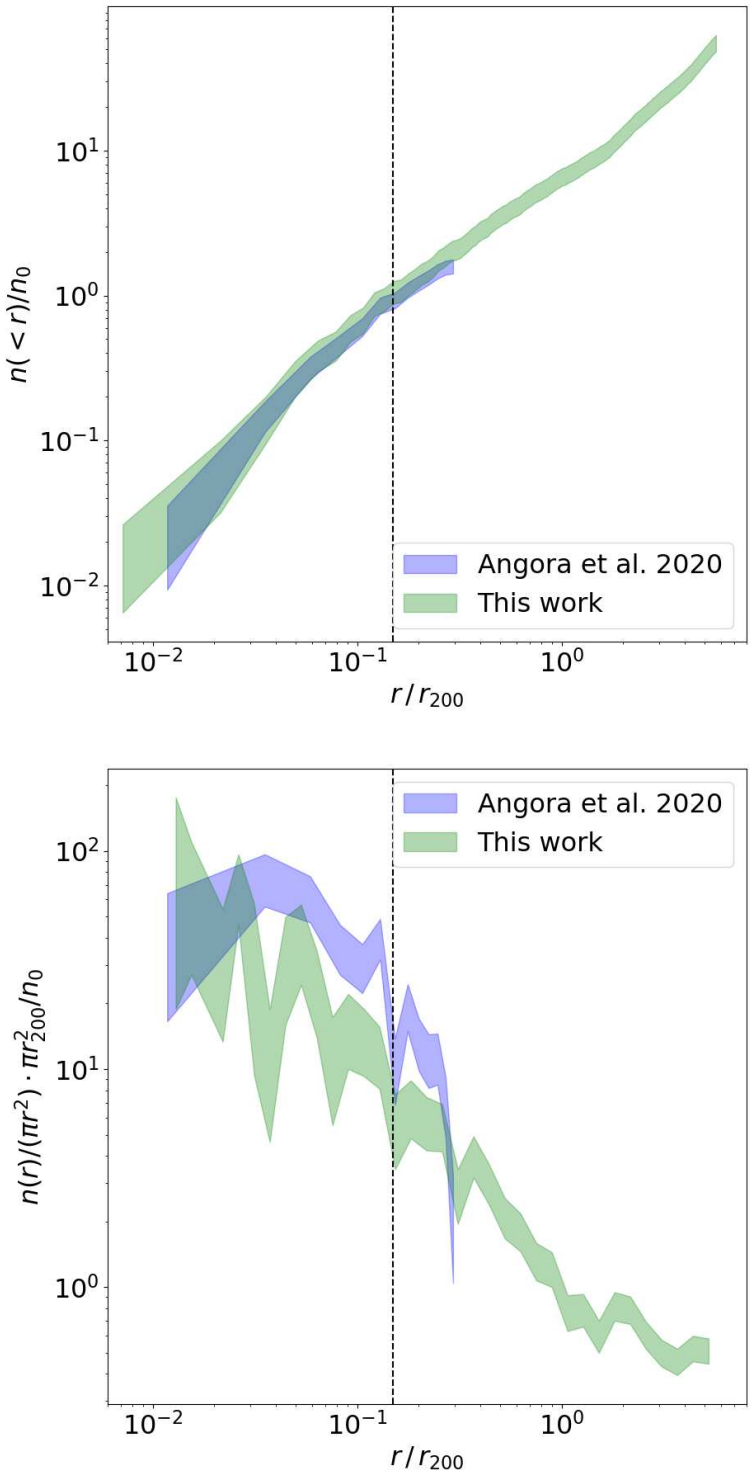


Figure 3.6: Cumulative (*top*) and differential (*bottom*) projected number density of M0416 members identified in our cluster interval (green), compared with results obtained by Angora et al. (2020) (blue). The areas correspond to the 68% confidence level regions. All profiles are normalized by the number density n_0 of members with r<0.15 r_{200}. The dashed line corresponds to r = 0.15 r_{200}.

3.4 Galaxy colors across the density field

3.4.1 Galaxy sample according to local environment

To characterize the local environment, I split the sample of all cluster members (regardless of color) into three density intervals: low ($\delta < 0$), medium ($0 < \delta < 1$), and high density ($\delta > 1$), according to the definition of density derived in section 3.3. The first interval corresponds to under-dense regions, with a local density (see Sec. 3.3) lower than the mean value over the field, $\bar{n} = 0.92$; the second interval corresponds to over-dense regions having a local density higher, up to two times the mean; and the third interval corresponds to regions with the highest density values over the field. The first density bin contains 1507 galaxies (37%), the second 1778 (44%), and the third 782 (19%).

The top panel of Fig. 3.7 shows the spatial distribution of galaxies and highlights the different density intervals, while the bottom panel of the same figure shows the distribution of galaxies as a function of the density value.

To analyze the properties of galaxies in the different environments I compare the location of the red sequence determined in Sec. 3.2 with the sample of galaxies in each of the three density intervals defined above. Figure 3.8 clearly shows the dependency of the color-magnitude diagram on the environment. The fraction of galaxies on (below) the red sequence increases (decreases) with density: the fraction of red galaxies goes from 21% $\pm$ 1% in the lowest density bin to 32 % $\pm$ 1% at intermediate densities to 65 % $\pm$ 2% in the highest density bin. Errors are binomial.

This behavior is also clearly visible in the normalized histograms (right panel of Fig. 3.8) of the $g - r$ color. All of them present a bimodal behavior with one peak for the galaxies populating the red sequence and the other one for the blue cloud, but the relative importance of the two peaks in the distribution strongly varies with the environment. The histogram also shows that the position of the red sequence slightly shifts with the environment: the peak of the distribution of red galaxies shifts toward redder colors with increasing density. Horizontal dashed lines in panel (d) of Fig. 3.8 show the position of the peak obtained by fitting a kernel density estimator (*kde*) to the red galaxies histogram.

3.4.2 Galaxy sample on outskirts substructures

Thanks to the large FoV of the VST and VISTA images I can investigate in detail also the outskirts of M0416. The left panel of Fig. 3.7 unveiled the presence of three overdense regions in the cluster outskirts at r$\sim$5r$_{200}$ with more than 25 galaxies each.

To better characterize their shape and galaxy content, I delimited the area of each region, defining a polygon that includes all the red-sequence galaxies with $\delta_{rs} > 1$. Regions A, B, and C are shown in the left panel of Fig. 3.9 and their properties are presented in table 3.1.

The middle panel of Fig. 3.9 shows the color-magnitude diagram for the galaxies inside the cluster core and outskirts overdensities, and the right panel the corresponding $g - r$ color histogram normalized. The most interesting feature of this color distribution is the presence of a similar fraction of red galaxies in the outskirts overdensities than in the

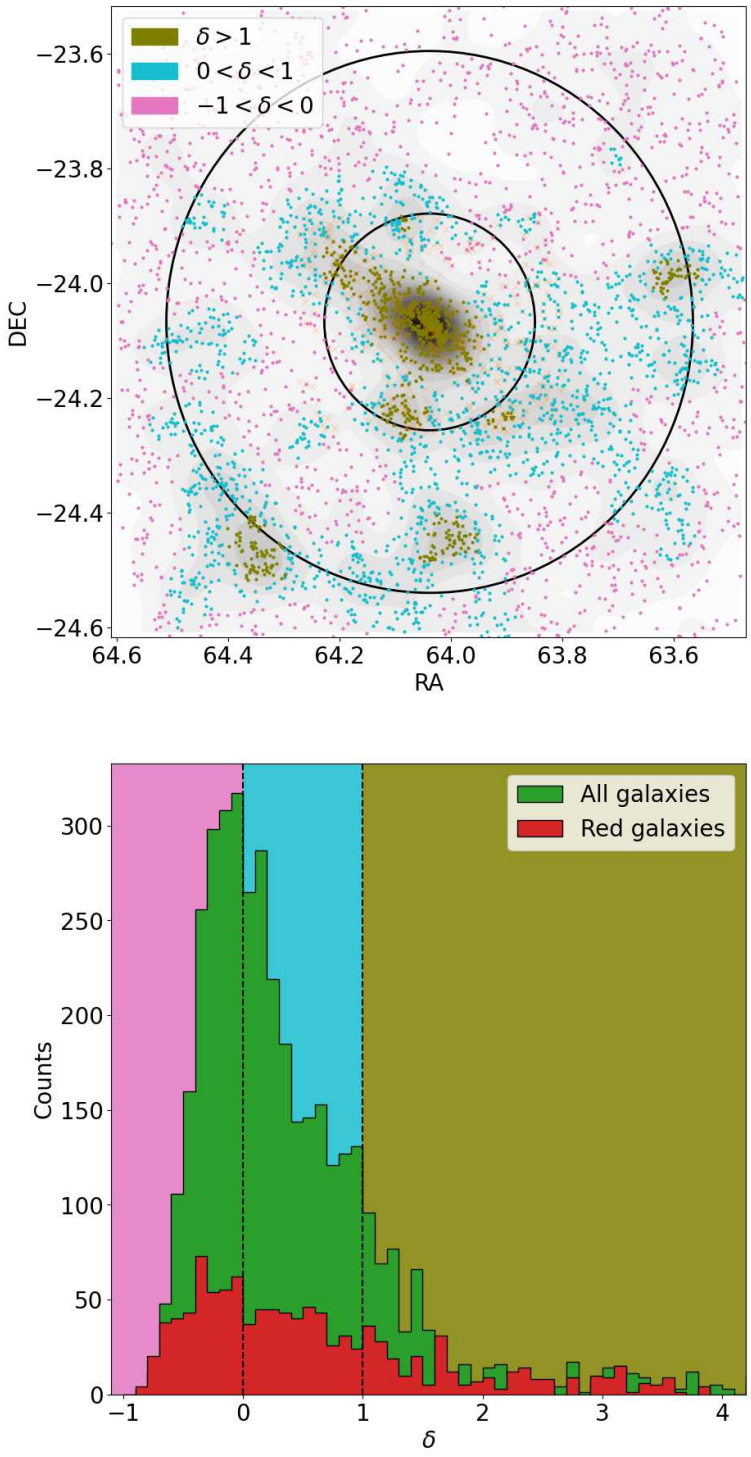


Figure 3.7: *Top*: Environment classification of galaxies with $0.348 < z_{phot} < 0.453$ across the VST-GAME MACS0416 field. Colors correspond to the three density intervals chosen for this work. Black circles indicate 3.64 Mpc and 9.1 Mpc which correspond to $2r_{200}$ and $5r_{200}$ respectively. *Bottom*: Distribution of galaxies according to their local density. The green histogram is computed with all the galaxies within the cluster interval, while the red histogram is made with red galaxies. Filled stripes identify the density intervals chosen as environment tracers.

Table 3.1: Properties of the three overdense regions in the cluster outskirts. RA and DEC coordinates indicate the geometrical center of the overdense region. Distance from the cluster center is also referred to as the geometrical center of each region.

	RA	DEC	Distance from core	Red galaxies	Blue galaxies
Region A	64:20:07.83	-24:28:06.09	9.3 Mpc	62	41
Region B	64:00:06.39	-24:27:14.00	6.7 Mpc	46	35
Region C	63:35:56.84	-23:58:30.23	9.1 Mpc	27	21

cluster core. In the figure, the three outskirts overdensities are shown together to increase the signal. Using a Kolmogorov -Smirnov test I can not reject the hypothesis that the color distribution of each of the three regions is generated from the same distribution. Nevertheless, an excess of blue galaxies is observed in region C (see table 3.1). The luminosity distribution of the three outer regions has a median of 22.34, 21.85, and 22.30 mag in the r band, respectively, showing that region B is slightly brighter. The mean luminosities of these three structures are similar to that of the cluster core, having a median luminosity of 22.07 r mag in the same band.

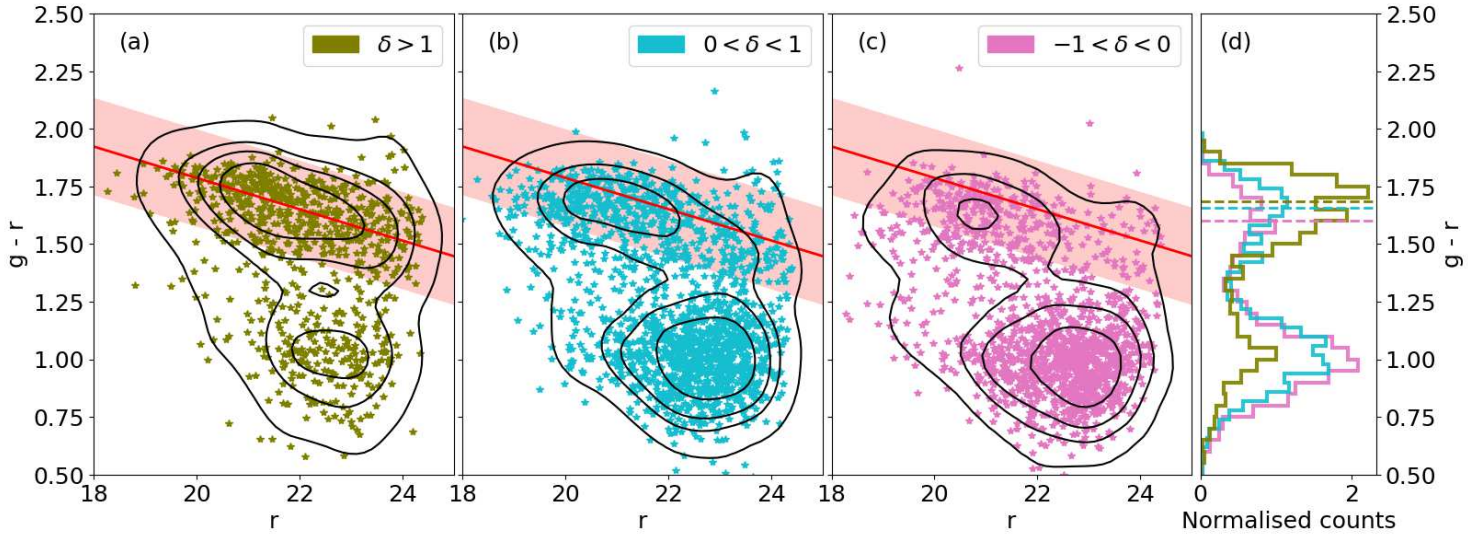


Figure 3.8: Analysis of the color-magnitude diagram dividing the galaxy sample (grey points of Fig. 3.2) according to the local environment. The color code follows that of Fig. 3.7. Panel (a) contains galaxies in the densest environments of the field, including the cluster core and outskirts overdensities. Panel (b) contains galaxies in denser environments up to two times the mean density. Panel (c) contains galaxies in environments less dense than the mean density of the field. Panel (d) shows a normalized histogram of the $g-r$ color for the three galaxy samples. Horizontal dashed lines represent the location of the peaks of the distributions (see text for details). The percentage of red galaxies is 65.2%, 32.8%, and 21.1% for the three density environments respectively, showing a clear abundance of red galaxies in the densest regions of the field. The black contours on each plot contain 5%, 25%, 50%, and 75% of the galaxies for each sample. The red line and the red area represent the sequence of red galaxies obtained with the spectroscopic cluster sample.

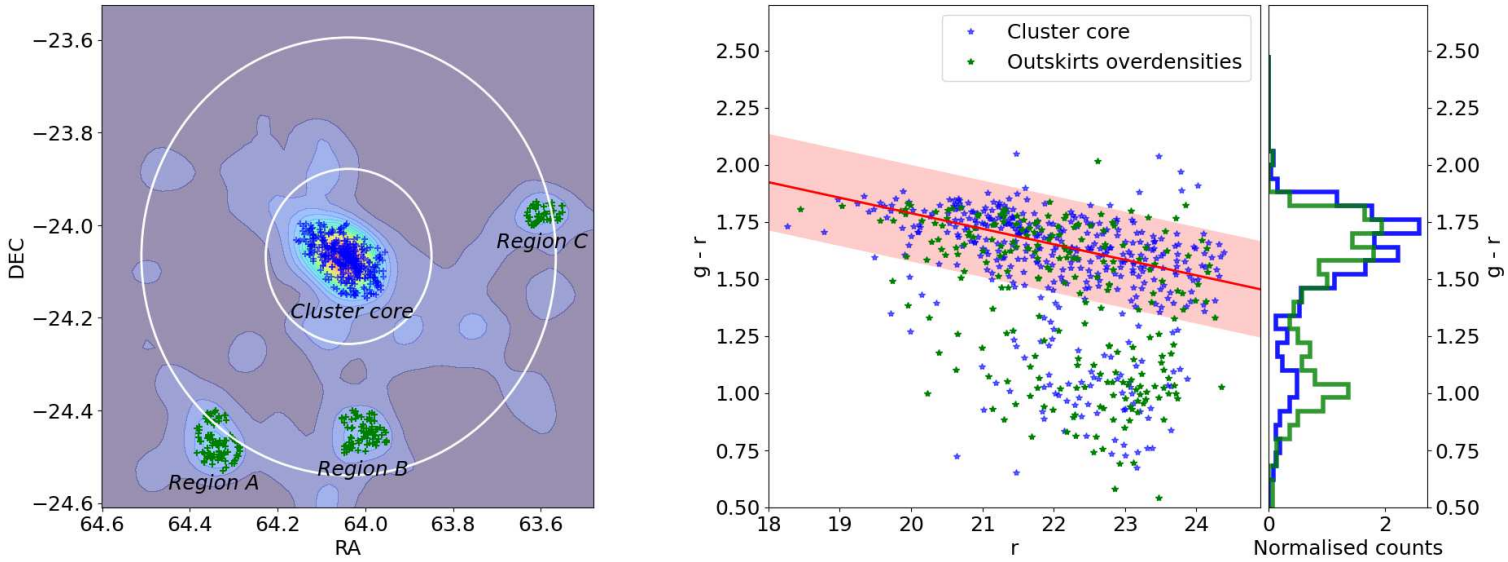


Figure 3.9: *Left*: Selection of the three overdense regions in the cluster outskirts r>2r$_{200}$. The background shows the density field computed with the red galaxy sample (bottom panel of Fig. 3.4). The crosses over the plot correspond to galaxies in overdense environments with $\delta > 1$. Regions A, B, and C are determined using the red-sequence density field. *Middle*: Detailed analysis of the color-magnitude diagram shown on panel (a) of Fig. 3.8. The red stripe represents the red sequence found with spectroscopic cluster members. Blue points are galaxies with $\delta > 1.5$, which corresponds to the cluster core. Green points are galaxies related to the three sub-structures (Regions A, B, and C) in the cluster outskirts at r>2r$_{200}$. *Right*: Histogram of the $g - r$ color for both galaxy samples shown in the central panel.

3.5 Discussion of the results on VSTGAME

As anticipated, M0416 has been already largely analyzed in the literature, however, most of these studies have focused on the central core r<r_{200} (Ebeling et al. 2001; Postman et al. 2012; Rosati et al. 2014; Ogrean et al. 2015; Merlin et al. 2016; Balestra et al. 2016; Castellano et al. 2016; Vulcani et al. 2016, 2017; Natarajan et al. 2017; Lotz et al. 2017; Annunziatella et al. 2017; Shipley et al. 2018) or were dedicated to characterizing the surface mass density from weak and strong-lensing analysis (Zitrin et al. 2013; Jauzac et al. 2015; Grillo et al. 2015; Bonamigo et al. 2017; Caminha et al. 2017; Gonzalez et al. 2020).

Only Balestra et al. (2016) and Olave-Rojas et al. (2018) examined the cluster outskirts. These studies identified spectroscopically confirmed substructures up to r $\sim$ 2 r_{200}, exploiting the Dressler & Shectman (1988) technique to isolate regions kinematically distinct from the main galaxy cluster (see also Dressler et al. 2013).

The results presented in this thesis show the importance of data covering regions up to 5r_{200} to investigate galaxy evolution in different cluster environments. Only the grasp and image quality of the optical VST data combined with the NIR VISTA images, allowed us the investigation of galaxy properties in the full range of environments, from the high-density cluster core to the outskirts. The importance of cluster outskirts up to these distances lies in the fact that they allow an environmental study of infalling populations without the contamination from backsplash galaxies. Those galaxies are cluster members that have undergone a first pericentric passage through the dense ICM and are currently located outside the virial radius. They have experienced environmental effects such as ram pressure stripping and tidal interactions during their infall and outflow trajectories, and will eventually fall back into the cluster core again (Jaffé et al. 2018; Haines et al. 2015). Backsplash galaxies pose a challenge for studying pre-processing effects in cluster outskirts, as they can be misidentified as infalling galaxies due to their location beyond the virial radius. However, unlike infalling galaxies, backsplash galaxies have already been affected by their previous encounter with the cluster core and its ICM. The $0.348 < z_{phot} < 0.453$ redshift interval is equivalent to a transverse distance of $\sim$ 360 Mpc, which remarks the importance of a spectroscopic follow-up.

In Sec. 3.3 I present, for the first time, the large-scale density field for M0416, up to $\sim$5r_{200}, using deep photometry (mag $r < 24.4$). Thanks to the density field shown in Fig. 3.5, I classified galaxies into three groups based on the local environment: low-, medium-, and high-density. (see Sect. 3.4.1). I study the dependence of galaxy colors as a function of local densities to obtain a characterization of the properties of galaxy populations as a function of the environments. I show that, as expected, galaxies in less dense regions are mainly blue (Fig. 3.8). I find that the peak of the distribution of red galaxies shows a shift toward redder colors with increasing density. This indicates the active role of the environment as a driver of galaxy evolution. The role of the masses in this process has still to be studied and is one of our next steps.

Additionally, in Sect. 3.4.2, I find the presence of three overdense regions (indicated as A, B, and C) at large distances ($\sim$ 5 r_{200}) from the cluster center and a large overden-

sity region aligned with the cluster core. I found that the three overdensities have mean densities and luminosities similar to the cluster core. Moreover, the galaxies populating these external overdensities are typically as red as galaxies in the cluster core (Fig. 3.9). It is possible to assume that these structures (e.g. dense groups or cluster cores) are in the processes to be incorporated into the cluster through filaments. The color distribution also suggests the presence of evolved galaxy populations, an insight into pre-processing phenomena over these substructures. In a forthcoming work, I will investigate complementary tracers for passive/star-forming galaxies, like the color-color UVJ diagram, to better characterize the quenching scenarios on the cluster outskirts.

To test this environment classification, at least in the inner region of the field, I make a juxtaposition of the obtained density field with the results of Olave-Rojas et al. (2018), where 15 substructures were identified using spectroscopic redshifts. All the substructures identified by Olave-Rojas et al. (2018) are placed in the overdense regions in our field ($0 < \delta < 1$ and $\delta > 1$). Fig 3.10 shows a comparison between the density classification of our field and Olave-Rojas et al. (2018).

The regions A, B, and C are outside the area covered in Olave-Rojas et al. (2018). The typical number of members in the substructures found in that work is 9, thanks to the dynamical analysis done with spectroscopic data. Overdense regions in this work typically contain ~ 70 galaxies due to the choice of the local environment (cell size, kernel length) influenced by uncertainties on photometric redshifts. The fraction of red galaxies in substructures on Olave-Rojas et al. (2018) is intermediate between that of the main cluster and the field, supporting the pre-processing picture. I find that the fraction of red galaxies on the overdensities is lower but similar to the cluster core, but higher than the field.

Schirmer et al. (2011) obtained a density field up to 10 Mpc from the cluster SCL2243 at z=0.45 using both spectroscopic and photometric data and including filaments detection using a weak lensing analysis. In the filaments, they observe a constant color, independent of the clustercentric distance across the field. Only in the cluster infall region (out to 1.5 Mpc outside r_{200}) do the filaments become noticeably redder, having the same average color as the supercluster center. Instead, in this work, I claim overdensities with the same average color as the cluster center up to larger distances (10 Mpc $\sim 5r_{200}$). This can be explained because the scale used to compute the density field in this work is spatially larger than the filament scale.

Verdugo et al. (2012) computed a density field for the cluster RX J1347.5-1145 at $z = 0.45$ on a scale of $\sim 20 \times 20$ Mpc. They also found dense regions with a low fraction of blue galaxies up to ~ 10 Mpc from the cluster core in perfect agreement with our work. Still, their optical photometry is shallower than ours (r mag = 23.5 vs mag = 24.4), which allows us to explore a wider range of masses. To exploit the depth of our data, in a forthcoming paper I will compute the stellar masses of our sample and I will study their dependence on different galaxy properties.

Lu et al. (2012) studied $\sim$100 galaxy clusters between $0.16 < z < 0.36$ finding that the fraction of optically blue galaxies is lower for the overdense galaxy population in the cluster outskirts compared to the average field value, at all stellar masses $M_* > 10^{9.8} M_\odot$

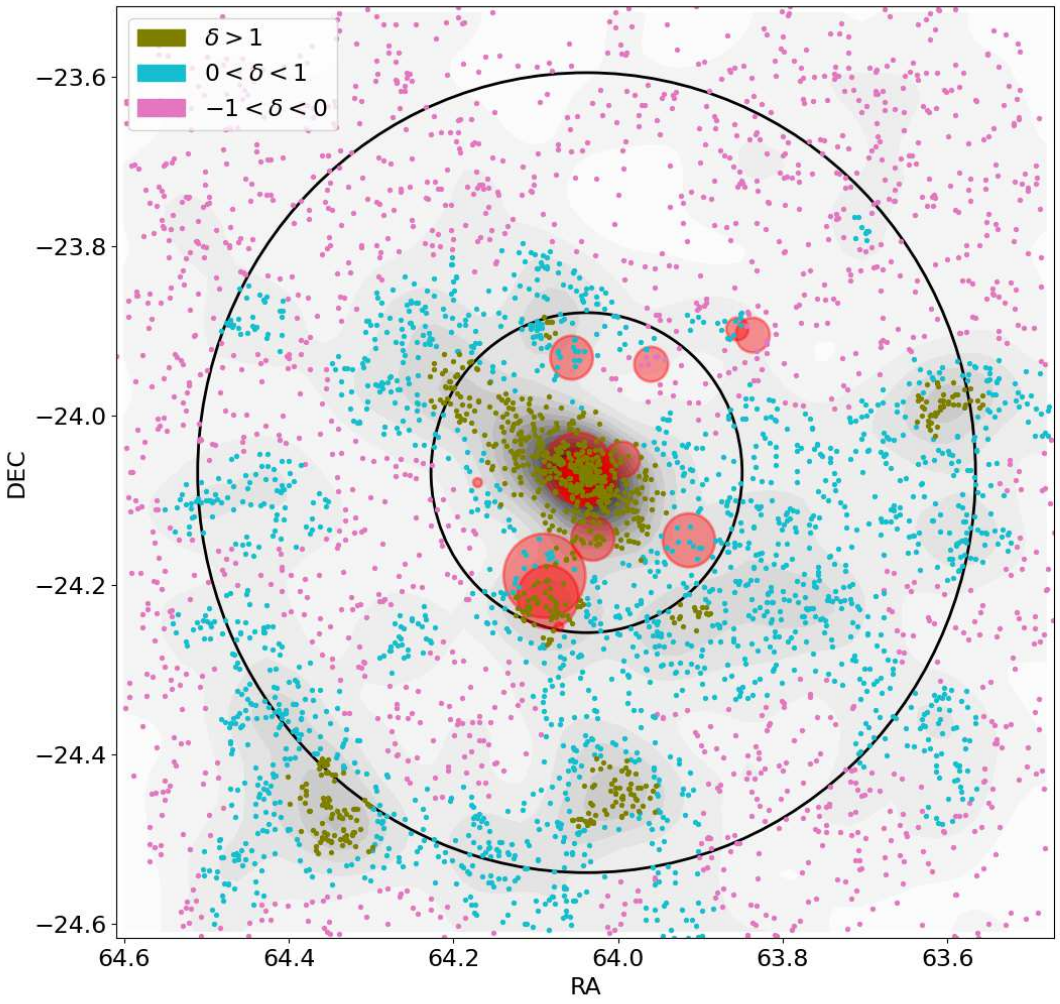


Figure 3.10: Same Figure as top panel of Fig. 3.7. Red circles are the 15 substructures identified dynamically by Olave-Rojas et al. (2018)

and at all radii out to at least 7 Mpc, in agreement with our results. Nevertheless, with VST-GAME data, I reach larger distances (∼10 Mpc at the same redshift). In forthcoming work, I intend to calculate the stellar masses of the galaxies to expand upon the findings presented in this thesis by incorporating a mass dependence. I aim to extend the results obtained by Lu et al. (2012) to less massive galaxies with stellar masses on the order of $10^9 M_\odot$, taking advantage of the greater depth of the VST-GAME photometry by 0.5 magnitudes in the r band.

Just et al. (2019) studied 21 galaxy clusters using spectroscopic data at $0.4 < z < 0.8$ up to r mag = 22.9, finding that galaxies in the infall regions, determined dynamically, show enhanced clustering in groups than field galaxies. They also find that the more highly clustered galaxies show an elevated red fraction, which is interpreted as preprocessing. Our analysis does not determine the infalling regions through spectroscopic data. Nevertheless, I also find that highly clustered galaxies, up to large radial distances, show an elevated red fraction.

Sarron et al. (2019) claims to detect cosmic filaments around galaxy clusters using photometric redshifts in the range $0.15 < z < 0.7$ and found that the fraction of passive galaxies is higher in filaments than in isotropically selected regions around clusters and that the passive fraction in filaments decreases with increasing distance to the cluster up to $\sim$ 5 Mpc. Our present work does not focus on filaments but current spectroscopic observations (see below) of our M0416 field will allow us to robustly identify filaments and to study their galaxy content up to $\sim$ 10 Mpc from the cluster core.

I have here reported various observational hints that have unraveled the unmistakable role of the environment in accelerating galaxy evolution at intermediate redshift. Indeed, several works have found that dense environments are populated with red galaxies, from the local universe up to $z \sim 1.5$ (e.g. Haines et al. 2007; Cooper et al. 2010; Pasquali et al. 2010; Peng et al. 2010, 2012; McGee et al. 2011; Sobral et al. 2011; Muzzin et al. 2012; Smith et al. 2012; Wetzel et al. 2012; La Barbera et al. 2014; Lin et al. 2014; Vulcani et al. 2015; van der Burg et al. 2020), and our results strengthen the role of the large scale structure in aging galaxies across cosmic times.

Unmasking the cosmic galaxy density field

Chapter 4

Cosmological parameters from the VIPERS density field

This chapter is the result of external collaboration with Prof. Luigi Guzzo and Dr. Benjamin Granett, using data from the VIPERS redshift survey. My contribution was the development of a Bayesian code to reconstruct the galaxy density field and compute joint probabilities on cosmological parameters. The results of this work are published in Estrada et al. 2022a.

The VIMOS Public Extragalactic Redshift Survey (VIPERS) is a galaxy survey carried out using the VIMOS spectrograph mounted on the telescope Melipal (UT3) at the Very Large Telescope (VLT), located at the Paranal Observatory in Chile. It has been developed to study galaxy clustering and redshift space distortions (RSD) within the interval $0.5 < z < 1$ Guzzo et al. (2014). This survey was designed to study a relatively unknown region of the universe, reaching a galaxy density comparable to surveys of the local universe.

4.1 Data description

4.1.1 The VLT telescope and the VIMOS instrument

The VLT is a telescope facility operated by ESO on Cerro Paranal, in the Atacama Desert of northern Chile. The Visible Multi-Object Spectrograph (VIMOS Le Fèvre et al. 2003) was a visible wide-field imager and multi-object spectrograph installed at Nasmyth B focus of *Melipal* telescope. This instrument was designed to execute deep astronomical surveys, by measuring nearly 100 spectra by pointing. VIMOS was composed of a 4-channel imaging spectrograph. Each channel, a "quadrant" on the left panel of Fig. 4.1, covered 7 x 8 $arcmin^2$ for a total field of view (a "pointing") of $\sim 218 arcmin^2$. Each channel was a complete spectrograph with the possibility to insert 30×30 cm^2 slit masks at the entrance focal plane, as well as broad-band filters. The wavelength range covered is between 5500 and 9500 Å.

The VIMOS instrument was able to operate in three different observation modes: direct imaging, multi-slit spectroscopy, and integral field spectroscopy. In the latter mode, a high-

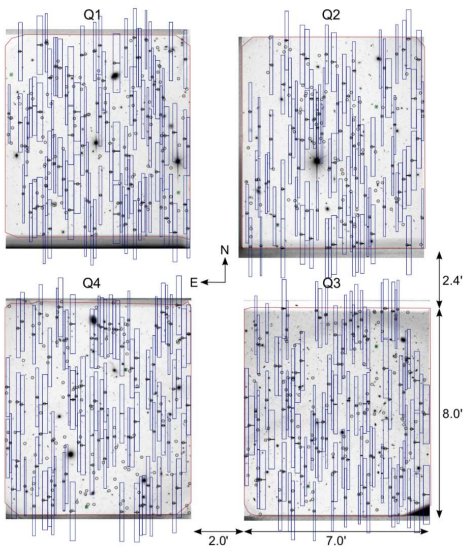


Figure 4.1: Description of the VIMOS mask. One pointing is composed of four quadrants, each of them nearly 7 x 8 $arcmin^2$. The red lines indicate the reconstructed boundaries of the illuminated area. The blue rectangles indicate the slices on the metal plate for measuring the spectra of only selected targets. Credit image: Guzzo et al. (2014).

precision laser-cutter drilled little slits of 1 width in a metal plate, at the exact position of the target galaxies. This metal plate, the *VIMOS mask*, was inserted in VIMOS so that only the light of the targets entered the instrument. The light of those objects was split into its color components by a prism-like device, yielding a spectrum. The VIMOS mask blocked the light from other sources, avoiding overlapping of spectra and improving the quality of the measurements. The spectra were recorded on the four-coupled device (CCD) cameras mounted at the end of the instrument. In this way, VIMOS collected nearly 220 spectra for each pointing in nearly one hour of exposition time. (See figure 4.1). VIMOS has been decommissioned in 2018. My work is directly related to the VIMOS mask since this mask covers up to 30% of the telescope FoV, and the procedure that I developed aims to recover information on the obscured regions.

4.1.2 VIPERS survey

The VIPERS is an ESO Large Program that has mapped in detail the spatial distribution of galaxies up to z$\sim$1. A detailed description of the survey is given by Guzzo et al. (2014); Scodeggio et al. (2018). VIPERS was conceived in 2007 with a focus on the clustering of galaxies and redshift-space distortions, nevertheless, the quality of the data enables broader goals involving large-scale structure and galaxy evolution. The survey design was also strongly driven by the specific features of the VIMOS spectrograph, which has a relatively small field of view compared to fiber positioners ($\sim$18 x 16 $arcmin^2$), but a larger yield in terms of redshifts per unit area. Observations were carried out from November 2008 to December 2014.

The survey targeted galaxies for medium-resolution spectroscopy using VIMOS within two regions of the W1 and W4 fields of the CFHTLS-Wide Survey (Canada-France-Hawaii Telescope Legacy Wide; Cuillandre et al. 2012). Targets were chosen with a limiting flux with $17.5 \leq i_{AB} \leq 22.5$ based upon a color selection which removes efficiently galaxies at $z < 0.5$ while yielding $> 98\%$ completeness for redshift between 0.5 and5 1.2. A precise calibration of this separation method was made possible by the location of the VVDS-Wide ($i_{AB} \leq 22.5$) and VVDS-Deep ($i_{AB} \leq 24$) samples within the W1 and W4 fields of CFHTLS, respectively. The overall footstep of VIPERS is 23.5 deg^2, corresponding to an effective sky coverage of 16.3 deg^2, with 97414 spectroscopically observed galaxies giving a peak number density of $\bar{n} \sim 6 \times 10^{-3} h^3 Mpc^{-3}$ at $z \sim 0.6$.

I used the second public data release of VIPERS (PDR2, Scodeggio et al. 2018). From the final catalog, I selected galaxies with an overall redshift confirmation rate of 98%, as defined on Guzzo et al. (2014). I selected galaxies in the redshift range $0.6 < z < 1.0$. The lower bound at $z = 0.6$ fully excludes the transition region produced by the nominal $z = 0.5$ color-color cut of VIPERS. The high redshift limit at $z = 1.0$ excludes the most sparse distant part of the survey, where shot noise dominates and so the effective volume is small. The total number of galaxies that I use in my analysis is 73572, and its redshift distribution is shown in Fig. 4.3.

4.1.3 VIPERS selection function

The VIPERS survey selection function is a representation of the probability that an object will be included in the survey. Studying the selection function helps to understand the biases and limitations of the survey.

In this work, I aim to remove the influence of the selection function on the galaxy density field, as described in Section 4.2. Therefore, in this section, I provide a detailed description of the selection function. Additional information on the construction of the VIPERS selection function can be found in Guzzo et al. (2014); Scodeggio et al. (2018).

The VIPERS selection function is the result of several angular completeness functions: i) binary masks, and ii) completeness weights for each galaxy. The average effect of the selection function is used therefore to weight the redshift distribution.

Binary mask

Scodeggio et al. (2018) describes two binary masks, the photometric mask, related to defects in the parent photometric sample, mostly areas masked by bright stars, and the VIMOS footprint mask, which corresponds to the specific footprint of VIMOS and how the different pointings are tailored together to mosaic the VIPERS area.

- **Photometric mask**: The photometric quality across the CFHTLS images is tracked with a set of masks that account for imaging artifacts and non-uniform coverage. Regions with corrupted source extraction or degraded photometric quality are excluded from the survey area. This mask consists primarily of patches around bright stars ($B_{Vega} < 17.5$) owing to the broad diffraction pattern and internal reflections

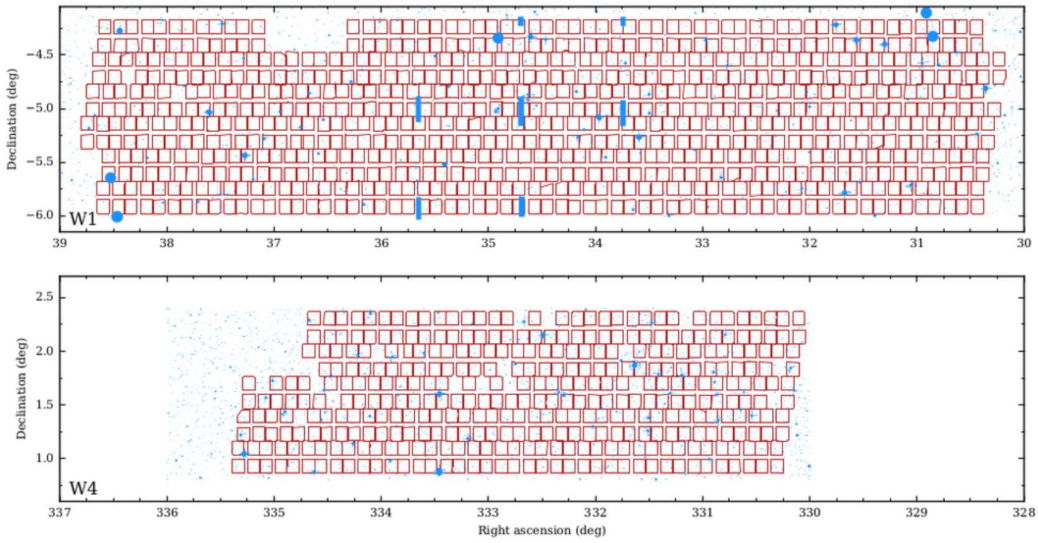


Figure 4.2: The image shows the footprint in the sky of the VIPERS W1 and W4 fields. The contours of each of the four quadrants that comprise all VIMOS pointings are shown in red. The blue (grey) areas in the background correspond to areas where the parent photometry is corrupted or observations are not possible due to the presence of foreground objects, particularly bright stars, and their diffraction spikes. Scodeggio et al. (2018)

in the telescope optics. At the core of a saturated stellar halo, there are no reliable detections, leaving a hole in the source catalog, while in the hale and diffraction spikes, spurious sources may appear in the catalog due to false detections. Extended galactic sources that may be fragmented into multiple detections or that may obscure potential VIPERS sources are also added to the mask. This mask corresponds to the blue regions in Fig. 4.2.

- **VIMOS footprint**: The general layout of VIMOS is well known, but the precise geometry of each quadrant has to be specified carefully for each observation, to perform precise clustering measurements with the VIPERS data. This footprint is composed of 288 VIMOS pointings, 192 over the W1 area, and 96 over the W4 area of the CFHTLA, overlapping a total sky area of about 23.5 square degrees. Due to the specific design of VIMOS, failed quadrants, and masked regions, this corresponds to an effectively covered area of 16.3 square degrees. This footprint is shown in red in Fig 4.2.

Completeness functions

The completeness function, as described on Guzzo et al. (2014); Pezzotta et al. (2017), is given by three factors: the color sampling rate (CSR), the target sampling rate (TSR), and the spectroscopic measurement success rate (SSR). The completeness value for each galaxy with $z > 0.6$ is given by the product of the TSR and the SSR, $c = c_{TSR}c_{SSR}$.

- **Color Sampling Rate (CSR)**: The pre-selection of VIPERS targets has been done through a color-color selection, applied to ideally isolate $z > 0.5$ galaxies using the CFHTLS corrected photometry. Using the data from the VVDS survey, the color sampling rate (CSR) was estimated as a function of redshift. In Figure 4 of Scodeggio et al. (2018) it is shown how the VIPERS catalog is virtually 100% complete above $z = 0.6$ when compared to a corresponding purely magnitude-limited sample.

- **Target Sampling Rate (TSR)**: A Multi-Object Spectrograph (MOS) survey inevitably has to deal with the limitation of MOS slits creating a shadow effect in the targeting of potential sources that are strongly density-dependent. In practice, the high-density peaks of the projected galaxy density field are under-sampled concerning the low-density regions, because the MOS slit length imposes a minimum angular pair separation in the spectroscopic target selection. The slit assignment for the VIMOS spectrograph was optimized using the SSPOC algorithm with the primary constraint that spectra cannot overlap on the focal plane, which maximizes the number of slits observed in each quadrant (Bottini et al. 2005). As a result, (a) very close pairs below a certain scale are practically unobservable; (b) the angular distribution of slits is more uniform than the underlying galaxy distribution. In VIPERS, the first effect suppresses angular clustering below a scale of 5 arcsec, producing a scale-dependent damping of the observed clustering below $\sim 1h^1$ Mpc; the second is instead responsible for a nearly scale-independent reduction of the two-point correlation function amplitude above this scale. This suppression is corrected by up-weighting galaxies based on the TSR with the factor $w_i = 1/c_{TSR,i}$ where $c_{TSR,i}$ depends on the local density of sources in the sky.

- **Spectroscopic Success Rate (SSR)**: The success rate depends primarily on the observational conditions and the flux of each source, and is defined as the ratio between the number of objects for which there is a successfully measured redshift N_{sucess}, and the number of objects targeted by the spectroscopic observations N_{target}. The SSR is computed using a nearest-neighbor algorithm in an N-dimensional space, using as parameters: apparent magnitude, rest-frame color, luminosity, and quadrant quality. For each object in this N-dimensional space, it is determined the distance R_K to its Kth nearest neighbor. N_{sucess} corresponds to the number of successful redshift measurements, defined based on the redshift quality flag, within the R_K distance. N_{target} corresponds to the number of nearest neighbors considered, which for VIPERS is K=100.

Redshift distribution

The combined galaxy weight $w = (c_{TSR}c_{SSR})^{-1}$ corrects for both the TSR and SSR effects. The weighted redshift distribution is shown in Fig. 4.3. The effective redshift of the sample is defined using the distribution of galaxy pairs which can be approximated based on the

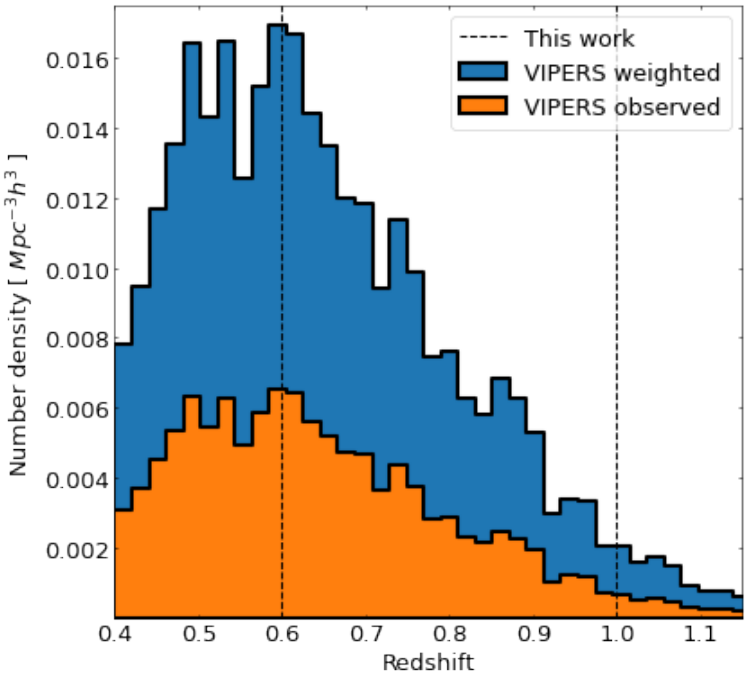


Figure 4.3: Mean spatial density of galaxies as a function of redshift for the final VIPERS sample used in this work. The lower histogram is the observed distribution, while the top one shows the effect of the completeness corrections including the spectroscopic success rate (SSR) and target sampling rate (TSR). The vertical dashed lines define the boundaries of the redshift range 0.6-1.0 used in this analysis.

weighted redshift distribution, $n(z)$, in the range $0.6 < z < 1.0$, as,

$$z_{\rm eff} = \frac{\sum_i n_i^2 z_i}{\sum_i n_i^2} \,. \tag{4.1}$$

with bin size $\Delta z = 0.02$ and where i indexes the redshift bins and z_i is the midpoint of the bin. The resulting effective redshift of the sample is $z_{\rm eff} = 0.71$.

4.1.4 VIPERS mocks

To quantify the level of systematic biases in the final results, I used a set of simulated VIPERS galaxy catalogs, produced by the VIPERS collaboration (de la Torre et al. 2013). These mock catalogs were built using the MultiDark N-body simulation (BigMD, Klypin et al. 2016) with the Planck cosmology (Ω_M, Ω_Λ, Ω_B, h, n_s, σ_8) = (0.307, 0.693, 0.0482, 0.678, 0.960, 0.823). A halo occupation distribution (HOD) prescription was applied to add galaxies to the dark matter haloes which were calibrated from the VIPERS data. In my work, I use 153 independent mock catalogs, taken from the VIPERS collaboration.

The mock catalogs match the number density of the VIPERS catalog after correcting for the TSR and SSR selection effects. They reproduce the angular survey mask but

do not include the spectrograph slit placement constraints. Therefore the TSR and SSR completeness weights are not applied in the mock analyses. The redshift measurement error was drawn from a Gaussian distribution with standard deviation $\sigma_z = 0.00054(1+z)$.

4.2 Bayesian Code for Unmasking Galaxy Surveys

The most important aim of a cosmological survey is the comparison of the observations with the predictions of a cosmological model, in particular, the ΛCDM model. The power spectrum is the main tool to evaluate this likelihood between theory and observations. The power spectrum estimation when the sky coverage is incomplete and non-homogeneous instrumental noise are problems of central importance for the extraction of cosmological information from a galaxy survey. I implemented a Bayesian algorithm to clean the masked data and recover cosmological information. This chapter is aimed to explore the theory that supports the numerical approach used in the main code.

Granett et al. (2015) carried out the first maximum-likelihood analysis of the VIPERS density field on a subset of the final data release. That analysis focused on optimizing the measurement of the matter power spectrum accounting for the dependence of the galaxy bias on galaxy luminosity and color. I developed a novel estimator that solves the maximum-likelihood problem by Monte Carlo sampling, improving the Granett et al. (2015) solution. My code marginalizes over the density field and samples the following parameters as free parameters: the matter density (sampled as $\Omega_M h$, with h fixed), the baryon fraction ($f_b = \Omega_B/\Omega_M$), the galaxy bias (b), the velocity dispersion (σ_v), and the combination of σ_8 and the growth rate ($f\sigma_8$). The priors for these parameters are provided in Table 4.1.

The first complete cosmological analysis of VIPERS was presented in Rota et al. (2017), which analyzed the galaxy power spectrum monopole using the FKP (Feldman et al. 1994). The FKP is based on the assumption that the observed galaxy positions are a random sampling of the underlying density field, and that the density field can be modeled as a Gaussian random field. The FKP estimator first computes the power spectrum of the observed density field and then deconvolves the noise effects induced by the selection function, my approach aims to denoise directly the observed density field, and then compute the power spectrum. I can evaluate the performance of both methods by comparing Ω_m and Ω_b in my results with Rota et al. (2017).

The growth rate in combination with the linear clustering amplitude, $f\sigma_8$, was measured from VIPERS based on the analysis of the correlation function multipoles (Pezzotta et al. 2017; Mohammad et al. 2018) and the void-galaxy correlation function (Hawken et al. 2017). Inferring the growth rate $f \equiv -d\log\sigma_8(z)/d\log(1+z)$ alone requires breaking the degeneracy between σ_8 and galaxy bias and was carried out with VIPERS through the joint analysis of the correlation function multipoles with galaxy-galaxy lensing (de la Torre et al. 2017) and three-point correlation statistics (Veropalumbo et al. 2021). Here the analysis is limited to two-point clustering, so there is a restriction to analyze the parameter combination $f\sigma_8$. Compared with the previous VIPERS studies based on the galaxy power spectrum and multipoles of the correlation function, the analysis presented

here has a greater freedom in the model since it simultaneously fits the matter density parameters and fσ_8.

4.2.1 Monte Carlo Markov Chains

A Monte Carlo Markov Chain (MCMC) is a class of algorithms for sampling from a probability distribution. It creates samples from a multi-dimensional continuous random variable, with probability density proportional to a known function. These samples can be used to evaluate an integral over that variable, as its expected value or variance.

An ensemble of chains is developed, starting from a set of points arbitrarily chosen and sufficiently distant from each other. These chains are stochastic processes of 'walkers' which move around randomly according to an algorithm that looks for places with a reasonably high contribution to the integral to move into next, assigning them higher probabilities. The random samples of the integrand used in an MCMC are autocorrelated. The algorithm creates Markov chains having an equilibrium distribution that is proportional to the function given.

The most important advantage of Bayesian data analysis is that it is possible to marginalize nuisance parameters. A nuisance parameter is required to model the process that generates the data but is otherwise of little interest. Marginalization is the process of integrating all possible values of the parameter and hence propagating the effects of uncertainty about its value into the final result. The exact result of marginalization is the probability function $p(\Theta|D)$ of the set of model parameters Θ given the set of observations D:

$$p(\Theta|D) = \int p(\Theta, \alpha|D) d\alpha, \tag{4.2}$$

where α is the set of nuisance parameters. Because the nuisance parameter set can be very large, this integral is often extremely daunting. In addition to the problem of marginalization, the MCMC could be very valuable obtaining the likelihood or the prior of an expensive simulation.

The general goal of MCMC algorithms is to draw M samples $\{\Theta_i\}$ from the posterior probability density

$$p(\Theta, \alpha|D) = \frac{1}{Z} p(D|\Theta, \alpha) p(\Theta, \alpha), \tag{4.3}$$

where the prior distribution $p(\Theta, \alpha)$ and the likelihood function $p(D|\Theta, \alpha)$ can be relatively easily computed for any particular value of (Θ_i, α_i). The normalization $Z = p(D)$ is independent of Θ and α once it is chosen the form of the generative model. This means that it is possible to sample from $p(\Theta, \alpha|D)$ without computing Z. This is important because Z is generally very expensive to compute.

Once the samples produced by MCMC are available, the marginalized constraints on Θ can be approximated by the histogram of the samples projected into the parameters subspace spanned by Θ. In particular, this implies that the expectation value of a function

of the model parameters $f(\Theta)$ is

$$\langle f(\Theta)\rangle = \int p(\Theta|D) f(\Theta) d\Theta \approx \frac{1}{M}\sum_{i=1}^{M} f(\Theta_i). \tag{4.4}$$

Generating the samples Θ_i is a non-trivial process unless $p(\Theta, \alpha, D)$ is a very specific analytic distribution. MCMC is a procedure for generating a random walk in the parameter space that, over time, draws a representative set of samples from the distribution. Each point in a Markov chain $X(t_i) = [\Theta_i, \alpha_i]$ depends only on the position of the previous step $X(t_{i-1})$.

4.2.2 Density field

To compute the density field reconstruction, I started by projecting the galaxy sample onto a comoving cartesian grid with cubic cells of 5 $Mpc\,h^{-1}$ side. The dimensions of the grid are $77 \times 16 \times 184$ cells, roughly corresponding to RA × DEC × redshift. I determined this box size to contain the largest possible light cone, corresponding to the VIPERS W1 field within the moving distances allowed by the prior. I use the same box size for both VIPERS fields, W1 and W4, to consistently compute the Fourier Transform with the same grid dimensions. The number of galaxies observed in a given grid cell is parametrized by

$$n_{obs,i} = \bar{n}_i(1 + \delta_{g,i}) + \epsilon_i, \tag{4.5}$$

where $\bar{n}_i$ is the expected number in the cell given the survey selection function and $\delta_{g,i}$ is the underlying galaxy over-density field in redshift space and ϵ_i is the noise contribution. $n_{obs,i}$ is computed inside the survey volume according to the nearest grid point assignment scheme, with zeros outside the survey area. I assumed that $n_{obs,i}$ is a Poisson process such that the variance is $\sigma^2_{\epsilon,i} = \langle \epsilon_i^2 \rangle = \bar{n}_i$. I generated the mean number of objects in the cell, $\bar{n}_i$, using a random catalog that was built following the VIPERS cumulative redshift distribution. The redshift distribution weighted by the inverse of the completeness was combined from the two fields W1 and W4 and smoothed with a Gaussian kernel with a width in the redshift of $\sigma = 0.07$. The matter density field is related to the galaxy field by a constant bias factor b: $\delta_g = b\delta$ which is evaluated at the effective redshift. The use of an effective redshift and bias to model the power spectrum (see Eq. 4.7) is warranted because the amplitude parameterized by $b(z)\sigma_8(z)$ does not evolve significantly in the VIPERS sample, as was shown in Rota et al. (2017).

The statistical properties of the density field δ are described by the power spectrum of the Fourier modes: $P_i = \langle |\tilde{\delta}_i|^2 \rangle$. The Fourier transform of δ is defined by

$$\tilde{\delta}(\vec{k}) = \frac{1}{2\pi}\int d^3\vec{x}\,\delta(\vec{x}) e^{-i\vec{k}\vec{x}}, \tag{4.6}$$

which is evaluated on a discrete coordinate grid using the fast Fourier transform (FFT) algorithm.

In the flat ΛCDM model, the linear matter power spectrum $P_m(k)$ depends on redshift, Ω_m, Ω_b, H_0 and σ_8. It adopted the non-linear model of the matter power spectrum using the CLASS code with the Halofit prescription (Blas et al. 2011; Takahashi et al. 2012).

To model the observations, the anisotropic redshift-space power spectrum is parametrized using the dispersion model, i.e. the Kaiser linear expression modified to empirically account for the velocity dispersion (Peacock & Dodds 1994):

$$P^s(\vec{k}) = \left(b + f\mu^2\right)^2 e^{-k_{||}^2\sigma_v^2} P_m(|\vec{k}|)\,. \tag{4.7}$$

The free parameters are the galaxy bias b, the logarithmic growth rate f, and the velocity dispersion σ_v, while the nonlinear matter power spectrum $P_m(k)$ depends on the cosmological parameters. The effective velocity dispersion term σ_v accounts for the sum of both the true pairwise dispersion along the line of sight and the redshift measurement error, which sum in quadrature. The parameter $\mu = k_{||}/k$ is the cosine of the angle of the wavevector $\vec{k}$ concerning the line of sight.

Since the density field is estimated on a discrete grid, the power spectrum is affected by aliasing (Jing 2005). Aliasing occurs when a signal is sampled at a rate that is lower than its Nyquist rate, which is the minimum sampling rate required to accurately reproduce the original signal. As a result, the sampled signal appears to have a lower frequency than it does. An aliasing correction is implemented by including in the power spectrum model the effect arising from the nearest-grid point mass assignment:

$$P_{grid}(\vec{k}) = \sum_n |W(\vec{k} + 2k_N\vec{n})|^2 P^s(\vec{k} + 2k_N\vec{n}), \tag{4.8}$$

where the summation includes the first harmonic so the three-dimension grid $\vec{n} = (n_x, n_y, n_z)$ where each element takes integer values -1, 0, and 1. The window function W corresponds to the *sinc* function for the nearest grid point mass assignment scheme. This treatment differs from Granett et al. (2015) which did not include an aliasing correction but reduced the effect by using super-sampling when binning galaxies onto the grid. Here computation time when building the grid is saved, and instead, evaluate the aliasing terms in the likelihood. A shot noise term is not included in the power spectrum model, but instead, a shot noise correction is included in the estimator (see Sec 4.2.5).

4.2.3 Estimator

An estimator for the model parameters is constructed starting from the observed galaxy number counts. The posterior distribution function of the set of cosmological parameters, $\vec{\Omega}$, is:

$$p(\vec{\Omega}|\vec{n}_{obs}) \propto p(\vec{n}_{obs}|\vec{\Omega})\; p(\vec{\Omega}). \tag{4.9}$$

The expectation of the posterior distribution provides an estimate of the parameters:

$$\langle\vec{\Omega}\rangle = \int \vec{\Omega} p(\vec{\Omega}|\vec{n}_{obs}) d\vec{\Omega}. \tag{4.10}$$

To evaluate this, the underlying density field is introduced as a latent variable. Marginalizing over this field $\vec{\delta}$, it can be written,

$$p(\vec{\Omega}|\vec{n}_{obs}) = \int p(\vec{\Omega}|\vec{n}_{obs},\vec{\delta})p(\vec{\delta})d\vec{\delta} \quad (4.11)$$

$$= \int p(\vec{\Omega},\vec{\delta}|\vec{n}_{obs})d\vec{\delta}. \quad (4.12)$$

The integrand can be written as

$$p(\vec{\Omega},\vec{\delta}|\vec{n}_{obs}) = p(\vec{n}_{obs}|\vec{\delta})\ p(\vec{\delta}|\vec{\Omega})p(\vec{\Omega}). \quad (4.13)$$

The first term $p(\vec{n}_{obs}|\vec{\delta})$ is the data likelihood, which will be described by a multivariate Gaussian distribution:

$$p(\vec{n}_{obs}|\vec{\delta}) = \left((2\pi)^N \Pi_i \bar{n}_i\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\sum_i \frac{\left[n_{obs,i} - \bar{n}_i(1+b\delta)\right]^2}{\bar{n}_i}\right). \quad (4.14)$$

The second term $p(\vec{\delta}|\vec{\Omega})$ represents the Bayesian prior. The power spectrum of δ is evaluated on the grid, $P_{grid,i} = \langle|\tilde{\delta}_i|^2\rangle$ which depends implicitly on the model parameters: $P_{grid,i} \equiv P_{grid,i}(\vec{\Omega})$. This formulation results in a Gaussian prior on δ:

$$p(\vec{\delta}|\vec{\Omega}) = \left((2\pi)^N \Pi_i P_{grid,i}\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\sum_{i<k_{max}} \frac{|\tilde{\delta}_i|^2}{P_{grid,i}}\right). \quad (4.15)$$

The Fourier transform of the density field $\tilde{\delta}$ was computed with by FFT without zero padding. In practice it is not evaluated the prior using all modes to the Nyquist frequency but applied a limit, k_{max}. In the analysis it is setted $k_{max} = 0.4hMpc^{-1}$ and the Nyquist frequency is $k_N = 0.6hMpc^{-1}$.

4.2.4 Gibbs sampler

A Gibbs sampler is a Monte Carlo Markov Chain algorithm for obtaining a sequence of observations that are approximated from a specified multivariate probability distribution function when direct sampling is difficult. This sequence can be used to approximate the marginal distribution of some subset of variables -unknown parameters or latent variables-. The algorithm is applicable when the joint distribution is not known explicitly or is difficult to sample from directly, but the conditional distribution of each variable is known and easy to sample from. The Gibbs sampling algorithm generates an instance from the distribution of each variable in turn, conditional on the current values of other variables. It can be shown that the sequence of samples constitutes a Markov chain, and the stationary distribution of that Markov chain is just the sought-after joint distribution.

The high dimensionality of the posterior distribution

$$p(\vec{\delta}, b, f\sigma_8, \sigma_v, \Omega_m h, f_b|\vec{n}_{obs}), \quad (4.16)$$

makes it challenging to sample directly. I employed the Gibbs sampling algorithm to divide the problem into two steps. My algorithm first draws a realization of the density field δ with the parameters fixed through the Wiener filter (i), and next sample the parameters with δ fixed through a secondary MCMC chain (ii). Iterating these two steps in sequence allows us to draw samples from the joint distribution. Schematically, the two steps are written below.

$$\delta_{\gamma+1} \leftarrow p(\delta|\vec{n}_{obs}, P_\gamma), \tag{i}$$

$$\begin{pmatrix} b_{\gamma+1} \\ f\sigma_{8\gamma+1} \\ \sigma_{v\gamma+1} \\ \Omega_m h_{\gamma+1} \\ f_{b\gamma+1} \end{pmatrix} \leftarrow p(b, f\sigma_8, \sigma_v, \Omega_m h, f_b|\delta_{\gamma+1}), \tag{ii}$$

where P_γ is the power spectrum computed at the step γ. This approach differs from the sampling scheme implemented in (Granett et al. 2015). In that work, the bins of the power spectrum were first sampled and only fit the cosmological parameters after sampling. Instead, here, the power spectrum is parametrized by Eq. 4.7 and evaluated with the CLASS code on every iteration. This approach avoids binning the power spectrum, since it is evaluated on the Fourier grid, and guarantees self-consistency of the power spectrum estimation and cosmological parameter constraints.

4.2.5 Wiener filter

The first Gibbs sampling step i corresponds to generating a density field that maximizes the posterior distribution $p(\delta|\vec{n}_{obs}, P_\gamma)$. The Wiener filter (Rybicki & Press 1992) is the tool used in this reconstruction of the density field.

The information given by VIPERS is incomplete in several regions due to the geometry of the VIMOS instrument, the density of targets, the presence of brilliant stars, or data-taking errors. For this reason, it is necessary to compute an *unmasked* density field.

The solution δ^{WF} is found by solving the linear equation:

$$\sum_j \left(S_{ij}^{-1} + N_{ij}^{-1}\right)\, \delta_j^{WF} = \sum_j N_{ij}^{-1}\, \delta_{g,j}, \tag{4.17}$$

here, S and N are the signal and noise covariance matrices. Under the model assumptions, the signal is diagonal in Fourier space and can be written in terms of the power spectrum: $\tilde{S}_{ii} = P_i$. Conversely, the noise matrix is diagonal in configuration space: $N_{ii} = \bar{n}^{-1}$. As P_i and $\bar{n}_i$ are diagonal matrices in different bases, it is computationally expensive to operate simultaneously with them and compute the inverse of their sum. Instead, here it is applied the iterative conjugate gradient solver to estimate the solution δ_i^{WF}. In practice, it is first computed the product $P_i^{-1}\, \delta_i^{WF}$ in Fourier space, and by using an inverse Fourier transform, return to configuration space, where $\bar{n}_i$ and $\delta_{g,i}$ are diagonal matrices.

While the solution δ^{WF} maximizes the posterior function and therefore carries the full information of the density field, it lacks power on small scales. To sample a full realization from the posterior distribution, it is constructed the constrained fluctuation field χ which is uncorrelated with δ^{WF}. The constrained realization $\delta = \delta^{WF} + \chi$ samples from the posterior distribution and recovers the target power spectrum. The constrained fluctuation field is found by solving the equation (Jewell et al. 2004):

$$\sum_j \left(S_{ij}^{-1} + N_{ij}^{-1}\right)\, \chi_j = \sum_j \left(S_{ij}^{-1/2}\xi_{1,j} + N_{ij}^{-1/2}\xi_{2,j}\right), \tag{4.18}$$

where ξ_1 and ξ_2 are uniform random fields between 0 and 1 in configuration space. It is solved Eq 4.18 using the conjugate gradient method.

4.2.6 Likelihood on Cosmological Parameters

The second Gibbs sampling step ii corresponds to obtaining a new set of parameters Ω based on a given realization of the density field δ. The posterior distribution of the parameters is

$$p(\Omega|\delta) = p(\delta|\Omega)p(\Omega), \tag{4.19}$$

where the conditional probability $p(\delta|\Omega)$ is computed with Eq. 4.15.

To integrate over the posterior distribution $p(\Omega|\delta)$, it is employed a Markov Chain Monte Carlo (MCMC) algorithm based on the `Emcee` ensemble sampler (Foreman-Mackey et al. 2013). The likelihood $p(\delta|\Omega)$ provided to the `Emcee` code is described in equation 4.15. A flat prior $p(\Omega)$ is used on the parameters, described in table 4.1, which allows the walkers to explore completely the physical configurations of the system. In this work, different from Granett et al. (2015), all parameters are sampled jointly.

To explore the five-dimensional parameter space, the `Emcee` ensemble sampler was configured with 20 walkers and 20 steps for each walker, starting from randomly selected points from within the prior. The final position of one of these walkers was employed as the input for the next step of the Gibbs chain, allowing the algorithm to explore the whole prior region.

4.3 Computational structure

The Gibbs sampling algorithm is applied to the masked data from VIPERS and corresponding mock catalogs. Both the real and mock catalogs are treated identically, except that the TSR and SSR completeness corrections are not applied to the mocks. The input data catalog includes the following galaxy parameters: right ascension, declination, redshift, quality flags (including TSR and SSR weights for real galaxies), and the polygon mask of the survey. The steps of the analysis are as follows:

- *Step 1*: Calculate the *masked* galaxy density field on the Cartesian comoving grid. A starting value for Ω_M is required to compute the comoving distances of the galaxies

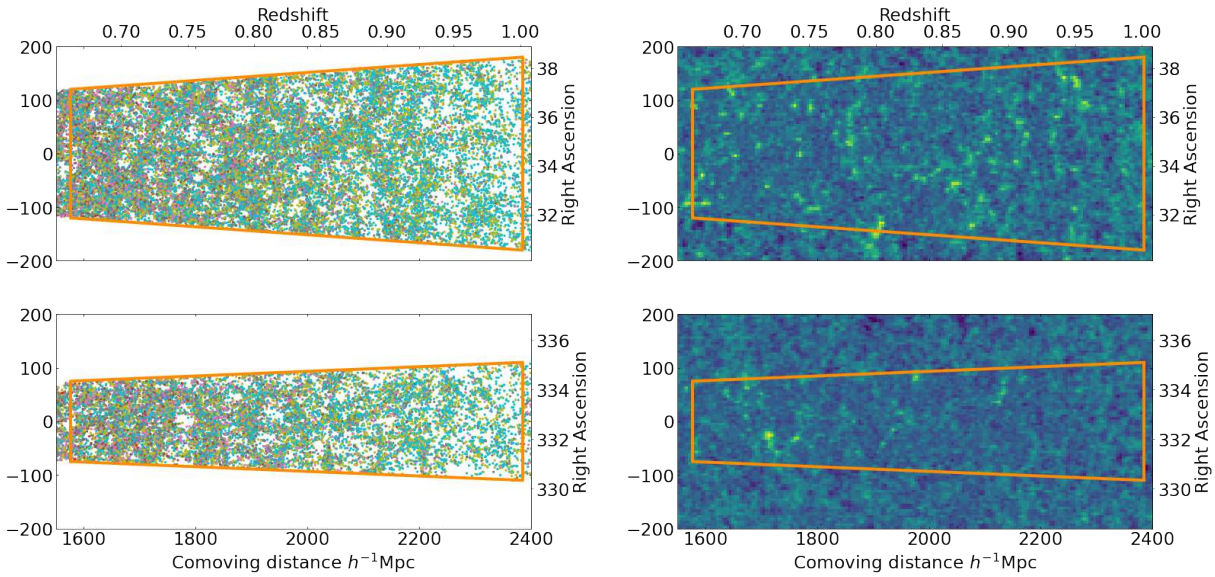


Figure 4.4: VIPERS cone diagrams for the fields W1 (top) and W4 (bottom). *Left*: redshift-space positions of observed galaxies. The orange line traces the field boundaries cut in the redshift direction at $0.6 < z < 1.0$. The graph includes all galaxies projected along the declination. Each galaxy is represented by a filled circle colored according to its i band luminosity. *Right*: a 10 $h^{-1}Mpc$ slice of the reconstructed density field taken from one step of the Gibbs sampler. It represents the anisotropic Wiener reconstruction from the weighted combination of galaxy tracers. The field is filled with a constrained Gaussian realization in the volume obscured by the survey mask.

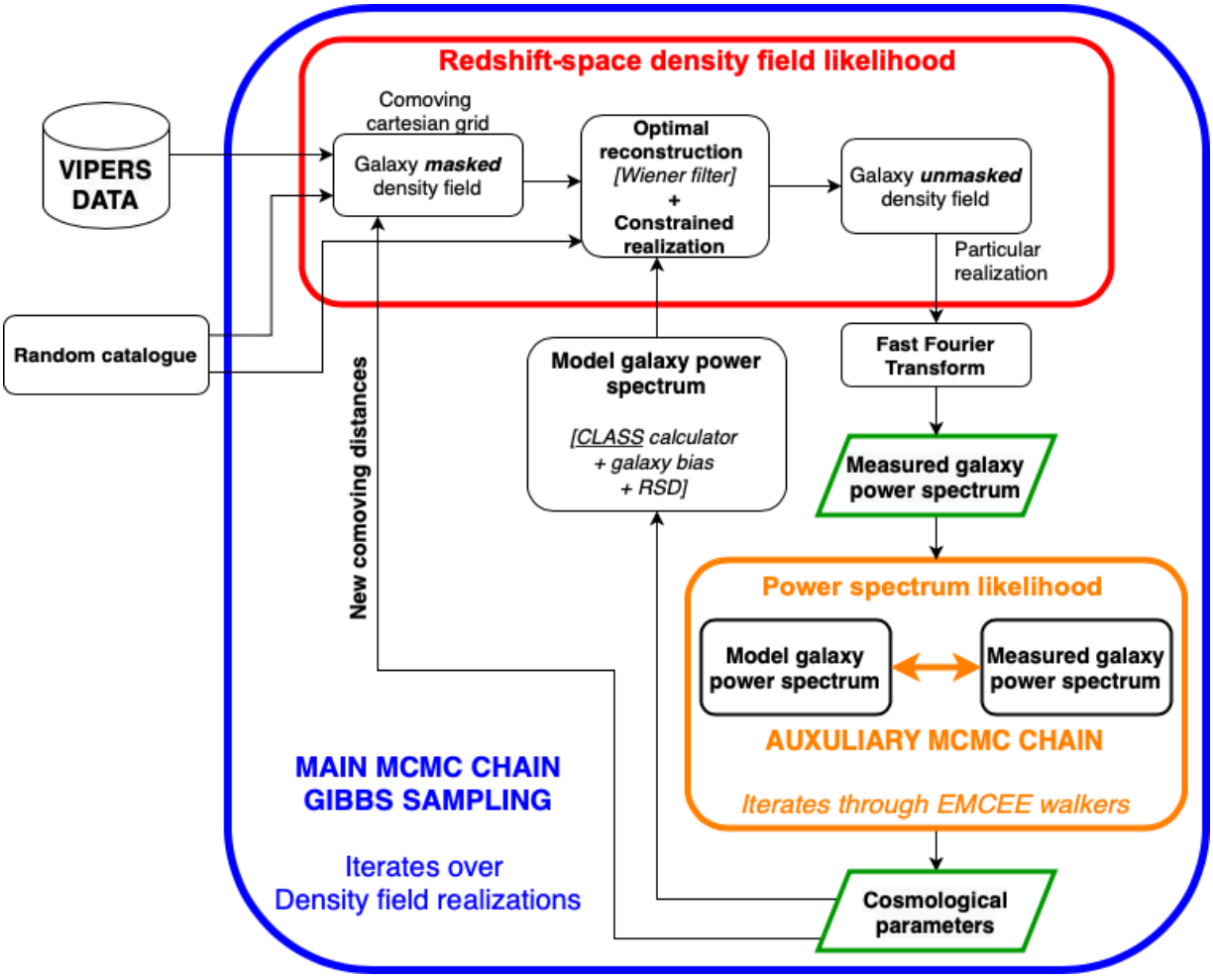


Figure 4.5: A flow chart summarising the structure of the code. The input includes VIPERS data (RA, DEC, z, TSR, and SSR) and the random catalog, which is generated using the angular mask and redshift distribution and accounts for the mean density by cell. The Gibbs sampler is represented by the external blue box. The first section of the chain, described by equation i is represented by the red box (Section 4.2.5). The second section of the chain, corresponding to equation ii is represented by the orange box (Section 4.2.6). The outputs of the code are the galaxy power spectrum and the set of cosmological parameters (internal green boxes).

Parameter	*Prior*
Matter density	$0.1 < \Omega_M h < 0.3$
Baryon fraction	$0 < f_B < 0.3$
Galaxy bias	$1 < b < 2$
Velocity dispersion	$1.5 < \sigma_v < 3.5$
Linear growth rate	$0.25 < f\sigma_8 < 0.55$

Table 4.1: Flat prior on the cosmological parameters sampled in this work. The walkers described in section 4.2.6 are free to move within these intervals.

using a ΛCDM model. In subsequent iterations, the value of Ω_M and the corresponding comoving distances for each galaxy are updated for consistency.

- *Step 2*: Generate a particular realization of the *unmasked* galaxy density field using the Wiener filter and constrained realization.
- *Step 3*: Calculate the power spectrum of the particular realization of the galaxy density field.
- *Step 4*: Using the MCMC `Emcee` sampler, generate samples of the model parameters (Eq. 4.15).

The output of the Gibbs sampler on each iteration step is the set of model parameters, the measured power spectrum, and the realization of the density field.

The analysis of the VIPERS data was iterated for 1000 steps, while 600 steps were used for the mock catalogs. The first 500 steps (200 for mocks) were discarded as the burn-in period. The main purpose of computing the algorithm on the mocks is to estimate the dispersion of the results, so it can be used in a lower number of steps, saving computational resources. In the left side of Fig. 4.4 is represented the VIPERS lightcone, which is the input of the code, while in the right side of the same figure is represented one reconstruction of the galaxy density field. In Fig. 4.5 is presented a full diagram of the code.

4.4 Results

4.4.1 Galaxy power spectrum

The left panel of Fig. 4.6 shows the galaxy power spectrum monopole in redshift space, measured from the unmasked galaxy density field. The shaded region indicates the 68% confidence interval estimated from the samples of the Gibbs chain. The results from the mocks and VIPERS data are shown in orange and blue, respectively. Excellent agreement is found between the mocks and VIPERS data at $k < 0.2$. At larger values of k the mocks show a higher power spectrum amplitude than the data. This divergence is attributed to a systematic enhancement in the velocity dispersion in the mocks compared with the

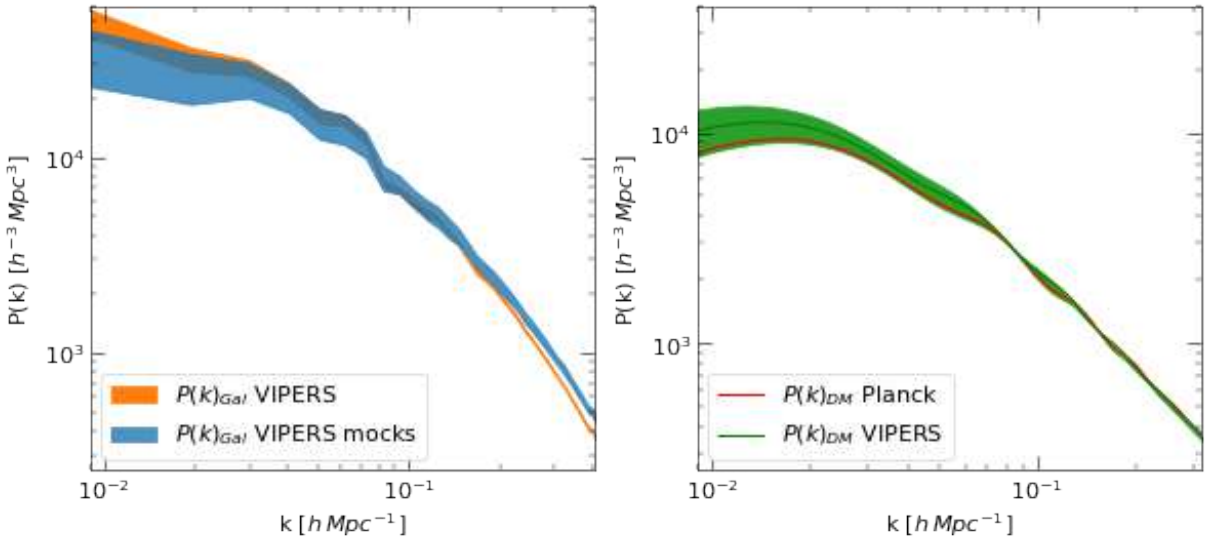


Figure 4.6: The power spectrum computed for the whole VIPERS sample, W1 and W4 fields using the redshift interval $0.6 < z < 1.0$. The left panel shows the power spectrum monopole in redshift space measured from the unmasked galaxy density field. The right panel shows the constraints on the model dark matter power spectrum in real space computed with the parameters from the Gibbs sampler.

data. This excess of velocities along the line of sight is reflected in the loss of power at small scales, where the velocity fluctuations are more significant. It is also noted that the confidence interval is significantly smaller in the data compared with the mocks; this is because the Gaussian likelihood underestimates the covariance which is instead captured by the scatter in the estimates from the 153 mock realizations. The same trend was observed in (Granett et al. 2015); however, in that work, the galaxy sample was divided into many sparser subsamples each with higher levels of shot noise that led to the overall more Gaussian behavior.

Next, it is considered the power spectrum model evaluated with the parameters obtained with the Gibbs sampler. The constraints on the parameters themselves will be presented in the following section. The right panel of Fig. 4.6 shows the posterior region of the non-linear matter power spectrum model in real space. The green shaded area shows the 68% confidence interval found with the VIPERS data, which shows agreement with the Planck model. The phase shift that can be seen in the Baryon acoustic oscillations is an artifact of the power spectrum model. The BAO signal is not detected in the VIPERS power spectrum, as seen in the left panel, due to the survey geometry (see also the discussion of the window function in Rota et al. 2017). Although the unmasked density field has been successfully evaluated in this analysis, the anisotropic shape of the window leaves a fundamental uncertainty in the estimate of the underlying power spectrum that washes out the BAO signal.

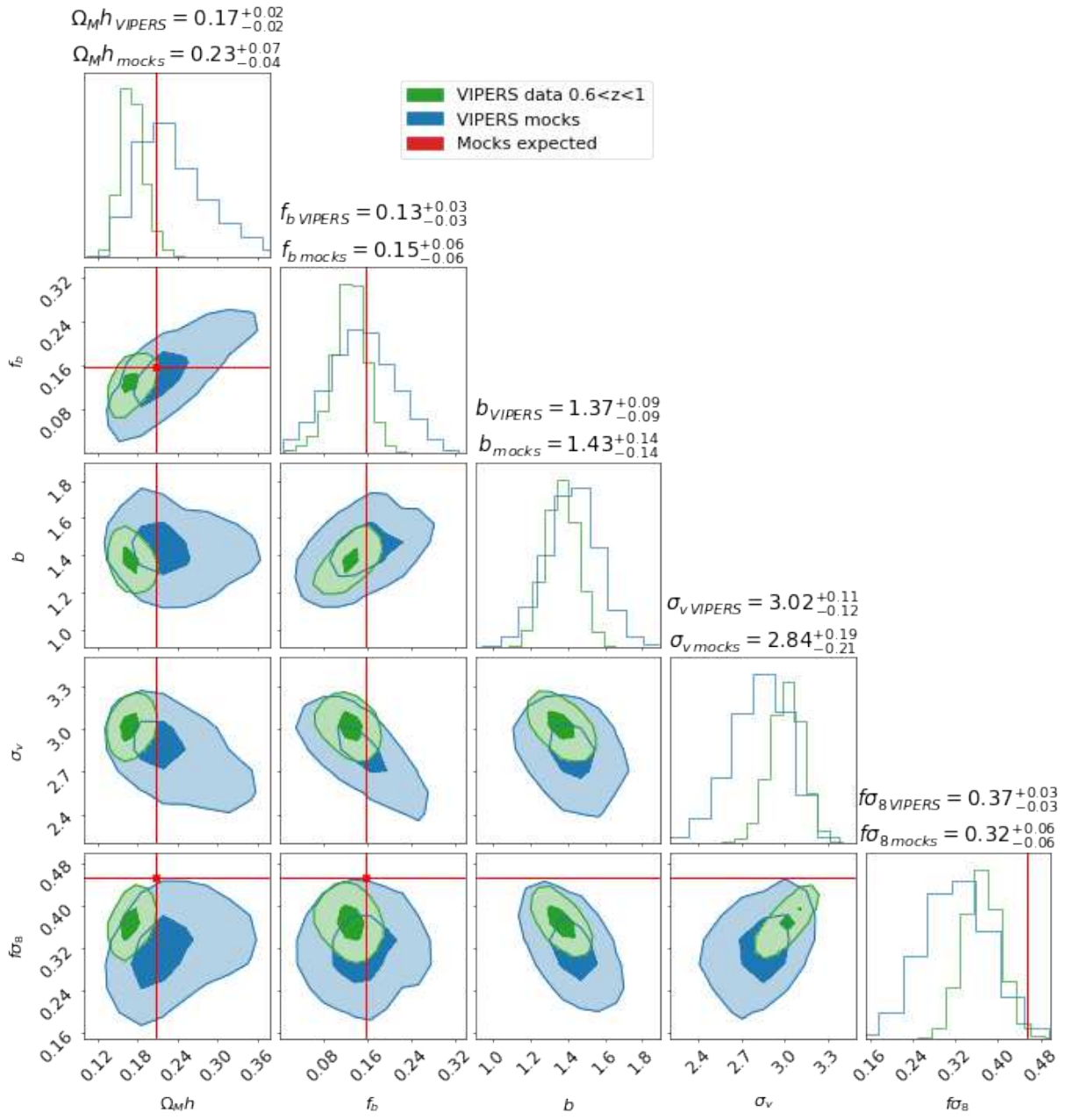


Figure 4.7: Joint probabilities obtained with the Gibbs sampler applied to the VIPERS data and mock catalogs. The green contours are the results obtained with the VIPERS data. The blue contours are obtained from the mock catalogs: the mean of the mocks distribution provides an estimate of the accuracy of the algorithm, while its dispersion gives a realistic estimate of the precision of the results. In this respect, the mean of the mock estimates is determined with a precision that is $\sim \sqrt{153}$ times smaller than the uncertainty shown by the blue contours (which represent a single realization). The red lines are the values used to generate the mocks according to the BigMD simulation using Planck values. The contours correspond to the 39th and 86th percentiles.

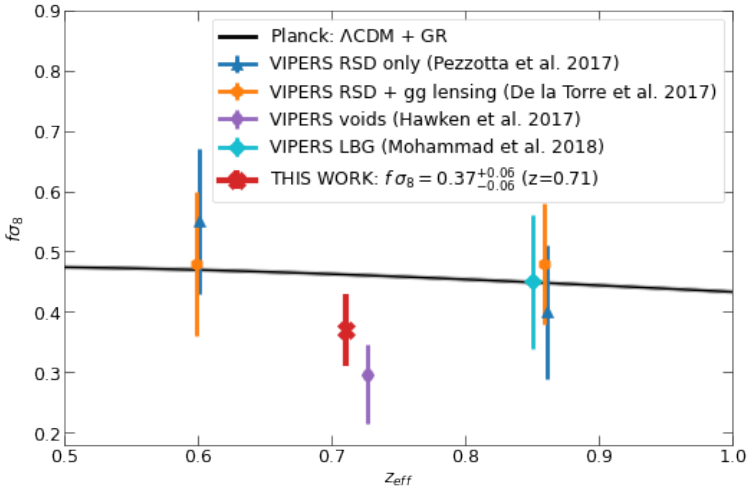


Figure 4.8: Estimate of the linear growth rate $f\sigma_8$ obtained here from the complete VIPERS galaxy sample (red marker), compared to other VIPERS measurements using different techniques. The black solid curve shows predictions of General Relativity with a ΛCDM model with parameters set to Plank 2018 (Planck Collaboration et al. 2020).

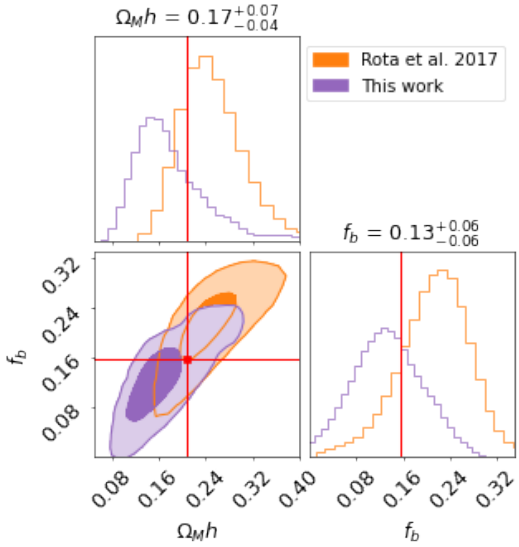


Figure 4.9: Comparison between the conditional probabilities on $\Omega_M h$ and f_b between this work and Rota et al. (2017). The shape of the contours of this work (magenta) is given by the mocks while the mean value is given by the real VIPERS data.. The Planck best-fit parameters are indicated by the horizontal and vertical lines. The contours correspond to the 39th and 86th percentiles.

4.4.2 Cosmological parameters

Fig. 4.7 shows the joint probabilities on cosmological parameters from the Gibbs chains computed on VIPERS (green) and mock catalogs (blue). The blue contours represent the concatenation of the chains from 153 mock catalogs after excluding the burn-in steps. The red lines are the values used to generate the mocks, following Planck's results. Reference values for the velocity dispersion, σ_v, and galaxy bias, b, are unknown because these parameters are determined by the N-body simulation, the HOD model, and the galaxy sample selection and cannot be precisely predicted analytically.

The analysis on the mocks recovers the values of the matter density, $\Omega_M h$, and Baryon fraction, f_b, demonstrating the soundness of the analysis. However, the growth rate of structure is underestimated at a level of 2.2σ.

Turning to the VIPERS data, it is constrained the matter density to $\Omega_M h = 0.17^{+0.07}_{-0.04}$. The error range given here was inferred from the scatter in the estimates from the mocks. The choice of sampling $\Omega_M h$ provides us with the opportunity of comparing directly the results with previous VIPERS measurements, in particular, those obtained by Rota et al. (2017) through the traditional FKP estimator: $\Omega_M h = 0.227^{+0.063}_{-0.050}$. This small tension at the 1σ level, visible in figure 4.9, will be discussed below.

The Baryon fraction obtained in this work is $f_B = 0.13^{+0.06}_{-0.06}$. Again it is found an agreement concerning Rota et al. (2017): $f_B = 0.220^{+0.058}_{-0.072}$. A comparison of the joint likelihood of the matter density and baryon fraction is plotted in Fig. 4.9. The size of the contour that is found is similar to Rota et al. (2017) but shifted by 1σ.

The galaxy bias is $b = 1.37^{+0.14}_{-0.14}$. The 1σ interval for this value overlaps with the b_{eff} found by Granett et al. (2015): $b_{eff} = 1.44^{+0.02}_{-0.02}$ using the same redshift interval and agrees with previous VIPERS analyses (Marulli et al. 2013; Di Porto et al. 2016). The velocity dispersion along the line of sight is measured to be $\sigma_v = 3.02^{+0.19}_{-0.21}$ which matches the value found with the mocks, when accounting for redshift measurement errors as seen in Fig. 4.7.

The anisotropic galaxy power spectrum can uniquely constrain the parameter combination $f(z)\sigma_8(z)$ in linear theory. It is found $f\sigma_8 = 0.37^{+0.06}_{-0.06}$ at $z = 0.71$, as shown in Fig. 4.8 together with the theory predictions in Planck cosmology. The measurement is lower than predicted and also lower than previous VIPERS measurements based on two-point correlation analysis (i.e. Pezzotta et al. 2017; de la Torre et al. 2017; Mohammad et al. 2018). This could be due to having modeled RSD using the dispersion model, which is known to underestimate the growth rate (Mohammad et al. 2018). The VIPERS reference RSD measurements were made on the correlation function, instead here it is fitted to the power spectrum model. A maximum scale $k_{max} = 0.4h\mathrm{Mpc}^{-1}$ is applied, which corresponds to a spatial scale of $\pi/0.4 \sim 0.8h^{-1}\mathrm{Mpc}$, although it is not equivalent to the minimum scale in the correlation which prevents us from making a direct comparison. However, that the estimate has a precision of $\Delta f\sigma_8/f\sigma_8 = 0.16$, in comparison to the previous VIPERS measurements from Pezzotta et al. (2017) in the redshift bin $0.5 < z < 0.7$, $\Delta f\sigma_8/f\sigma_8 = 0.22$ and $0.7 < z < 1.2$, $\Delta f\sigma_8/f\sigma_8 = 0.27$. This analysis gives better precision despite having greater freedom in the model because there has been jointly estimated the matter density parameter and it is not subject to geometric distortions (Marulli et al.

2012; Ballinger et al. 1996).

To test the robustness of the constraints, two additional analyses were performed on VIPERS with narrow priors imposed on selected parameters. First, to investigate whether the low value of $f\sigma_8$ is related to the freedom of the modeling, narrow priors on Ω_m and f_b were placed, leaving the bias and RSD parameters $f\sigma_8$ and σ_v free. The results of this more constrained configuration are equal to the original findings with $f\sigma_8 = 0.37^{+0.03}_{-0.03}$ and have consistent values of bias and σ_v. From this exercise can be concluded that the low value of $f\sigma_8$ is robust to the modeling of the shape of the power spectrum.

As a second robustness test, it was imposed the prior $0.43 < f\sigma_8 < 0.46$ and left the other parameters free. This allows us to investigate the 1σ tension concerning Rota et al. (2017) in the estimates of $\Omega_m h$ and f_b seen in Fig. 4.9. This configuration better matches the analysis of Rota which only used the power spectrum monopole.

In this configuration it was found a higher value of the matter density $\Omega_m h = 0.19$ to be compared with $\Omega_M h = 0.227^{+0.063}_{-0.050}$ from Rota et al. (2017); thus, the narrow prior reduces but does not solve the tension. This shows that the results of the Gibbs sampler are stable within 1σ; however, the joint fit of redshift-space distortions and the shape of the power spectrum prefers lower values of both $\Omega_m h$ and $f\sigma_8$.

Final remarks

Chapter 5

Summary and conclusions

Throughout this thesis I aimed at reconstructing and characterizing the galaxy density field on large scales, from the outskirts of galaxy clusters up to cosmological scales, at intermediate redshift $0.4 < z < 1$, to address one of the most puzzling questions in astrophysics: the role of environment in galaxy evolution and cosmology.

I studied the galaxies in a broad range of environments (from the dense core to under- and over-dense outskirts of clusters) up to redshift 1, finding for the first time environmental effects on galaxy properties on extreme outskirts of clusters ($\sim 5r_{200}$ at redshift 0.4, and improving the accuracy for ΛCDM model parameters determination from a spectroscopic survey. The original contribution of this work relies in the use of the density field as a bridge between the two scales, and therefore, remarking the importance of establishing an unified definition which can improve a systematic understanding of the Universe.

Here, I summarize the work done by matching each chapter as an answer to a given question:

Which kind of data can expand the frontier of galaxy evolution, pushing up the edges of the state-of-the-art of environmental dependence on galaxy evolution?

In chapter 2.1, I introduced the VST-GAME survey, a project aimed at gathering observations at optical wavelengths for six massive galaxy clusters at $0.2 \lesssim z \lesssim 0.6$ to investigate cluster assembly and galaxy evolution as a function of stellar mass and environment. The environmental effects on galaxy evolution are well known on the cluster cores, mainly up to $z \sim 1.5$ (Dressler et al. 1997; Boselli & Gavazzi 2006; Poggianti et al. 2017). Studying the cluster outskirts at intermediate-high redshift is a key element to push the frontier of galaxy evolution (Cornwell et al. 2022).

In particular, my Ph.D. work was focused on the VST and VISTA photometry of the cluster M0416. By assessing the photometric properties of this cluster, I developed a pipeline that will be used to analyze the other clusters of the survey.

The work presented in detail in this thesis was published in Estrada et al. (2022b), which is the first refereed publication of the VST-GAME survey. In my analysis, I adopted a highly conservative approach, such as applying strict criteria for catalog construction and selecting a redshift interval that is certain to contain a substantial number of cluster mem-

bers, despite uncertainties. As a result, my findings are robust and effectively demonstrate the potential of the data to explore cosmic regions that have largely remained unexplored thus far.

How to convert the deep photometry of the VST-GAME survey into a high-quality catalog, ready for a scientific use?

Chapter 2.2 presented the whole photometric assessment of the VST-GAME images centered in the cluster M0416, starting from the reduced u, g, r, i, Y, J, Ks images up to the assembly of a complete multi-band catalog.

Section 2.2.1 was focused on the extraction of sources from each of the bands using *SExtractor*, optimizing the parameter choice to maximize the number of detected sources and setting the completeness limit for each band. Section 2.2.2 described the star/galaxy separation according to the luminosity profiles and tested the separation using colors.

Deep photometry ($r < 24.4$ mag) has the advantage of revealing a large number of faint sources but simultaneously entails the presence of optical artifacts such as ghosts, haloes, and spikes in the field of view, generated by bright stars. I also developed an ad-hoc procedure, presented in section 2.2.3, based on the positions and magnitudes of the bright stars responsible for such spurious features, to flag sources whose photometry is potentially affected. This procedure is sufficiently automatic and general and could be applied also to other images.

In section 2.2.4, I explained the procedure to compute photometric redshifts applying *LePhare* to the multi-band catalog. Spectroscopic redshifts in the central part of the cluster from Balestra et al. (2016) were used to calibrate the *LePhare* setup and assess the quality of the obtained photometric redshifts. Overall, I obtained the following statistics to quantify the accuracy of the photo-z's: $bias = 0.0364$, $\sigma = 0.0296$, $\sigma_{NMAD} = 0.0425$, $\eta = 3.75\%$.

The paper Estrada et al. (2022b) presents the catalog described in Sect. 2.2.5 containing the main photometric parameters of the cluster galaxies. This catalog was the result of a prudent process of source extraction and band matching, which guarantees its quality, penalizing the total number of objects, 74114 in total. Single band catalogs, available under request, contain a larger number of objects, v.g. 237094 in the r band, but over the completeness limit or without a clear counterpart in the other bands.

How to determine the galaxy's local environment from the cluster core up to several virial radii?

In chapter 3 I used the multi-band catalog to obtain insights on galaxy properties according to their environment.

In section 3.1 I presented the criteria adopted to define the cluster membership and discuss the reliability of this selection. Galaxies with a S/N>5 in the NIR bands, were considered cluster members if their photo-z is within $0.348 < z_{phot} < 0.453$. I estimated the completeness and purity of the photometric sample on that interval, with respect to the whole spectroscopic sample, which are 63.3 % and 55.8 %, respectively.

In section 3.2 I used the spectroscopic cluster members to define the sequence of red galaxies, using $g - r$ colors, across the whole field. The completeness and the purity of the red galaxy sample on the cluster interval increased to 100.0% and 71.1% respectively with respect to the red spectroscopic sample.

In section 3.3, I performed an environmental analysis to identify regions of different densities. Considering only cluster members, I estimated a density field following a count-in-cell nearest grid point (NGP) approach using squared cells of a comoving distance of 300 kpc at the cluster redshift. Then, I computed the fluctuations over the mean value, $\delta_i^* = n_i/\bar{n} - 1$, and made a convolution of such density field with a Gaussian kernel to obtain a more solid environment tracer, δ (Fig. 3.5). This smoothed density field was presented in Fig. 3.7 and was used to individuate the local environment. Thanks to the wide-field coverage and the depth of our multi-band data, I was able to unveil the presence of overdense regions, also at large distances ($\sim 5 r_{200}$) from the cluster center, and a large overdensity region aligned with the cluster core. As additional validation of the cluster membership selection, I analyzed the radial distribution of galaxy density for M0416 (Fig. 3.6). The obtained profile follows the same distribution found by Angora et al. (2020) using four CLASH-VLT clusters, including M0416. This confirmed that the density field is dominated by the cluster core at the center of the field, but also supported the presence of structure up to 5 r_{200}.

What are the main findings on the evolution of galaxies at redshift 0.4 from the current study, and how does the environment in the outskirts of large clusters affect this evolution?

In section 3.4.1, I divided the sample of cluster members into three density intervals and compare the location of the red sequence determined in Sec. 3.2 with the sample of galaxies in different environments. I found that there is a dependency of the color-magnitude diagram on the environment: the fraction of galaxies on (below) the red sequence increases (decreases) with density: the fraction of red galaxies goes from 21% ± 1% in the lowest density bin to 32 % ± 1% at intermediate densities, to 65 % ± 2% in the highest density bin, considering binomial errors. This behavior was also clearly visible in the normalized histograms on panel (d) of Fig. 3.8, where the relative importance of the two peaks at red and blue colors strongly varies with the environment. The histogram also showed that the peak of the distribution at the reddest colors overall shifts toward redder colors with increasing density.

Then, in section 3.4.2, I selected the sample of galaxies on high-density regions and divided that sample into two: core galaxies and outskirts galaxies. I found that there are substructures up to 5 r_{200}, and in particular in Fig. 3.9 is evidenced that the composition of those galaxy populations is as evolved as the cluster core, giving strong constraints on environmental effects up to large cluster distances.

The advancement of knowledge that my thesis brings with respect to the literature is discussed in section 3.5. In particular, this is the first time that overdensities are found at large outskirts ($r \sim 5 r_{200}$) of an evolving cluster at intermediate redshift ($z \sim 0.4$). The deepness of those observations makes M0416 a perfect candidate for filament identification

beyond the local Universe. Additionally, the shift toward redder colors with increasing density, in particular in these outskirts overdensities, confirms a scenario where galaxies in denser environments evolve faster than in the field.

How to extract parameters of the cosmological model by using data from a spectroscopic survey?

Chapter 4 offered a complete vision of the work made in collaboration with Prof. Luigi Guzzo and Dr. Benjamin Granett in the context of the VIPERS survey. During this collaboration, I carried out a complete analysis of the final data release of VIPERS, based on a maximum likelihood estimator for the density field and cosmological parameters. The primary aim was to improve the accuracy of previous VIPERS results (Rota et al. 2017), through the use of an estimator that optimally corrects for the complex survey geometry. I reconstructed the un-masked redshift-space galaxy density field using the Wiener filter, which is the maximum posterior estimator in the case of a Gaussian field. With the Gibbs sampling approach, I estimated the joint posterior probabilities on the parameters of the ΛCDM model as shown in Fig. 4.7.

In section 4.1, I introduced the VIPERS survey, and in particular the selection function of the survey, which generates some 'noise' effects on the observations. My work started with the final data release of the survey. My work aimed to clean the observations from those effects induced by the selection function.

In section 4.2, I presented the theoretical construction of the new analysis pipeline that I developed, starting from the Bayesian formalism described in Granett et al. (2015), to study the VIPERS survey. The pipeline estimates the joint posterior probabilities of the density field in the redshift-space along with the cosmological parameters, using data from a spectroscopic redshift survey. In comparison with previous VIPERS analysis on cosmological parameters, I allowed a larger number of degrees of freedom, which here include the shape of the power spectrum and redshift-space distortions. The analysis does not suffer from Alcock-Pacynzki geometric distortions since the density field was recomputed consistently on every iteration.

In section 4.3, I gave a detailed description of the structure of the code and its implementation. The pipeline was applied to the VIPERS data and the 153 mock realizations of VIPERS. The algorithm yields the unmasked density field in redshift space and its power spectrum monopole. The uncertainty is computed from the dispersion in the mocks.

In section 4.4, I presented the results of this work. The precision on the cosmological parameters is competitive with previous VIPERS analyses (Rota et al. 2017; Pezzotta et al. 2017, the former for Ω_M and f_b, the latter for $f\sigma_8$). The results on the matter density and baryon fraction parameters are in agreement with measurements at lower redshift from 2dFGRS at $z = 0.2$ Cole et al. (2005), SDSS LRG at $z = 0.35$ Tegmark et al. (2004), and WiggleZ at $0.2 < z < 0.8$ Parkinson et al. (2012) as well as with Planck determinations (Planck Collaboration et al. 2020). From the results on mocks, I found the matter and baryon density parameter estimates to be unbiased. In both the mocks and data, the estimates of the growth-of-structure parameter $f\sigma_8$ were systematically low. This systematic bias is attributed to the use of the dispersion model for the anisotropic

power spectrum. It was also found $f\sigma_8$ and $\Omega_M h$ to be correlated and the analysis of the VIPERS data also preferred a low value of $\Omega_M h$ with a 1-σ tension with Rota et al. (2017). Fixing $f\sigma_8$ with a tight prior decreased this tension. However, given the volume of the VIPERS data set, these results are consistent with the statistical error, so a larger data set is needed to investigate the issue further.

Future prospects

This thesis leaves several open questions that can be discussed to keep characterizing the role of galaxy density field as an environment tracer that shapes galaxy properties.

Galaxy evolution

First of all, I plan to further expand the photometric analysis and density characterization to all of the clusters of the VST-GAME survey to have a more significant sample and discover possible new outskirt overdensities. I started investigating the radial profiles of the single outskirts overdensities. I wish to compute galaxy masses using Ks luminosities, to study whether the distribution of galaxy masses depends mainly on the local environment or the clustercentric distance. I aim to understand the environmental dependence of galaxy properties on outskirts substructures, e.g. star-formation rate, age, or metallicities computed through SED fitting. Furthermore, I would like to compute the morphology by using machine learning, exploiting labeled data, e.g. from the GalaxyZoo (Lintott et al. 2011; Dieleman et al. 2015), or deep learning techniques (Huertas-Company et al. 2015), to compute the morphology-density and star formation-density relations, to explore assembly processes at large clustercentric distances. I wish to extract from hydro-dynamical simulations, e.g. The Three Hundred project (Cui et al. 2018) or the Magneticum simulation (*Dolag et al., in prep*, Kimmig et al. 2022), similar overdensities at that clustercentric distances to explore the possible accretion mechanisms into the cluster core. The purpose of this comparison is to evaluate the accuracy of simulations in predicting the results obtained from current and future data, as well as to assess their ability to predict the properties of high-redshift systems.

As an improvement on the photo-z measurements over the whole VST field, it was proposed and approved a spectroscopic follow-up for the M0416 with 2dF at AAOmega (P.I. A. Mercurio, observations in semester 2021B). This follow-up aimed to obtain redshifts of $\sim$1000 galaxies, enabling the study of spectral properties (i.e. only absorption or also the presence of emission lines) and their association with the cluster environments: substructures, infalling filaments. The selection of the targets has been done using colors and photo-z derived in this thesis. I am working on improving the photo-z accuracy over the whole field with this new sample of spectroscopic redshifts, and I am integrating the spectroscopic information with the photometric one, to reconstruct filaments over the whole field.

Cosmology

The maximum-likelihood analysis presented here represents a forward model of the galaxy survey based on a multivariate Gaussian likelihood and prior for the density field. I would like to improve this model to account for the non-Gaussian distribution of the density field, for example with a log-normal distribution (Kitaura et al. 2010) and non-Poissonian sampling models (Ata et al. 2015). It has been shown that more cosmological information can be unlocked from the field with the log-normal model than is available in the standard two-point statistics of the density field (Carron & Szapudi 2014).

However, limiting this work to the Gaussian case, I found that the error estimates were underestimated concerning the dispersion of the mock catalogs. This aspect of the algorithm may be improved by adopting a prior and likelihood that better represent the galaxy distribution on the quasi-linear scales that were considered. Even so, the discrepancy in the error analysis is likely to be less significant for larger surveys for which coarser grids can be employed for the density field reconstruction. This is a motivation for me to apply this approach to a larger galaxy survey, e.g., Euclid (Laureijs et al. 2011). To optimize the use of upcoming surveys, I would like to it will be necessary to carry out joint analyses between multiple observables and account for a multitude of observational systematic effects. Forward-modeling approaches at the level of the density field provide a promising approach. However, more experience is needed to build and apply forward modeling to the next generation of galaxy surveys. The application to VIPERS is one step in this direction.

The results of this thesis have highlighted the significance of environmental effects on galaxy evolution in the previously unexplored outskirts of intermediate-redshift clusters. This discovery opens the possibility for new lines of observational inquiry with the new generation of ground- and space-based telescopes. Additionally, the utilization of the density field in this study has also contributed to the refinement of cosmological parameter determinations in redshift surveys. This supports the ongoing development of a more comprehensive theoretical framework for galaxy evolution, which aims to unify the understanding of the universe on a large scale.

As a final remark of this thesis, the concept of cosmic density has not been fully exploited in astrophysics and cosmology. A standard definition for this concept could facilitate the unification of frameworks across local and cosmological scales. The advent of new generation instruments, capable of unprecedentedly detailed observations of a vast number of galaxies, clusters, and filaments, offers the potential for a holistic understanding of the Universe by connecting these various scales.

Appendix A

Masking SHARKS

The procedure developed for masking the VST-GAME photometry of M0416 is versatile and can be applied to other VST and VISTA surveys, as demonstrated by the work of *Lorenzon et al. (in prep.)* on the SHARKS survey.

The Southern H-ATLAS Regions in Ks-band (SHARKS) is a deep Ks-band imaging survey conducted with the VISTA telescope (PI H. Dannerbauer; Dannerbauer et al. 2022). It covers ∼300 deg^2, including parts of the South Galactic Plane, GAMA-12h, and GAMA-15h fields from the H-ATLAS survey, the largest Herschel program. Its main goals are to identify counterparts to ∼90% of the sources detected by H-ATLAS (Eales et al. 2010), ASKAP (Johnston et al. 2007), and LOFAR (van Haarlem et al. 2013) in the redshift range $0 < z < 3$, to create a sample of strong lenses for cosmography studies, and to study the evolution of massive structures in the Universe. SHARKS DR1 includes calibrated images and single-band source catalogs from observations taken between March 2017 and January 2019, covering a total area of about 20 deg^2. It reaches a depth of Ks magnitude ~ 22.7 (AB, 5σ) with a mean seeing of ~ 1.

In *Lorenzon et al. (in prep.)*, the authors used the procedure that I presented in section 2.2.3, to mask the Ks-band photometry of the SHARKS survey, covering an area of 50 deg^2 Examples of the masking procedure for halos and ghosts are shown in the left and right panels of Figure A.1, respectively. This technique is useful for accurately masking the spurious photometry of galaxies in these surveys, as it allows for the removal of artifacts that may affect the measurements.

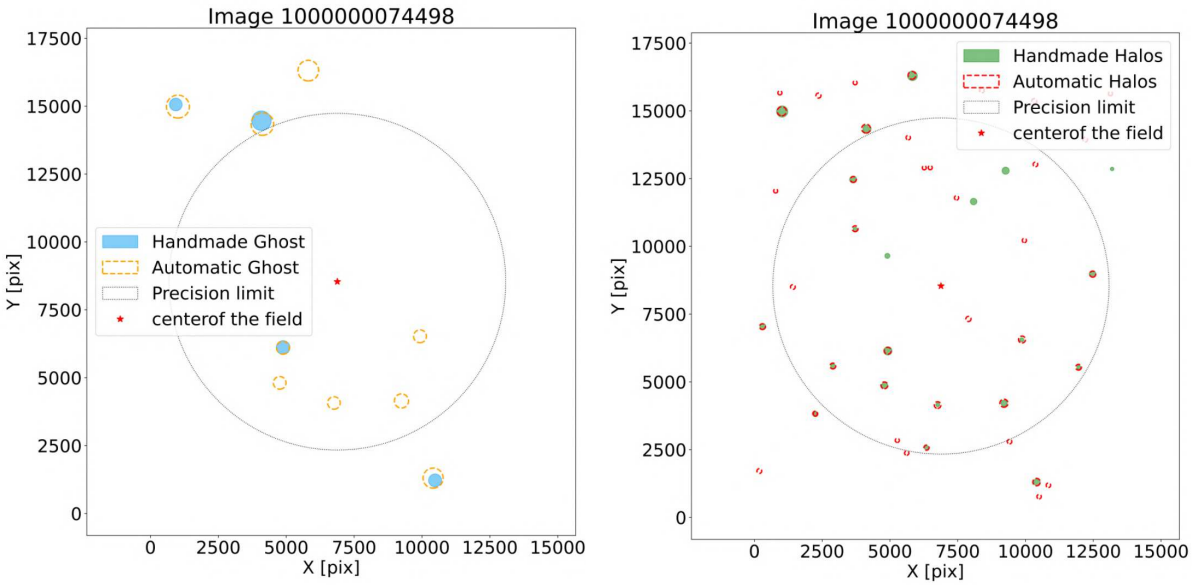


Figure A.1: Masking procedure implemented by *Lorenzon et al. (in prep.)* to mask ghosts (left) and halos (right) in the Ks-band photometry of the SHARK survey.

List of Figures

List of Tables

Bibliography

Abbas, U. & Sheth, R. K. 2005, MNRAS, 364, 1327

Abbott, T. M. C., Abdalla, F. B., Alarcon, A., et al. 2018, , 98, 043526

Abell, G. O. 1958, ApJS, 3, 211

Abraham, R. G. & van den Bergh, S. 2001, Science, 293, 1273

Alcock, C. & Paczynski, B. 1979, Nature, 281, 358

Angora, G., Rosati, P., Brescia, M., et al. 2020, A&A, 643, A177

Annunziatella, M., Biviano, A., Mercurio, A., et al. 2014, A&A, 571, A80

Annunziatella, M., Bonamigo, M., Grillo, C., et al. 2017, ApJ, 851, 81

Annunziatella, M., Mercurio, A., Biviano, A., et al. 2016, A&A, 585, A160

Annunziatella, M., Mercurio, A., Brescia, M., Cavuoti, S., & Longo, G. 2013, PASP, 125, 68

Aragón-Calvo, M. A., Jones, B. J. T., van de Weygaert, R., & van der Hulst, J. M. 2007, A&A, 474, 315

Aragón-Calvo, M. A., van de Weygaert, R., & Jones, B. J. T. 2010, MNRAS, 408, 2163

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

Ata, M., Kitaura, F.-S., Lee, K.-G., et al. 2021, MNRAS, 500, 3194

Ata, M., Kitaura, F.-S., & Müller, V. 2015, MNRAS, 446, 4250

Bahe, Y. M., McCarthy, I. G., Balogh, M. L., & Font, A. S. 2013, MNRAS, 430, 3017

Bai, L., Rieke, G. H., Rieke, M. J., Christlein, D., & Zabludoff, A. I. 2009, ApJ, 693, 1840

Baldry, I. K., Balogh, M. L., Bower, R., Glazebrook, K., & Nichol, R. C. 2004, in American Institute of Physics Conference Series, Vol. 743, The New Cosmology: Conference on Strings and Cosmology, ed. R. E. Allen, D. V. Nanopoulos, & C. N. Pope, 106–119

Baldry, I. K., Balogh, M. L., Bower, R. G., et al. 2006, MNRAS, 373, 469

Baldry, I. K., Robotham, A. S. G., Hill, D. T., et al. 2010, MNRAS, 404, 86

Balestra, I., Mercurio, A., Sartoris, B., et al. 2016, ApJS, 224, 33

Ballinger, W. E., Peacock, J. A., & Heavens, A. F. 1996, MNRAS, 282, 877

Balogh, M. L., Baldry, I. K., Nichol, R., et al. 2004, ApJ, 615, L101

Barnes, J. E. & Hernquist, L. 1996, ApJ, 471, 115

Battaglia, N., Bond, J. R., Pfrommer, C., & Sievers, J. L. 2012, The Astrophysical Journal, 758, 74

Benson, A. J., Frenk, C. S., Baugh, C. M., Cole, S., & Lacey, C. G. 2001, MNRAS, 327, 1041

Bergamini, P., Grillo, C., Rosati, P., et al. 2022, arXiv e-prints, arXiv:2208.14020

Bergamini, P., Rosati, P., Mercurio, A., et al. 2019, A&A, 631, A130

Bertin, E. 2006, in Astronomical Society of the Pacific Conference Series, Vol. 351, Astronomical Data Analysis Software and Systems XV, ed. C. Gabriel, C. Arviset, D. Ponz, & S. Enrique, 112

Bertin, E. 2011, in Astronomical Society of the Pacific Conference Series, Vol. 442, Astronomical Data Analysis Software and Systems XX, ed. I. N. Evans, A. Accomazzi, D. J. Mink, & A. H. Rots, 435

Bertin, E. & Arnouts, S. 1996, A&AS, 117, 393

Bertin, E., Mellier, Y., Radovich, M., et al. 2002, in Astronomical Society of the Pacific Conference Series, Vol. 281, Astronomical Data Analysis Software and Systems XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley, 228

Bianconi, M., Smith, G. P., Haines, C. P., et al. 2018, MNRAS, 473, L79

Biffi, V., Planelles, S., Borgani, S., et al. 2018, MNRAS, 476, 2689

Binggeli, B. 1982, A&A, 107, 338

Biviano, A., Katgert, P., Thomas, T., & Adami, C. 2002, A&A, 387, 8

Biviano, A., Murante, G., Borgani, S., et al. 2006, A&A, 456, 23

Blanton, M., Cen, R., Ostriker, J. P., & Strauss, M. A. 1999, ApJ, 522, 590

Blanton, M. R., Eisenstein, D., Hogg, D. W., Schlegel, D. J., & Brinkmann, J. 2005, ApJ, 629, 143

Blas, D., Lesgourgues, J., & Tram, T. 2011, , 2011, 034

Blitz, L. 1993, in Protostars and Planets III, ed. E. H. Levy & J. I. Lunine, 125

Bonamigo, M., Grillo, C., Ettori, S., et al. 2017, ApJ, 842, 132

Bonamigo, M., Grillo, C., Ettori, S., et al. 2018, ApJ, 864, 98

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

Boselli, A. & Gavazzi, G. 2006, PASP, 118, 517

Bottini, D., Garilli, B., Maccagni, D., et al. 2005, PASP, 117, 996

Braglia, F. G., Pierini, D., Biviano, A., & Böhringer, H. 2009, A&A, 500, 947

Brescia, M. & Cavuoti, S. 2014, PhotoRApToR: PHOTOmetric Research APplication TO Redshifts, Astrophysics Source Code Library, record ascl:1408.022

Brüggen, M. & De Lucia, G. 2008, MNRAS, 383, 1336

Brunetti, G. & Jones, T. W. 2014, International Journal of Modern Physics D, 23, 1430007

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

Budavari, T. 2009, ApJ, 695, 747

Bureau, M. & Carignan, C. 2002, AJ, 123, 1316

Butcher, H. & Oemler, A., J. 1978, ApJ, 219, 18

Butcher, H. & Oemler, A., J. 1984, ApJ, 285, 426

Cai, Y.-C., Bernstein, G., & Sheth, R. K. 2011, MNRAS, 412, 995

Caminha, G. B., Grillo, C., Rosati, P., et al. 2017, A&A, 600, A90

Caminha, G. B., Rosati, P., Grillo, C., et al. 2019, A&A, 632, A36

Cappellari, M., Scott, N., Alatalo, K., et al. 2013, MNRAS, 432, 1709

Carlstrom, J. E., Holder, G. P., & Reese, E. D. 2002, ARA&A, 40, 643

Carrasco Kind, M. & Brunner, R. J. 2013, MNRAS, 432, 1483

Carrasco Kind, M. & Brunner, R. J. 2014, MNRAS, 438, 3409

Carron, J. & Szapudi, I. 2014, MNRAS, 439, L11

Castellano, M., Amorín, R., Merlin, E., et al. 2016, A&A, 590, A31

Castignani, G., Combes, F., Jablonka, P., et al. 2022, A&A, 657, A9

Cautun, M., van de Weygaert, R., & Jones, B. J. T. 2013, MNRAS, 429, 1286

Cautun, M., van de Weygaert, R., Jones, B. J. T., & Frenk, C. S. 2014, MNRAS, 441, 2923

Cavuoti, S., Brescia, M., De Stefano, V., & Longo, G. 2015, Experimental Astronomy, 39, 45

Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv e-prints, arXiv:1612.05560

Colberg, J. M., Sheth, R. K., Diaferio, A., Gao, L., & Yoshida, N. 2005, MNRAS, 360, 216

Colberg, J. M., White, S. D. M., Jenkins, A., & Pearce, F. R. 1999, MNRAS, 308, 593

Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, MNRAS, 362, 505

Colless, M., Dalton, G., Maddox, S., et al. 2001, MNRAS, 328, 1039

Colless, M., Peterson, B. A., Jackson, C., et al. 2003, arXiv e-prints, astro

Cooper, M. C., Coil, A. L., Gerke, B. F., et al. 2010, MNRAS, 409, 337

Cornwell, D. J., Kuchner, U., Aragón-Salamanca, A., et al. 2022, MNRAS, 517, 1678

Cortese, L., Catinella, B., & Smith, R. 2021, , 38, e035

Cortese, L., Gavazzi, G., Boselli, A., et al. 2006, A&A, 453, 847

Couch, W. J., Barger, A. J., Smail, I., Ellis, R. S., & Sharples, R. M. 1998, ApJ, 497, 188

Cowie, L. L. & Songaila, A. 1977, Nature, 266, 501

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

Cox, T. J., Jonsson, P., Somerville, R. S., Primack, J. R., & Dekel, A. 2008, MNRAS, 384, 386

Cucciati, O., Iovino, A., Kovač, K., et al. 2010, A&A, 524, A2

Cucciati, O., Iovino, A., Marinoni, C., et al. 2006, A&A, 458, 39

Cui, W., Knebe, A., Yepes, G., et al. 2018, MNRAS, 480, 2898

Cui, Y., Xiang, Y., Rong, K., Feris, R., & Cao, L. 2014, 213

Cuillandre, J.-C. J., Withington, K., Hudelot, P., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8448, Observatory Operations: Strategies, Processes, and Systems IV, 84480M

Dannerbauer, H., Carnero, A., Cross, N., & Gutierrez, C. M. 2022, arXiv e-prints, arXiv:2212.09471

Danovich, M., Dekel, A., Hahn, O., & Teyssier, R. 2012, MNRAS, 422, 1732

Darvish, B., Mobasher, B., Sobral, D., et al. 2015, ApJ, 814, 84

Davidzon, I., Cucciati, O., Bolzonella, M., et al. 2016, A&A, 586, A23

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371

de la Torre, S., Guzzo, L., Peacock, J. A., et al. 2013, A&A, 557, A54

de la Torre, S., Jullo, E., Giocoli, C., et al. 2017, A&A, 608, A44

de Lapparent, V., Kurtz, M. J., & Geller, M. J. 1986, ApJ, 304, 585

De Lucia, G., Boylan-Kolchin, M., Benson, A. J., Fontanot, F., & Monaco, P. 2010, MNRAS, 406, 1533

De Lucia, G., Poggianti, B. M., Aragón-Salamanca, A., et al. 2007, MNRAS, 374, 809

De Lucia, G., Springel, V., White, S. D. M., Croton, D., & Kauffmann, G. 2006, MNRAS, 366, 499

De Lucia, G., Weinmann, S., Poggianti, B. M., Aragón-Salamanca, A., & Zaritsky, D. 2012, MNRAS, 423, 1277

Dekel, A., Birnboim, Y., Engel, G., et al. 2009, Nature, 457, 451

Demarco, R., Rosati, P., Lidman, C., et al. 2007, ApJ, 663, 164

Desai, S., Armstrong, R., Mohr, J. J., et al. 2012, ApJ, 757, 83

Desai, V., Dalcanton, J. J., Aragón-Salamanca, A., et al. 2007, ApJ, 660, 1151

Di Porto, C., Branchini, E., Bel, J., et al. 2016, A&A, 594, A62

Dieleman, S., Willett, K. W., & Dambre, J. 2015, MNRAS, 450, 1441

Diemer, B., Mansfield, P., Kravtsov, A. V., & More, S. 2017, ApJ, 843, 140

Dressler, A. 1980, ApJ, 236, 351

Dressler, A., Oemler, Augustus, J., Couch, W. J., et al. 1997, ApJ, 490, 577

Dressler, A., Oemler, Augustus, J., Poggianti, B. M., et al. 2013, ApJ, 770, 62

Dressler, A. & Shectman, S. A. 1988, AJ, 95, 985

Dressler, A., Smail, I., Poggianti, B. M., et al. 1999, ApJS, 122, 51

Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, MNRAS, 413, 971

Dubois, Y., Peirani, S., Pichon, C., et al. 2016, MNRAS, 463, 3948

Eales, S., Dunne, L., Clements, D., et al. 2010, PASP, 122, 499

Ebeling, H., Barrett, E., & Donovan, D. 2004, ApJ, 609, L49

Ebeling, H., Edge, A. C., & Henry, J. P. 2001, ApJ, 553, 668

Efstathiou, G., Sutherland, W. J., & Maddox, S. J. 1990, Nature, 348, 705

Eggen, O. J., Lynden-Bell, D., & Sandage, A. R. 1962, ApJ, 136, 748

Einstein, A. 1916, Annalen der Physik, 354, 769

Eke, V. R., Frenk, C. S., Baugh, C. M., et al. 2004, MNRAS, 355, 769

Eke, V. R., Navarro, J. F., & Frenk, C. S. 1998, ApJ, 503, 569

Ellingson, E., Lin, H., Yee, H. K. C., & Carlberg, R. G. 2001, ApJ, 547, 609

Estrada, N., Granett, B. R., & Guzzo, L. 2022a, MNRAS, 512, 2817

Estrada, N., Mercurio, A., Vulcani, B., et al. 2022b, arXiv e-prints, arXiv:2212.00841

Euclid Collaboration, Desprez, G., Paltani, S., et al. 2020, A&A, 644, A31

Evrard, A. E., Metzler, C. A., & Navarro, J. F. 1996, ApJ, 469, 494

Farouki, R. & Shapiro, S. L. 1981, ApJ, 243, 32

Fasano, G., Poggianti, B. M., Bettoni, D., et al. 2015, MNRAS, 449, 3927

Feldman, H. A., Kaiser, N., & Peacock, J. A. 1994, ApJ, 426, 23

Flewelling, H. A., Magnier, E. A., Chambers, K. C., et al. 2020, ApJS, 251, 7

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Frenk, C. S. 1995, Nuclear Physics B Proceedings Supplements, 43, 333

Fujita, Y. 1998, ApJ, 509, 587

Fujita, Y. 2004, , 56, 29

Fukugita, M., Hogan, C. J., & Peebles, P. J. E. 1998, ApJ, 503, 518

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1

Gargiulo, A., Cucciati, O., Garilli, B., et al. 2019, A&A, 631, A15

Garilli, B., Maccagni, D., & Andreon, S. 1999, A&A, 342, 408

Gavazzi, G., Boselli, A., Cortese, L., et al. 2006, A&A, 446, 839

Gavazzi, G., Fumagalli, M., Cucciati, O., & Boselli, A. 2010, A&A, 517, A73

Girardi, M., Giuricin, G., Mardirossian, F., Mezzetti, M., & Boschin, W. 1998, ApJ, 505, 74

Gnedin, O. Y. 2003, ApJ, 582, 141

Gómez, P. L., Nichol, R. C., Miller, C. J., et al. 2003, ApJ, 584, 210

Gonzalez, E. J., Chalela, M., Jauzac, M., et al. 2020, MNRAS, 494, 349

Granett, B. R., Branchini, E., Guzzo, L., et al. 2015, A&A, 583, A61

Grillo, C., Suyu, S. H., Rosati, P., et al. 2015, ApJ, 800, 38

Guglielmo, V., Poggianti, B. M., Moretti, A., et al. 2015, MNRAS, 450, 2749

Guglielmo, V., Poggianti, B. M., Vulcani, B., et al. 2018, A&A, 620, A15

Gunn, J. E. & Gott, J. Richard, I. 1972, ApJ, 176, 1

Guth, A. H. 1981, , 23, 347

Guzzo, L., Pierleoni, M., Meneux, B., et al. 2008, Nature, 451, 541

Guzzo, L., Scodeggio, M., Garilli, B., et al. 2014, A&A, 566, A108

Haines, C. P., Gargiulo, A., Mercurio, A., et al. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 379, Cosmic Frontiers, ed. N. Metcalfe & T. Shanks, 206

Haines, C. P., Pereira, M. J., Smith, G. P., et al. 2015, ApJ, 806, 101

Haines, C. P., Pereira, M. J., Smith, G. P., et al. 2013, ApJ, 775, 126

Haines, C. P., Smith, G. P., Egami, E., et al. 2009, ApJ, 704, 126

Hawken, A. J., Granett, B. R., Iovino, A., et al. 2017, A&A, 607, A54

Henriksen, M. & Byrd, G. 1996, ApJ, 459, 82

Hong, S., Zou, Z., Luo, A. L., et al. 2022, MNRAS[arXiv:2211.04260]

Hoyle, B. 2016, Astronomy and Computing, 16, 34

Hubble, E. P. 1926, ApJ, 64, 321

Huchra, J., Davis, M., Latham, D., & Tonry, J. 1983, ApJS, 52, 89

Huchra, J. P. & Geller, M. J. 1982, ApJ, 257, 423

Huchra, J. P., Macri, L. M., Masters, K. L., et al. 2012, ApJS, 199, 26

Huertas-Company, M., Gravet, R., Cabrera-Vives, G., et al. 2015, ApJS, 221, 8

Ichikawa, K., Matsushita, K., Okabe, N., et al. 2013, ApJ, 766, 90

Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841

Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111

Jaffé, Y. L., Poggianti, B. M., Moretti, A., et al. 2018, MNRAS, 476, 4753

Jarrett, T. 2004, 21, 396

Jarvis, M. J., Bonfield, D. G., Bruce, V. A., et al. 2013, MNRAS, 428, 1281

Jauzac, M., Jullo, E., Eckert, D., et al. 2015, MNRAS, 446, 4132

Jewell, J., Levin, S., & Anderson, C. H. 2004, ApJ, 609, 1

Jing, Y. P. 2005, ApJ, 620, 559

Johnston, S., Bailes, M., Bartel, N., et al. 2007, , 24, 174

Joo, H. & Jee, M. J. 2023, Nature, 613, 37

Jouvel, S., Host, O., Lahav, O., et al. 2014, A&A, 562, A86

Just, D. W., Kirby, M., Zaritsky, D., et al. 2019, ApJ, 885, 6

Kaiser, N. 1987, MNRAS, 227, 1

Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. 1999, MNRAS, 303, 188

Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003, MNRAS, 346, 1055

Kauffmann, G., White, S. D. M., Heckman, T. M., et al. 2004, MNRAS, 353, 713

Keller, B. W., Wadsley, J., Benincasa, S. M., & Couchman, H. M. P. 2014, MNRAS, 442, 3013

Kennicutt, Robert C., J. 1998, ARA&A, 36, 189

Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, MNRAS, 363, 2

Kimmig, L. C., Remus, R.-S., Dolag, K., & Biffi, V. 2022, arXiv e-prints, arXiv:2209.09916

Kitaura, F. S., Jasche, J., Li, C., et al. 2009, MNRAS, 400, 183

Kitaura, F.-S., Jasche, J., & Metcalf, R. B. 2010, MNRAS, 403, 589

Klypin, A. & Shandarin, S. F. 1993, ApJ, 413, 48

Klypin, A., Yepes, G., Gottlöber, S., Prada, F., & Hess, S. 2016, MNRAS, 457, 4340

Knebe, A., Gill, S. P. D., Gibson, B. K., et al. 2004, ApJ, 603, 7

Knollmann, S. R. & Knebe, A. 2009, ApJS, 182, 608

Koester, B. P., McKay, T. A., Annis, J., et al. 2007, ApJ, 660, 239

Kovač, K., Lilly, S. J., Cucciati, O., et al. 2010, ApJ, 708, 505

Kron, R. G. 1980, ApJS, 43, 305

La Barbera, F., Pasquali, A., Ferreras, I., et al. 2014, MNRAS, 445, 1977

Lahav, O., Lilje, P. B., Primack, J. R., & Rees, M. J. 1991, MNRAS, 251, 128

Larson, R. B., Tinsley, B. M., & Caldwell, C. N. 1980, ApJ, 237, 692

Lau, E. T., Kravtsov, A. V., & Nagai, D. 2009, ApJ, 705, 1129

Lau, E. T., Nagai, D., Avestruz, C., Nelson, K., & Vikhlinin, A. 2015, ApJ, 806, 68

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv e-prints, arXiv:1110.3193

Le Fèvre, O., Saisse, M., Mancini, D., et al. 2003, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes, ed. M. Iye & A. F. M. Moorwood, 1670–1681

Lemaux, B. C., Gal, R. R., Lubin, L. M., et al. 2012, ApJ, 745, 106

Lemaux, B. C., Tomczak, A. R., Lubin, L. M., et al. 2019, MNRAS, 490, 1231

Levi, M., Bebek, C., Beers, T., et al. 2013, arXiv e-prints, arXiv:1308.0847

Lewis, I., Balogh, M., De Propris, R., et al. 2002, MNRAS, 334, 673

Libeskind, N. I., van de Weygaert, R., Cautun, M., et al. 2018, MNRAS, 473, 1195

Lin, L., Jian, H.-Y., Foucaud, S., et al. 2014, ApJ, 782, 33

Linde, A. D. 1982, Physics Letters B, 108, 389

Linder, E. V. 2005, , 72, 043529

Lintott, C., Schawinski, K., Bamford, S., et al. 2011, MNRAS, 410, 166

Longhetti, M. & Saracco, P. 2009, MNRAS, 394, 774

Lotz, J. M., Koekemoer, A., Coe, D., et al. 2017, ApJ, 837, 97

Loveday, J., Peterson, B. A., Efstathiou, G., & Maddox, S. J. 1992, ApJ, 390, 338

Lu, T., Gilbank, D. G., McGee, S. L., Balogh, M. L., & Gallagher, S. 2012, MNRAS, 420, 126

Lubin, L. M., Gal, R. R., Lemaux, B. C., Kocevski, D. D., & Squires, G. K. 2009, AJ, 137, 4867

Malavasi, N., Pozzetti, L., Cucciati, O., et al. 2017, MNRAS, 470, 1274

Mann, A. W. & Ebeling, H. 2012, MNRAS, 420, 2120

Mansfield, P., Kravtsov, A. V., & Diemer, B. 2017, ApJ, 841, 34

Martínez, H. J., Muriel, H., & Coenda, V. 2016, MNRAS, 455, 127

Marulli, F., Bianchi, D., Branchini, E., et al. 2012, MNRAS, 426, 2566

Marulli, F., Bolzonella, M., Branchini, E., et al. 2013, A&A, 557, A17

McCarthy, I. G., Frenk, C. S., Font, A. S., et al. 2008, MNRAS, 383, 593

McGee, S. L., Balogh, M. L., Bower, R. G., Font, A. S., & McCarthy, I. G. 2009, MNRAS, 400, 937

McGee, S. L., Balogh, M. L., Wilman, D. J., et al. 2011, MNRAS, 413, 996

Mercurio, A., Merluzzi, P., Busarello, G., et al. 2015, MNRAS, 453, 3685

Mercurio, A., Rosati, P., Biviano, A., et al. 2021, A&A, 656, A147

Mercurio, A. & VST-GAME Team. 2018, in VST in the Era of the Large Sky Surveys, 23

Merlin, E., Amorín, R., Castellano, M., et al. 2016, A&A, 590, A30

Mestric, U., Vanzella, E., Zanella, A., et al. 2022, MNRAS, 516, 3532

Mihos, J. C. 2004, in Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution, ed. J. S. Mulchaey, A. Dressler, & A. Oemler, 277

Mihos, J. C. & Hernquist, L. 1996, ApJ, 464, 641

Mirakhor, M. S. & Walker, S. A. 2021, MNRAS, 506, 139

Mo, H. J., Mao, S., & White, S. D. M. 1998, MNRAS, 295, 319

Mo, H. J. & White, S. D. M. 1996, MNRAS, 282, 347

Mohammad, F. G., Granett, B. R., Guzzo, L., et al. 2018, A&A, 610, A59

Monaco, P. & Efstathiou, G. 1999, MNRAS, 308, 763

Montero-Dorta, A. D., Abramo, L. R., Granett, B. R., de la Torre, S., & Guzzo, L. 2020, MNRAS, 493, 5257

Montes, M. 2022, Nature Astronomy, 6, 308

Moore, B., Katz, N., Lake, G., Dressler, A., & Oemler, A. 1996, Nature, 379, 613

Moore, B., Lake, G., Quinn, T., & Stadel, J. 1999, MNRAS, 304, 465

Muldrew, S. I., Croton, D. J., Skibba, R. A., et al. 2012, MNRAS, 419, 2670

Muzzin, A., Wilson, G., Lacy, M., Yee, H. K. C., & Stanford, S. A. 2008, ApJ, 686, 966

Muzzin, A., Wilson, G., Yee, H. K. C., et al. 2012, ApJ, 746, 188

Nagai, D. & Lau, E. T. 2011, ApJ, 731, L10

Natarajan, P., Chadayammuri, U., Jauzac, M., et al. 2017, MNRAS, 468, 1962

Oemler, Augustus, J. 1974, ApJ, 194, 1

Ogrean, G. A., van Weeren, R. J., Jones, C., et al. 2015, ApJ, 812, 153

Olave-Rojas, D., Cerulo, P., Demarco, R., et al. 2018, MNRAS, 479, 2328

Paccagnella, A., Vulcani, B., Poggianti, B. M., et al. 2016, ApJ, 816, L25

Park, C. & Hwang, H. S. 2009, ApJ, 699, 1595

Parkinson, D., Riemer-Sørensen, S., Blake, C., et al. 2012, , 86, 103518

Pasquali, A., Gallazzi, A., Fontanot, F., et al. 2010, MNRAS, 407, 937

Peacock, J. A., Cole, S., Norberg, P., et al. 2001, Nature, 410, 169

Peacock, J. A. & Dodds, S. J. 1994, MNRAS, 267, 1020

Pearson, D. W., Samushia, L., & Gagrani, P. 2016, MNRAS, 463, 2708

Peebles, P. J. E. 1980, The large-scale structure of the universe (Address: Publisher)

Peebles, P. J. E. 1982, ApJ, 263, L1

Peebles, P. J. E. & Yu, J. T. 1970, ApJ, 162, 815

Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010, ApJ, 721, 193

Peng, Y.-j., Lilly, S. J., Renzini, A., & Carollo, M. 2012, ApJ, 757, 4

Penzias, A. A. & Wilson, R. W. 1965, ApJ, 142, 419

Percival, W. J., Baugh, C. M., Bland-Hawthorn, J., et al. 2001, MNRAS, 327, 1297

Percival, W. J., Verde, L., & Peacock, J. A. 2004, MNRAS, 347, 645

Percival, W. J. & White, M. 2009, MNRAS, 393, 297

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565

Pezzotta, A., de la Torre, S., Bel, J., et al. 2017, A&A, 604, A33

Pimbblet, K. A., Shabala, S. S., Haines, C. P., Fraser-McKelvie, A., & Floyd, D. J. E. 2013, MNRAS, 429, 1827

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6

Poggianti, B. M., Moretti, A., Gullieuszik, M., et al. 2017, ApJ, 844, 48

Poggianti, B. M., Smail, I., Dressler, A., et al. 1999, ApJ, 518, 576

Poggianti, B. M., von der Linden, A., De Lucia, G., et al. 2006, ApJ, 642, 188

Pogosyan, D., Bond, J. R., Kofman, L., & Wadsley, J. 1996, in American Astronomical Society Meeting Abstracts, Vol. 189, American Astronomical Society Meeting Abstracts, 13.03

Polletta, M., Tajer, M., Maraschi, L., et al. 2007, ApJ, 663, 81

Postman, M., Coe, D., Benítez, N., et al. 2012, ApJS, 199, 25

Postman, M. & Geller, M. J. 1984, ApJ, 281, 95

Postman, M., Lubin, L. M., Gunn, J. E., et al. 1996, AJ, 111, 615

Razim, O., Cavuoti, S., Brescia, M., et al. 2021, MNRAS, 507, 5034

Renzini, A. 2006, ARA&A, 44, 141

Rhee, J., Smith, R., Choi, H., et al. 2017, ApJ, 843, 128

Roncarelli, M., Ettori, S., Borgani, S., et al. 2013, MNRAS, 432, 3030

Rosati, P., Balestra, I., Grillo, C., et al. 2014, The Messenger, 158, 48

Rosati, P., Borgani, S., & Norman, C. 2002, ARA&A, 40, 539

Rota, S., Granett, B. R., Bel, J., et al. 2017, A&A, 601, A144

Rudd, D. H. & Nagai, D. 2009, ApJ, 701, L16

Rybicki, G. B. & Press, W. H. 1992, ApJ, 398, 169

Salerno, J. M., Martínez, H. J., & Muriel, H. 2019, MNRAS, 484, 2

Salerno, J. M., Martínez, H. J., Muriel, H., et al. 2020, MNRAS, 493, 4950

Sánchez-Blázquez, P., Jablonka, P., Noll, S., et al. 2009, A&A, 499, 47

Sarazin, C. L. 1988, X-ray emission from clusters of galaxies

Sarron, F., Adami, C., Durret, F., & Laigle, C. 2019, A&A, 632, A49

Schawinski, K., Virani, S., Simmons, B., et al. 2009, ApJ, 692, L19

Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521

Schirmer, M., Hildebrandt, H., Kuijken, K., & Erben, T. 2011, A&A, 532, A57

Schlafly, E. F. & Finkbeiner, D. P. 2011, ApJ, 737, 103

Scodeggio, M., Guzzo, L., Garilli, B., et al. 2018, A&A, 609, A84

Scodeggio, M., Vergani, D., Cucciati, O., et al. 2009, A&A, 501, 21

Scudder, J. M., Ellison, S. L., Torrey, P., Patton, D. R., & Mendel, J. T. 2012, MNRAS, 426, 549

Sheth, R. K. & Tormen, G. 2002, MNRAS, 329, 61

Shipley, H. V., Lange-Vagle, D., Marchesini, D., et al. 2018, ApJS, 235, 14

Sifón, C. 2017, Nature Astronomy, 1, 0181

Smith, R. J., Lucey, J. R., Price, J., Hudson, M. J., & Phillipps, S. 2012, MNRAS, 419, 3167

Smith, S. 1936, ApJ, 83, 23

Sobral, D., Best, P. N., Smail, I., et al. 2011, MNRAS, 411, 675

Solanes, J. M., Sanchis, T., Salvador-Solé, E., Giovanelli, R., & Haynes, M. P. 2002, AJ, 124, 2440

Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv e-prints, arXiv:1503.03757

Spitzer, Lyman, J. & Baade, W. 1951, ApJ, 113, 413

Springel, V. 2010, MNRAS, 401, 791

Springel, V., Frenk, C. S., & White, S. D. M. 2006, Nature, 440, 1137

Springel, V. & Hernquist, L. 2003, MNRAS, 339, 312

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629

Sunyaev, R. A. & Zeldovich, Y. B. 1970, , 7, 3

Sunyaev, R. A. & Zeldovich, Y. B. 1972, Comments on Astrophysics and Space Physics, 4, 173

Takahashi, R., Sato, M., Nishimichi, T., Taruya, A., & Oguri, M. 2012, ApJ, 761, 152

Taylor, M. B. 2006, in Astronomical Society of the Pacific Conference Series, Vol. 351, Astronomical Data Analysis Software and Systems XV, ed. C. Gabriel, C. Arviset, D. Ponz, & S. Enrique, 666

Tegmark, M. 1997, , 79, 3806

Tegmark, M., Blanton, M. R., Strauss, M. A., et al. 2004, ApJ, 606, 702

Teyssier, R., Chapon, D., & Bournaud, F. 2010, ApJ, 720, L149

Tomczak, A. R., Lemaux, B. C., Lubin, L. M., et al. 2019, MNRAS, 484, 4695

Toomre, A. & Toomre, J. 1972, ApJ, 178, 623

Tully, R. B. 1987, ApJ, 321, 280

Umetsu, K., Medezinski, E., Nonino, M., et al. 2014, ApJ, 795, 163

Valluri, M. 1993, ApJ, 408, 57

van de Weygaert, R. 2016, in The Zeldovich Universe: Genesis and Growth of the Cosmic Web, ed. R. van de Weygaert, S. Shandarin, E. Saar, & J. Einasto, Vol. 308, 493–523

van de Weygaert, R. & Bond, J. R. 2008, in A Pan-Chromatic View of Clusters of Galaxies and the Large-Scale Structure, ed. M. Plionis, O. López-Cruz, & D. Hughes, Vol. 740, 24

van der Burg, R. F. J., Rudnick, G., Balogh, M. L., et al. 2020, A&A, 638, A112

van Haarlem, M. & van de Weygaert, R. 1993, ApJ, 418, 544

van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, A&A, 556, A2

van Weeren, R. J., de Gasperin, F., Akamatsu, H., et al. 2019, Space Sci. Rev., 215, 16

Vanzella, E., Calura, F., Meneghetti, M., et al. 2019, MNRAS, 483, 3618

Vanzella, E., Caminha, G. B., Rosati, P., et al. 2021, A&A, 646, A57

Vazza, F., Brunetti, G., Kritsuk, A., et al. 2009, A&A, 504, 33

Verdugo, M., Lerchster, M., Böhringer, H., et al. 2012, MNRAS, 421, 1949

Verdugo, M., Ziegler, B. L., & Gerken, B. 2008, A&A, 486, 9

Veropalumbo, A., Sáez Casares, I., Branchini, E., et al. 2021, MNRAS, 507, 1184

Vikhlinin, A., Kravtsov, A., Forman, W., et al. 2006, ApJ, 640, 691

Villalobos, Á., De Lucia, G., & Murante, G. 2014, MNRAS, 444, 313

Vogelsberger, M., Genel, S., Springel, V., et al. 2014, MNRAS, 444, 1518

Vogelsberger, M., Marinacci, F., Torrey, P., & Puchwein, E. 2020, Nature Reviews Physics, 2, 42

von der Linden, A., Wild, V., Kauffmann, G., White, S. D. M., & Weinmann, S. 2010, MNRAS, 404, 1231

Vulcani, B., Poggianti, B. M., Fasano, G., et al. 2012, MNRAS, 420, 1481

Vulcani, B., Poggianti, B. M., Fritz, J., et al. 2015, ApJ, 798, 52

Vulcani, B., Poggianti, B. M., Smith, R., et al. 2022, ApJ, 927, 91

Vulcani, B., Treu, T., Nipoti, C., et al. 2017, ApJ, 837, 126

Vulcani, B., Treu, T., Schmidt, K. B., et al. 2016, ApJ, 833, 178

Walker, S., Simionescu, A., Nagai, D., et al. 2019, Space Sci. Rev., 215, 7

Wang, D., Zhang, Y.-X., Liu, C., & Zhao, Y.-H. 2008, , 8, 119

Wang, L. & Steinhardt, P. J. 1998, ApJ, 508, 483

Way, M. J., Foster, L. V., Gazis, P. R., & Srivastava, A. N. 2009, ApJ, 706, 623

Welker, C., Bland-Hawthorn, J., van de Sande, J., et al. 2020, MNRAS, 491, 2864

Wetzel, A. R., Deason, A. J., & Garrison-Kimmel, S. 2015, ApJ, 807, 49

Wetzel, A. R., Tinker, J. L., & Conroy, C. 2012, MNRAS, 424, 232

Wetzel, A. R., Tinker, J. L., Conroy, C., & van den Bosch, F. C. 2013, MNRAS, 432, 336

White, S. D. M. & Rees, M. J. 1978, MNRAS, 183, 341

Wild, V., Kauffmann, G., Heckman, T., et al. 2007, MNRAS, 381, 543

Williams, J. P., Blitz, L., & McKee, C. F. 2000, in Protostars and Planets IV, ed. V. Mannings, A. P. Boss, & S. S. Russell, 97

Willott, C. J., Doyon, R., Albert, L., et al. 2022, PASP, 134, 025002

York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, AJ, 120, 1579

Zabludoff, A. I. & Mulchaey, J. S. 1998, ApJ, 496, 39

Zinger, E., Dekel, A., Kravtsov, A. V., & Nagai, D. 2018, MNRAS, 475, 3654

Zitrin, A., Meneghetti, M., Umetsu, K., et al. 2013, ApJ, 762, L30

Zwicky, F. 1933, Helvetica Physica Acta, 6, 110

Zwicky, F. 1937, ApJ, 86, 217

Zwicky, F. 1951, PASP, 63, 61