Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

ASTRONOMIA

Ciclo XXIX

Settore Concorsuale di afferenza: 02/C1

Settore Scientifico disciplinare: FIS/05

THE ROLE OF THE ENVIRONMENT IN GALAXY EVOLUTION

Presentata da: **Nicola Malavasi**

Coordinatore Dottorato

Prof. Francesco R. Ferraro

Relatore

Prof. Andrea Cimatti

Correlatori

Dott.ssa Lucia Pozzetti
Dott. Sandro Bardelli
Dott.ssa Olga Cucciati

Esame finale anno 2016

A Milena

"Che fai tu, Luna, in ciel? Dimmi, che fai,
silenziosa Luna?"

— Giacomo Leopardi,
Canto notturno di un pastore errante dell'Asia

ABSTRACT

Although the general picture of how galaxies have formed and evolved through cosmic time is already in place, still many details are missing. An example is represented by a quantitative characterization of the role that environment plays during the evolution of galaxies in determining their observed properties in the local Universe. Local galaxy environment has been thoroughly explored through observations at low redshift and it has recently started to being investigated also in the distant Universe. Moreover, general attention has been increasingly focused on the relation between galaxy evolution and galaxy positions with respect to the large-scale structures of the cosmic web.

This work explores how environment (defined both in terms of local overdensities on the Mpc scale around galaxies, as well as in terms of the distribution of galaxy distances from $\sim 10^2$ Mpc long cosmic filaments) is observationally related to galaxy properties such as their stellar mass, their star-formation activity, or the presence of Active Galactic Nucleus phenomena. The analysis reported here is performed in a large redshift range, going from the local Universe ($z \gtrsim 0.2$) to intermediate ($z \sim 0.7$) and high-redshifts ($z \lesssim 3$). To perform this study, both spectroscopic redshift data sets as well as photometric redshift surveys have been deployed and (particularly for the photometric redshift data sets) a thorough analysis of the effect of their observational issues on the measurement of local and global environment has been executed.

The picture that emerges is that high-density environments are favoured places for galaxy evolution, showing an enhanced fraction of massive galaxies which have stopped forming stars and belong to the quiescent population. Galaxies in high-density environments have an increased probability of hosting radio-AGNs (belonging to the ensemble of processes that turn off the star-formation in massive galaxies) and these effects are visible since the start of the decline in the cosmic star-formation history. Large-scale structures, such as the filaments of the cosmic web, have also a role in shaping galaxy evolution and they may represent key places for the pre-processing of galaxies before they enter higher density regions such as the cluster or group environment.

As an increasing amount of observational evidence on the relation between the evolution of galaxies and their environment is gathered, a complete solution to the problem of how the local galaxy population has evolved approaches, and the history of galaxies from their formation at the dawn of time to the present day gets closer to being traced.

PUBLICATIONS

The work in this thesis is based on the following publications:

Malavasi, N., Arnouts, S., Vibert, D., et al., *The VIMOS Public Extragalactic Redshift Survey (VIPERS): Galaxy Segregation Inside Filaments At* $z \simeq 0.7$, MNRAS accepted, 2016.

Malavasi, N., Pozzetti, L., Cucciati, O., et al., *Reconstructing The Galaxy Density Field With Photometric Redshifts: II. The Environment-Dependent Evolution Of Massive Galaxies To* $z \simeq 3$, submitted to MNRAS, 2016.

Malavasi, N., Bardelli, S., Ciliegi, P., et al., *The Environment Of Radio Sources In The VLA-COSMOS Survey Field*, 2016, ASSP, Volume 42, Page 109.

Malavasi, N., Pozzetti, L., Cucciati, O., et al., *Reconstructing The Galaxy Density Field With Photometric Redshifts: I. Methodology And Validation On Stellar Mass Functions*, 2015, A&A, Volume 585, Page A116.

Malavasi, N., Bardelli, S., Ciliegi, P., et al., *The Environment Of Radio Sources In The VLA-COSMOS Survey Field*, 2016, A&A, Volume 576, Page A101.

Focardi, P., Malavasi, N., *The Effect Of The Environment On The Faber-Jackson Relation*, 2012, ApJ, Volume 756, Page 117.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

2dFGRS Two-degree Field Galaxy Redshift Survey

2MASS Two-Micron All-Sky Survey

AGN Active Galactic Nucleus

ALHAMBRA Advanced Large Homogeneous Area Medium Band Redshift Astronomical

BOSS Baryon Oscillation Spectroscopic Survey

CANDELS Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey

CFHTLS Canada-France-Hawaii Telescope Legacy Survey

CL Confidence Level

CMB Cosmic Microwave Background

COSMOS Cosmic Evolution Survey

CW Cosmic Web

CWD Cosmic Web Detachment

CWI Cosmic Web Imager

dec declination

DisPerSE Discrete Persistent Structure Extractor

DM Dark Matter

DTFE Delaunay Tessellation Field Estimator

DTFI Delaunay Tessellation Field Interpolator

FoG Finger of God

FWHM Full Width Half Maximum

GAMA Galaxy and Mass Assembly

GOODS Great Observatories Origins Deep Survey

GSMF Galaxy Stellar Mass Function

HETDEX Hobby-Eberly Telescope Dark Energy Experiment

ICM Intra-Cluster Medium

IFU Integral Field Unit Spectrograph

IGM Inter-Galactic Medium

IMF Initial Mass Function

IR Infra-red

J-PAS Javalambre Physics of the accelerating universe Astronomical Survey

KiDS Kilo-Degree Survey

KS Kolmogorov-Smirnov

LAE Lyman-α Emitter

ΛCDM Λ-Cold Dark Matter

LCRS Las Campanas Redshift Survey

LSS Large Scale Structure of the Universe

MAGPHYS Multi-wavelength Analysis of Galaxy Physical Properties

MLS Multi-λ Survey

MMF Multiscale Morphology Filter

MOONS Multi-Object Optical and Near-infrared Spectrograph

MOS Multi-Object Spectrograph

NIR Near Infra-red

NN Nearest Neighbour

NUV Near Ultra-violet

PDF Probability Distribution Function

PDR Public Data Release

PEP PACS Evolutionary Probe

PFS Prime Focus Spectrograph

PRIMUS PRIsm MUlti-object Survey

QSO Quasi-Stellar Object

RA Right Ascension

SAM Semi-Analytical Model

SDSS Sloan Digital Sky Survey

SED Spectral Energy Distribution

SFH Star Formation History

SFR Star Formation Rate

SI Système International d'unités (International System of Units)

SKA Square Kilometer Array

SMBH Super-Massive Black Hole

S/N signal-to-noise

SN Supernova

sSFR specific Star Formation Rate

SWML Step-Wise Maximum-Likelihood

TSR Target Sampling Rate

UDS Ultra-Deep Survey

UKIDSS UKIRT Infrared Deep Sky Survey

UKIRT UK Infra-red Telescope

ULIRG Ultra-Luminous Infrared Galaxy

UV Ultra-violet

VIDEO VISTA Deep Extragalactic Observations

VIKING VISTA Kilo-degree Infrared Galaxy survey

VIMOS VIsible Multi-Object Spectrograph

VIPERS VIMOS Public Extragalactic Redshift Survey

VISTA Visible and Infrared Survey Telescope for Astronomy

VLA Very Large Array

VLT Very Large Telescope

VUDS VIMOS Ultra-Deep Survey

VVDS VIMOS-VLT Deep Survey

WFIRST Wide Field Infra-red Survey Telescope

Part I

THE STUDY OF GALAXY EVOLUTION IN RELATION TO ENVIRONMENT IN MODERN ASTROPHYSICS

1

INTRODUCTION

Galaxies are vast structures, composed essentially of Dark Matter (DM), stars, and variable amounts of gas and dust. Ever since the Shapley-Curtis "Great Debate" (1920), in which the true dimensions of our galaxy, the Milky Way, have been understood and the existence of galaxies other than our own has been recognised ("Island Universes", see e. g. Hubble, 1925), astronomers have started to investigate how such structures could have formed and evolved, trying to trace their history from the early Universe to the present times. Starting with the work of Hubble (1926), which proposed a first classification of galaxies based on differences in their morphologies (see also Hubble, 1936; Sandage, 1961), extragalactic astronomy has quickly progressed in understanding many aspects of the process of galaxy formation and evolution.

It is now widely recognised that galaxies can be distinguished into two broad classes, an aspect which it is often referred to as "galaxy bimodality" (see e. g. Baldry et al., 2004; Blanton et al., 2003; Strateva et al., 2001). This distinction has evolved from being purely based on colour and/or morphology to a real characterisation of galaxies in two distinct types: a class of red galaxies, with elliptical morphologies, low Star Formation Rates (SFRs), and almost no gas or dust content and a complementary class of blue galaxies, with disk-like (showing various degrees of spiral patterns) or irregular morphologies, active star-formation, and large amounts of gas and interstellar dust (see e. g. Blanton and Moustakas, 2009, and references therein for a review on nearby galaxies).

Moreover, together with the understanding that galaxies were objects located out of the Milky Way, came the discovery that their spatial distribution is not uniform and isotropic, but that they can cluster together (see e. g. Hubble, 1934; Peebles, 1975; Seldner et al., 1977; Shapley and Ames, 1932), forming structures such as clusters and groups, and tracing a complex network of filaments, walls, and voids (Davis et al., 1982; Geller and Huchra, 1989; de Lapparent, Geller, and Huchra, 1986). Astronomers have then started to investigate whether the position of a galaxy in space, its proximity to other galaxies and its being located inside a structure like a cluster or a void region (i. e. its environment) is related to its other properties such as colour, SFR, morphology, or in general to its evolution. Starting with early works on the relation between galaxy environment and morphology (see e. g. Davis and Geller, 1976; Dressler, 1980; Oemler, 1974; Post-

man and Geller, 1984; Sandage and Visvanathan, 1978), the connection between galaxy environment and galaxy evolution has acquired an important role in the study of galaxy formation, becoming an open problem in modern astrophysics.

In recent years, the advent of wide-area, deep spectroscopic and photometric redshift surveys has offered an incredible opportunity for the research of a solution to the problem of how galaxies have formed and evolved, with particular regard to the role played by galaxy environment. Large galaxy samples and powerful statistical methods have been deployed in synergy with complex numerical simulations so that the history of galaxy formation has been traced from the present time ($z = 0$) up to high redshift, with objects detected well into the reionisation era ($z \sim 10$, see e. g. Oesch et al., 2016). Nevertheless, many details of the process that leads to galaxy formation remain undiscovered. An example is constituted by the determination of the importance of the role that environment (both on small and large scales) plays in determining aspects of galaxy evolution (both locally and at high redshift), such as the transition of galaxies from blue, disk-like, star-forming systems to red, passively evolving, elliptical objects. This thesis aims at shedding new light on the problem of studying how galaxy environment impacts galaxy evolution from an observational point of view.

The goal of this Chapter is to provide the framework in which this thesis is located. In particular, Section 1.1 briefly summarises the state of the art in the understanding of the galaxy formation process. Section 1.2 is more focused on the role of the environment in galaxy formation and the observational evidence regarding this topic is summarised. In Section 1.3 some theoretical results useful for the interpretation of observational data are presented, while in Section 1.4 the open questions to which this thesis aims at providing answers are introduced.

1.1 A BRIEF HISTORY OF GALAXY FORMATION

Approximately 380 000 years after the Big Bang ($z \sim 1100$) the protons and electrons composing the ionised medium which filled the Universe combined to form hydrogen atoms, and allowed photons to travel freely across space, them being no more subject to scattering with the particles of the hot plasma. The surface of last scattering for the photons represents the farthest object visible in the Universe and it has been detected as an emission background in the microwaves domain (Penzias and Wilson, 1965), covering the whole sky, with a Black Body emission profile characteristic of a temperature of $T \sim 2.73$ K. It is usually referred to as the Cosmic Microwave Background (CMB) and an example of this emission can be found in Figure 1 where it

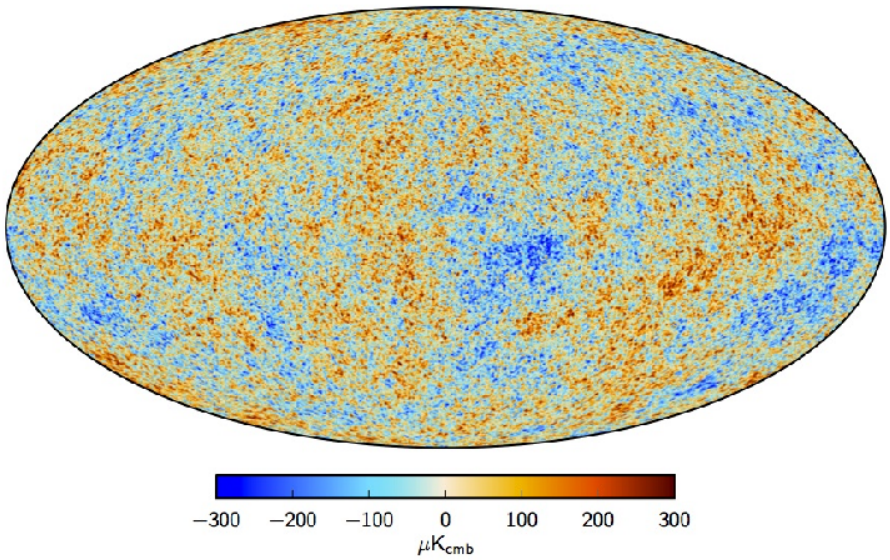


Figure 1: *Angular sky map of the CMB from the Planck satellite*. Temperature deviations from the Black Body profile (anisotropies) are marked as red and blue regions. Figure from Planck Collaboration et al. (2016).

is shown an angular map of the sky in the microwaves taken by the Planck satellite[1] (Planck Collaboration et al., 2016).

Although the Spectral Energy Distribution (SED) for the CMB can be perfectly described by a Black Body emission, still temperature anisotropies of the order of

$$\frac{\Delta T}{T} \sim 10^{-5} \tag{1}$$

can be identified in the map. These temperature anisotropies give direct evidence of the amplitude of the matter density fluctuations at the time of recombination. Matter density fluctuations originate from quantum gravitational fluctuations in the early Universe and represent the seeds of matter overdensities which can grow in time under the effect of gravity. It is from these initial overdensities that galaxies originate (see e. g. White and Rees, 1978).

1.1.1 *Gravitational collapse*

The matter density fluctuations can be described in terms of overdensities relative to the critical density of the Universe $\langle\rho\rangle$

$$\delta = \frac{\rho - \langle\rho\rangle}{\langle\rho\rangle} \tag{2}$$

While the condition $\delta << 1$ is met, perturbations are said to remain in the linear regime. When $\delta \simeq 1$ non linearity is introduced and

1 http://www.esa.int/Our_Activities/Space_Science/Planck

the gravitational collapse of the perturbations can take place. As DM is the first component to decouple from radiation, the only processes which contrast the collapse of DM haloes are the expansion of the Universe and the free streaming of DM particles. When the perturbations grow out of the linear regime they are detached from the Hubble flow and they can grow through dissipationless spherical collapse, to become virialised haloes. Small perturbations are the first to collapse and larger haloes are subsequently created by the merger of smaller ones (see e. g. Kauffmann and White, 1993; Lacey and Cole, 1993). In particular, Press and Schechter (1974) derived a prediction for the mass function of DM haloes based on the gravitational collapse of an initial Gaussian distribution of density fluctuations. A detailed description of the process of structure formation can be found e. g. in Coles and Lucchin (2002).

1.1.2 *Protogalaxy formation*

Figure 2 shows the evolution of density fluctuations as a function of cosmic time for the DM component, the baryonic component, and radiation. The DM fluctuations are the first ones to grow and when baryonic matter finally decouples from radiation the Universe is already populated by DM haloes. These haloes form a distribution of potential wells in which gas can fall. As gas can dissipate gravitational energy trough the emission of radiation, it quickly collapses to the centre of the potential wells, while DM (whose collapse is dissipationless) remains in the form of a diffuse halo.

When baryonic matter enters the DM halo it is shocked and heated up to a temperature of $T \gtrsim 10^5$ K, and the halo is filled with hot gas. A protogalaxy is formed when this gas cools down and reaches the centre of the halo, and star-formation is onset. Intense bursts of star-formation (of the order of 10^2 and up to $10^3 M_\odot yr^{-1}$) at high redshift ($z \gtrsim 2-3$) can be sustained by the infall of fresh gas through cold flows (see e. g. Dekel and Birnboim, 2006; Dekel et al., 2009; Kereš et al., 2005; Martin et al., 2016) or by mergers (see e. g. Welker et al., 2015). The relative importance of gas cooling, smooth accretion through cold flows, and galaxy mergers has not yet been fully understood.

1.1.3 *Star-formation quenching*

At high redshift ($z \gtrsim 2-3$) the Universe is essentially dominated by star-forming galaxies, some of which, like Ultra-Luminous Infrared Galaxies (ULIRGs), can have SFRs up to $1000 M_\odot yr^{-1}$. Figure 3 shows the evolution with redshift of the global SFR density, as derived from Ultra-violet (UV) and Infra-red (IR) observations. The SFR density steadily rises from $z \sim 8$ as more and more galaxies are formed and start form-

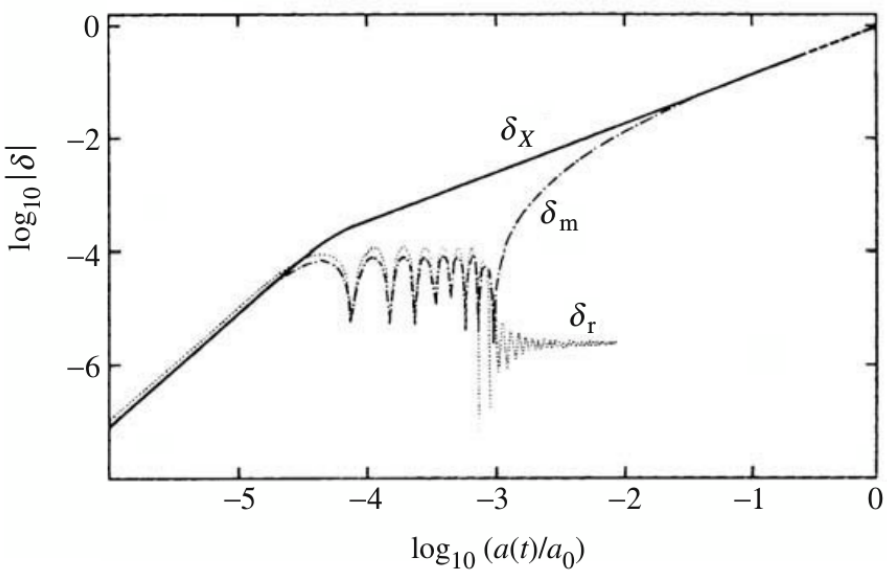


Figure 2: *Growth of structures*. Evolution of the density fluctuations as a function of the expansion factor ($a(t)$) for the DM component (δ_X), the baryonic component (δ_m), and the radiation component (δ_r). Figure from Coles and Lucchin (2002).

ing stars. The SFR density reaches a peak at $z \sim 2$, after which it rapidly declines. This decline in the cosmic SFR density is the cumulative effect of individual galaxies shutting off their star-formation on a short timescale, a phenomenon often referred to as "quenching" .

Star-formation quenching happens at earlier epochs for more massive galaxies, a trend often referred to as "downsizing" (see e. g. Cimatti, Daddi, and Renzini, 2006, and references therein). More massive galaxies ($M^* \geqslant 10^{11-12} M_\odot$) are the first ones to shut off their star-formation and assemble their final mass, followed by progressively lower mass galaxies at later epochs (see also Renzini, 2006; Thomas et al., 2005).

Although several mechanisms for star-formation quenching have been proposed (from stellar and AGN feedback to galaxy-galaxy and galaxy-gas interactions, see Section 1.3) a complete and consistent picture is still missing. The degree to which the proposed physical processes contribute to the cease of star-formation in galaxies has not yet been understood. Nevertheless, the quenching of galaxies has effects on observational properties that are linked to star-formation and it is at the origin of their bimodal trends in colours, masses, and morphologies.

Moreover, an important unresolved problem of modern astrophysics is represented by the lack of understanding of what is the main driver of galaxy evolution and whether the determination of a galaxy final properties depends only on one or a few possible observables. An example of this is constituted by the lack of agreement on whether

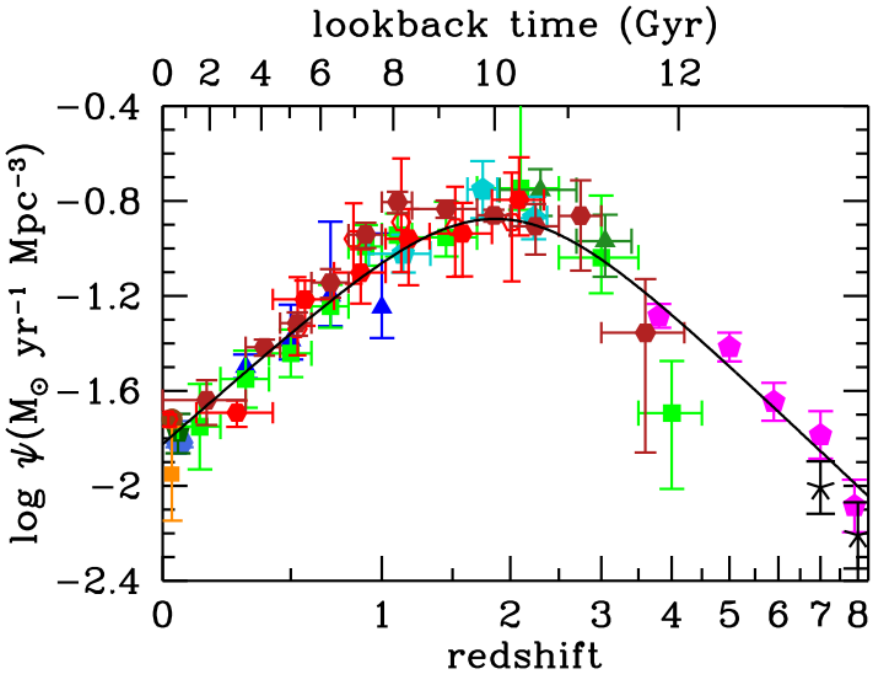


Figure 3: *Evolution of the global SFR density*. Global SFR density (Ψ) as a function of redshift and lookback time. Data points come from UV (magenta, cyan, blue, black, and green points) and IR (orange, dark and light red points) measurements, the black line is the best-fit to the data. Figure from Madau and Dickinson (2014).

star-formation quenching depends only on the environment a galaxy lives in, only on its stellar or host DM halo mass or on both these quantities.

1.2 THE ROLE OF THE ENVIRONMENT IN GALAXY FORMATION

The role of galaxy environment, both on the local and on the global scale, in galaxy evolution is an hot topic in modern astrophysics. Many works have investigated this problem, and a complete review is beyond the goal of this thesis. Nevertheless in this Section the problem of the study of the relation between galaxy environment and galaxy evolution is introduced, together with some of the results achieved in this field from the observational point of view.

One of the first discoveries was that galaxy bimodality was indeed related to galaxy environment. In particular early morphological types (elliptical galaxies) are more frequent in high-density environments, compared to late morphological types (spiral galaxies, see e. g. Dressler, 1980; Giuricin et al., 2001; Guzzo et al., 1997; Hashimoto and Oemler, 1999; Postman and Geller, 1984). As the capability to obtain precise data sets increased and observational campaigns were undertaken to increase the size of statistical samples used and the red-

shift range explored, a more complete picture has started to emerge. In particular many correlations have been found, among galaxy observational properties (e. g. colour, morphology, size), physical properties (e. g. stellar mass, SFR, AGN presence, halo mass) and environmental density, both on the local and on the global scale (Large Scale Structure of the Universe).

These relations have been studied in many different ways, and the various methods to define environment surely have an impact on the strength of the relations recovered. Many of the studies conducted in the literature have measured either the local galaxy density field (through a variety of estimators, see Chapter 2) or they have compared the properties of galaxies in clusters, voids, and the field. Recently also other Large Scale Structure (LSS) features such as filaments and walls have started to be identified and correlated with galaxy properties.

As an extensive and exhaustive review of all the observational results on galaxy environment and galaxy properties is beyond the intent of this Chapter, in the following the main focus will be on those relations between galaxy properties and environment that have been also addressed in this thesis, in particular: the SFR-density relation and the Galaxy Stellar Mass Function (GSMF) for galaxies in different environments (see Chapter 4), the relation between galaxy properties and position within the Large Scale Structure of the Universe (LSS) (see Chapter 5), and the relation between local environment and AGN presence (see Chapter 6). More extensive information can be found in reviews such as Blanton and Moustakas (2009), Boselli and Gavazzi (2006, 2014), and Miley and De Breuck (2008).

1.2.1 *Star Formation Rate-density relation*

Much observational evidence has been gathered on the existence of a relation between galaxy colour and environment. For example, Balogh et al. (2004) and Blanton et al. (2005) studied the optical properties of galaxies up to $z = 0.2$ in the Sloan Digital Sky Survey (SDSS) (York et al., 2000), finding that $(g - r)$ colour is predictive of galaxy environment and that the galaxy colour bimodality tends to disappear in denser environments, which are populated by an enhanced fraction of red galaxies. This result has been confirmed also up to $z = 1$ by e. g. Cooper et al. (2006) using the DEEP2 redshift survey (Davis et al., 2003) and up to $z = 1.5$ by e. g. Cucciati et al. (2006) using the VIMOS-VLT Deep Survey (VVDS) (Le Fèvre et al., 2005). Figure 4 shows an example of the colour-density relation for galaxies in the VVDS survey up to $z = 1.5$. It can be seen how the fraction of red galaxies increases with overdensity (while the fraction of blue galaxies shows a complementary decrease). The monotonic increase of the fraction of red galaxies with overdensity is the most visible for luminous galax-

ies (bottom panels of each column) and at low redshift (left panels of each row). In some of the highest redshift bins ($1.2 \leqslant z \leqslant 1.5$) hints of a reversal of the relation are visible.

The colour-density relation has its roots in the SFR-density relation. In fact, a well known trend relates galaxy star-formation with environment. This relation has been the subject of intense study (see e. g. Balogh et al., 1998; Christlein and Zabludoff, 2005; Gómez et al., 2003; Lewis et al., 2002; Norberg et al., 2002). In particular, Kauffmann et al. (2004) studied the problem using SDSS data and found that denser environments are dominated by low-SFR and redder galaxies. This can be seen in Figure 5, which shows the specific Star Formation Rate (sSFR) ($\frac{\mathrm{SFR}}{M^*}$) distribution for galaxies in different environments. Denser environments (red line) show a higher fraction of low-SFR galaxies, as opposed to low-density environments (cyan line).

When the evolution of the SFR-density relation is explored, the well defined trends present in the local Universe are generally confirmed at higher redshifts. Nevertheless, Elbaz et al. (2007) found a possible reversal of the SFR-density relation at $z \sim 1$, based on data from the Great Observatories Origins Deep Survey (GOODS) (Dickinson, Giavalisco, and GOODS Team, 2003). According to their analysis at redshift $z \sim 1$ galaxy SFR would be enhanced in high-density environments, instead of being suppressed. This result is not unexpected, as it is predicted that at an early epoch gas-rich mergers happening at the scale of groups trigger starburst events, thus resulting in a higher fraction of star-forming galaxies in high-density environments (see e. g. Cucciati et al., 2012, and references therein).

Still, processes happening in high-density environments are also expected to contribute to the star-formation quenching in galaxies and this scenario is observationally supported by many works (including this thesis, see Chapter 4) which do not find a reversal of the SFR-density relation at $z \gtrsim 1$. For example, the SFR-density relation is at the base of methods for finding clusters and protoclusters in astrophysical surveys by looking at concentrations of red galaxies (see e. g. Strazzullo et al., 2015, and references therein). Moreover, the existence of the SFR-density relation well correlates with results that indicate that galaxies in clusters and denser regions have a lower hydrogen gas content (Cayatte et al., 1994; Chung et al., 2008; Giovanelli and Haynes, 1983; Haynes, Giovanelli, and Chincarini, 1984; Levy et al., 2007; Vogt et al., 2004; Warmels, 1988). It is for this reason that theoretical explanations seek for processes that remove gas from galaxies in high-density environments, thus ending their star-formation (see Section 1.3).

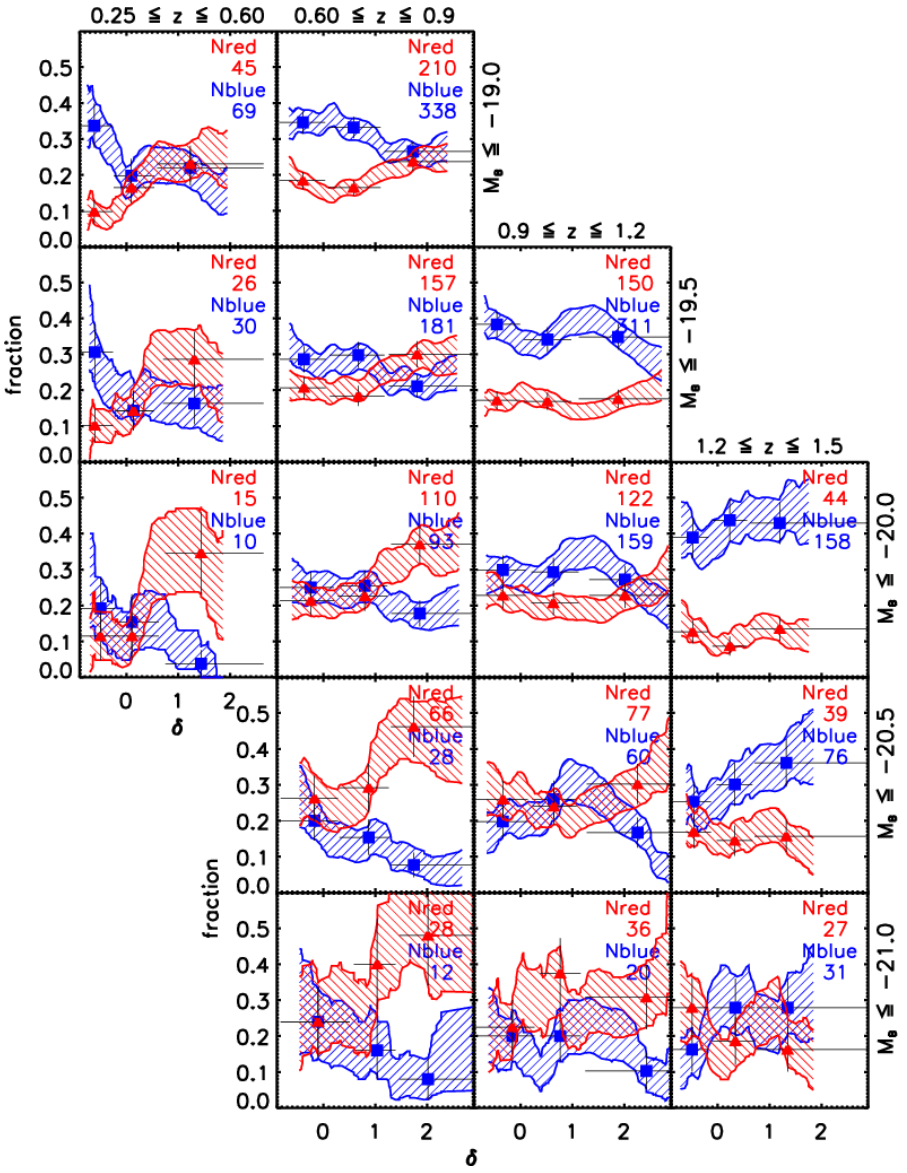


Figure 4: *Colour-density relation for VVDS galaxies*. Fraction of blue galaxies (blue points and shaded area) and red galaxies (red points and shaded area) as a function of overdensity for different redshift bins (columns) and magnitude ranges (rows). The numbers of red and blue galaxies in each redshift and magnitude bin are marked in the top-right corner of each plot. Figure from Cucciati et al. (2006).

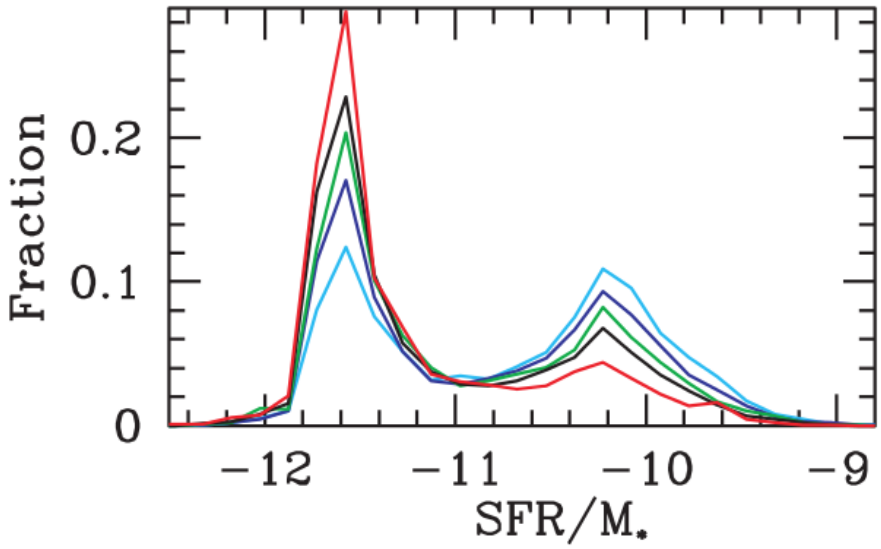


Figure 5: *specific Star Formation Rate distribution in different environments*. Fraction of the total stellar mass in the local Universe contained in galaxies as a function of $\frac{SFR}{M^*}$. Lines are colour-coded according to the number of neighbours of each galaxy: 0 or 1 neighbour (cyan); $2-3$ neighbours (blue); $4-6$ neighbours (green); 7–11 neighbours (black); more than 12 neighbours (red). Figure from Kauffmann et al. (2004).

1.2.2 *The Galaxy Stellar Mass Function in different environments*

The Galaxy Stellar Mass Function (GSMF) is a powerful statistical tool, as it allows to summarise in a single distribution function the dependence on mass of the number density of galaxies in a given volume of space (often defined by the volume explored by a given survey). By studying the evolution of the GSMF with redshift or the shape of the GSMF when it is calculated only for specific galaxy populations (as e. g. only for quiescent or star-forming galaxies or only for galaxies in high- or low-density environments) important conclusions on the processes that regulate galaxy evolution can be drawn. For its great importance and capability of depicting in a simple way complex relationships among various variables, the GSMF has been thoroughly studied by many works.

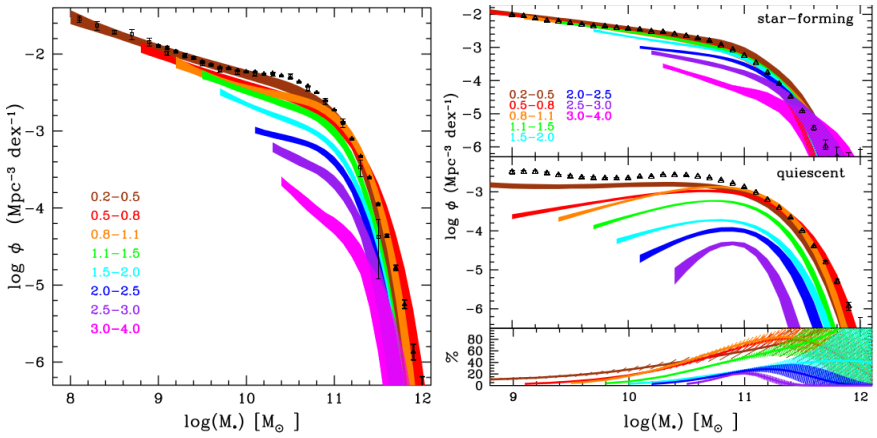
In the local Universe and up to $z \sim 1-1.5$ spectroscopic surveys have been mainly used. In particular, in the local Universe, Baldry, Glazebrook, and Driver (2008) and Baldry et al. (2004) derived the GSMF for galaxies in the SDSS (York et al., 2000), Bell et al. (2003) and Kochanek et al. (2001) relied on both the Two-Micron All-Sky Survey (2MASS) (Jarrett et al., 2000) and the SDSS, Cole et al. (2001) used both the 2MASS and the Two-degree Field Galaxy Redshift Survey (2dFGRS) (Colless et al., 2001), while Baldry et al. (2012) used data from the Galaxy and Mass Assembly (GAMA) survey (Driver et

al., 2011). All these works are consistent in finding that the massive end of the GSMF is dominated by red galaxies, while blue galaxies dominate the low-mass end. At $z = 1$, the GSMF has been analysed in detail, through a wide variety of spectroscopic surveys. For example, Fontana et al. (2004) measured the GSMF for the K20 survey (Cimatti et al., 2002), Pozzetti et al. (2007, 2010) relied on the VVDS (Le Fèvre et al., 2005) and the zCOSMOS surveys (Lilly et al., 2007), Moustakas et al. (2013) used data from the PRIsm MUlti-object Survey (PRIMUS) (Coil et al., 2011), while Davidzon et al. (2013) relied on the VIMOS Public Extragalactic Redshift Survey (VIPERS) (Guzzo et al., 2014). All these works consistently find that the high-mass end of the GSMF is dominated by red, quiescent galaxies, in complete agreement to what found in the local Universe. Moreover, many of these works also found that the mass at which the quiescent GSMF starts to dominate over the star-forming one (a quantity often referred to as M_{cross}) is monotonically increasing with redshift. This provides further observational support to the downsizing scenario, in which star-formation quenching affects progressively lower mass galaxies with cosmic time.

The GSMF has been explored also at higher redshifts (up to $z = 4$) (see e. g. Ilbert et al., 2010, 2013; Moutard et al., 2016a; Muzzin et al., 2013), using photometric redshift surveys, such as the Cosmic Evolution Survey (COSMOS) and the UltraVISTA survey (McCracken et al., 2012; Scoville et al., 2007) and the VIPERS-Multi-λ Survey (MLS) (Moutard et al., 2016b).

As an example, Figure 6 shows the GSMF for all galaxies (left panel) and for quiescent and star-forming galaxies (right panels) as derived in the redshift range $0 \leqslant z \leqslant 4$ using photometric redshifts from the UltraVISTA survey. This figure shows how the evolution of the GSMF depends on mass, with the low-mass end of the GSMF for the total and star-forming galaxy populations evolving more rapidly than the high-mass end ($M^* \gtrsim 10^{11} M_\odot$). For the GSMF of quiescent galaxies, instead, a significant evolution is visible only at $z \gtrsim 1$.

The GSMF in different environments has been thoroughly investigated in several works, again relying both on spectroscopic surveys up to $z \sim 1-1.5$ and with photometric redshift surveys up to $z \sim 3$. For example, in the local Universe, Bundy et al. (2006) using DEEP2 data (Davis et al., 2003), McNaught-Roberts et al. (2014) relying on GAMA data (Driver et al., 2011), and Balogh et al. (2001), who used the 2MASS (Jarrett et al., 2000) and the Las Campanas Redshift Survey (LCRS) (Shectman et al., 1996), all found consistent results of an enhancement of the massive end of the GSMF (and in particular of the GSMF of quiescent galaxies) in high-density environments. Spectroscopic surveys have allowed the study of the GSMF also up to $z \sim 1-1.5$ (see e. g. Annunziatella et al., 2016; Bolzonella et al., 2010; Cooper et al., 2010; Davidzon et al., 2016; Giodini et al., 2012; Hahn et



(a) (b)

Figure 6: *GSMF evolution from the UltraVISTA survey*. (a): GSMFs for the total galaxy population. (b): GSMFs for star-forming galaxies only (top), GSMFs for quiescent galaxies only (middle), and fraction of quiescent galaxies as a function of stellar mass in each redshift bin (bottom). In each panel GSMF measurements from the UltraVISTA survey are colour coded according to the corresponding redshift bin (marked in colour in each panel). Triangles and squares are local GSMF measurements from Moustakas et al. (2013) and Baldry et al. (2012), respectively. Shaded areas correspond to 68% uncertainty intervals. Figure from Ilbert et al. (2013).

al., 2015; Kodama and Bower, 2003; Vulcani et al., 2012; van der Burg et al., 2013). For example, Bolzonella et al. (2010) studied the GSMF in different environments in the COSMOS field (using the zCOSMOS survey, see Lilly et al., 2007, 2009) up to $z = 1$, finding a difference between the GSMF of high- and low-density environments, with the massive end of the GSMF being more enhanced in high-density environments (see Figure 7).

This result has been confirmed also by Davidzon et al. (2016), by means of the VIPERS survey (Garilli et al., 2014; Guzzo et al., 2014). On the other hand, Vulcani et al. (2012) find no difference for the GSMF of galaxies in clusters and in the field, while van der Burg et al. (2013) reports a general difference in the GSMF for galaxies in clusters and in the field but the same shape for the GSMF of quiescent and star-forming galaxies in each environment. Although these differences could be due to the fact that both these works study a GSMF derived from a compilation of clusters and not from a single system, still the problem of determining whether local environment has an impact on the shape of the GSMF remains open.

The GSMF has been studied in different environments also up to $z \sim 3$, with photometric redshift surveys (see e. g. Darvish et al., 2015; Mortlock et al., 2015; Scoville et al., 2013). The difficulty in performing environmental studies at high redshift relies mainly in

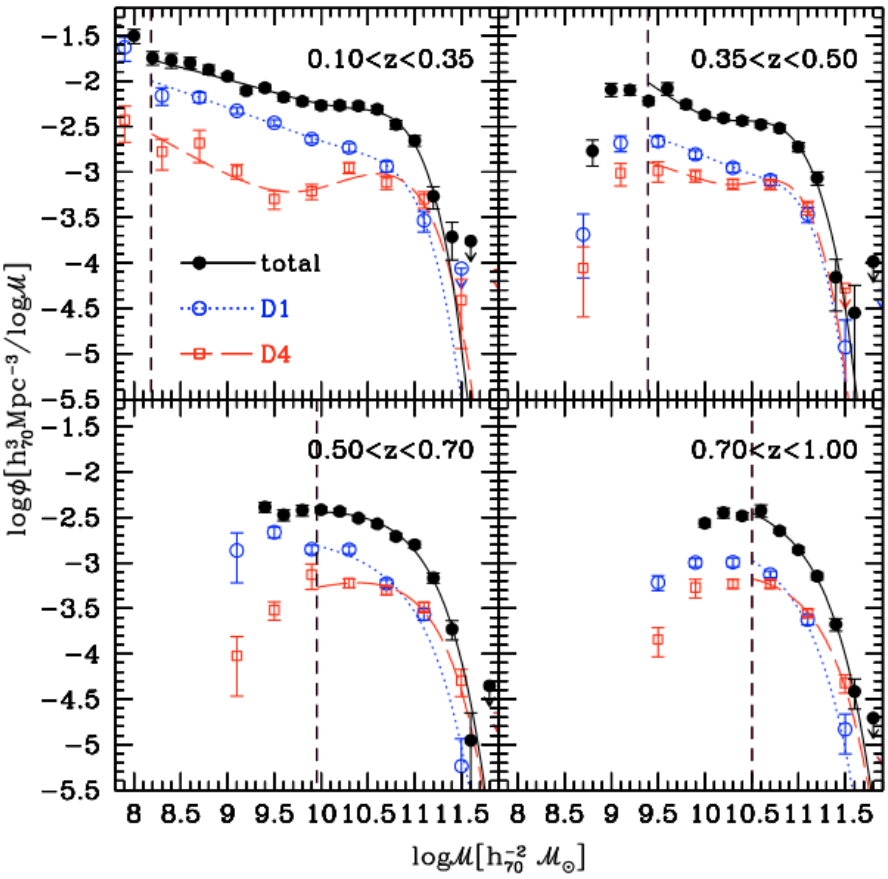


Figure 7: *GSMF in different environments for the zCOSMOS survey*. This figure shows the GSMF for the total population (black points) and separately for high-density (red points) and low-density environments (blue points) in four redshift bins (marked in each panel). Vertical dashed lines correspond to the mass completeness limits for the survey. Figure from Bolzonella et al. (2010).

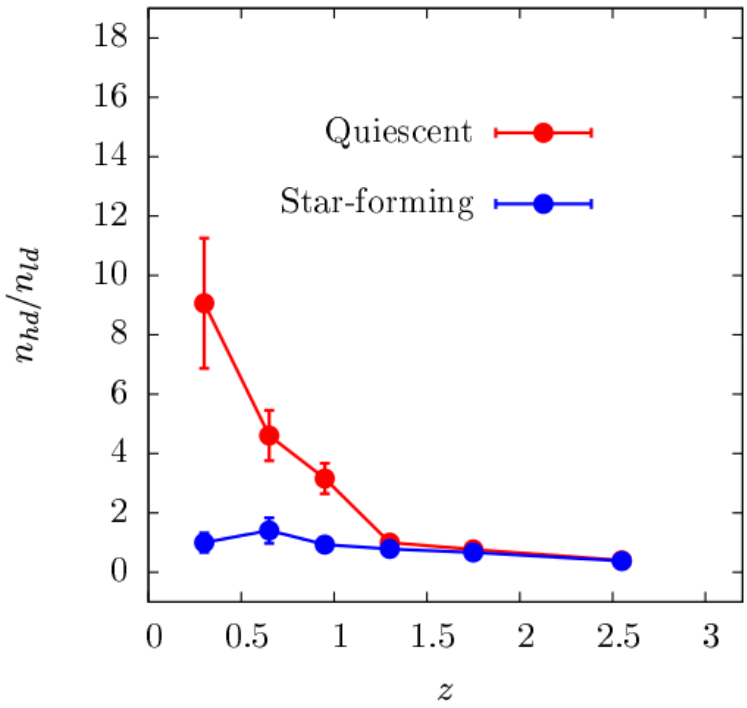


Figure 8: *Galaxy number densities in different environments from the UltraVISTA survey*. Ratio of the number density of galaxies (integral of the GSMF) in high-density environments (n_{hd}) and low-density environments (n_{ld}) for quiescent galaxies (red) and star-forming galaxies (blue) as a function of redshift. Figure from Darvish et al. (2015).

the scarce availability of spectroscopic surveys of all galaxy types that sample a large enough volume (wide area and deep limiting magnitude) with large enough statistical samples. Using photometric redshift surveys performed in the COSMOS field (Ilbert et al., 2013; McCracken et al., 2012), Darvish et al. (2015) found strong evidence that massive ($M > 10^{11} M_{\odot}$), quiescent galaxies increase their number density by an order of magnitude in high-density environments with respect to low-density ones from $z \sim 1.5$ to $z = 0.2$ (see Figure 8). Also using photometric redshifts, in particular the UKIRT Infrared Deep Sky Survey (UKIDSS)-Ultra-Deep Survey (UDS) and the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS) (Galametz et al., 2013; Guo et al., 2013), Mortlock et al., 2015 found that the GSMF is different in high- and low-density environments, with high-density regions dominated by massive ($M > 10^{10} M_{\odot}$) galaxies up to $z \sim 1.5$.

The picture that emerges from all of these considerations is that galaxies in high-density environments are less star-forming, redder and more massive, leading to the conclusion that mass and Star Formation History (SFH) appear to be the two variables that mostly depend on the environment. Moreover, as galaxy properties correlate

with environment in the same way from rich clusters to poor groups, the processes that quench galaxies appear to be the same in all kind of high-density environments, but they occur more frequently in denser regions.

1.2.3 *Galaxy environment and the influence of the Active Galactic Nucleus*

Active Galactic Nuclei (AGNs) are a broad and diversified class of astrophysical sources that have been known and studied since a long time. AGNs are now recognised to be the result of a massive output of radiation from material accreting onto a Super-Massive Black Hole (SMBH). Aside from the study and the classification of these sources, AGNs are thought to be deeply related to galaxy evolution (e. g. a tight relation exists between the mass of the SMBH at the centre of galaxies and the mass of the bulge, see e. g. Ferrarese and Merritt, 2000; Gebhardt et al., 2000; Magorrian et al., 1998). In particular, AGNs are commonly invoked as a possible mechanism (often referred to as AGN feedback) that causes the quenching of star-formation.

Many works have tried to link the different classes of AGNs to the properties of their host galaxies, such as stellar mass, type, SFR, and local environment. In particular, observations indicate that AGNs classified as such due to their emission at radio wavelengths (hereafter radio-AGNs) are commonly located in massive, early-type galaxies (see e. g. Bardelli et al., 2009; Best et al., 2005; Mandelbaum et al., 2009; Matthews, Morgan, and Schmidt, 1964; Sadler, Jenkins, and Kotanyi, 1989; Sadler et al., 2002; Yee and Green, 1987). Therefore, a correlation between radio-AGN presence and environment is somewhat expected. This was studied by many works in the literature and significant evidence was found (see e. g. Bardelli et al., 2010; Croston et al., 2008; Wake et al., 2008, and references therein). Moreover, radio-AGNs are also used as beacons to locate and identify distant protoclusters, thanks to their property of being located in high-density environments (see e. g. Chiaberge et al., 2010; Hatch et al., 2011, 2014, and references therein). Nevertheless, although well explored at low redshift, the relation between environment and radio-AGN presence still lacks a thorough and complete characterisation in the distant Universe.

1.2.4 *Galaxy properties and Large Scale Structure of the Universe*

For long time, the study of the role of environment in galaxy evolution has been been mainly restricted to the study of local density or to the comparison of the properties of galaxies in clusters to galaxies in the field. Only recently, a statistical comparison of the characteristics of galaxies in other LSS features (such as e. g. filaments) has begun (see e. g. Chincarini, 2015, for a review).

The various structures of the Cosmic Web (CW) arise during the collapse of the DM (see Section 1.1.1). In this process, matter departs from underdense regions and flows through dense sheets that wind up, forming filaments along which matter drifts and progressively gets accreted onto high-density peaks. This leads to a CW pattern where dense nodes are connected by filaments, framing walls separated by large voids (see e. g. Bond, Kofman, and Pogosyan, 1996). The baryonic gas follows the gravitational potential gradients imposed by the DM distribution, then is shocked (see Section 1.1.2), forming (among other structures) tenuous ionised hydrogen filaments, the Inter-Galactic Medium (IGM), in which galaxies can form. These filaments are regions where gas, momentum and energy are exchanged through the complex processes of infall and outflow.

Although the CW and the properties of galaxies in relation to their positions with respect to the LSS have been studied in simulations (see Section 1.3), the study of the role of the global environment in galaxy evolution is a new and evolving field, still largely unexplored. The main results in this field have been generally obtained through the use of large statistical data sets obtained thanks to wide-area spectroscopic surveys.

For example, Beygu et al. (2013), Brough et al. (2006), Mahajan, Raychaudhury, and Pimbblet (2012), and Porter et al. (2008) investigated the enhanced SFR of galaxies infalling into clusters through filaments, while Ricciardelli et al. (2014) and Rojas et al. (2004, 2005) using the SDSS (York et al., 2000) found that galaxies in voids have a higher SFR than galaxies located in the walls that surround empty regions. Moreover, Alpaslan et al. (2015, 2016) using the GAMA survey (Driver et al., 2011) found that galaxies close to the centre of filaments are less star-forming and more massive than galaxies located in the filament periphery. On the other hand, Darvish et al. (2014) and Fadda et al. (2008) found an increased SFR for galaxies in filaments compared to clusters. In recent years, a few studies have also started to investigate the CW in great detail, considering all the possible structures (clusters, filaments, walls, and voids) and studying the statistical distribution of galaxy properties with respect to the CW (Brouwer et al., 2016; Eardley et al., 2015; Martínez, Muriel, and Coenda, 2016).

An important aspect of galaxy evolution with respect to CW features is the correlation between the spin of galaxies and their filaments, as recently detected in the SDSS (Tempel, Stoica, and Saar, 2013; Trujillo, Carretero, and Patiri, 2006; Zhang et al., 2013). This result confirms the role played by the large scale dynamical environment in the evolution of galaxies.

The study of galaxy properties in relation to their position in the CW is a hot topic in astrophysics which opens the possibility to study environment under a new perspective and to determine what is the importance of the large scale environment in galaxy evolution.

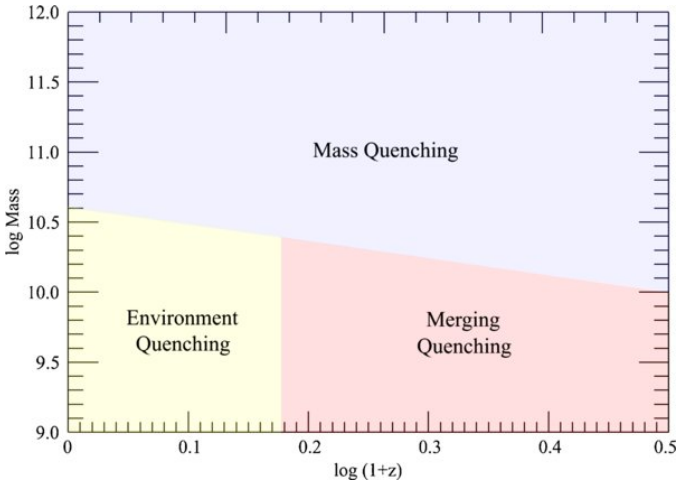
1.3 A THEORETICAL PERSPECTIVE ON THE ROLE OF ENVIRONMENT IN QUENCHING STAR-FORMATION IN GALAXIES

In order to interpret the wealth of observational evidence discussed above, many theoretical and numerical models have been proposed. N-body and hydrodynamical simulations, together with Semi-Analytical Models (SAMs) of galaxy formation, have been extensively used to simulate the formation of galaxies and to try to understand the complex details of their evolution (for a recent review see e. g. Somerville and Davé, 2015).

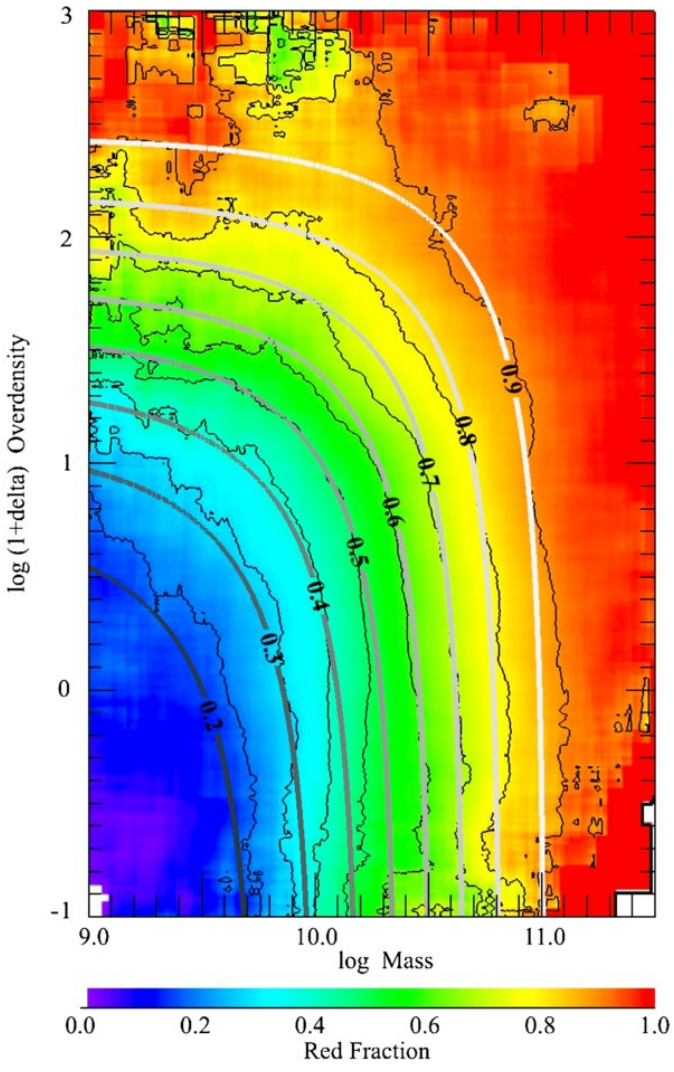
As this thesis is aimed at studying the role of environment in galaxy evolution from an observational point of view, a complete review of all the theoretical works that have been performed is beyond the goal of this work. Nevertheless, in the following a description will be given of a few theoretical and empirical models that have improved the understanding of the role of the environment in galaxy evolution and that have been directly used to interpret the observational results described in the rest of this thesis.

Based on observations of the SDSS and the zCOSMOS GSMFs, Peng et al. (2010) proposed an empirical model in which the quenching process has a separate dependence on mass and environment, as a function of redshift. Figure 9 (top panel) shows a schematic representation of this model. In particular, according to Peng et al. (2010), galaxies above a certain mass ($\sim 10^{10-10.5} M_{\odot}$) are quenched by processes that only depend on galaxy mass (the so-called "mass quenching", commonly identified with AGN feedback), independently of their environment and of their redshift. For lower mass systems, at low redshift quenching happens because of processes that depend only on the environment that galaxies live in, independently of their mass (the so-called "environmental quenching" commonly identified with effects due to galaxies infalling in larger haloes, also called "satellite quenching") while at higher redshift quenching happens because of merging. This empirical model is based on the observation that the fraction of red galaxies at fixed mass is almost independent of environment, while at fixed environment it does not depend on mass (see bottom panel of Figure 9) and on the assumption that the mass quenching efficiency is proportional to the SFR of galaxies.

Interestingly, Gabor and Davé (2015) managed to reproduce the dependence of the red fraction on mass and environment observed by Peng et al. (2010, see Figure 10) by means of an hydrodynamical simulation in which galaxies are quenched when they enter a hot gas-dominated halo. In the model proposed by Gabor and Davé (2015), DM haloes filled with hot gas ($T \geqslant 10^{5.4}$ K, which is always the case for haloes with masses $M_h \geqslant 10^{12} M_{\odot}$) are able to interrupt the flow of cold gas to galaxies and to deprive them of the cold gas reservoir necessary to sustain star-formation. As a result, galaxies that enter a



(a)



(b)

Figure 9: *Separability of the quenching process between mass and environment*. (a): Schematic representation of the empirical model by Peng et al. (2010), which proposes the separability of quenching processes between mass and environment. (b): Fraction of quiescent galaxies as a function of mass and environment in the SDSS and zCOSMOS surveys. Figures from Peng et al. (2010).

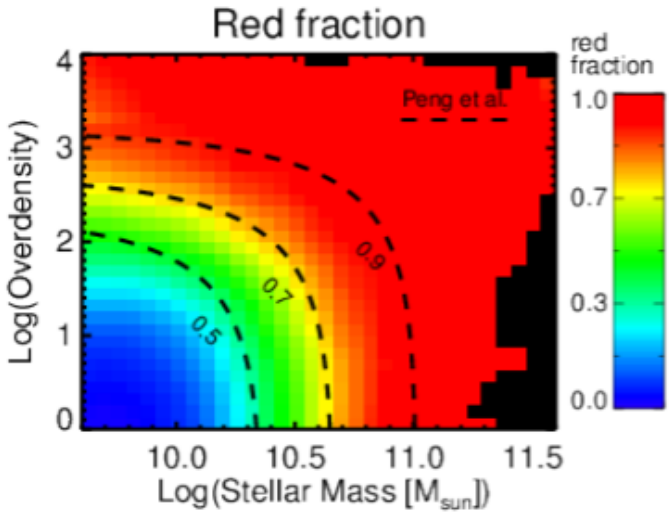


Figure 10: *Fraction of red galaxies reproduced with a quenching model based on hot gas in DM haloes*. Fraction of passive galaxies as a function of mass and environment as derived by hydrodynamical simulations in which galaxies are quenched when they enter a hot gas-dominated halo. The black lines correspond to the observations by Peng et al. (2010). Figure from Gabor and Davé (2015).

hot-gas dominated halo are quenched and become part of the passive population. Galaxies can enter hot gas-dominated haloes independently of their mass either as centrals of their own massive halo or as satellites of other, more massive galaxies.

Regarding the possible physical mechanisms that can quench star-formation in galaxies, a distinction is often operated among internal and external processes. Internal processes are those related to the fact that galaxy formation proceeds differently in overdense regions (the so-called "nature" hypothesis). External processes, instead, come into play only after galaxies have become part of a high-density environment, like a group or a cluster (the so-called "nurture"hypothesis). Nevertheless, it has been recently pointed out that the distinction between mass- and environment-driven quenching or between nature and nurture scenarios is ill-posed, as there exists a tight correlation between galaxy mass and environment (Section 1.2.2).

Among internal processes, stellar feedback (e. g. Supernova explosions and stellar winds) and AGN feedback are the most probable and this agrees well with AGNs being present in massive early-type galaxies (see e. g. the discussion in Hickox et al., 2009). Regarding external processes, they can be classified in two broad classes (see e. g. De Lucia, 2011; De Lucia et al., 2012; Treu et al., 2003), namely interactions between galaxies and other galaxies or the cluster gravitational potential and interactions between galaxies and the hot Intra-Cluster Medium (ICM). Interactions between galaxies and other galaxies or the cluster gravitational potential include:

GALAXY MERGERS: Low-speed interactions between galaxies that happen in relatively high-density regions (infalling groups, outskirts of galaxy clusters) as in the centre of clusters the velocity dispersion is too high for mergers to be efficient. Mergers onset strong dynamical transformations in the galaxies involved and they can significantly affect their SFR (see e. g. Barnes and Hernquist, 1991; Barnes and Hernquist, 1992, 1996; Bekki, 1998; Icke, 1985; Mihos, 1995; Mihos and Hernquist, 1994, 1996, and references therein).

HARASSMENT: High-speed interactions between galaxies that can happen in high-density regions, as the centre of galaxy clusters. It produces less strong transformations than mergers. (see e. g. Farouki and Shapiro, 1981; Moore, Lake, and Katz, 1998; Moore et al., 1996, 1999; Richstone, 1976; Spitzer and Baade, 1951, and references therein).

TIDAL INTERACTIONS: The tidal field of galaxy clusters in which galaxies are located can affect galaxies, compressing them or removing their outer regions. Tidal compression can enhance star-formation in galaxies (see e. g. Byrd and Valtonen, 1990; Fujita, 1998; Henriksen and Byrd, 1996), tidal truncation can suppress it (see e. g. Ghigna et al., 1998; Merritt, 1983, 1984).

Interactions between galaxies and the hot ICM phase, include:

RAM PRESSURE STRIPPING: Stripping of cold gas due to galaxies travelling with high-velocity in a dense and hot medium. It can efficiently suppress star-formation in a galaxy and its efficiency increases at the centre of clusters, where the gas density is higher (see e. g. Abadi, Moore, and Bower, 1999; Cowie and Songaila, 1977; Fujita, 2001; Fujita and Nagashima, 1999; Gunn and Gott, 1972; Nulsen, 1982; Quilis, Moore, and Bower, 2000; Toniazzo and Schindler, 2001, and references therein).

STRANGULATION/STARVATION: When a galaxy enters a hot gas-dominated environment the hot ICM medium can prevent the cooling of fresh gas on the galaxy. The galaxy then rapidly exhausts its cold gas reservoir and the star-formation is quenched. It may have a long timescale (see e. g. Balogh, Navarro, and Morris, 2000; Larson, Tinsley, and Caldwell, 1980).

In particular, by performing a detailed study of a single cluster at various radii and using simulations to understand the typical timescales of the various processes, Treu et al. (2003) managed to obtain information on the range of distances from the cluster centre at which these processes are efficient (reported in Figure 11 for reference).

In recent times also the LSS environment has gained an increased importance in models of galaxy evolution. For example, Pichon et al.

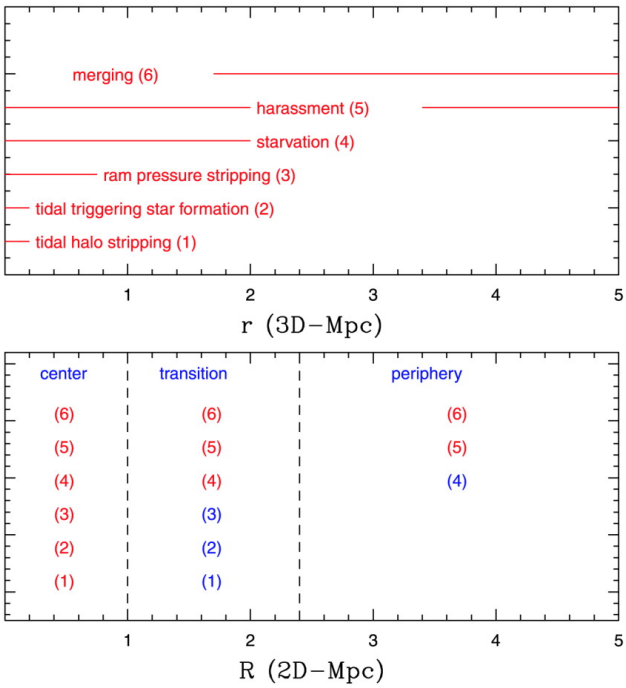


Figure 11: *Distance at which each physical quenching process is effective inside a cluster*. This figure shows the range of distances from the cluster centre in 3D (top panel) and in projection (bottom panel) at which the various physical mechanisms that can quench star-formation in a galaxy are effective. Figure from Treu et al. (2003).

(2011) proposed that the filamentary flows advect angular momentum onto the disks of galaxies and that the spin of newly formed galaxies tends preferentially to be parallel to the axis of their closest filament. Codis et al. (2012) quantified a mass transition, with the most massive haloes ending up with a spin perpendicular to the filaments as a result of successive mergers along the filaments. These results have been extended to galaxies by Dubois et al. (2014) with the state-of-the-art hydrodynamical simulation Horizon-AGN (see also Codis, Pichon, and Pogosyan, 2015, for a theoretical motivation for this transition based on constrained tidal torque theory). Moreover, new models like the recently developed Cosmic Web Detachment (CWD) model (Aragon-Calvo, Neyrinck, and Silk, 2016) propose a unification of the processes described above, based on the connection of the galaxy and the features of the CW.

1.4 OPEN PROBLEMS

Although the general picture of galaxy formation and evolution is already in place, still many details are missing. It is currently known that galaxies form from the gravitational collapse of baryonic gas in the potential wells set by DM, but the exact mechanism that brings gas to cool and form disks with intense bursts of star-formation is still

not well constrained. Moreover, although several mechanisms that can interrupt the formation of stars in a galaxy have been proposed, still it has not been well understood what role is played at a given redshift by local environment and by the Large Scale Structure of the Universe. Particularly interesting open problems are the following:

1. How does the percentage of quiescent and star-forming galaxies changes as a function of mass and environment?
2. At what redshift do processes linked to local environment start to become effective?
3. Do LSS features (such as filaments) have any role in the mass assembly and star-formation quenching of galaxies?
4. Can environment be responsible also for other quenching mechanisms than those commonly associated with galaxy-galaxy and galaxy-ICM interaction (e. g. AGN feedback)?

The aim of this thesis is to provide answers to the questions above. This work is an investigation of the role of environment in quenching the star-formation, in determining the build up of galaxy mass, and the transition from a star-forming galaxy population to a quiescent one. Environment is defined both in terms of the local density field and in terms of LSS features, and the study is performed considering both the small overdensities of galaxies (at typical sizes of a few hundreds of kpc) and some of the largest known structures (tens to hundreds of Mpc long Cosmic Web filaments). Environmental effects on the evolution of galaxies are traced from the nearby Universe ($z \sim 0.7$ where wide-area spectroscopic surveys have allowed us to obtain an extended knowledge of the physical processes happening in galaxies and cosmic structures) up to high-redshift ($z \sim 2-3$) where a large fraction of structures have yet to form and environmental effects are less explored.

In particular, to study the transition from the star-forming to the quiescent galaxy population as a function of mass, redshift and environment, the Galaxy Stellar Mass Function, a powerful statistical tool for the study of galaxy evolution is used. In order to explore the effect of environment up to high redshift high-precision photometric redshifts are employed and the effect of their uncertainty on the measurement of local environment is carefully tested on mock galaxy catalogues. Photometric redshifts (albeit less precise ones) are also used to study the relation between radio-AGN presence and local environment. Instead, precise spectroscopic measurements are used to accurately recover the complex network of CW filamentary structures in a large volume of space and to study their impact on galaxy evolution.

This thesis is structured as follows: in Chapter 2 the majority of the methods available in the literature to compute local and global

environment are reviewed and those used for the research presented in this work are introduced. In Chapter 3 and 4 a study of the GSMF in different environments, for both quiescent and star-forming galaxies up to $z \sim 3$ is presented. In particular in Chapter 3 the extensive tests that have been performed to study whether it is still possible to study the GSMF in different environments if the local density field is measured with photometric redshifts are described. In Chapter 5 galaxy properties are studied in relation to their position with respect to filaments of the CW. Finally, in Chapter 6 an analysis of the relation between local environment and radio-AGN presence, performed up to $z \sim 2$, is described. In Chapter 7 results are summarised and conclusions are drawn.

2 METHODS

The study of the role of the environment in galaxy evolution is based on the observation of existing correlations among the physical properties of galaxies and a measure of their position with respect to cosmic structures or of the local density around them. All these informations can be extracted from photometric and spectroscopic redshift surveys.

Many techniques and mathematical methods can be used to measure the local and global environment of galaxies, while important statistical tools (such as the GSMF, which allows us to summarize in a single distribution function how the galaxy number density changes as a function of mass, redshift, galaxy type, and environment) can be used to obtain a complete view of galaxy properties with respect to their position in the density field.

In this Chapter the main mathematical and statistical methods employed in the rest of this thesis are described. As a plethora of different techniques have been proposed in the literature to give a mathematical description of the galaxy density distribution, an exhaustive list of all the possible methods is beyond the goal of this Chapter. Nevertheless, those estimators that are linked to the analysis that is presented here, either because they have been directly employed or because they have been used by works whose results are compared to those presented in this thesis, are reviewed. Moreover, in the last part of this Chapter, particular attention is given to the description of the GSMF, the main statistical tool to study galaxy properties that has been adopted in the rest of this thesis.

This Chapter is structured as it follows: in Section 2.1 the methods through which local environment can be measured are reviewed. In Section 2.2 the problem of measuring global environment is presented and a brief list of a few existing algorithms to identify LSS (Section 2.2.1) is provided, together with a thorough description of the method used to perform the analysis presented in Chapter 5 (Section 2.2.2). Finally, in Section 2.3 the algorithm used to calculate GSMFs and employed for the analysis presented in Chapters 3 and 4 is introduced.

2.1 LOCAL ENVIRONMENT

The study of local environment aims at providing a description of the matter distribution in the Universe using the galaxy spatial distribution (an example coming from the VIPERS survey is shown in

Figure 12[1]) as a (biased) tracer of the total (luminous + dark) matter disposition.

Therefore, the simplest description of local environment is performed through counting the number of sources N at a given position in space $\vec{r}$, ($N(\vec{r})$). From these source counts a measurement of the local density field ($\rho(\vec{r})$, usually expressed in units of Mpc^{-3}) can be derived

$$\rho(\vec{r}) = N(\vec{r})/V(\vec{r}) \tag{3}$$

where $V(\vec{r})$ is a given volume of space. Often, local environment is quantified in terms of the density contrast ($\delta(\vec{r})$), defined as

$$\delta(\vec{r}) = \frac{\rho(\vec{r}) - \bar{\rho}}{\bar{\rho}} \tag{4}$$

where $\bar{\rho}$ is the mean density of the Universe. If local environment is not measured in 3D, but only in projection, then a surface density of sources ($\Sigma(\vec{r})$, expressed in units of Mpc^{-2} or deg^{-2}) is sometimes used, defined as

$$\Sigma(\vec{r}) = N(\vec{r})/S(\vec{r}) \tag{5}$$

where $S(\vec{r})$ is a given surface.

The simplest measurement of local environment consists in counting the number of sources present at a given position in space. For the sake of clarity, in the rest of this thesis I will often refer to the following definitions:

TARGETS: the positions $\vec{r}$ *at* which local environment (in terms of $N(\vec{r})$, $\rho(\vec{r})$, $\Sigma(\vec{r})$, or $\delta(\vec{r})$) is measured. These positions can coincide with the position of a galaxy (in this case the galaxy's local environment is measured), with random positions in space (see e. g. Cucciati et al., 2016), or they can be distributed on a regular grid (see e. g. Kovač et al., 2010).

TRACERS: the sample of $N_{tracers}$ sources that are considered by the density estimator, *through* which local environment is measured. They are used to measure local environment at the considered target position and the tracer sample may be constituted by all galaxies or by only a selected galaxy population (e. g. only quiescent galaxies, see e. g. Strazzullo et al. 2015, only galaxies above a given mass or luminosity threshold, see e. g. Kovač et al. 2010, or only peculiar galaxies, like Lyman-α Emitters (LAEs), see e. g. Dey et al. 2016). For example, if all galaxies are used as tracers down to the magnitude limit of a survey (i. e. a flux-limited sample of tracers is adopted) the environment can be studied to much smaller scales than by using only galaxies above a given

1 Image taken from the VIPERS press release page: http://vipers.inaf.it/pr/

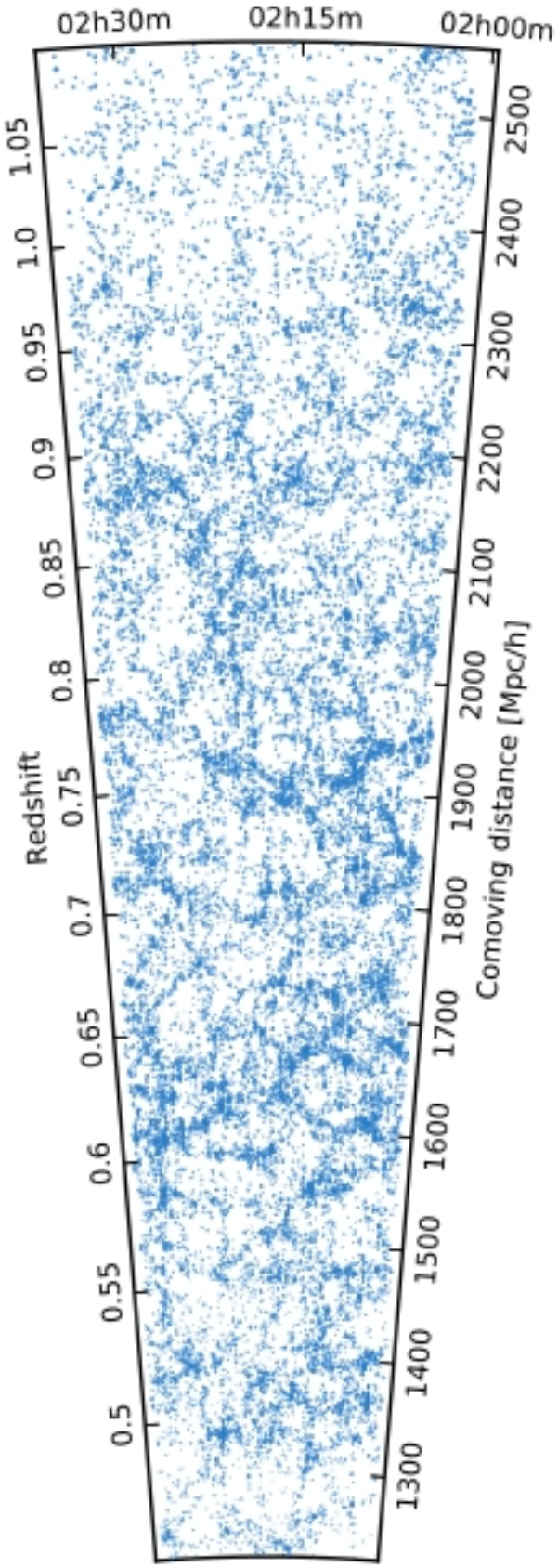


Figure 12: Galaxy spatial distribution as traced by the VIPERS spectroscopic survey (Garilli et al., 2014; Guzzo et al., 2014) in the W1 field of the CFHTLS-Wide imaging survey.

luminosity threshold (which can be chosen evolving with redshift, in order to obtain a volume-limited sample of tracers). In contrast, the choice of a volume-limited tracer sample ensures that the same galaxy population is considered at every redshift. In the following, the position of the i-th tracer in the sample will be indicated with the symbol $\vec{\mathbf{r}}_i$.

SMOOTHING FILTER: it is the spatial "window" function (also referred to as "kernel" function) inside which the source counts ($N(\vec{\mathbf{r}})$) are obtained. The smoothing filter can be of fixed size (as in the case of the fixed aperture and Gaussian kernel estimators, Sections 2.1.1 and 2.1.2) or adaptive (as in the case of Nearest Neighbour or tessellation methods, Sections 2.1.3 and 2.1.4). Fixed size filters are local environment estimators which allow the measurement of the density field in a volume of fixed physical dimensions and therefore will be useful to trace those environmental processes that happen on the chosen scale size. The drawback is, of course, the excessive smoothing of the densest environments and the loss of information on processes happening at smaller scales than the fixed one. In particular the fixed aperture method is the one that has been used to perform the environmental measurements described in Chapters 3, 4, and 6[2]. On the other hand, an adaptive filter means that the scale on which local environment is measured changes according to the density at the target position. Adaptive methods trace smaller scales in denser environments, without performing any sort of smoothing of the density field. This allows these methods to accurately trace local density features in high-density environments, without losing information. The drawback is that it is not possible to study local density on a given fixed physical scale and high-density and low-density environments will be traced by different scale sizes. An adaptive tessellation method will be used extensively for the analysis presented in Chapter 5.

EDGE EFFECTS: as galaxy surveys are performed inside specific Right Ascension (RA) and declination (dec) limits, galaxies placed close to the survey boundaries may have their environment underestimated due to part of the smoothing filter volume falling outside of the survey. Standard approaches consist in rejecting galaxies for which the smoothing filter volume is outside of the survey edges by a threshold fraction and correcting the environmental measure for the rest of the galaxies by dividing $\rho(\vec{\mathbf{r}})$ by the fraction of the smoothing filter volume inside the survey boundaries. These standard approaches have been adopted to take

2 The measurements of local environment described in these Chapters have been performed with a *Python* routine, that I have expressly written to implement a fixed aperture method and suitably optimised to run on large data sets characterised by photometric redshifts.

into account edge effects on the environmental measurement performed in Chapter 4. More sophisticated approaches can be adopted, depending on the survey specific layout (see e. g. Cucciati et al., 2014).

In order to obtain a precise measurement of local environment, both target and tracer positions have to be accurate. Usually, positions are expressed in terms of their angular coordinates RA (α) and dec (δ), and in terms of their redshift z (therefore $\vec{\mathbf{r}} = (\alpha, \delta, z)$). In the following, especially in Chapter 5, these will be converted to Cartesian positions ($\vec{\mathbf{r}} = (p_x, p_y, p_z)$).

In the rest of this Section, a description of some of the available estimators for the measurement of local environment (namely counts in an aperture, Section 2.1.1, Gaussian Kernel estimator, Section 2.1.2, Nearest Neighbour (NN) identification, Section 2.1.3, tessellation methods, Section 2.1.4) will be provided. In Section 2.1.5 the use of photometric redshifts to measure local environment, a topic which will be thoroughly addressed in Chapter 3, will be briefly discussed.

2.1.1 *Fixed aperture*

The fixed aperture method, (see e. g. Gallazzi et al., 2009; Grützbauch et al., 2011) consists in counting galaxies inside a smoothing filter with a volume of given shape ($V(\vec{\mathbf{r}})$, see Equation (3)). The volume can be taken of spherical shape (see e. g. Abbas and Sheth, 2006; Croton et al., 2005, this is especially done to compute environment in numerical simulations where galaxy positions are exactly known), but usually it is taken in the form of a parallelepiped or a cylinder, with a square or circular shape on the RA-dec plane of side l or radius R, and a much longer dimension in the redshift direction (h). This is done to take into account the scatter of galaxies in the radial direction due to uncertainties in the redshift measurement. In fact, the redshift is a measurement of galaxy velocity (used as a proxy for galaxy distance under the assumption that the velocity which originates the redshift is due to the Hubble flow). Therefore, additional velocity components (such as those due to the peculiar velocities of galaxies inside LSS systems) may affect the redshift measurement and, together with other sources of redshift uncertainty due to the survey observational properties, may scatter galaxies in the redshift direction, an effect known as Finger of God (FoG).

In the fixed aperture method, local environment is estimated through the formula

$$N(\vec{\mathbf{r}}) = \sum_{i=1}^{N_{tracers}} W_i(\vec{\mathbf{r}} - \vec{\mathbf{r}}_i) \tag{6}$$

where $W_i(\vec{\mathbf{r}} - \vec{\mathbf{r}}_i)$ is the so-called window or kernel function. In the case of the fixed aperture method $W_i = 1$ if the i-th tracer is inside

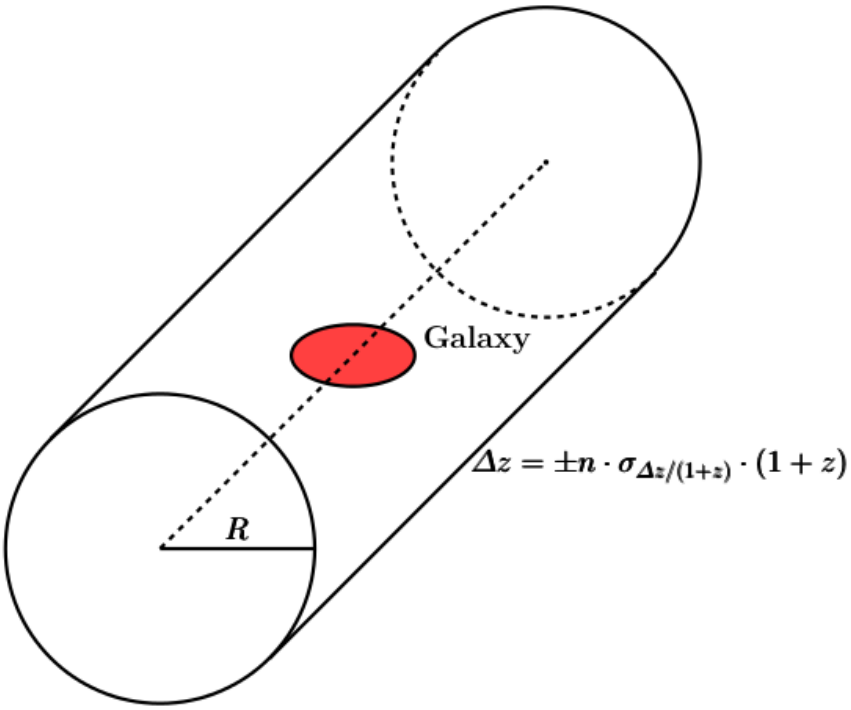


Figure 13: Schematic representation of the fixed aperture method with a cylindrical volume of radius R and height $h = \pm\Delta z$ (with Δz from Equation (8)).

the volume and zero otherwise. Common choices for fixed aperture radii (or side lengths l) are between $R = 1 - 10$ Mpc (as these are the typical sizes of galaxy groups and clusters), while typical values for the length in the redshift direction usually depend on the accuracy of the redshift measurement. In Chapter 3 it has been considered a height $h = \pm\Delta z$ proportional to both a velocity cut $dv = 1000$ km s^{-1}, through the formula

$$\Delta z = \frac{dv}{c}(1+z) \tag{7}$$

and to the uncertainty $\sigma_{\Delta z/(1+z)}$ on the redshift measurement (through a factor n), parametrized as

$$\Delta z = n \cdot \sigma_{\Delta z/(1+z)} \cdot (1+z) \tag{8}$$

A schematic representation of the fixed aperture method is shown in Figure 13. This figure shows a cylindrical aperture, of radius R and height $h = \pm\Delta z$ (with Δz from Equation (8)) centred on a target galaxy. Tracer galaxies that define the environmental density are counted inside the cylinder.

2.1.2 *Gaussian kernel*

To perform some smoothing of the density field, without using a sharp threshold to decide whether a galaxy has to be included or not in the density measurement, a different window function can be used to weight galaxy counts. Among the possible choices (see e. g. Kovač et al., 2010; Muldrew et al., 2012), in Chapter 5 it has been used a Gaussian kernel. This approach is widely used in the literature (see e. g. Cucciati et al., 2006) and it consists in choosing a window function in the shape of a Gaussian curve $W_i = G_i$ with $\sigma = R_G$ and defined by

$$G_i\left(\frac{|\Delta\vec{\mathbf{r}}|}{R_G}\right) = \frac{1}{(2\pi R_G^2)^{3/2}} \exp\left[-\frac{1}{2}\left(\frac{|\Delta\vec{\mathbf{r}}|}{R_G}\right)^2\right] \tag{9}$$

where $\Delta\vec{\mathbf{r}} = \vec{\mathbf{r}} - \vec{\mathbf{r}}_i$. Common choices for the value of the dispersion of the Gaussian kernel are $R_G = 1 - 10$ Mpc as in the fixed aperture case. Again if the i-th tracer is outside the volume, then the window function is taken to be zero.

2.1.3 *Nearest neighbours*

The Nearest Neighbour (NN) method (see e. g. Li et al., 2011; Muldrew et al., 2012) relies on the measurement of the distance to the n-th nearest tracer to a selected target position. Usually a number of neighbours between 3 and 10 are used, and the distance to the n-th NN can be directly taken as a proxy of the density field at the target position. The distance to the NNs can be measured in 3D (therefore defining a sphere of radius equal to the distance to the n-th nearest neighbour) or in 2D in projection on the plane of the sky. The distance to the projected NN can be used to define the radius R of a cylinder (a sort of adaptive aperture), whose height h is taken in the redshift direction as in the fixed aperture case (see e. g. Equations (7) and (8)), to minimise the FoG effect. In this case, the number counts inside the adaptive aperture are computed by counting tracer galaxies inside the volume, as in the fixed aperture case (see Equation 6).

A schematic representation of the NN method can be found in figure 14. In this Figure, the five NNs with distances measured in 3D to a target galaxy (called "seed" in the figure) are shown. A cylinder is constructed based on this neighbours, with the radius defined by the projected 2D distance to the 5-th nearest neighbour and the height determined by the maximum redshift difference between the neighbours.

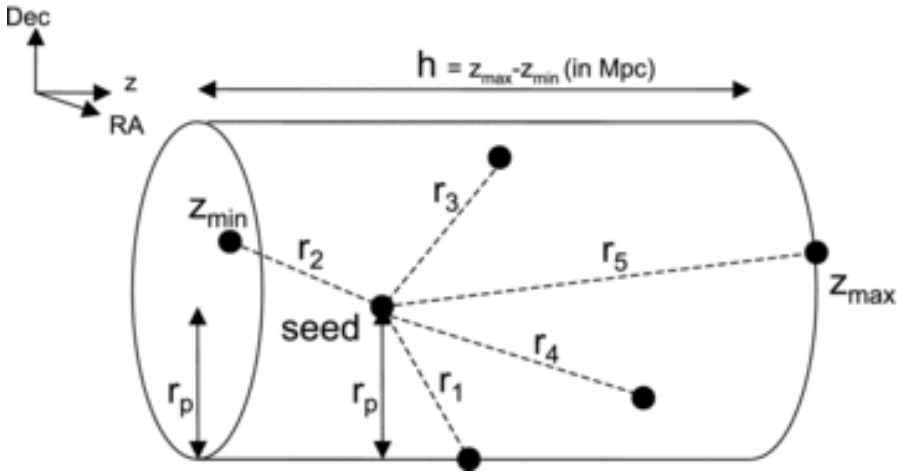


Figure 14: An schematic representation of an implementation of the NN method. The five NNs with distances measured in 3D to a target galaxy (seed) are marked as r_i, with $i = 1, 2, ..., 5$. A cylinder is constructed based on this neighbours, with radius r_p defined by the maximum projected distance of all the five NNs and height h determined by the maximum redshift difference of the same galaxies. Figure from Li et al. (2011).

2.1.4 *Tessellation methods*

The term "tessellation" refers to the partition of the space populated by a discrete set of points into a set of polygons (in 2D) or polyhedra (in 3D). The tessellation of a space populated by a discrete set of points can provide a measurement of the spatial density of the points themselves. The two most used methods are the Voronoi tessellation (Voronoi, 1908) and its mathematical dual, the Delaunay triangulation (Delaunay, 1934). They can be applied to a set of discrete points (e. g. target positions) and are widely used in astrophysics to study the local environment of galaxies (see e. g. Marinoni et al., 2002, for a first application of these methods). In this case, the target and the tracer samples coincide.

Given a set of points (e. g. a galaxy catalogue or a set of target positions), Voronoi tessellation partitions space in an ensemble of polyhedra, each of which contains only one point of the discrete set and has its faces equidistant from all the others. Delaunay triangulation, instead, builds all the possible tetrahedra which have points from the discrete set as vertices. The inverse of the volume of each Voronoi polyhedra (or the mean at each point position of the inverse of the volumes of the Delaunay tetrahedra that have the point as a vertex) provides an estimate of the local density field ($\rho(\vec{r})$). In this work, space tessellation built using Delaunay triangulation is used to detect CW features, as explained in Section 2.2.2[3].

3 The code used to implement Delaunay tessellation in the VIPERS survey volume is the DTFE (Schaap, 2007; Schaap and van de Weygaert, 2000).

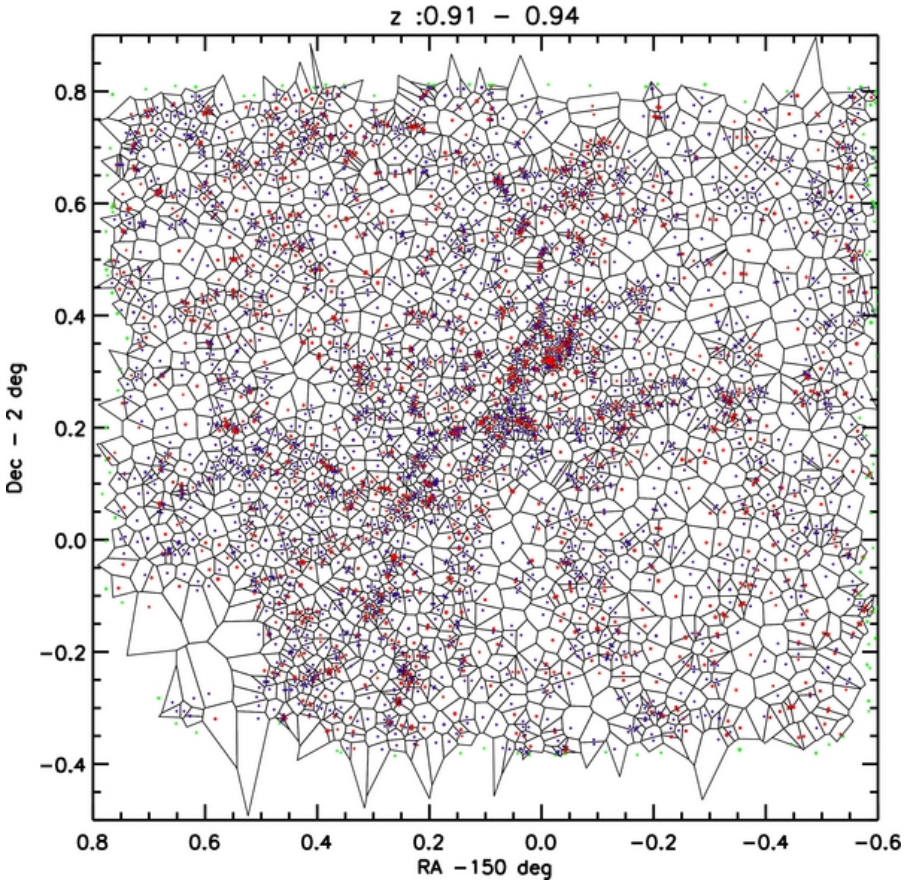


Figure 15: 2D Voronoi tessellation for galaxies in the UltraVISTA survey (red and blue points, early-type and late-type galaxies, respectively) performed in the redshift range $0.91 \leqslant z \leqslant 0.94$. Green points mark galaxies close to the survey edges for which Voronoi polygons are not closed. Figure from Scoville et al. (2013).

An example of 2D Voronoi tessellation of a sky field can be seen in Figure 15. In this Figure, Voronoi tessellation has been applied to the galaxies in the UltraVISTA survey (Ilbert et al., 2013; McCracken et al., 2012), to measure local environment. Red and blue points mark the positions of galaxies (early-type and late-type, respectively), while green points mark the position of galaxies for which the Voronoi polygons are not closed. An example of 2D Delaunay triangulation can instead be seen in Figure 16, applied to a sample of galaxies from the SDSS survey (York et al., 2000).

2.1.5 *Local environment with photometric redshifts*

As mentioned in Section 2.1, in order to measure local environment accurate galaxy positions are needed. While galaxy angular positions (RA and dec) are usually known with great accuracy, the redshift (z) precision strongly relies on the method used for its determination. In particular, spectroscopic redshifts can be very precise (of the order of

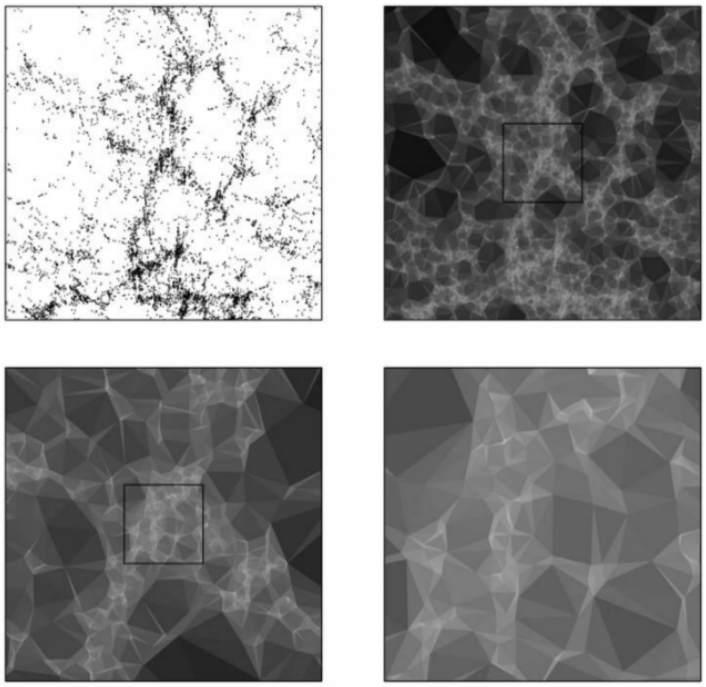


Figure 16: This figure shows a sample of galaxies from the SDSS (top left panel), as well as the density field measured through Delaunay triangulation (top right panel). Bottom panels show two zoomed-in regions of the same field. Figure from Aragón-Calvo (2007).

$\sigma_{\Delta z/(1+z)} \sim 10^{-3} - 10^{-4}$), but have the drawback of being difficult to obtain for large samples of galaxies due to the excessive telescope time that would be required. This limits wide-area spectroscopic redshift surveys to low redshift ($z \lesssim 1.5$, e. g. 2dFGRS, Colless et al. 2001, DEEP2, Newman et al. 2013, SDSS, York et al. 2000, GAMA, Driver et al. 2011, zCOSMOS-Bright, Lilly et al. 2009) and to incomplete sampling rates (e. g. as low as $\sim 40\%$, VIPERS, Guzzo et al. 2014). Spectroscopic redshift surveys are available at high redshift ($z > 1.5$), but only for small sky areas and small galaxy samples ($\lesssim 1\,\mathrm{deg}^2$, e. g. VVDS, Le Fèvre et al. 2005, VIMOS Ultra-Deep Survey (VUDS), Le Fèvre et al. 2015, zCOSMOS-Faint, Lilly et al. 2007). To overcome these limitations and obtain large samples of galaxies with measured redshifts, on wide sky areas, and out to high redshifts, we are often forced to rely on redshift measurements coming from photometry. Photometric redshifts can be easily obtained for large samples of galaxies, but at the price of a larger uncertainty (going from $\sigma_{\Delta z/(1+z)} \lesssim 0.01$ for high-precision photometric redshifts to $\sigma_{\Delta z/(1+z)} \sim 0.1 - 0.3$ for less precise ones). These error values are mainly due to the number and width of filters available to measure galaxy photometry. As photometric redshifts are often derived through the fitting of galaxy SED templates to photometric data points obtained with sets of broad- and narrow-band filters, their uncertainty depends on how well specific

features of a galaxy SED (located at precise restframe wavelengths, such as the Balmer and Lyman breaks) are sampled by the available filters given their shift due to the galaxy redshift. For this reason, photometric redshift uncertainty is a strong function of redshift and magnitude (see e. g. the discussion of Figure 9 in Ilbert et al., 2009).

As the redshift information is necessary for the study of the density field, it is important to understand how much the measurement of local environment is dependent on the redshift accuracy (for the effect of photometric redshifts on the galaxy spatial distribution see e. g. Figure 18). For the goal of this thesis this topic is of high importance, as most of the results presented here (see Chapters 4 and 6) rely on a measurement of the density field made using photometric redshifts. The dependence of the local environment measurement on redshift accuracy has been addressed in various ways in the literature, both using galaxy surveys and mock galaxy catalogues. These works, whose methods differ from one another, have been largely motivated by the need to study the impact of redshift precision on the measurement of local environment with a particular analysis in mind (e. g. recovering the colour-density relation, Etherington and Thomas 2015; Lai et al. 2016, simulate the possibility of studying local environment within the context of a particular survey, Cooper et al. 2005; Cucciati et al. 2016, identifying central and satellite galaxies in a DM halo, Fossati et al. 2015).

The picture that seems to emerge is that environment can still be studied with photometric redshifts provided that their uncertainty is not too large ($\sigma_{\Delta z/(1+z)} < 0.02$, see e. g. Cooper et al., 2005). In particular, environmental estimators can be tuned to find the best combination of parameters (like R and h for the fixed aperture method or the number of neighbours used in the NN method) that minimises the error on local density measurements when dealing with photometric redshifts (see e. g. Etherington and Thomas, 2015; Fossati et al., 2015; Lai et al., 2016). Even if photometric redshifts do not allow a detailed environmental analysis, high-density regions can still be securely distinguished from low-density ones. Moreover, several methods can be identified to increase the accuracy of the environmental measurement when dealing with photometric redshifts (e. g. mass ranking inside fixed apertures to identify central and satellite galaxies, see Fossati et al. 2015, or using spectroscopic information to increase the photometric redshift accuracy, like the ZADE method, see Cucciati et al. 2016).

In this thesis, it has been explored if it is possible to study the GSMF in different local environments if the density field is measured using a fixed aperture method and a photometric redshift survey. A detailed description of the results obtained and their application to the study of the GSMF in the UltraVISTA survey are found in Chapters 3 and 4, respectively.

2.2 GLOBAL ENVIRONMENT

The Universe is characterised at large scales by a complex network of different structures (galaxy clusters, filaments, walls and voids, often referred to as the Cosmic Web), which can be identified through a variety of methods, most of which have as a starting point a measurement of the local environment. In particular, galaxy groups and clusters have been extensively investigated, both with dedicated observations (see e. g. Boselli and Gavazzi, 2006, 2014, for review) and by applying structure-finding algorithms to galaxy surveys (see Castignani et al., 2014a; Cucciati et al., 2010, for a review of methods for finding galaxy clusters). Together with voids (see Colberg et al. 2008 for a review on void-finding algorithms, Beygu et al. 2016; van de Weygaert and Platen 2011, and references therein for results on the study of properties of galaxies in voids), galaxy clusters have been by far the most-studied LSS features and have been recently detected also at very high redshift (see e. g. Toshikawa et al., 2016).

In recent years, other features of the CW such as filaments and walls (already known to be present in the galaxy distribution, see e. g. Davis et al., 1982; Geller and Huchra, 1989; de Lapparent, Geller, and Huchra, 1986) have received renewed attention, and statistical algorithms to detect the filamentary network in galaxy redshift surveys have been developed. As for part of the work described in this thesis (see Chapter 5) the main focus has been on the detection of filaments and the study of galaxy properties in relation to their distance to LSS, in the following a brief and non exhaustive description of a few of the methods that have been proposed to detect the CW starting from a discrete catalogue of galaxy positions (Section 2.2.1) will be given. Many of these methods are not limited to the sole detection of filaments, but they are able to recover the complete variety of global environments, from galaxy clusters (hereafter often referred to also as "nodes"), to filaments, walls and voids. In Section 2.2.2 the focus will be more on the algorithm that has been used to perform the analysis described in Chapter 5.

2.2.1 *Exploring global environment*

In this section, a brief description of a few of the main methods that have been proposed to recover global environment starting from a galaxy redshift survey will be provided. For a more extensive review see e. g. Martínez and Saar (2002).

HESSIAN MATRIX: based on the theory by Zel'dovich (1970) and proposed by Hahn et al. (2007) (see also Eardley et al., 2015), this method uses the eigenvalues of the tidal tensor T_{ij} to classify a region of space (usually a cell of a few Mpc of side of a uniform grid) based on the number of its collapsing dimensions.

In particular, given a target position in space ($\vec{r}$), the tidal tensor T_{ij} is defined as the Hessian of the gravitational potential Φ

$$T_{ij} = \frac{\partial^2 \Phi}{\partial r_i \partial r_j} \tag{10}$$

Given a density contrast $\delta(\vec{r})$ (which can be computed from the starting galaxy catalogue, see Equation (4)), the gravitational potential can be obtained from the Poisson equation

$$\nabla^2 \Phi = 4\pi G \bar{\rho} \delta \tag{11}$$

where $\bar{\rho}$ is the mean density of the Universe and G is the gravitational constant. The number of eigenvalues of the tidal tensor above a given threshold (which can be defined according to the starting data set, see Eardley et al., 2015, and references therein) is related to the number of dimensions of the region of space considered which are collapsing under the effect of gravity. In this way, the region of space can be classified as a void (zero collapsing dimensions), a wall (1 collapsing dimension), a filament (2 collapsing dimensions), or a node (3 collapsing dimensions).

WATERSHED TRANSFORM: this method has been first proposed by Beucher and Lantuejoul (1979) and Beucher and Meyer (1993) and a thorough description can be found in Platen, van de Weygaert, and Jones (2007) (see also Sousbie, Colombi, and Pichon, 2009). The Watershed transform is able detect regions of space (composed e. g. by cells of a few Mpc of side distributed on a uniform grid) around local density minima and maxima of a given galaxy distribution. Starting from a measured density field $\rho(\vec{r})$ (see Equation (3)), local minima are found and a discrete set of N_l levels of increasing density ($\rho_{l,k}$ with $k = \{1, 2, ..., N_l\}$) is created, uniformly binning the density field from its absolute minimum to its absolute maximum. Each density level of the discrete set is then considered, monotonically increasing the density. Starting from a local minima (m) and given a density level ($\rho_{l,k}$), each spatial position $\vec{r}$ on which the density field has been computed which satisfies the condition $\rho(\vec{r}) < \rho_{l,k}$ is said to belong to the region of space relative to the minimum m. The procedure is then repeated for all the local minima. If a region of space belongs to two different minima it is said to belong to their boundary. The procedure continues until the absolute maximum is reached. This divides the space in a series of regions belonging to local minima (identified as voids), separated one from the other by boundary regions (interpreted as filaments and walls), while local maxima of the density field are interpreted as nodes.

MULTISCALE MORPHOLOGY FILTER: the Multiscale Morphology Filter (MMF) has been developed by Aragón-Calvo (2007) and Aragón-Calvo et al. (2007) and it is based on the Scale Space theory by Florack et al. (1992) and Lindeberg (1998). Depending on a set of conditions imposed during the algorithm application, the MMF is able to identify only one kind of LSS feature at a time. It has, thus, to be used recursively in order to completely recover all the possible features of the CW (e. g. on a first application regions belonging to nodes are identified and removed from the sample, the algorithm is then run again to identify filaments and remove them, then again to identify walls). Given the measurement of a density field $\rho(\vec{\mathbf{r}})$, the MMF method builds a series of N copies with decreasing resolution, e. g. smoothing the original density field with a Gaussian kernel with increasing R_G (see Equation 9). On each of the N density maps, structures are identified (e. g. by means of the Hessian matrix method) and each point $\vec{\mathbf{r}}$ in space is identified as either belonging to or not belonging to the LSS feature considered (possibly with some measurement of the reliability of the detection). All the N maps are then stacked and a final map is created in which space points are assigned to the considered LSS feature based on the persistence of the structure at different resolutions.

Although powerful, these methods present the drawback of requiring a smoothing of the local environment measurement before application. The density field (calculated e. g. through Delaunay tessellation) is thus often resampled on a uniform grid, with cells of a few Mpc side, which provides some sort of uniform "pixelization"of the volume of space analysed. Moreover, both the MMF and the Hessian matrix methods depend on the choice of some free parameters (like the threshold condition for the study of the eigenvalues). Nevertheless, they are able to assign a region of space to a given LSS feature, thus enabling a possible way for the characterisation of the physical dimensions of walls, filaments and voids. To perform the analysis described in Chapter 5 a method similar to the ones described above has been used, which relies on the study of the critical points of the density field. A description of the method will be given in the following section.

2.2.2 *The Discrete Persistent Structure Extractor (DisPerSE)*

The Discrete Persistent Structure Extractor (DisPerSE) code[4] implements an algorithm for the detection of CW features based on the discrete Morse theory (Forman, 1998, 2002) and the persistent homology theory (Edelsbrunner, Letscher, and Zomorodian, 2002; Robins, 2000).

4 http://www2.iap.fr/users/sousbie/web/html/indexd41d.html

The algorithm is fully described in Sousbie (2011), while Sousbie, Pichon, and Kawahara (2011) presents a first application to both data and numerical simulations. Although the code and its theoretical background are fairly complex, in this Section a brief description of its structure is provided. The DisPerSE algorithm relies on the measurement of the (discrete) gradient of the density field ($\vec{\nabla}\rho(\vec{\mathbf{r}})$) and on the detection of its critical points (maxima and minima), as well as the field lines connecting them. The structure of the algorithm can be summarised as follows:

1. *Density field.* In order to apply the DisPerSE algorithm, a measurement of the density field is necessary. To perform the analysis described in Chapter 5 $\rho(\vec{\mathbf{r}})$ has been measured using the Delaunay Tessellation Field Estimator (DTFE) (Schaap, 2007; Schaap and van de Weygaert, 2000). This provides a continuous measurement of the density field, with an accurate description of both the highest and the lowest density regions, thanks to the adaptive nature of the Delaunay tessellation.

2. *Gradient.* Given a density field $\rho(\vec{\mathbf{r}})$ as input, DisPerSE measures its gradient $\vec{\nabla}\rho(\vec{\mathbf{r}})$.

3. *Critical points and field lines.* The measured gradient of the density field is used to find critical points, which are identified as those points where $\vec{\nabla}\rho(\vec{\mathbf{r}}) = \vec{\mathbf{0}}$. Critical points are assigned an order k which identifies them as maxima, minima, and saddle points based on the number k of negative eigenvalues of the Hessian matrix

$$\mathcal{H}\rho(\vec{\mathbf{r}}) = \frac{\partial^2 \rho(\vec{\mathbf{r}})}{\partial r_i \partial r_j} \tag{12}$$

   Field lines ($L(t)$) connect critical points (which are said to be their origin or destination based on the sign of $\vec{\nabla}\rho(\vec{\mathbf{r}})$) and are characterised by the equation

$$\frac{dL(t)}{dt} = \vec{\nabla}\rho(\vec{\mathbf{r}}) \tag{13}$$

4. *Identification of n-manifolds.* All points of space that belong to a given field line that has its origin or destination in a critical point of order k are said to belong to a n-manifold relative to that critical point. Manifolds are classified according to their number n, which depends on the order of their corresponding critical point and on whether the analysis is performed in 1D, 2D, or 3D. In practical terms, supposing the analysis is performed in 3D, a 0-manifold identifies a point, a 1-manifold is a line, a 2-manifold is a surface, and a 3-manifold corresponds to a volume. These manifolds are put in relation to LSS features as described below.

5. *Persistence.* Due to the fact that LSS features are identified from discrete data sets, whose elements are associated to an uncertainty on their position, it is possible that some of the critical points identified are spurious detections. The theory of persistence is therefore applied, in order to eliminate CW elements that are purely due to noise. Critical points (maxima, minima and saddle points) are coupled in so-called "critical pairs" and the persistence provides a measurement of the relative density contrast between the elements of the pair (see Sousbie, 2011, for a detailed description of the theory application). The distribution of persistence values is compared to the same distribution calculated on a random density field. A threshold can therefore be user-defined (hereafter expressed in terms of the number of σ the distribution of persistence values for the data has to differ from the same distribution for the random field) in order to eliminate from the final compilation of LSS features all the spurious ones. Typical values for the persistence threshold are $P_{threshold} = 3\sigma$, but in Chapter 5 this value has been increased up to $P_{threshold} = 5 - 6\sigma$.

6. *Interpretation.* Once the map of critical points and n-manifolds has been created and filtered through persistence it is interpreted to produce a map of the CW from the given data set. Maxima of the density field correspond to nodes, density peaks in the galaxy distribution. Minima correspond to the centres of voids and saddle points (depending on their order k) are local density minima (low density regions) inside filaments or walls. Manifolds connecting maxima to saddle points (or saddle points between them) are interpreted as filaments (or walls). Therefore a map of the LSS is created, where voids (empty regions) centred on minima of the density field are surrounded by walls (sheets of matter) framed by filaments (ridges of the density field) which connect to maxima (peaks of the density field) where nodes (clusters) are located.

An example of filamentary structures reconstructed using DisPerSE is shown in Figure 17. This figure shows the spatial distribution of particles from a N-body DM simulation (top left panel) and the corresponding density field measured using Delaunay triangulation (middle top panel, top right panel is a zoom-in on a DM halo). Middle row shows filaments as reconstructed by DisPerSE adopting a 0σ (middle left panel), 2σ (central panel), and a 4σ (middle right panel) persistence threshold. Bottom row shows the filament reconstruction with the same persistence thresholds as in the panels of the middle row for the zoomed-in halo.

The advantage of the DisPerSE algorithm is that it is able to work directly on the density field measured through Delaunay tessellation (or any other method), without requiring any further smoothing or

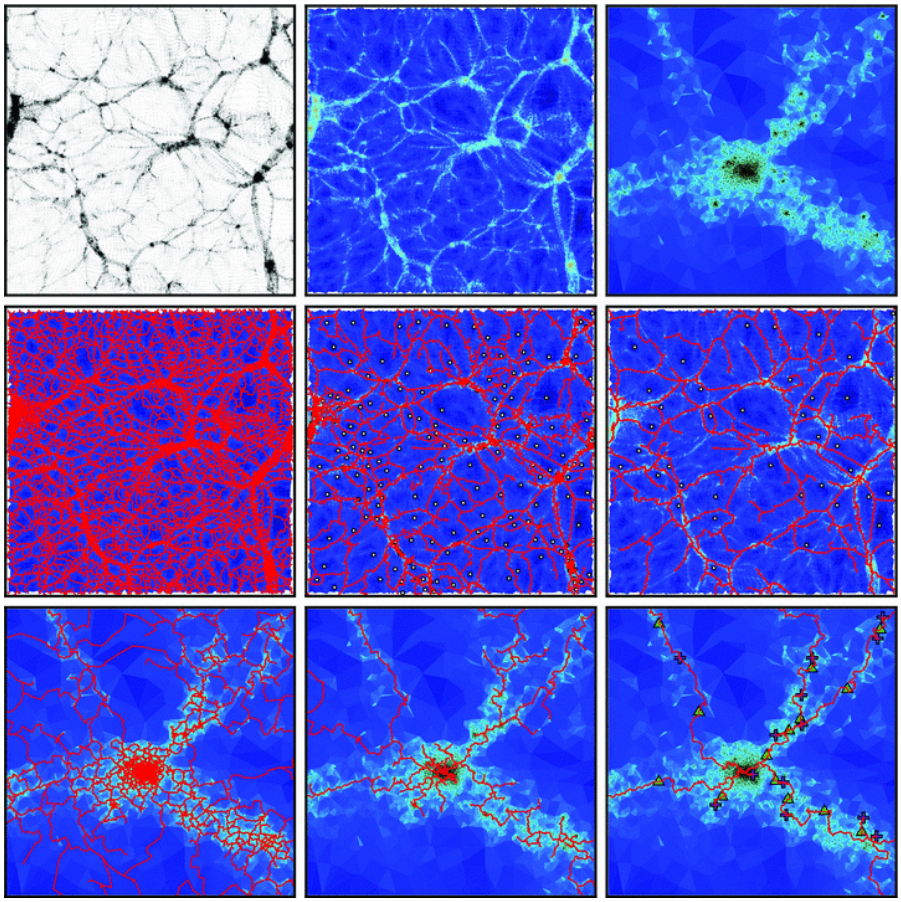


Figure 17: This figure shows the performance of the DisPerSE algorithm in reconstructing filaments starting from a N-body DM simulation. Top row shows the spatial distribution of particles from a N-body DM simulation (top left panel) and the corresponding density field measured using Delaunay triangulation (middle top panel, top right panel is a zoom-in on a DM halo). Middle row shows filaments as reconstructed from DisPerSE adopting a 0σ (middle left panel), 2σ (central panel), and a 4σ (middle right panel) persistence threshold. Bottom row shows the filament reconstruction with the same persistence thresholds as in the panels of the middle row for the zoomed-in halo. The white points, green triangles and purple crosses reported in some of the figures represent the minima, saddle points, and maxima of the density field, respectively. Figure from Sousbie (2011).

resampling. This is a powerful approach, which allows the code to accurately trace the position of LSS features having as input only a discrete set of galaxy positions. The drawback of this method is that it is not able to assign a physical dimension to the structures it recovers, them being just geometrical points (nodes), unidimensional lines (filaments) or 2D surfaces (walls). Nevertheless, the ability to work on unsmoothed density fields and the capability to filter spurious structures through persistence has determined the choice of this method to perform the analysis described in Chapter 5.

2.3 MASS FUNCTIONS

In this last section the algorithm that has been used to calculate GSMFs is described (see Chapter 3 and 4, see also Figure 6 for an example of GSMFs). The GSMF is a measurement of the mass distribution of the number of galaxies per unit volume. The methods commonly found in the literature to derive the GSMF can be separated in two classes, depending on whether they do not describe the GSMF with a specific model, but simply compute the number of galaxies per unit volume in a given mass interval (called "non-parametric" methods) or whether they assume a functional form for the GSMF in order to derive it (called "parametric" methods). The most common parametrization of the GSMF is the Schechter (1976) function (constituted by a power-law at low masses and an exponential cut-off at high masses)

$$\Phi(\mathcal{M})\mathrm{d}\mathcal{M} = \Phi^*\left(\frac{\mathcal{M}}{M^*}\right)^{\alpha} e^{-\frac{\mathcal{M}}{M^*}}\mathrm{d}\left(\frac{\mathcal{M}}{M^*}\right) \tag{14}$$

which can be theoretically derived from the Press and Schechter (1974) formalism. In Equation (14), Φ^* represents the normalisation of the function (which gives the number of objects per unit volume in the given mass interval considered), α is the (negative) slope of the power-law that characterises the function at low masses, and M^* is the mass at which the exponential cut-off starts to dominate on the power-law, which roughly represents the typical mass of the galaxies at the considered redshift. These parameters usually depend on galaxy type, environment, colour, SFR and other properties (see e. g. Bolzonella et al., 2010; Davidzon et al., 2013, 2016; Pozzetti et al., 2010, and references therein). For some galaxy populations (like red, quiescent galaxies), a modelling with two Schechter (1976) components is often adopted (see e. g. Baldry et al., 2012; Moustakas et al., 2013), in the form of

$$\Phi(\mathcal{M})\mathrm{d}\mathcal{M} = \Phi_1^*\left(\frac{\mathcal{M}}{M^*}\right)^{\alpha_1} e^{-\frac{\mathcal{M}}{M^*}}\mathrm{d}\left(\frac{\mathcal{M}}{M^*}\right) + \Phi_2^*\left(\frac{\mathcal{M}}{M^*}\right)^{\alpha_2} e^{-\frac{\mathcal{M}}{M^*}}\mathrm{d}\left(\frac{\mathcal{M}}{M^*}\right) \tag{15}$$

Many methods can be found in the literature to compute the GSMF (e. g. the parametric STY method, Sandage, Tammann, and Yahil 1979,

or the non-parametric C^- method, Lynden-Bell 1971, and Step-Wise Maximum-Likelihood (SWML) method, Efstathiou, Ellis, and Peterson 1988). To perform the analysis described in this work, the $1/V_{max}$ estimator by Avni and Bahcall (1980) (see also Bolzonella, Pelló, and Maccagni, 2002, and references therein) has been used. This estimator is non-parametric, meaning that it does not assume a predefined shape for the GSMF. The $1/V_{max}$ estimator basically consists in directly counting the number of galaxies per unit volume and in a given mass bin. Considering a mass bin B of extremes $B = [M_{bin}, M_{bin} + dM_{bin}]$, the mass function can be derived by taking the sum

$$\Phi(\mathcal{M})d\mathcal{M} = \sum_{\forall \mathcal{M} \in B} \frac{1}{V_{max}(\mathcal{M})} \tag{16}$$

where $V_{max}(\mathcal{M})$ is the maximum comoving volume in which the galaxy of mass $\mathcal{M}$ and apparent magnitude m can be observed, given the apparent magnitude limits of the survey. $V_{max}(\mathcal{M})$ is calculated by computing the integral

$$V_{max} = \int_{max(z1, z_{min})}^{min(z2, z_{max})} \frac{dV}{dz} dz \tag{17}$$

where $z1$ and $z2$ are the lower and upper extremes of the redshift bin considered when computing the GSMF. The values z_{min} and z_{max} are the redshift that the galaxy would have if it had the same absolute magnitude, but apparent magnitude corresponding to the faint and bright apparent magnitude limits (m_{max} and m_{min}) of the survey, respectively. These redshift values are computed solving the system

$$\begin{cases} M = m_{max} - 5 \log d_L(z_{max}) - 25 - k(z_{max}) \\ M = m_{min} - 5 \log d_L(z_{min}) - 25 - k(z_{min}) \end{cases} \tag{18}$$

where M is the absolute magnitude of the galaxy, d_L is the luminosity distance and k is the k-correction. The quantities z_{min} and z_{max} are usually computed through SED fitting (note that the galaxy k-correction depends on the adopted templates). The $1/V_{max}$ method is subject to Poissonian errors which can be easily computed through the sum

$$\sigma^2_{\Phi(\mathcal{M})} = \sum_{\forall \mathcal{M} \in B} \frac{1}{V^2_{max}(\mathcal{M})} \tag{19}$$

Among other sources of uncertainty in the GSMF calculation, cosmic variance and mass completeness are strongly connected to the survey observational properties. In particular, cosmic variance is related to the differences in the galaxy distributions between different sky fields and to the presence of LSS inside a survey volume. Cosmic variance will mostly affect surveys characterised by a small area and there

exists various methods to quantify its impact on the measurement of the GSMF (see e. g. Trenti and Stiavelli, 2008, and references therein).

The mass completeness limit of a survey is the characteristic mass at which galaxies start to become too faint to be detected, given the survey limiting magnitude. Measuring the GSMF for masses below the mass completeness limit results in a systematic underestimation of the galaxy number density and a drop in the GSMF. A common approach for computing mass completeness limits is the method described in Pozzetti et al. (2010). In particular, the limiting stellar mass ($\mathcal{M}_{lim}$, i. e. the mass a galaxy would have at its redshift if it had an apparent magnitude equal to the limiting magnitude of the survey m_{max}) is computed for each galaxy through the formula

$$\log(\mathcal{M}_{lim}) = \log(\mathcal{M}) + 0.4(m - m_{max}) \tag{20}$$

The result is a distribution of limiting stellar masses that reflects the distribution of stellar mass-to-light ratios at each redshift. $\mathcal{M}_{lim}$ is calculated for the 20% faintest galaxies at each redshift and the mass completeness limit is defined as the upper envelope below which lie 95% of the $\mathcal{M}_{lim}$ values in each redshift bin. The $1/V_{max}$ method to calculate the GSMF, together with mass completeness limits derived as in Pozzetti et al. (2010), has been successfully applied for the analysis presented in Chapters 3, 4, and 6[5].

2.4 SUMMARY

This Chapter is aimed at giving an overview of the possible methods that have been developed to solve the problem of measuring local and global environment as well as to giving an overview of the estimator used to measure the GSMF, starting from a large astrophysical data set like the ones typically output from wide-area, deep redshift surveys. The measurement of galaxy environment is important to constrain the processes that shape galaxy formation and evolution and it is a demanding problem which requires an accurate investigation and whose solution is often not univocal. The dependence of the environmental measurement on the chosen method and on the observational properties of the data set that is used have to be carefully understood to provide an unbiased view of the spatial distribution of galaxies in the Universe. The GSMF is a powerful statistical tool which allows us to summarise the many dependencies of the galaxy number density on galaxy physical properties such as stellar mass, redshift, galaxy type, and environment in a single distribution function.

In this Chapter several methods to measure a galaxy local and global environment have been reviewed and the main algorithms that have been used to perform the local environment measurements of

5 To apply the $1/V_{max}$ method to the observational data sets I wrote and optimised a *Python* routine.

Chapters 3, 4, and 6 (fixed aperture) and the global environment measurement of Chapter 5 (Delaunay tessellation and DisPerSE) have been introduced. Moreover the algorithm that has been used in Chapters 3, 4, and 6 to measure GSMFs has been described.

Part II

NEW FINDINGS ON THE RELATION BETWEEN LOCAL AND GLOBAL ENVIRONMENT AND GALAXY EVOLUTION

3

MEASURING THE GALAXY STELLAR MASS FUNCTION IN DIFFERENT ENVIRONMENTS WITH PHOTOMETRIC REDSHIFTS SURVEYS

Each survey is conceived with an observational strategy, which includes a given layout, depth, and accuracy on the redshift measurement. Understanding the limitations that these characteristics impose on the measurement of local environment is necessary to quantify the accuracy of the density field estimate and to understand if this accuracy is sufficient for a specific scientific goal.

As said in Section 2.1.5, among the various observational properties of a galaxy sample the uncertainty on the redshift plays a major role in determining the accuracy with which local environment is measured. With the advent of wide-area, deep photometric redshift surveys it has become possible to study galaxy environment at higher redshifts, but the effect introduced in the environment measurement by the increased redshift uncertainty of photometric redshifts (together with the dependence of their uncertainty on redshift and magnitude) needs to be studied. Moreover, it is useful to tune the analysis with respect to a specific task, as the local environment measured with a sample of photometric redshifts with a given accuracy may be sufficient to study a certain problem (e. g. recovering the colour-density relation) and not others (e. g. studying the GSMF in different environments).

The main effect of photometric redshift uncertainty on the galaxy distribution is to scatter galaxies from their original position by an amount which can be very large. For reference, Table 1 reports the conversion in Mpc of typical values of photometric redshift errors. These values show an example of the kind of distances at which galaxies can be scattered in the radial direction at different redshifts due to the uncertainty of the photometric redshift measurement. For example, already at $z = 0.5$ a photometric redshift uncertainty of $\sigma_{\Delta z/(1+z)} = 0.01$ scatters galaxies in the radial direction by more than 100 Mpc and the scatter reaches almost 200 Mpc at $z = 2.5$. In the case of photometric redshifts with an uncertainty of $\sigma_{\Delta z/(1+z)} = 0.1$, the scatter is > 1 Gpc at all redshifts.

In this Chapter the effects of the accuracy of photometric redshifts on the measurement of local environment with a fixed aperture method and on the estimate of the GSMF in different environments are analysed in detail. In particular, it is known that the GSMF is different in high- and low-density environments at $z \lesssim 1$. The goal of this Chapter

Δz	$z = 0.5$	$z = 1$	$z = 1.5$	$z = 2$	$z = 2.5$
$dv = 1000$ km/s	68.02	109.00	149.72	190.35	231.11
$\sigma_{\Delta z/(1+z)} = 0.001$	13.60	16.35	17.97	19.03	19.81
$\sigma_{\Delta z/(1+z)} = 0.01$	136.05	163.49	179.66	190.35	198.09
$\sigma_{\Delta z/(1+z)} = 0.05$	680.05	817.35	898.25	951.73	990.45
$\sigma_{\Delta z/(1+z)} = 0.1$	1358.97	1634.16	1796.23	1903.33	1980.81

Table 1: Lengths in Mpc of common photometric redshift errors. For comparison, the length in Mpc corresponding to a $dv = 1000$ km/s (see Equation (7)) is also reported. Note, that the values of $\sigma_{\Delta z/(1+z)}$ are not multiplied by the $(1+z)$ term (see Equation (8)).

is to explore whether the local environment measured using photometric redshifts is sufficiently accurate to recover a difference in the GSMFs in high- and low-density regions also at $z > 1$.

To this aim, mock galaxy catalogues are used to simulate astrophysical surveys with photometric redshifts of various precisions and the uncertainties in their use to measure local environment and GSMFs are quantified. This work is fully described in the paper Malavasi et al. (2016a).

This Chapter is structured as follows: in Section 3.1 the data set used to perform this work is introduced and it is briefly summarised how environment has been measured. In Section 3.2 a first overview of the results is presented. In Section 3.3 the performance of the fixed aperture method in measuring local environment for simulated photometric redshifts with various uncertainties is shown. Section 3.4 is, instead, focused on the results obtained with photometric redshifts having the same accuracy of those used in Chapter 4. Results regarding GSMFs in different environments are presented in Section 3.5 and conclusions are summarised in Section 3.6. In particular, Sections 3.4 and 3.5 are important as they constitute an analysis directly aimed at studying the feasibility of the work described in Chapter 4, where the results obtained here with mock galaxy catalogues are directly applied to a real data set. The cosmology adopted for the analysis presented in this Chapter is a standard cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$.

3.1 MOCK DATA AND LOCAL ENVIRONMENT MEASUREMENT

In order to study the impact of photometric redshifts on the measurement of local environment, the first step is to choose a reference measurement of the density field. As stated in section 2.1.5, many works have explored different methods. A possible way of quantifying the impact of photometric redshifts on the measurement of local envi-

ronment would be to compare the measurement of the density field for a sample of galaxies with both measured spectroscopic redshifts and measured or simulated photometric redshifts, taking the environmental estimate in the case when spectroscopic redshifts are used as a reference. This has been done e. g. by Etherington and Thomas (2015) using SDSS galaxies. Here results obtained with a different approach are shown, i. e. relying on mock galaxy catalogues, similar to what done e. g. by Cucciati et al. (2016) and Lai et al. (2016). With this approach, it has been possible to extend the analysis without restrictions on the redshift range and the area explored, thus enabling an investigation of the problem up to $z \geqslant 2$ and on a sufficiently large area to have a significant statistical sample. Moreover, by comparing the density field and GSMFs when using each galaxy's true and photometric redshifts, it has been possible to derive results that do not completely depend on how well the mock galaxy catalogues reproduce the spatial distribution of real galaxies. In fact, although Cucciati et al. (2016) tested whether the mock catalogues adopted here represent a suitable reproduction of the observed spatial distribution of galaxies, the approach adopted in this Chapter is focused on determining whether a given reference sample of different environments created starting from the spatial distribution of mock galaxies is recovered if the redshift information is degraded. As such, the results described in this Chapter can also be applied to real data, as shown in Chapter 4.

3.1.1 *Mock data*

The data set used to perform the analysis presented in this chapter is composed of the mock galaxy catalogues of Merson et al. (2013)[1]. These catalogues cover an area of $100\,\mathrm{deg}^2$ in the redshift range $z \in [0, 3]$, and are limited to a maximum apparent magnitude of $H \leqslant 27$[2]. They were constructed using the *Millennium Run* DM simulation (Springel et al., 2005) and the `galform` semi-analytic model of galaxy formation (Bower et al., 2006; Cole et al., 2000; Lagos et al., 2012). A complete description of the mock lightcones properties and construction can be found in Merson et al. (2013). Table 2 shows a summary of the lightcone geometry and physical parameters.

For the purposes of this work, a square area of $8\,\mathrm{deg}^2$ has been extracted from the whole $100\,\mathrm{deg}^2$. The redshift range was kept unaltered, but a further cut in K-band magnitude to $K \leqslant 24$ has been introduced to match at best the UltraVISTA Survey (see McCracken et al. 2012 for the survey overview and data reduction process, and Ilbert et al. 2013 for the photometric redshift calculation, see also Chapter 4)

1 The mock catalogues are freely available for download at `http://astro.dur.ac.uk/~d40qra/lightcones/EUCLID/`

2 The covered area, redshift range, and magnitude limit are defined by the Euclid survey.

PARAMETER	VALUE
Ω_m	0.25
Ω_Λ	0.75
Ω_b	0.045
H_0 $(km \cdot s^{-1} \cdot Mpc^{-1})$	73
Redshift range	0.0-3.0
Sky coverage (deg^2)	100
Field centre (RA,dec) (deg)	(0.0,0.0)
Maximum H-band magnitude	27

Table 2: Lightcone geometry and cosmology.

and to be consistent with the magnitude limit expected for the Euclid Survey. The final sample is composed of 1 054 752 mock galaxies.

For each mock galaxy, several parameters were available, but in particular two redshift values have been used in the estimate of the density field. One is the true (cosmological) redshift of each galaxy (z_{true}) and the other is the same redshift to which the peculiar motion of each galaxy was added (z_{obs}). In order to create a photometric redshift measure for each source, values from a Gaussian distribution with dispersion $\sigma_{\Delta z/(1+z)} \times (1 + z_{obs})$ have been randomly extracted and added to each galaxy's z_{obs}. In the following, it will be referred to these constructed photometric redshifts as z_{phot}.

3.1.2 *The measurement of local environment*

To measure each mock galaxy's local environment, a fixed aperture method has been adopted (see Section 2.1.1), as with this approach it is possible to choose a scale for the environment parametrization that is independent of redshift. Volume densities have been measured, following Equation (3), where $V(\vec{r})$ is the volume of the cylindrical fixed aperture considered. In the following, it will be referred to the true environment (ρ_{true}) when the z_{true} of each galaxy is used for the environmental estimate and to the reconstructed environment (ρ_{rec}) in case the z_{phot} of each galaxy is used. The true environment is taken as the reference environment and the comparison with the reconstructed environment is used to estimate how much photometric redshifts impact the density field measurement. It is important to keep in mind that even the definition of the true environment as the reference one is not unambiguous, as different methods for measuring local environment explore different scales (see e. g. the distinction between fixed scale and adaptive methods in Chapter 2) and

different scales are characterised by different physical processes (see e. g. Figure 11).

The fixed aperture has been taken of cylindrical shape, with base radius R and height h, centred on each galaxy of the sample (targets), using all other galaxies as tracers. The volume $V(\vec{r})$ is therefore given by $V(\vec{r}) = \pi R^2 h$. The number counts $N(\vec{r})$ are calculated as in Equation (6). Several values of R and h were adopted, so as to have an exploration of the dependence of the environment reconstruction precision on these two parameters. In particular, R has been varied between 0.3 Mpc and 2 Mpc, comoving. The height of the cylinder (i. e. its length on the radial direction) has been chosen proportional to redshift in two different ways, according to whether the environmental estimate was performed using the z_{true} or z_{phot} of galaxies. In particular, h is defined as $h = \pm\Delta z$. When using z_{true}, a Δz corresponding to a $dv = 1000 km/s$, through Equation (7), has been adopted. This cylinder length was chosen to be consistent with other definitions of local environment from the literature, as this is the value that is generally adopted to estimate the density field when using spectroscopic redshifts. In the z_{phot} case, this small height for the cylinder is useless as the errors on the redshift are much larger than that. An environmental estimate with photometric redshifts and Δz from Equation (7) has been performed as well, but it is only used as a reference for what happens when the cylinder length is very small compared to the photometric redshift uncertainty. Therefore, in the z_{phot} case, Δz has been chosen proportional to the error on the photometric redshift as in Equation (8).

The parameter n was varied as $n = 0.5, 1.5, 3$ to have a total length of the cylinder (h) ranging from $\pm 0.5\sigma$ to $\pm 1.5\sigma$ and $\pm 3\sigma$ photometric redshift error. In order to simulate different photometric redshift surveys, the photometric redshift error has been progressively increased from $\sigma_{\Delta z/(1+z)} = 0.003$ to $\sigma_{\Delta z/(1+z)} = 0.06$. The exact values adopted can be found in Table 3.

The analysis has been performed in the redshift range $0 \leqslant z \leqslant 3$, considering redshift bins of size $dz = 0.25$. In every redshift bin, high-density and low-density environments have been defined with the percentiles of the density distributions. Galaxies residing in environments denser than the 75th percentile of the distribution of volume densities were defined as belonging to high-density environments, and, conversely, galaxies whose environment is less dense than the 25th percentile were defined as belonging to low-density environments. In the following, it will also be referred to these environments as D_{75} and D_{25} for high-density and low-density environments, respectively. As a test, the analysis has also been performed by choosing more extreme environments, using the 10% and 90% of the distribution. Nevertheless, the results found with quartiles are more stable and more significant because of the larger statistics of the samples of

PROPERTY	ρ_{true}	ρ_{rec}
Redshift Used	z_{true}	z_{phot}
R	$0.3, 0.6, 1, 2$ Mpc	$0.3, 0.6, 1, 2$ Mpc
h	$dv = 1000 km/s$	$n = 0.5, 1.5, 3$ & $dv = 1000 km/s$
$\sigma_{\Delta z/(1+z)}$	none	$0.003, 0.01, 0.03, 0.06$

Table 3: Environmental reconstruction parameters. Here, R and h refer to the radius and length of the volume. Parameters dv and n are introduced in Equations (7) and (8), respectively. When the parameter n is used, h is defined as $h = \pm n \cdot \sigma_{\Delta z/(1+z)} \cdot (1+z)$

galaxies constituting the various environments. Therefore, the discussion will be focused on high-density and low-density environments derived using the quartile distinction. At high redshift (generally at $z > 2.5$) it may become difficult to define the 25th percentile of the volume density distribution, as the reduced size of statistical samples in the highest redshift bins leads the smallest volume density recovered to be shared by more than 25% of galaxies. For this reason, it has been decided to limit the analysis at redshifts lower than $z \leqslant 2.5$ in order to not introduce a bias in our results. A summary of the various environmental reconstructions can be found in Table 3.

3.2 LOCAL ENVIRONMENT WITH PHOTOMETRIC REDSHIFTS

The three panels of Figure 18 each show three slices of the analysed sky field, with RA on the abscissas and redshift on the ordinate. In each panel, from left to right, the redshift on the vertical axis changes from z_{true} to z_{obs} and z_{phot}. The three panels correspond to three large redshift ranges, namely $z \in [0, 1]$, $z \in [1, 2]$, and $z \in [2, 3]$.

It can be easily seen how the overdense regions that are so sharply defined in the z_{true} case become more elongated when z_{obs} is used. This behaviour is the already mentioned FoG effect, due to the introduction of peculiar velocities in the redshift measurement. High-density regions appear elongated in the redshift direction and (in particularly dense areas) structures that were distinct in the z_{true} case begin to overlap. When relying on z_{phot}, the galaxy spatial distribution becomes definitively blurred, as all the fine details that were visible in the left and central panels get cancelled out. Nevertheless, it can be seen how high-density regions are still recognisable as distinct from the mean density field and low-density areas, also in the case of z_{phot} with $\sigma_{\Delta z/(1+z)} = 0.01$.

Figure 19 shows an example of the comparison between ρ_{true} and ρ_{rec} (obtained with with $n = 1.5$ and $\sigma_{\Delta z/(1+z)} = 0.01$) both estimated on a scale of $R = 1$ Mpc. A correlation between the two

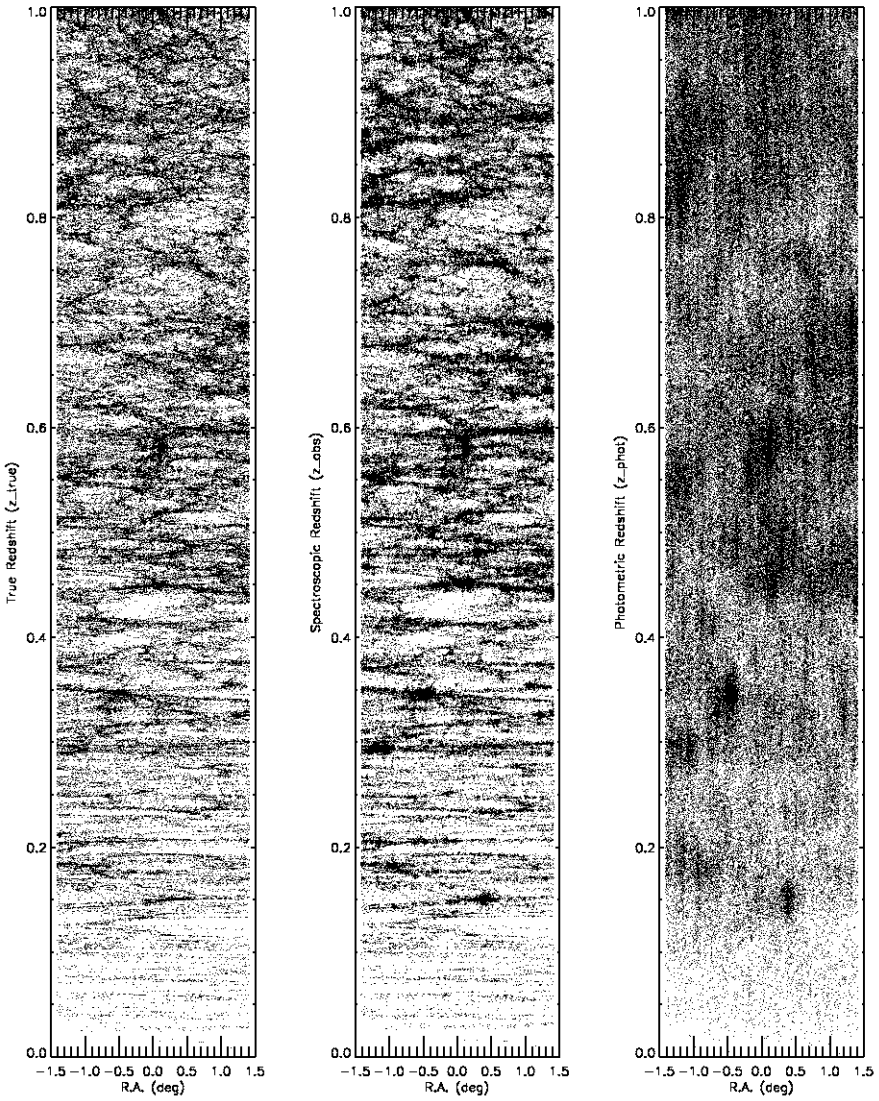


Figure 18: *Lightcones*. In each panel, left plot shows how galaxies are distributed when using z_{true}, middle plot when using z_{obs} and right plot when using z_{phot} with $\sigma_{\Delta z/(1+z)} = 0.01$. For the sake of clarity, the lightcones have been limited in dec to the central 1.2 deg ($-0.6 < \mathrm{dec} < 0.6$). $z \in [0, 1]$.

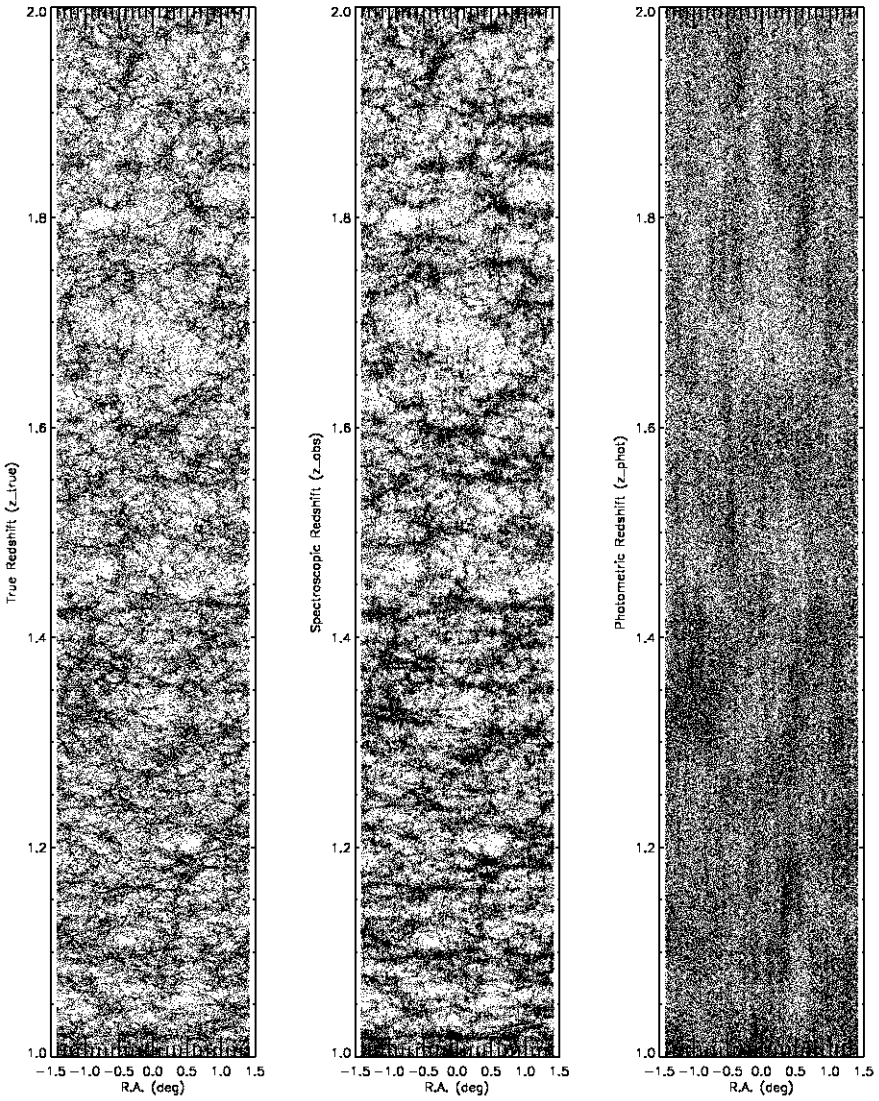


Figure 18: continued, $z \in [1, 2]$.

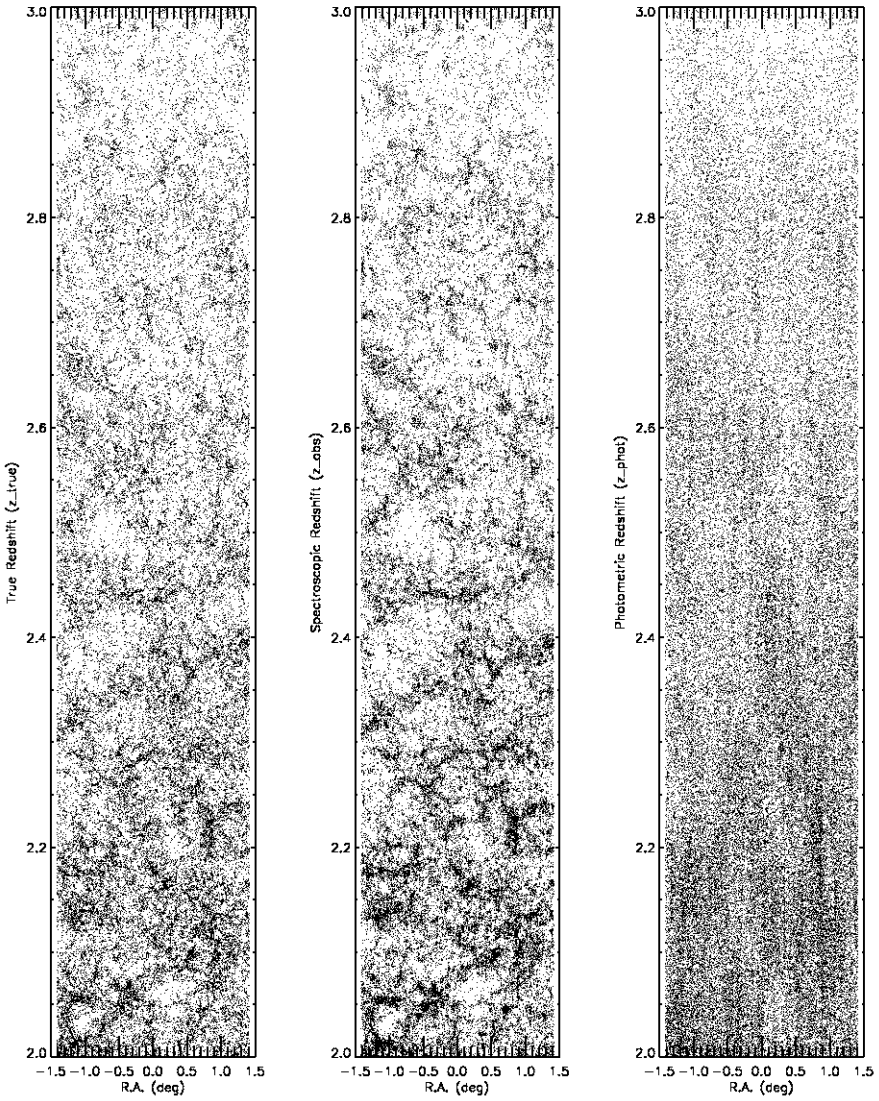


Figure 18: continued, $z \in [2, 3]$.

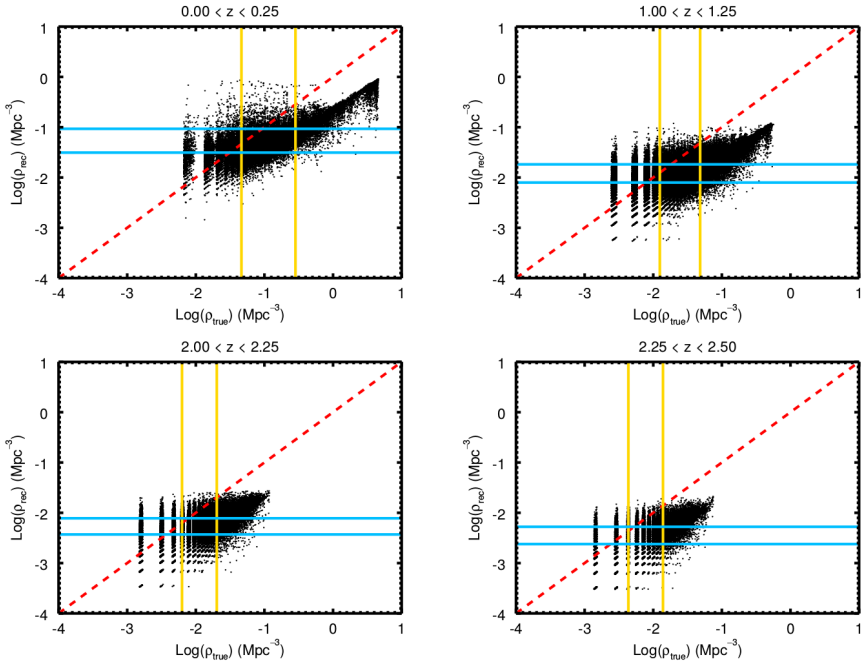


Figure 19: *Environment comparison*. The figure shows the comparison between ρ_{true} and ρ_{rec} (black points) in four representative redshift bins, labelled on top of each plot. Cyan horizontal lines represent the quartiles of the ρ_{rec} distribution, while yellow vertical lines are the quartiles of the ρ_{true} distribution. The $1:1$ relation is reported as a red dashed line for reference. Parameters used in the environmental estimate are $R = 1$ Mpc, $n = 1.5$ and $\sigma_{\Delta z/(1+z)} = 0.01$.

environments is present, although the scatter is large and the points result tilted and displaced from the $1:1$ relation. In fact, in each panel (which correspond to four representative redshift bins), ρ_{true} spans a wider range of volume densities (going from $0.005-5\,\mathrm{Mpc}^{-3}$ at $z\sim0$ to $0.001-0.1\,\mathrm{Mpc}^{-3}$ at $z\sim2$) compared to ρ_{rec} (which ranges from $0.005-1\,\mathrm{Mpc}^{-3}$ at $z\sim0$ to $0.002-0.02\,\mathrm{Mpc}^{-3}$ at $z\sim2$).

At very low densities the distribution of the points begins to show discrete bands, with points clustering at precise density values. This is due to the process of environmental estimate: as the density field is computed from galaxy number counts $N(\vec{r})$, volume densities can only assume discrete values. At high densities discretization effects are less visible and the volume density distribution becomes more continuous. At very low densities, the effect of discretization is more visible as there is only a small and finite amount of galaxies inside each volume. This results in a loss of continuity in the density values in low-density environments, which become progressively less discrete going towards high-densities. As the size of statistical samples is reduced in high-redshift bins, the volume densities recovered will

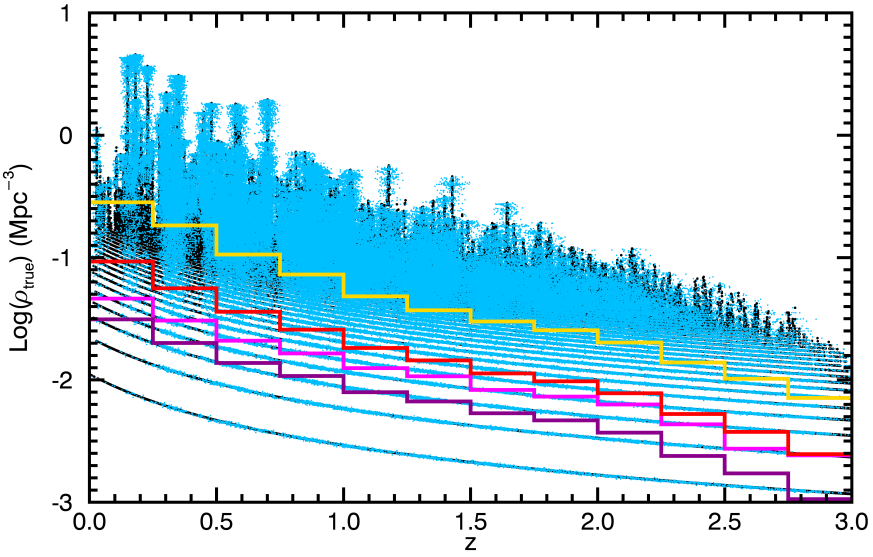


Figure 20: *Density-redshift relation*. This figure shows ρ_{true} as a function of redshift (black points). Cyan points highlight galaxies in high-density regions according to ρ_{rec}. The parameters used for the environmental reconstruction are $R = 1$ Mpc, $n = 1.5$ and $\sigma_{\Delta z/(1+z)} = 0.01$. The yellow and magenta lines refer to true environment D_{75} and D_{25}, respectively, while the red and purple lines refer to the reconstructed D_{75} and D_{25}, respectively.

be lower, resulting in correspondingly lower values for the quartiles of the volume density distributions. As the discretization effect is due to the lower amount of galaxies counted inside each cylinder it will be visible at higher density values moving at high redshifts.

The effects on the environment parametrization of passing from the true to the reconstructed estimate, are shown in Figure 20. This figure shows ρ_{true} as a function of redshift. Objects that are placed in high-density reconstructed environments, for the parameter combination $n = 1.5$, $R = 1$ Mpc and $\sigma_{\Delta z/(1+z)} = 0.01$ are highlighted. High-density reconstructed environments are contaminated by many galaxies coming from low-density true environments, and true D_{75} environments are not fully recovered. This figure also shows the effects of discretization at low densities as a lack of continuous density values, which progressively disappears moving at high-density values.

3.2.1 *Recovery and contamination fractions*

Several parameters play a role in determining the accuracy of the density field reconstruction. An exploration of the parameter space is needed to study the degeneracies between R and Δz of the cylin-

ders and the effect induced by choosing redshifts with worsening precision. In particular, in the following it will be studied how these parameters affect two distinct quantities, namely the "recovery" and the "contamination".

Recovery (f_{Rec}) is the fraction of galaxies that are correctly placed in either high-density or low-density regions according both to ρ_{true} and ρ_{rec}. In particular, N^{True}_{High} and N^{Rec}_{High} are the number of galaxies in high-density environments according to ρ_{true} and ρ_{rec}, respectively (and, correspondingly, N^{Rec}_{Low} and N^{True}_{Low} for low-density environments). So, if N_{HH} is the number of galaxies that are placed in a high-density environment according both to ρ_{true} and ρ_{rec}, and, correspondingly, N_{LL} for low-density environments, then the recovery fraction is defined as

$$f_{Rec} = \begin{cases} \frac{N_{HH}}{N^{True}_{High}} & \text{for } D_{75} \\ \frac{N_{LL}}{N^{True}_{Low}} & \text{for } D_{25} \end{cases} \tag{21}$$

Therefore, a f_{Rec} of 1 means that all the galaxies that are in the high-density (or low-density) reconstructed environments are placed in the correct true density quartile.

It is then useful to calculate the contamination fraction, f_{Con}. This quantity is the fraction of galaxies that are placed in a density quartile according to ρ_{rec}, which actually come from the opposite quartile according to ρ_{true}. If N_{HL} is the number of galaxies that are placed in a high-density environment when relying on ρ_{rec}, but that actually come from a low-density environment when relying on ρ_{true} (and, conversely, N_{LH} is the number of galaxies that are placed in a low-density environment when relying on ρ_{rec}, but that actually come from a high-density environment when relying on ρ_{true}), then the contamination fraction (f_{Con}) is expressed as

$$f_{Con} = \begin{cases} \frac{N_{HL}}{N^{Rec}_{High}} & \text{for } D_{75} \\ \frac{N_{LH}}{N^{Rec}_{Low}} & \text{for } D_{25} \end{cases} \tag{22}$$

Both f_{Rec} and f_{Con} vary with redshift as the reconstruction of the environment is more difficult for high-redshift galaxies because of the lower accuracy of the photometric redshift estimate. A good way to visualise f_{Rec} and f_{Con} is shown in Fig. 21. Only four redshift bins are reported for reference.

In this figure it is shown how the high-density sample (objects above the 75th percentile) and low-density sample (objects below the 25th percentile), which are identified using ρ_{rec}, with $n = 1.5$, $R = 1$ Mpc and $\sigma_{\Delta z/(1+z)} = 0.01$, are distributed according to ρ_{true}. The distribution of ρ_{true} is also reported for reference, together with vertical dashed lines corresponding to 25% and 75% of the ρ_{true} distribution. At very high densities and very low densities, the D_{75} and

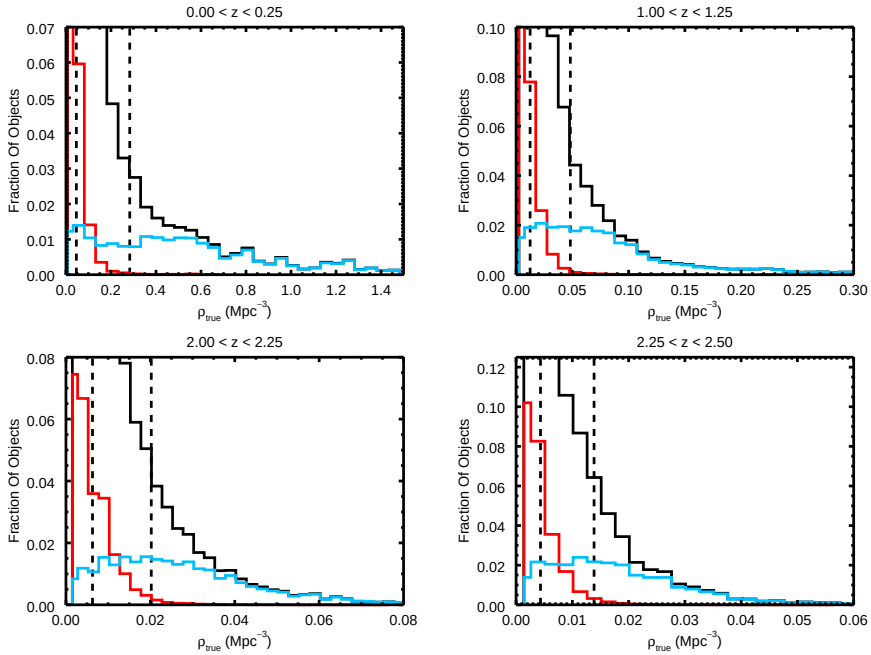


Figure 21: *Distributions comparison*. This figure shows the distribution in terms of ρ_{true} of the high-density and low-density samples, identified using ρ_{rec} in four representative redshift bins (labelled on top of each plot). In particular, the black histogram is the ρ_{true} distribution shown as a reference, the cyan histogram refers to D_{75} ρ_{rec}, while the red histogram refers to D_{25} ρ_{rec}. The environmental estimate has been performed with $n = 1.5$, $R = 1$ Mpc and $\sigma_{\Delta z/(1+z)} = 0.01$. The two vertical dashed lines are the values of environmental density corresponding to the 25th and 75th percentile of the ρ_{true} distribution.

D_{25} ρ_{rec} distributions closely follow the corresponding parts of the ρ_{true} distribution. The better a reconstructed distribution follows the true distribution, the higher its recovery fraction f_{Rec}. The ideal case of perfect reconstruction would imply that the reconstructed distribution of the high-density environments followed the ρ_{true} distribution down to the line of the 75th percentile and then dropped to zero, or that the low-density reconstructed distribution rose following the ρ_{true} distribution up to the 25th percentile line and then dropped to zero as well.

However, it can be seen how the distributions of the ρ_{rec} D_{25} and D_{75} environments have tails extending to ρ_{true} values of the opposite quartile. This means that a fraction of objects identified as high- or low-density objects according to ρ_{rec} actually comes from low- or high-density regions according to ρ_{true}. This fraction contributes to the contamination fraction of the sample (f_{Con}).

3.3 DEPENDENCE OF THE RESULTS ON THE PHOTOMETRIC REDSHIFT UNCERTAINTY $\sigma_{\Delta z/(1+z)}$

In the previous Section, a photometric redshift error of $\sigma_{\Delta z/(1+z)} = 0.01$ has been adopted as reference. In the overview of the photometric redshift surveys, this is a fairly optimistic value, as very few surveys can reach this kind of precision (e. g. the expected value of the photometric redshift precision for the Euclid survey is $\sigma_{\Delta z/(1+z)} = 0.03 - 0.05$). The choice for such a small value has been determined by the fact that this is the value of the uncertainty of photometric redshifts in the COSMOS-UltraVISTA survey sample (see Ilbert et al., 2013; McCracken et al., 2012) to which the results described here have been applied (see Chapter 4).

In this Section the dependence of f_{Rec} and f_{Con} on several values for the photometric redshift error, as well as on both R and h, is explored. In particular, it has been chosen to vary the error as $\sigma_{\Delta z/(1+z)} = 0.003, 0.01, 0.03, 0.06$. These values have been chosen as representative of various future and ongoing surveys, in particular, $\sigma_{\Delta z/(1+z)} = 0.003$ is the value expected for the Javalambre Physics of the accelerating universe Astronomical Survey (J-PAS)[3] (PI: Benítez, see Benítez et al., 2015), $\sigma_{\Delta z/(1+z)} = 0.03$ is the minimum error expected for the Euclid Survey[4] (PI: Mellier, see Laureijs et al., 2011), $\sigma_{\Delta z/(1+z)} = 0.06$ is the error on photometric redshifts derived for the sources in previous releases of the COSMOS[5] (PI: Scoville, see Capak et al., 2007; Ilbert et al., 2009; Scoville et al., 2007) and used for example in Malavasi et al. (2015) and in Chapter 6 of this thesis.

3 http://www.j-pas.org/
4 http://www.euclid-ec.org/
5 http://www.cosmos.astro.caltech.edu/

For clarity, in the following it will be distinguished between R_T, when it is referred to the radius used to estimate ρ_{true}, and R_R, when it is referred to the radius used to estimate ρ_{rec}.

A first, expected result is that the environmental reconstruction is more difficult when $\sigma_{\Delta z/(1+z)}$ is large. Figure 22 shows how f_{Rec} decreases and f_{Con} increases with increasing $\sigma_{\Delta z/(1+z)}$ both in the high-density and low-density situations. In particular, for high-density environments f_{Rec} is always above 75% (and close to 90% at $z \sim 0$) in the high-accuracy $\sigma_{\Delta z/(1+z)} = 0.003$ case, and progressively decreases to values slightly above 40% at all redshifts for $\sigma_{\Delta z/(1+z)} = 0.06$. Contamination fractions, instead, range from below 5% ($\sigma_{\Delta z/(1+z)} = 0.003, 0.01$) to between 10% and 20% ($\sigma_{\Delta z/(1+z)} = 0.03, 0.06$). For low-density environments, f_{Rec} is lower, ranging from around 75% at all redshifts for $\sigma_{\Delta z/(1+z)} = 0.003$ and progressively decreasing to slightly above 35% with increasing $\sigma_{\Delta z/(1+z)}$. Contamination fractions are also lower, but show the same trend in $\sigma_{\Delta z/(1+z)}$ as for high-density environments, which are below 5% for $\sigma_{\Delta z/(1+z)} = 0.003, 0.01$ and progressively reaching 10-15% for $\sigma_{\Delta z/(1+z)} = 0.03, 0.06$.

3.3.1 *The impact of* n *in the case of varying* $\sigma_{\Delta z/(1+z)}$

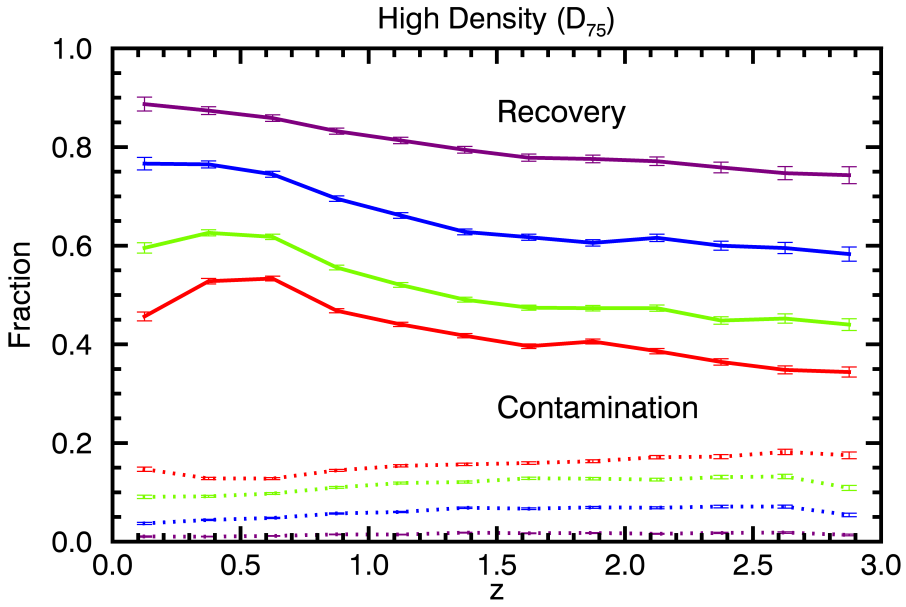
It has also been investigated how f_{Rec} and f_{Con} vary with R and n for each value of $\sigma_{\Delta z/(1+z)}$ discussed above. It has been found that R and n have a great impact on the measurement of the density field also in the case of large $\sigma_{\Delta z/(1+z)}$ values.

Figure 23 shows f_{Rec} and f_{Con} for three redshift bins ($1.50 \leqslant z \leqslant 1.75$, $2.00 \leqslant z \leqslant 2.25$ and $2.50 \leqslant z \leqslant 2.75$) as a function of n for various values of $\sigma_{\Delta z/(1+z)}$. In the high-density case, f_{Rec} is generally higher for $n = 1.5$, a trend shared by all redshift bins. Only for really large values of the photometric redshift uncertainty ($\sigma_{\Delta z/(1+z)} = 0.06$) is $n = 0.5$ a viable solution too, but the difference in f_{Rec} between this value and the one obtained with $n = 1.5$ is negligible. Instead, f_{Con} is always lower in the case of $n = 1.5$ independent of redshift and photometric redshift uncertainty.

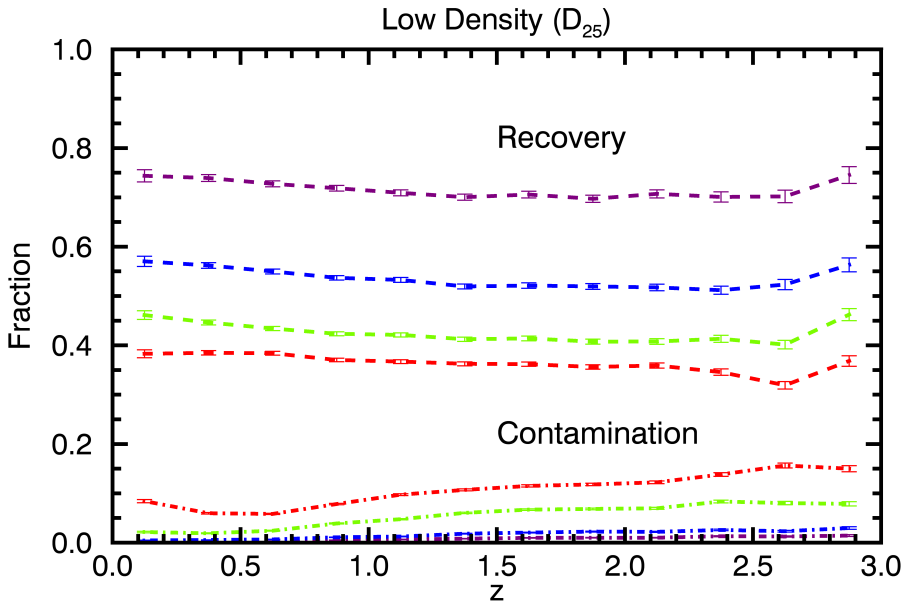
In the low-density case, shown in the bottom panel of the same figure, f_{Rec} is always higher and f_{Con} is lower when a value of $n = 1.5$ is adopted. This demonstrates that environmental reconstruction is better performed when it is chosen a value of the cylinder length of a similar size as the $\pm 1.5\sigma$ error on the photometric redshifts.

3.3.2 *The effect of the fixed aperture radius* R

Figures 24 and 25 show f_{Rec} and f_{Con} for one redshift bin ($1.50 \leqslant z \leqslant 1.75$) as a function of the ratio R_R/R_T, for various values of $\sigma_{\Delta z/(1+z)}$. All values of f_{Rec} and f_{Con} have been normalised to their value at $R_R/R_T = 1$, separately for every case of $\sigma_{\Delta z/(1+z)}$ considered, to bet-

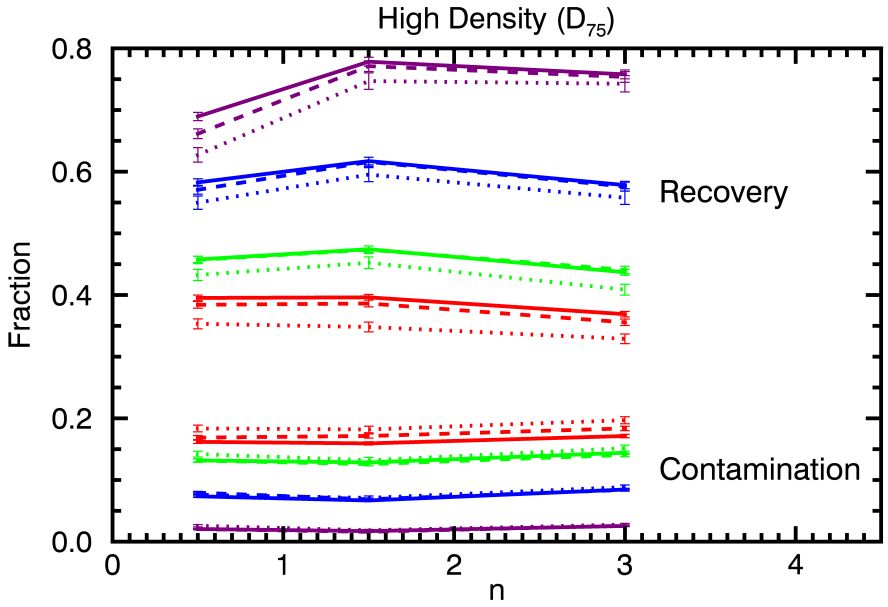


(a)

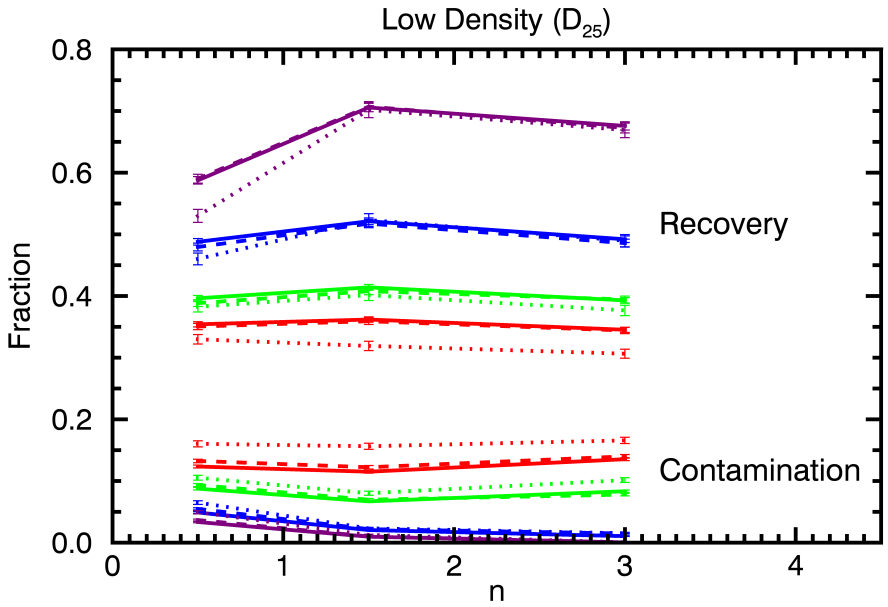


(b)

Figure 22: *Recovery and contamination fractions - varying* $\sigma_{\Delta z/(1+z)}$. This figure shows f_{Rec} (solid and dashed lines) and f_{Con} (dotted and dot-dashed lines) as a function of redshift. (a): high-density environments. (b): low-density environments. The various curves are colour-coded according to the various values of $\sigma_{\Delta z/(1+z)}$ (purple: $\sigma_{\Delta z/(1+z)} = 0.003$; blue: $\sigma_{\Delta z/(1+z)} = 0.01$; green: $\sigma_{\Delta z/(1+z)} = 0.03$; and red: $\sigma_{\Delta z/(1+z)} = 0.06$). The volume parameters for environmental reconstruction have been kept fixed to $R_R = R_T = 1$ Mpc and $n = 1.5$.



(a)



(b)

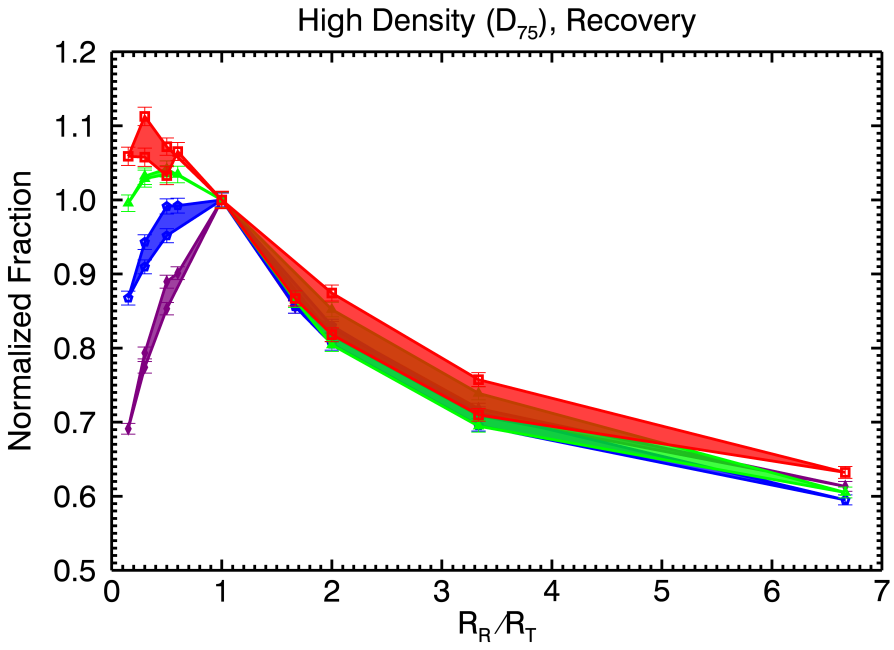
Figure 23: *Varying* $\sigma_{\Delta z/(1+z)}$ - *effect of* n. This figure shows f_{Rec} and f_{Con} as a function of the n parameter in three redshift bins, namely $1.50 < z < 1.75$ (solid lines), $2.00 < z < 2.25$ (dashed lines), and $2.50 < z < 2.75$ (dotted lines). (a): high-density environments. (b): low-density environments. The various curves are colour-coded according to the various values of $\sigma_{\Delta z/(1+z)}$ (purple: $\sigma_{\Delta z/(1+z)} = 0.003$; blue: $\sigma_{\Delta z/(1+z)} = 0.01$; green: $\sigma_{\Delta z/(1+z)} = 0.03$; and red: $\sigma_{\Delta z/(1+z)} = 0.06$). The aperture radius has been kept fixed to $R_R = R_T = 1$ Mpc.

ter visualise the trend in the fractions with R_R/R_T. In particular, while exploring the dependence on R (see also the procedure followed in Section 3.4), at first it has been chosen to fix R_T and to test several values of R_R. Afterwards, the procedure has been repeated varying R_T, resulting in the discovery that smaller R_T are better reconstructed than larger ones. For this reason, it has been chosen to normalise the f_{Rec} and f_{Con} showed in Figures 24 and 25 to a common value, thus reducing the intrinsic dispersion in the data (because smaller scales are better reconstructed than large scales) for the sake of clarity. Residual scatter is shown as shaded regions for each curve. The points at $R_R/R_T < 1$ are given by apertures whose ratio is lower than 1, for example, $R_R = 1$ Mpc and $R_T = 2$ Mpc, which yield $R_R/R_T = 0.5$ and so on. Moreover a fixed value of R_R/R_T could be given by more than one combination of R_R and R_T, for example, $R_R/R_T = 0.5$ could be given by $R_R = 0.3$ Mpc and $R_T = 0.6$ Mpc or by $R_R = 1$ Mpc and $R_T = 2$ Mpc. For this reason at a given value of R_R/R_T more than one point may be visible. The only exception is the point at $R_R/R_T = 1$ where all the curves have been normalised to unity. Hence, the normalisation of each curve corresponding to each value of $\sigma_{\Delta z/(1+z)}$ is lost, but the shapes and the trends in R_R/R_T can be better studied.

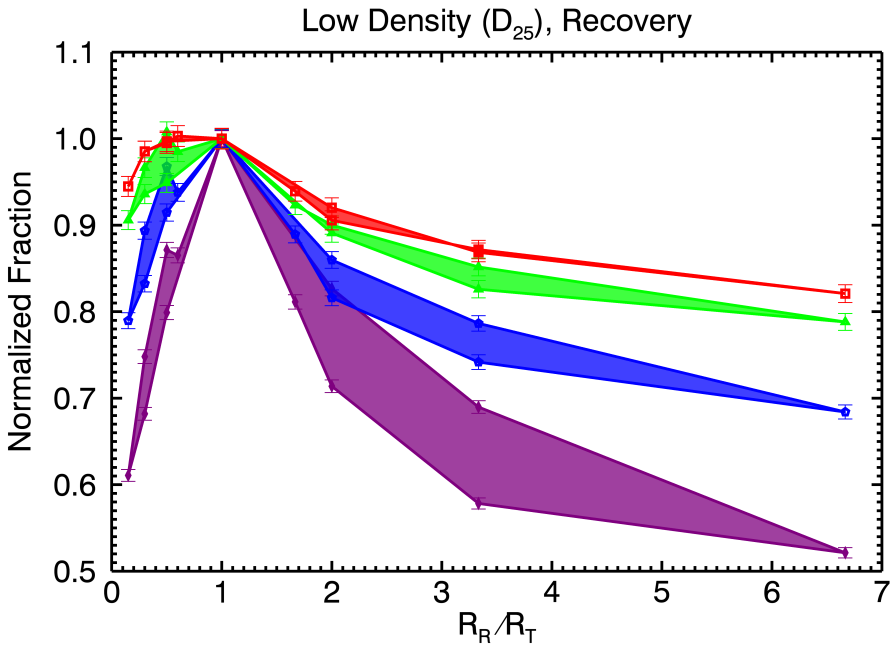
Recovery fractions (Figure 24) clearly show a steep decrease for $R_R/R_T > 1$ for every value of $\sigma_{\Delta z/(1+z)}$. This decrease is similar for all values of photometric redshift uncertainty in the D_{75} case and it is shallower for larger values of $\sigma_{\Delta z/(1+z)}$ in the D_{25} case. For values of $R_R/R_T < 1$, instead, an increase in the fraction values is present for values of $\sigma_{\Delta z/(1+z)} < 0.06$ in the D_{25} case, and for values of $\sigma_{\Delta z/(1+z)} \leqslant 0.01$ in the D_{75} case. For these values of $\sigma_{\Delta z/(1+z)}$, a maximum in the recovery fraction can be clearly individuated at $R_R/R_T = 1$, which translates in the best value for the fixed aperture radius to obtain an accurate environmental reconstruction. For larger $\sigma_{\Delta z/(1+z)}$ values, instead, smaller apertures than the physical scale that is to be investigated should be considered.

Contamination fractions (Figure 25) show a minimum in $R_R/R_T = 1$ for all values of $\sigma_{\Delta z/(1+z)}$ only in the D_{25} case, while the minimum is clearly recognisable in the D_{75} case only for values of $\sigma_{\Delta z/(1+z)} \leqslant 0.01$. For larger values of the photometric redshift uncertainty, f_{Con} in the high-density case are rather flat or slightly increasing. Thus, for small values of $\sigma_{\Delta z/(1+z)}$ the trend is confirmed of $R_R/R_T = 1$ as the best choice for the environmental reconstruction, while for larger values, an aperture radius smaller than the physical scale that is to be investigated is probably preferable.

The fact that the best environmental reconstruction is obtained for $R_R \simeq R_T$ is not an obvious result, as for large $\sigma_{\Delta z/(1+z)}$ fixed aperture radii smaller than the physical scale of the studied environment may be the best option to limit the number of contaminating interlopers. Here it is shown that this is indeed the case for $\sigma_{\Delta z/(1+z)} = 0.03, 0.06,$

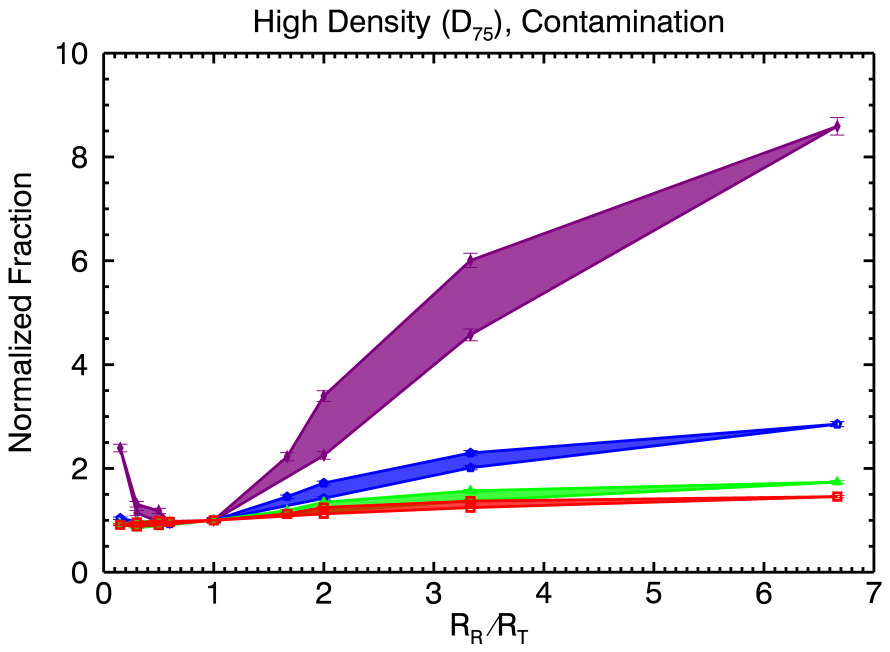


(a)

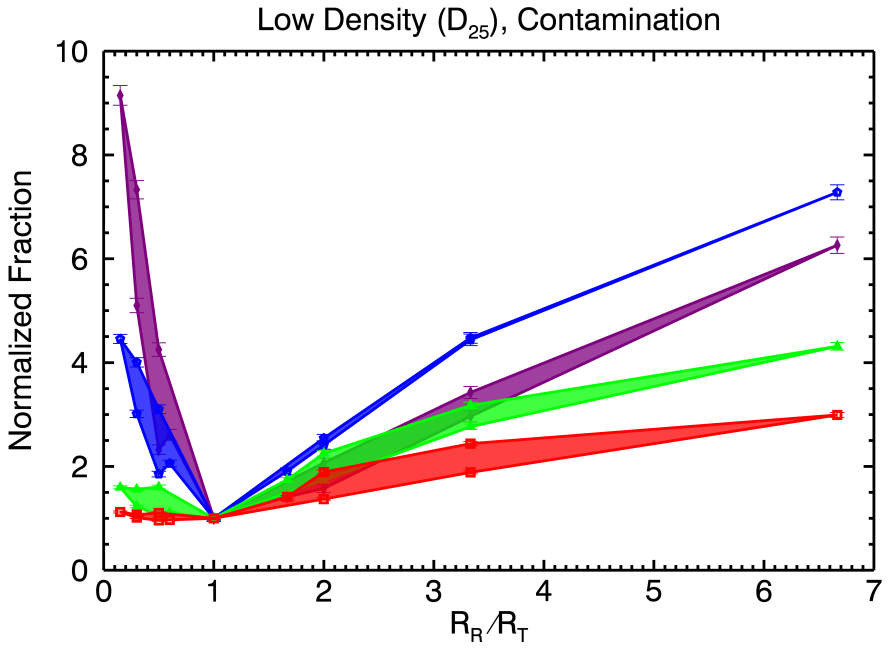


(b)

Figure 24: *Varying* $\sigma_{\Delta z/(1+z)}$ - *effect of* R_R/R_T, *recovery*. This figure shows f_{Rec} (normalised to the value of f_{Rec} at $R_R/R_T = 1$ separately for each value of $\sigma_{\Delta z/(1+z)}$) as a function of R_R/R_T in the redshift bin $1.50 < z < 1.75$. (a): high-density environments. (b): low-density environments. The various curves are colour-coded according to the various values of $\sigma_{\Delta z/(1+z)}$ (purple: $\sigma_{\Delta z/(1+z)} = 0.003$; blue: $\sigma_{\Delta z/(1+z)} = 0.01$; green: $\sigma_{\Delta z/(1+z)} = 0.03$; and red: $\sigma_{\Delta z/(1+z)} = 0.06$). Shaded regions show the dispersion in the fraction values at fixed R_R/R_T given by different R_T. The length of the cylinder has been kept fixed, with $n = 1.5$.



(a)



(b)

Figure 25: *Varying* $\sigma_{\Delta z/(1+z)}$ - *effect of* R_R/R_T, *contamination*. This figure shows f_{Con} (normalised to the value of f_{Con} at $R_R/R_T = 1$ separately for each value of $\sigma_{\Delta z/(1+z)}$) as a function of R_R/R_T in the redshift bin $1.50 < z < 1.75$. (a): high-density environments. (b): low-density environments. The various curves are colour-coded according to the various values of $\sigma_{\Delta z/(1+z)}$ (purple: $\sigma_{\Delta z/(1+z)} = 0.003$; blue: $\sigma_{\Delta z/(1+z)} = 0.01$; green: $\sigma_{\Delta z/(1+z)} = 0.03$; and red: $\sigma_{\Delta z/(1+z)} = 0.06$). Shaded regions show the dispersion in the fraction values at fixed R_R/R_T given by different R_T. The length of the cylinder has been kept fixed, with $n = 1.5$.

while for lower values of $\sigma_{\Delta z/(1+z)}$ the situation where $R_R/R_T = 1$ optimises the environmental reconstruction. For all curves, shaded regions show the residual dispersion (after normalisation) in the fraction values at fixed R_R/R_T. This dispersion occurs because smaller scales ($R_R = R_T = 0.3$ Mpc) are reconstructed better than large scales ($R_R = R_T = 2$ Mpc). Therefore, even at fixed R_R/R_T smaller scales have higher f_{Rec} and lower f_{Con} compared to larger scales.

To summarise, it is possible to reconstruct environment in an accurate way only if the photometric redshift uncertainty is small ($\sigma_{\Delta z/(1+z)} \leqslant 0.01$), otherwise the environment is not sufficiently well recovered ($f_{Rec} < 50\% \div 60\%$) and it becomes too contained ($f_{Con} > 10\%$). Moreover, for uncertainty values $\sigma_{\Delta z/(1+z)} \leqslant 0.01$, the best environmental reconstruction is obtained for $n = 1.5$ and for $R_R/R_T = 1$. This also remains generally true for values of $\sigma_{\Delta z/(1+z)} = 0.03$, although the recovery fraction is lower and the contamination fraction is higher. For values of $\sigma_{\Delta z/(1+z)} > 0.03$, the recovery and contamination fractions show that it is not possible to reconstruct environment in an accurate fashion and volume parameters $n = 0.5$ and $R_R/R_T < 1$ optimise the measurement of the density field as they limit the number of contaminating interlopers.

3.4 A DETAILED STUDY OF THE $\sigma_{\Delta z/(1+z)} = 0.01$ CASE

The ultimate goal of this analysis is to understand the effect of photometric redshift uncertainties on local environment, in order to measure the density field and to study the GSMF in different environments for galaxies in the UltraVISTA survey. The application of what found in this chapter to the real data set will be fully described in Chapter 4. Nevertheless, in this Section a specific analysis is performed, assuming the same photometric redshift uncertainty as the UltraVISTA survey, in order to carefully investigate the biases and systematics which can be encountered when relying on a real photometric redshift survey to measure local environment. Moreover, the analysis is performed at all redshifts and not only at selected ones as before.

In order to mimic with this analysis the UltraVISTA data (McCracken et al., 2012), a value for the photometric redshift error of $\sigma_{\Delta z/(1+z)} = 0.01$ has been adopted, in agreement with the value reported in Figure 1 of Ilbert et al. (2013) and with the mean of the error values reported in Table 1 of Ilbert et al. (2013, see also Figure 32 and Table 6 of this thesis), weighted by the number of sources in each spectroscopic sample used to determine the error (see also the discussion in Section 4.1).

Figure 26 shows the effect of varying n on the environmental reconstruction. For all values of n the reconstruction of high-density environments is fairly accurate and, in any case, better than that of low-density environments. The recovery fraction in the high-density

case is above 55% up to $z \sim 2$, and above 60% at $z < 1$. Contamination fractions are always below 10% at all redshifts, although the analysis is not extended farther than $z \sim 2.5$, as the reduced sample size in the farthest redshift bins may have a predominant role in creating the trends observed in the data. The situation is slightly worse for low-density environments, which are reconstructed in a less precise way, because low number counts have a higher error. Contamination from high-density interlopers is low, but f_{Rec} is never above 60%.

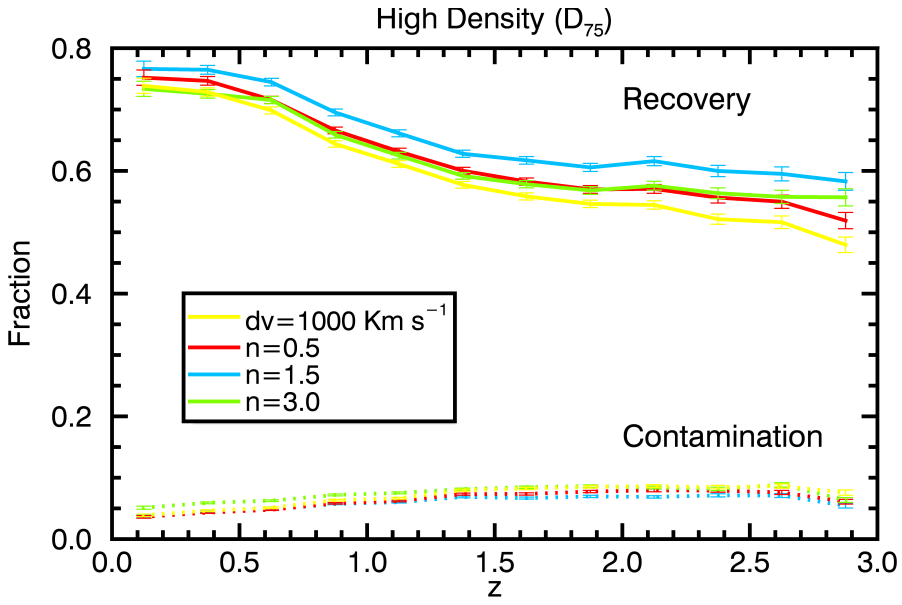
As stated in the previous Section, the length of the cylinder used for environmental reconstruction has indeed an effect on how accurately the environment is recovered. In fact, both cylinder heights that are too small (such as $n = 0.5$ or $dv = 1000 km/s$) or too large (such as $n = 3$) compared to the 3σ error of the photometric redshifts have the effect of worsening the reconstruction of the environment, increasing f_{Con} and decreasing f_{Rec}. Therefore, when dealing with high-precision photometric redshifts ($\sigma_{\Delta z/(1+z)} = 0.01$) a fixed aperture volume with a length roughly of the scale of $\pm 1.5\sigma$ error around z_{phot} grants the best environmental reconstruction.

The height of the cylinder used has then been fixed, to check the effect that a varying aperture radius from $R = 0.3$ Mpc to $R = 2$ Mpc may have on the process. Results are shown in Figure 27. Again it can be seen that the reconstructed environment is not too different from the true environment, with f_{Rec} always above 55% up to redshift $z \sim 2$ (above 60% at redshift $z \lesssim 1$) and f_{Con} always below 10% at all redshifts. Again, the environmental reconstruction is better for high-density environments than for low-density environments.

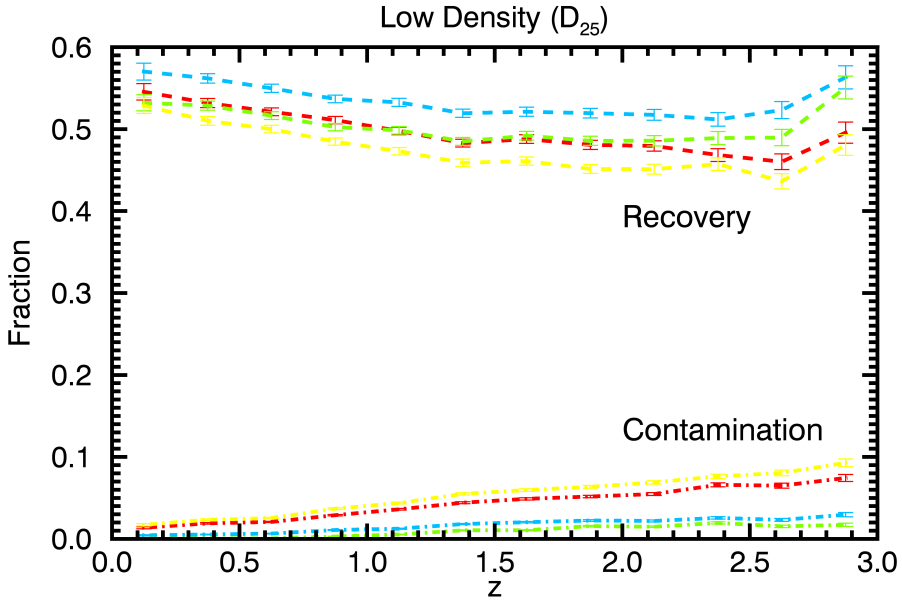
Figure 27 shows that also the aperture radius has an effect on the environmental reconstruction when dealing with high-precision photometric redshifts ($\sigma_{\Delta z/(1+z)} = 0.01$). In particular, both radii that are too large (e. g. $R_R = 2$ Mpc) and radii that are too small (e. g. $R_R = 0.3$ Mpc), compared to R_T (in this case $R_T = 1$ Mpc), have the effect of lowering the accuracy of the reconstructed environment, increasing f_{Con} and decreasing f_{Rec}. Thus, it is possible to conclude that increasing or decreasing the fixed aperture size too much has the effect of worsening the precision of the environmental reconstruction. Therefore, as stated in the previous section, a value of $R_R \simeq R_T$ optimises the environmental reconstruction. This remains true at all redshifts.

3.5 THE RECONSTRUCTION OF THE GSMF FOR THE BEST-CASE $\sigma_{\Delta z/(1+z)} = 0.01$

On the basis of the analysis described above, it has been investigated whether the accuracy of the environmental reconstruction with photometric redshifts has consequences for the differential study of GSMFs in different environments. It is known from spectroscopic surveys that the GSMF of galaxies in different environments have a different

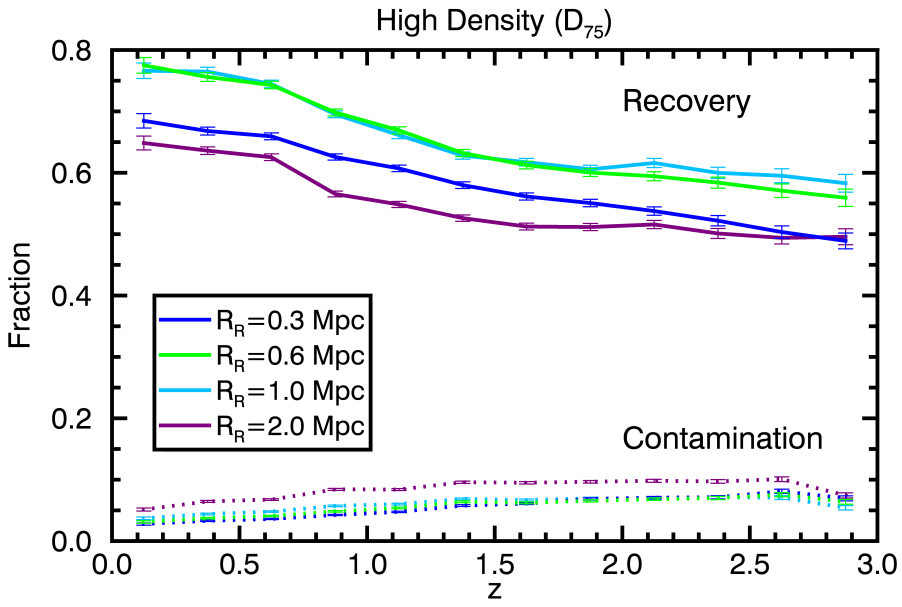


(a)

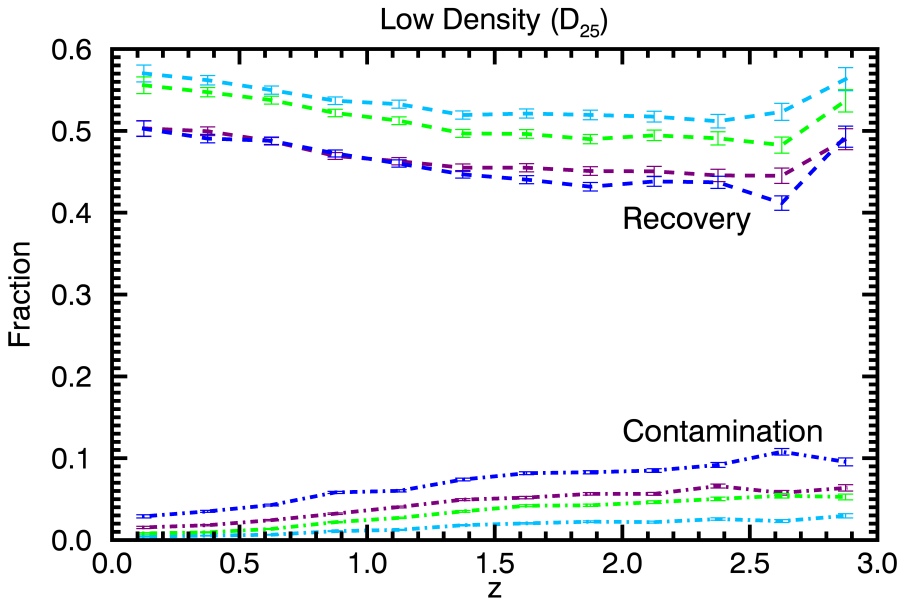


(b)

Figure 26: *Recovery and contamination fractions - varying* n. This figure shows f_{Rec} (solid and dashed lines) and f_{Con} (dotted and dot-dashed lines) as a function of redshift. (a): high-density environments. (b): low-density environments. The various curves are colour-coded according to the various values of n (yellow: $dv = 1000 km/s$; red: $n = 0.5$; cyan: $n = 1.5$; and green: $n = 3$). The radius has been fixed to $R_T = R_R = 1$ Mpc and the uncertainty in the photometric redshift is $\sigma_{\Delta z/(1+z)} = 0.01$. Note the different scale on the ordinate axis between high-density and low-density environments.



(a)



(b)

Figure 27: *Recovery and contamination fractions - varying* R_R. This figure shows f_{Rec} (solid and dashed lines) and f_{Con} (dotted and dot-dashed lines) as a function of redshift. (a): high-density environments. (b): low-density environments. The various curves are colour-coded according to the various values of R_R (blue: $R_R = 0.3$ Mpc; green: $R_R = 0.6$ Mpc; cyan: $R_R = 1$ Mpc; and purple: $R_R = 2$ Mpc). The cylinder height has been fixed to $n = 1.5$, the uncertainty in the photometric redshift is $\sigma_{\Delta z/(1+z)} = 0.01$ and the true environment has been estimated on a scale $R_T = 1$ Mpc. Note the different scale on the ordinate axis between high-density and low-density environments.

shape at least up to $z \leqslant 1$ (see Chapter 1). This is caused by the different formation and evolution scenarios of galaxies within clusters and in the field. For this reason, the contamination of interlopers from different environments (the f_{Con} defined above) along with the dilution of the signal of environmental segregation (quantified by values of $f_{Rec} < 100\%$) might have the effect of changing the shape of the GSMF in a given environment. In the following, the degree of accuracy with which an analysis of the GSMF in different environments is possible if the density field is reconstructed using photometric redshifts with an error of $\sigma_{\Delta z/(1+z)} = 0.01$ is quantified.

The GSMFs of high-density and low-density environments have been computed with the environment reconstructed with $n = 1.5$, $\sigma_{\Delta z/(1+z)} = 0.01$, and $R_T = R_R = 0.3$ and $R_T = R_R = 2$ Mpc. Only the smallest and largest radii have been considered, since they grant the best and worst environmental reconstruction. All other values of R grant intermediate f_{Rec} and f_{Con}. The GSMFs of the mock catalogues are shown in Fig. 28 for the $R_R = R_T = 0.3$ Mpc case and in Fig. 30 for the $R_R = R_T = 2$ Mpc case.

The GSMFs have been derived using the non-parametric $1/V_{max}$ estimator (see Section 2.3), considering all the galaxies down to $K \leqslant 24$. As the z_{max} information was not available, the total volume in the considered redshift range was used. In particular, not taking the z_{max} into account mainly affects the low-mass end of the GSMF, where the volume occupied by each source is more likely to be overestimated. For this reason, mass completeness limits have also been derived, following Pozzetti et al. (2010), but at the upper boundary of each redshift bin (instead of at the lower boundary, which is the case when z_{max} values are available).

The error bars shown in the plots represent only the Poissonian errors. As the UltraVISTA Survey field is smaller compared to the area of the mock catalogues considered here (1.48 deg^2 for UltraVISTA compared to the 8 deg^2 used in this work), GSMFs have also been calculated for galaxies in five independent areas of 1.48 deg^2 to simulate the real UltraVISTA data. It has been found that, despite the larger error bars due to the lower number of galaxies present in the smaller fields, the results found with the larger area, and detailed below, hold up to $z \sim 2$. At higher redshifts, massive galaxies of $M^* \gtrsim 10^{11} M_\odot$ (which carry most of the signal of environmental difference) are too few in every redshift bin, because of the smaller area, in the true environment case already; therefore no environmental difference is recoverable. Nevertheless this may not be a limit in the analysis of the UltraVISTA data, as different mass or redshift bins may be applied to increase statistics, and because mock galaxy catalogues may underestimate the number of massive galaxies in comparison to reality (see Figures 14 and 15 of Ilbert et al., 2013). For these reasons, here GSMFs calculated with an area of 8 deg^2 are shown.

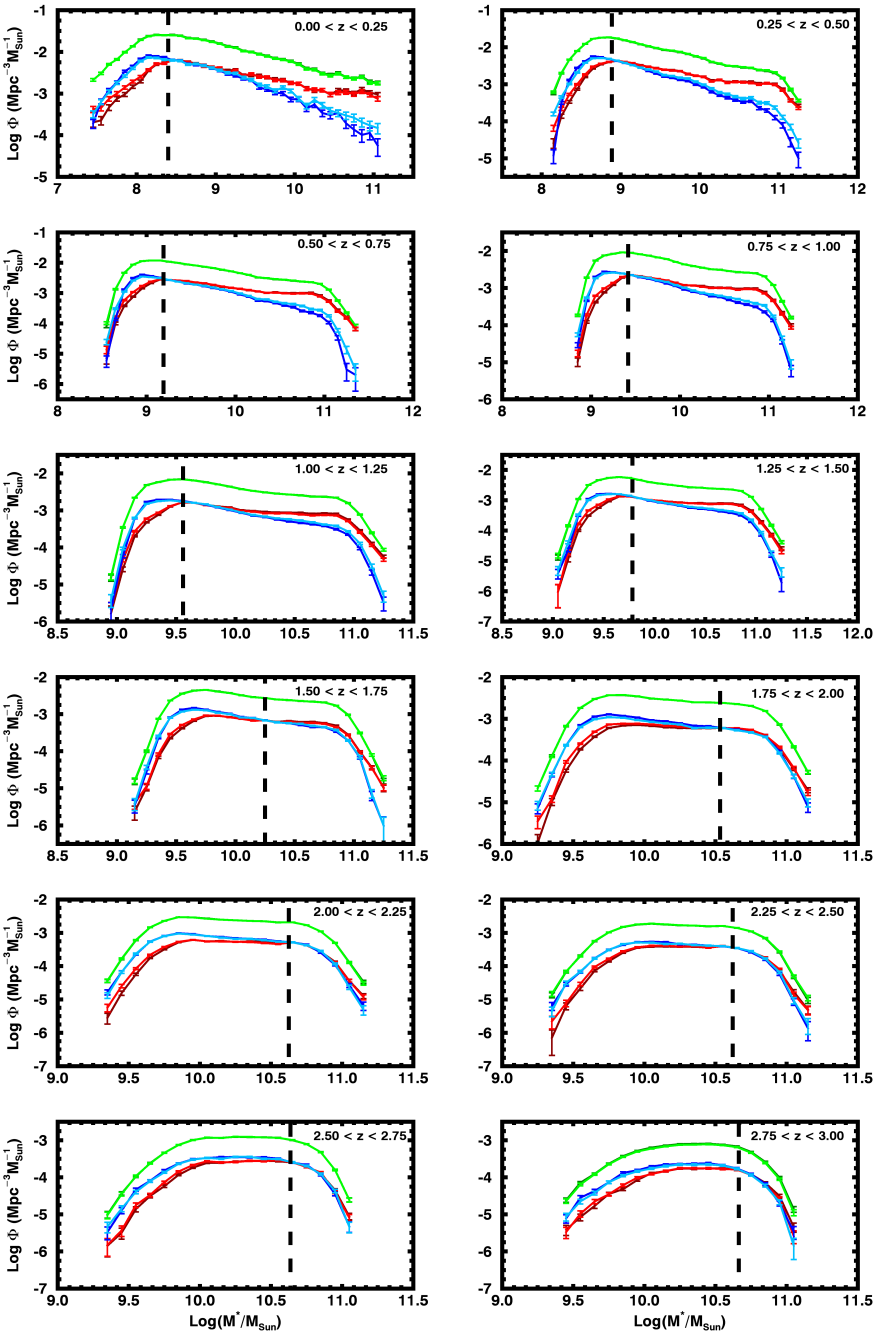


Figure 28: *Mock catalogues* GSMFs - $R_R = R_T = 0.3$ *Mpc*. The dark green and light green curves refer to the total GSMF, respectively, using z_{true} and z_{phot} for the calculation of the mass function. Dark red and light red curves refer to high-density environments, respectively, using ρ_{true}, dark red, and using ρ_{rec}, light red. Dark blue and light blue curves refer to low-density environments, respectively, using ρ_{true}, dark blue, and using ρ_{rec}, light blue. The black dashed lines are the mass completeness limits described in the text. The parameter values for the aperture are set to $R_R = R_T = 0.3$ and $n = 1.5$, with $\sigma_{\Delta z/(1+z)} = 0.01$. Error bars refer to $1/V_{max}^2$ estimates.

It is known that galaxies in high-density environments occupy a smaller volume than galaxies in low-density environments (see e. g. Davidzon et al., 2016). To account for this fact and to be able to compare GSMFs in different environments, GSMFs in high-density and low-density environments have been normalised to 1/4 of the value that the total GSMF has at the mass limit, both in the true and reconstructed environments. Hence, although the information on the GSMF normalisation is lost, it is still possible to compare their shapes.

From Figure 28 it is possible to see how the GSMFs of true D_{75} and D_{25} environments are intrinsically different in the $R_R = R_T = 0.3$ Mpc case. This difference is a function of mass and redshift and for masses $M \gtrsim 10^{11} M_\odot$ ranges between $\gtrsim 1$ dex at $z < 0.5$ and ~ 0.3 dex at $z \sim 2.5$. At higher redshifts the two GSMFs become indistinguishable. This trend with mass and redshift of the differences between true environment GSMFs can be better appreciated in the ratio between D_{75} and D_{25} GSMFs, shown in Figure 29.

The difference is smaller when considering $R_R = R_T = 2$ Mpc (Figures 30 and 31). At low redshifts the difference between the true D_{75} and D_{25} GSMFs is $\lesssim 0.5$ dex for masses below $10^{11} M_\odot$. This difference is of ~ 0.1 dex at $z \sim 2.5$ for $M \sim 10^{11} M_\odot$. The ratio of the two GSMFs is significantly different from one below $z \sim 2$ and only for masses above $10^{11} M_\odot$.

In the same four figures GSMFs in different reconstructed environments together with their ratios are also reported. In the $R_R = R_T = 0.3$ Mpc case (Figure 28), it is possible to see how the reconstructed environment GSMF closely follows the true environment, since D_{75} is well distinguishable from the D_{25} at low redshifts and shows the same trends in mass and redshift. Also, the ratio of D_{75} to D_{25} GSMFs (Figure 29) follows the true case, although it is possible to see that resulting differences are somewhat damped when passing from ρ_{true} to ρ_{rec}. To quantify this decrease of the difference between high-density and low-density environments, in Table 4 it is reported the maximum decrease of the ratio of the GSMFs in extreme environments between the ρ_{true} and the ρ_{rec} case, defined as

$$\xi = \max \left[\log \left(\frac{\Phi(D75)}{\Phi(D25)} \right) \Bigg|_{True} - \log \left(\frac{\Phi(D75)}{\Phi(D25)} \right) \Bigg|_{Rec} \right] \tag{23}$$

The reduction in the differences is always below ~ 0.7 dex. This value is obtained at high masses ($M \sim 10^{11} M_\odot$) where the true environment GSMFs are intrinsically different. Table 5 shows that this is true also for the $R_R = R_T = 2$ Mpc case. Although, from Figures 30 and 31 it is possible to see that already the true environment GSMFs are less different between high- and low-density than in the $R_R = R_T = 0.3$ Mpc case.

Thus, it has been shown that if the GSMFs in high- and low-densities are different for ρ_{true}, they will still be different for ρ_{rec}, although

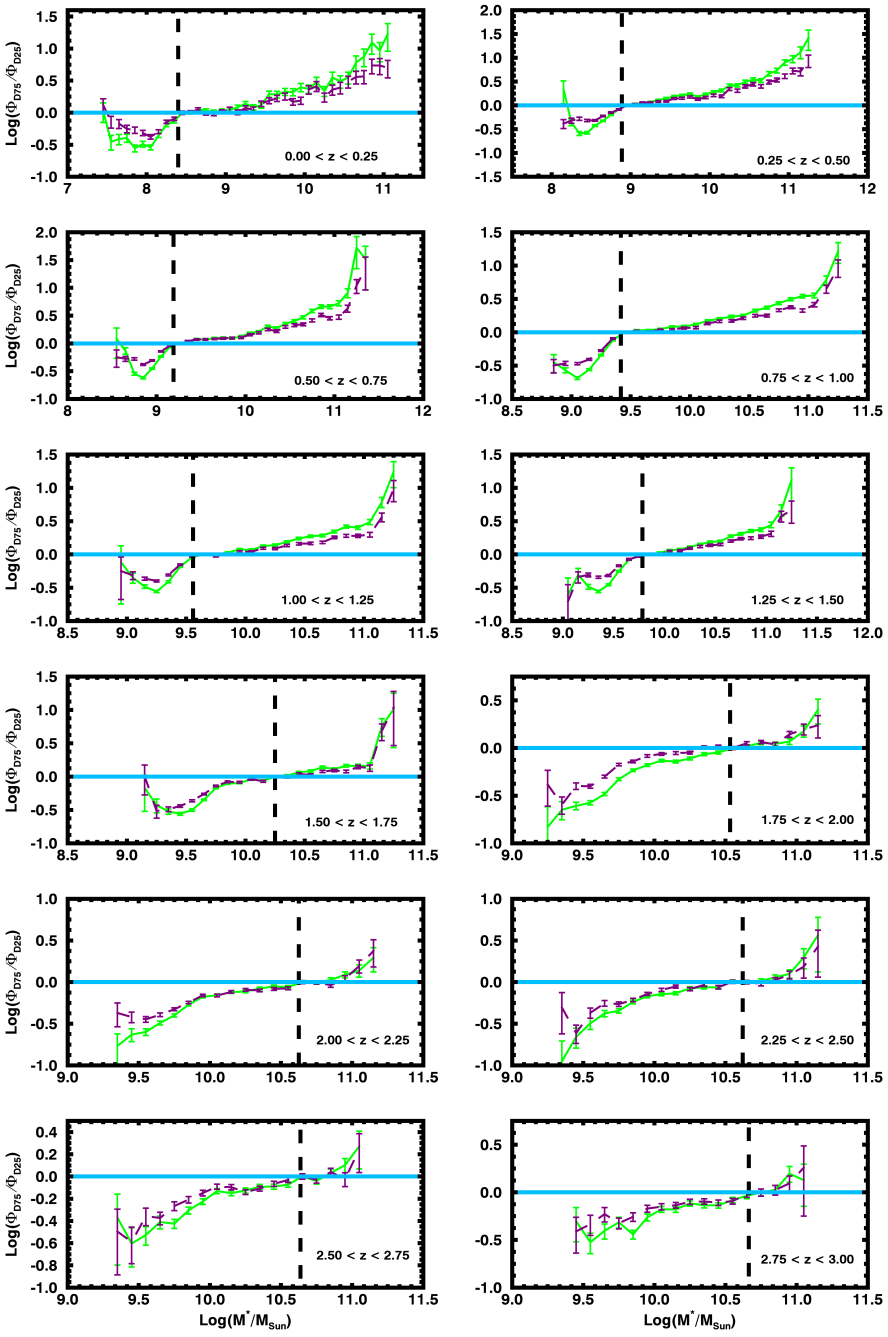


Figure 29: *Ratio of high-density and low-density GSMFs* - $R_R = R_T = 0.3$ *Mpc*. Ratio of the high-density mass function and the low-density mass function (Φ_{D75}/Φ_{D25}) in the true (light green curve) and reconstructed (purple curve) environments. The black dashed lines are the mass completeness limits described in the text. The parameter values for the aperture are set to $R_R = R_T = 0.3$ Mpc and $n = 1.5$, with $\sigma_{\Delta z/(1+z)} = 0.01$. Error bars refer to $1/V^2_{max}$ estimates.

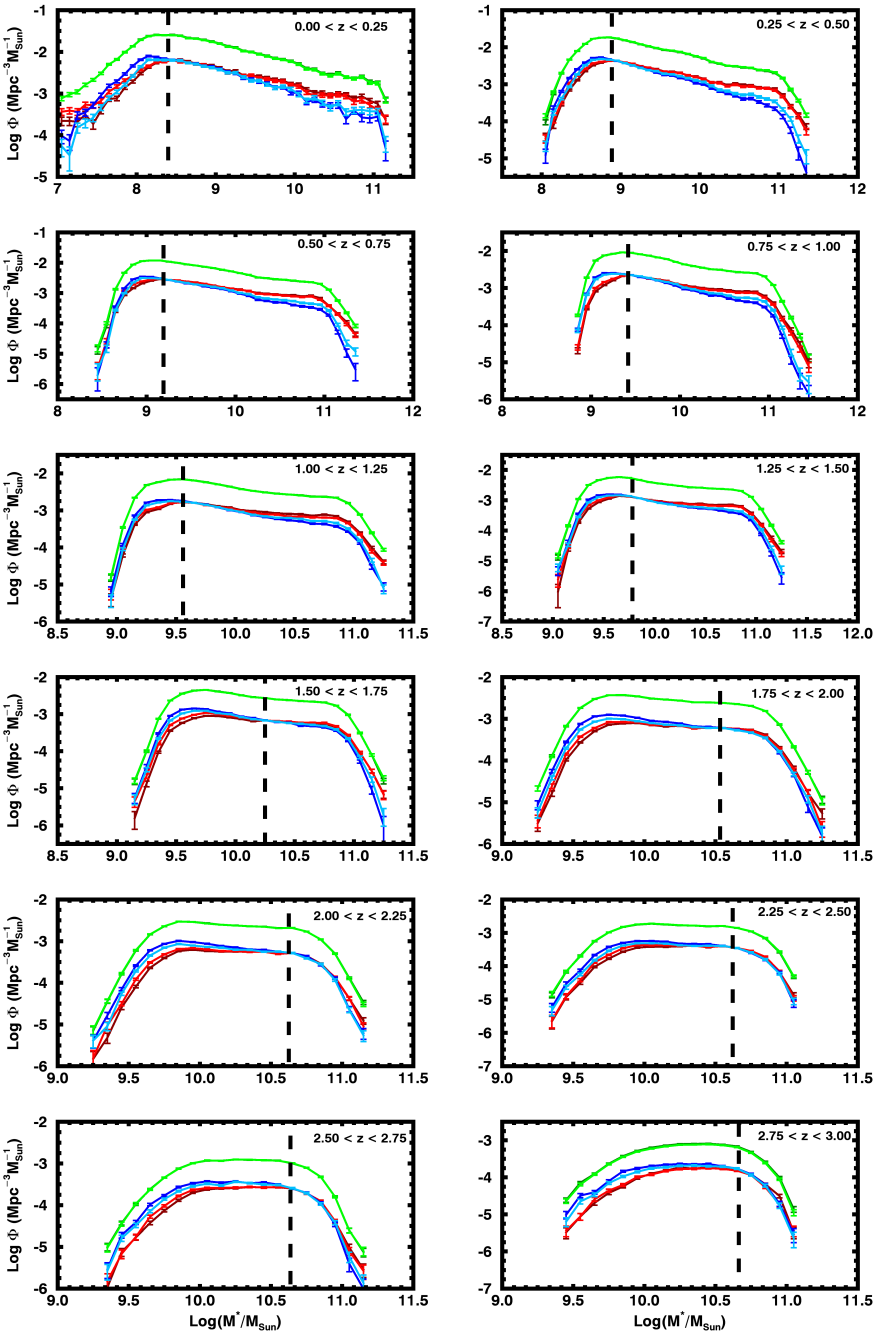


Figure 30: *Mock catalogues* GSMFs - $R_R = R_T = 2$ *Mpc*. The dark green and light green curves refer to the total GSMF, respectively, using z_{true} and z_{phot} for the calculation of the mass function. Dark red and light red curves refer to high-density environments, respectively, using ρ_{true}, dark red, and using ρ_{rec}, light red. Dark blue and light blue curves refer to low-density environments, respectively, using ρ_{true}, dark blue, and using ρ_{rec}, light blue. The black dashed lines are the mass completeness limits described in the text. The parameter values for the aperture are set to $R_R = R_T = 2$ Mpc and $n = 1.5$, with $\sigma_{\Delta z/(1+z)} = 0.01$. Error bars refer to $1/V^2_{max}$ estimates.

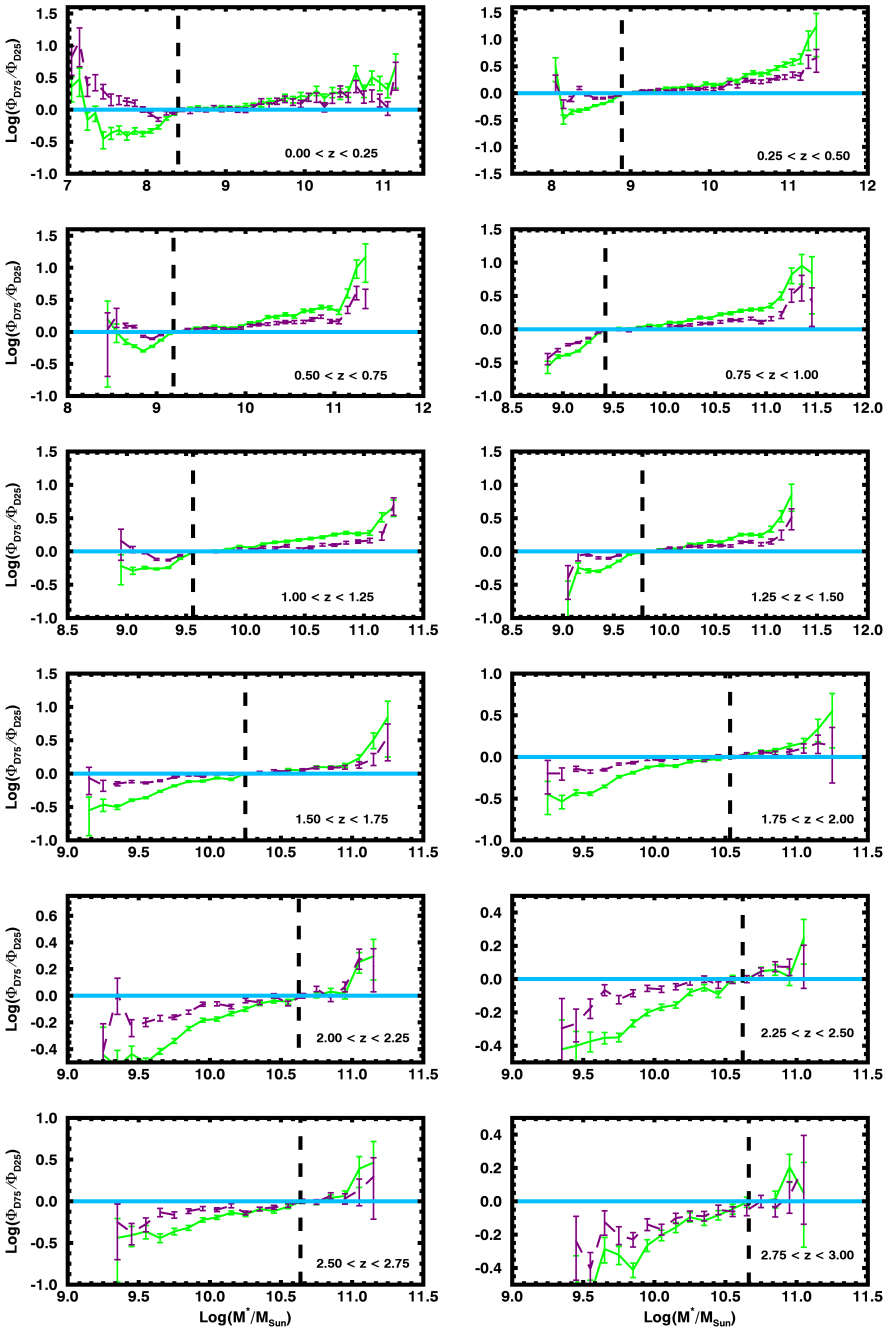


Figure 31: *Ratio of high-density and low-density GSMFs* - $R_R = R_T = 2$ *Mpc*. Ratio of the high-density mass function and the low-density mass function (Φ_{D75}/Φ_{D25}) in the true (light green curve) and reconstructed (purple curve) environments. The black dashed lines are the mass completeness limits described in the text. The parameter values for the aperture are set to $R_R = R_T = 2$ Mpc and $n = 1.5$, with $\sigma_{\Delta z/(1+z)} = 0.01$. Error bars refer to $1/V^2_{max}$ estimates.

REDSHIFT	ξ (DEX)	$\log(M_\xi/M_\odot)$
$0.00 < z < 0.25$	0.53	11.05
$0.25 < z < 0.50$	0.48	11.25
$0.50 < z < 0.75$	0.68	11.25
$0.75 < z < 1.00$	0.24	11.25
$1.00 < z < 1.25$	0.26	11.25
$1.25 < z < 1.50$	0.45	11.25
$1.50 < z < 1.75$	0.09	10.85
$1.75 < z < 2.00$	0.16	11.15
$2.00 < z < 2.25$	0.06	10.85
$2.25 < z < 2.50$	0.14	11.15
$2.50 < z < 2.75$	0.13	10.95
$2.75 < z < 3.00$	0.10	10.95
Average	0.28	

Table 4: Maximum decrease in the ratio of D_{75} and D_{25} GSMFs between ρ_{rec} and ρ_{true}. The ratio ξ is calculated as in Equation. (23), M_ξ is the mass at which ξ is located. These values refer to $R_R = R_T = 0.3$ Mpc.

REDSHIFT	ξ (DEX)	$\log(M_\xi/M_\odot)$
$0.00 < z < 0.25$	0.28	10.85
$0.25 < z < 0.50$	0.59	11.35
$0.50 < z < 0.75$	0.63	11.35
$0.75 < z < 1.00$	0.42	11.45
$1.00 < z < 1.25$	0.28	11.15
$1.25 < z < 1.50$	0.34	11.25
$1.50 < z < 1.75$	0.29	11.25
$1.75 < z < 2.00$	0.41	11.25
$2.00 < z < 2.25$	0.08	11.15
$2.25 < z < 2.50$	0.15	11.05
$2.50 < z < 2.75$	0.26	11.05
$2.75 < z < 3.00$	0.17	10.95
Average	0.33	

Table 5: Maximum decrease in the ratio of D_{75} and D_{25} GSMFs between ρ_{rec} and ρ_{true}. The ratio ξ is calculated as in Equation. (23), M_ξ is the mass at which ξ is located. These values refer to $R_R = R_T = 2$ Mpc.

the difference is somewhat reduced. If, instead, there are no differences in the true environment GSMFs then no spurious differences are introduced when using reconstructed environments. It can be then concluded that an analysis of the GSMF in different environments is possible, even when relying on photometric redshifts for the environmental reconstruction. In particular, this is generally true at all redshifts (up to $z \sim 2.5$), at all masses, and for both large and small scales.

When investigating the differences between GSMFs calculated for galaxies in high-density and low-density environments with high-precision photometric redshifts, all differences found may be considered as lower limits of the true differences. In fact, this analysis shows that the effect of the error of photometric redshifts on the GSMFs of galaxies in different environments is to damp differences between high-density and low-density regions, while nevertheless not deleting them completely.

3.6 CONCLUSIONS

In this Chapter, the validated mock galaxy catalogues by Merson et al. (2013) have been used to investigate how the galaxy environment can be reconstructed using photometric redshifts with varying precision. The mock galaxy sample has been selected to have $K \leqslant 24$ and an area of $8\,\mathrm{deg}^2$ has been extracted from the original $100\,\mathrm{deg}^2$ of the catalogue. Using each galaxy's cosmological redshift (z_{true}) a set of photometric redshifts (z_{phot}) with varying precision has been simulated by adding a Gaussian error to each galaxy's observed redshift (i. e. the cosmological redshift to which the contribution of the galaxy peculiar velocity has been added).

Galaxy environments have been estimated through the use of a fixed aperture method (see Section 2.1.1), by counting objects inside a cylinder of base radius R and length h proportional to the photometric redshift uncertainty through the parameter n as $h = \pm n \cdot \sigma_{\Delta z/(1+z)} \cdot (1+z)$. High-density ($D_{75}$) and low-density ($D_{25}$) environments have been defined using the 75th and 25th percentiles of the volume density distribution. For each galaxy, both a true environment estimate (ρ_{true}, using each galaxy's z_{true}) and a reconstructed estimate (ρ_{rec} using each galaxy's z_{phot}) have been derived. Several combinations of both the fixed aperture volume parameters n and R and of the photometric redshift uncertainty $\sigma_{\Delta z/(1+z)}$ have been explored. The derived ρ_{true} and ρ_{rec} have then been compared to study how the density field can be reconstructed with photometric redshifts. These results have also been applied to the study of the GSMF in different environments for the best-case photometric redshift uncertainty $\sigma_{\Delta z/(1+z)} = 0.01$. What found can be summarised as follows:

1. It is possible to reconstruct galaxy environments in an accurate way only with the use of high-precision photometric redshifts ($\sigma_{\Delta z/(1+z)} \leqslant 0.01$). In particular, to well recover high- and low-density environments (with $f_{Rec} \geqslant 60\% \div 80\%$) with a low level of contaminating interlopers ($f_{Con} \leqslant 10\%$), it is necessary to carefully tune the parameters of the cylinder used for the estimate of the density field. In this case, the best environmental reconstruction is obtained considering a cylinder with length $\propto \pm 1.5\sigma$ error on the photometric redshift and a radius $R_R = R_T$. A volume with a length too large or too small compared to the $\pm 1.5\sigma$ error and with a base area too large or too small compared to the size of the physical scale of the studied environment leads to an inaccurate environmental reconstruction, with lower f_{Rec} and higher f_{Con}.

2. Even if all the cylinder parameters are tuned to have the best case of environmental reconstruction, recovery fractions are still higher ($f_{Rec} \geqslant 70\%$) and contamination fractions are lower $f_{Con} \leqslant 5\%$ for smaller physical scales ($R_R = R_T = 0.3 - 0.6$ Mpc) compared to larger scales ($R_R = R_T = 1 - 2$ Mpc).

3. For $\sigma_{\Delta z/(1+z)} \geqslant 0.03$ recovery fractions lower rapidly to $f_{Rec} < 50\%$ and contamination fractions increase up to $f_{Con} \sim 20\%$. This result is reasonable if it is considered that the typical velocity dispersion inside the richest galaxy clusters is on the order of $\Delta z \simeq \pm \frac{1500 km/s}{c} \cdot (1+z) \simeq \pm 0.005 \cdot (1+z)$, which is comparable to the best-case photometric redshift uncertainty considered here. Moreover, for photometric redshift errors on the order of $\sigma_{\Delta z/(1+z)} \leqslant 0.01$ recovery fractions are still higher and contamination fractions lower for $n = 1.5$ and $R_R/R_T = 1$, while for higher values of $\sigma_{\Delta z/(1+z)}$ values of $R_R/R_T < 1$ are preferable.

4. When applying these results to the GSMF (in the best-case of $\sigma_{\Delta z/(1+z)} = 0.01$) it is found that differences between the GSMFs in different environments can be recovered accurately if present, although some reduction (which reaches at most ~ 0.7 dex, with an average of ~ 0.3 dex) is inevitable. Nevertheless, spurious differences between the GSMFs in high- and low-density environments do not seem to be created. Therefore any environmental segregation found in real data may be regarded as a lower limit of what would be found if a measure of the intrinsic density field were available.

With this study, it has been found that an analysis of the GSMF in different environments is only possible with high-precision $\sigma_{\Delta z/(1+z)} \leqslant 0.01$ photometric redshifts, provided that the fixed aperture length and radius are optimised to give the best measurement of the density field. GSMFs in different environments can be studied while keeping

in mind that photometric redshifts, even high-precision redshifts, reduce differences between high-density and low-density environments by as much as ~ 0.7 dex. In Chapter 4 the results found here are applied to the galaxies in the UltraVISTA survey.

4

THE ENVIRONMENT-DEPENDENT EVOLUTION OF MASSIVE GALAXIES TO $z \simeq 3$

The Galaxy Stellar Mass Function (GSMF) is a powerful statistical tool to study how galaxies are formed and evolve, as it allows to summarize in a single Probability Distribution Function (PDF) the number of galaxies per unit volume as a function of mass. By studying how the GSMF changes as a function galaxy properties, redshift, and local environment it is possible to study in detail the physical processes which regulate the galaxy mass assembly. The study of the shape of the GSMF is a strong indicator of how the build-up of galaxy mass happens throughout cosmic history. Moreover, theory and numerical simulations can make predictions for the shape of the GSMF to be compared with observations and therefore understand the physical processes responsible for galaxy evolution.

In particular, the GSMF in different environments has been thoroughly investigated in several works (see Section 1.2.2), relying on spectroscopic surveys of both local galaxies (like SDSS and GAMA) and up to $z \sim 1 - 1.5$. As said in Chapter 2, very few spectroscopic redshifts surveys allow us to study the local environment at high redshift with a sufficient statistic, therefore the GSMF has been studied at $z \sim 3$ mainly using photometric redshift surveys. There is broad agreement in the literature over the fact that the shape of the GSMF is different in high- and low-density environments, at least up to $z \sim 1 - 1.5$ (see Section 1.2.2). Moreover, high-density environments are dominated by a massive and quiescent galaxy population, while lower mass, star-forming galaxies are more frequent in low-density environments. Nevertheless, the exact galaxy stellar mass at which quiescent galaxies start to dominate high-density environments and the redshift and density ranges in which environmental trends are visible are still not well constrained. Moreover, the wide use of photometric redshifts for the study of the GSMF in different environments at $z \geqslant 1.5$ creates a discrepancy in the precision of the results obtained with respect to measurements performed using spectroscopic redshift surveys.

In this Chapter the results of an analysis of the GSMF in different environments performed in the redshift range $0.2 \leqslant z \leqslant 3$, using the high-precision photometric redshift sample of the UltraVISTA survey are reported. This study makes large use of the results exposed in Chapter 3, by measuring local environment with a fixed cylindrical aperture with $R = 0.3 - 2$ Mpc and $n = 1.5$. The results exposed in this Chapter are fully described in Malavasi et al. (2016b).

This Chapter is structured as follows: in Section 4.1 the data set used is introduced, while in Section 4.2 the methods employed to measure local environment and calculate GSMFs are described. In Section 4.3 the main results on the GSMFs in the UltraVISTA field and their shape in different environments are presented, while a discussion is provided in Section 4.4. A comparison with other works carried out in the same field or redshift range is performed in Section 4.5. Conclusions are summarised in Section 4.6. The cosmology adopted for the analysis presented in this Chapter is a standard cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$.

4.1 THE ULTRAVISTA SURVEY

The UltraVISTA Survey (McCracken et al., 2012) is a Near Infra-red (NIR) survey performed in the COSMOS field. It covers an area of $\sim 1.5\,\mathrm{deg}^2$ using the Visible and Infrared Survey Telescope for Astronomy (VISTA) in Y, J, H, K_S bands, to an apparent magnitude limit of $K_S \leqslant 24$. Stellar masses, photometric redshifts and physical parameters for the galaxies and AGNs have been measured by Ilbert et al. (2013), using (in addition) the wealth of multiwavelength data collected in the context of the COSMOS survey (see Table 1 of Ilbert et al., 2009).

The sample that has been used for this work comprises galaxies from the UltraVISTA Survey. The total sample has been selected in K_S band and is composed of 339 384 objects. After the removal of X-ray sources, stars and objects in masked areas[1], the number of objects is reduced to 209 758 galaxies with photometric redshift between $0.2 \leqslant z \leqslant 4$, $K_S \leqslant 24$, and measured stellar mass. These objects constitute the final sample on which the analysis has been performed. The K_S-band and redshift cuts have been introduced to be consistent with Ilbert et al. (2013) and to be able to compare the GSMF for the total, quiescent and star-forming populations with what derived by Ilbert et al. (2013).

The galaxies of the final sample have been divided in quiescent and star-forming following the colour-colour diagram ($NUV - r^+$ *vs* $r^+ - J$) as in Ilbert et al. (2013). The $NUV - r^+ - J$ colour-colour diagram has been chosen instead of a UVJ selection to be consistent with Ilbert et al. (2013). Moreover, in Ilbert et al. (2010, 2013) it is stated that the $NUV - r^+ - J$ colour-colour diagram grants a better distinction between star-forming and quiescent galaxies, as the $NUV - r^+$ colour is a better indicator of the current *vs* past star-formation activity (see e. g. Arnouts et al., 2007; Martin et al., 2007). Moreover, the NUV rest-

1 X-ray sources have been removed from the sample as their photometric redshifts are less reliable and in order to be consistent with other works who used the same data set (Darvish et al., 2015; Scoville et al., 2013). Masked areas are regions of the field close to the survey edges or close to saturated stars, where the photometry has lower quality and it has not been possible to calculate reliable photometric redshifts.

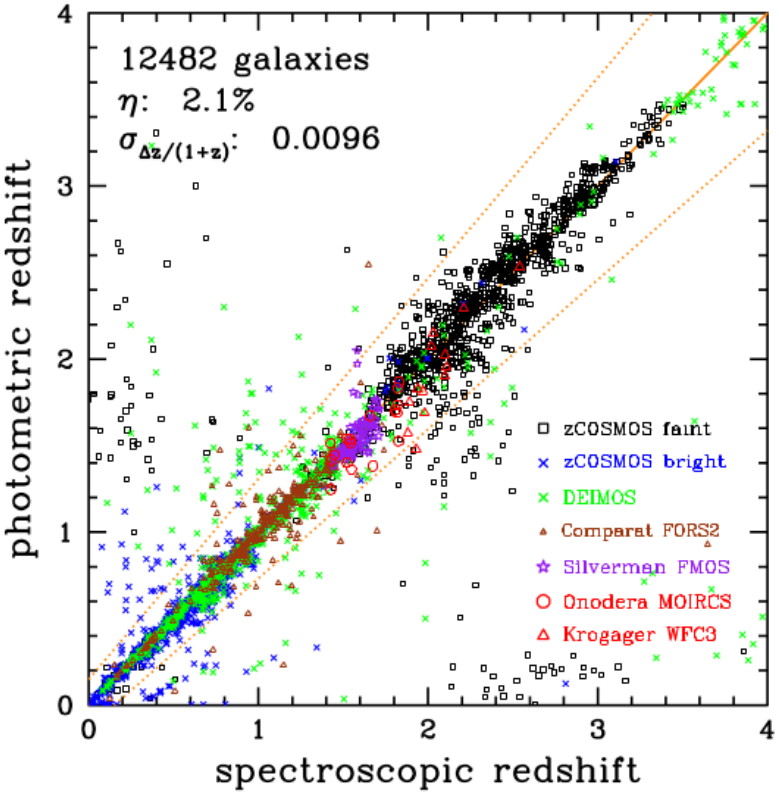


Figure 32: Comparison of the photometric redshifts from the UltraVISTA survey with a selection of spectroscopic redshifts from different surveys, listed in the bottom right corner of the plot. The photometric redshift error for galaxies limited at $K_S \leqslant 24$ is reported in the top left corner, together with the fraction η of catastrophic failures, and the total number of galaxies in the spectroscopic samples. Figure from Ilbert et al. (2013).

frame band is still sampled by optical data at $z > 2$ which does not happen for the rest-frame U-band. According to the $NUV - r^+ - J$ selection, $\sim$ 10% of the galaxies between $0.2 \leqslant z \leqslant 4$ are quiescent and the remaining fraction of $\sim$ 90% are star-forming.

The analysis has been performed in 8 redshift bins from $z = 0.2$ to $z = 4$. Following Ilbert et al. (2013) it has been assumed a photometric redshift uncertainty of $\sigma_{\Delta z/(1+z)} = 0.01$ for all the galaxies in the sample (see also Section 3.4). This value of $\sigma_{\Delta z/(1+z)}$ has been chosen to be consistent with Figure 1 of Ilbert et al. (2013, reported in Figure 32 for reference). This figure shows a comparison between the photometric redshifts for the UltraVISTA survey and spectroscopic redshifts from a set of various samples limited at $K_S \leqslant 24$. The photometric redshift error for the total sample at $K_S \leqslant 24$ is reported in the top left corner, together with the fraction η of catastrophic failures.

The sizes of the spectroscopic samples used for the comparison are listed in Table 1 of Ilbert et al. (2013), reported in Table 6 for reference. A value of $\sigma_{\Delta z/(1+z)} = 0.01$ is in agreement with the average

SAMPLE	SIZE	z_{med}	I_{med}	$\sigma_{\Delta z/(1+z)}$	η (%)
zCOSMOS-Bright	9389	0.50	21.4	0.0080	0.6
Kartaltepe et al. (2015)	570	0.73	22.0	0.0105	3.2
Comparat et al. (2015)	382	1.12	22.6	0.0163	4.7
Capak et al., in prep.	631	1.15	23.5	0.0213	9.5
Onodera et al. (2012)	17	1.55	23.9	0.0446	0.0
Silverman et al. (2015)	88	1.58	23.2	0.0259	1.1
Krogager et al. (2014)	13	2.02	24.8	0.0708	7.7
zCOSMOS-Faint	1392	2.15	23.6	0.0297	7.7

Table 6: Spectroscopic samples used for determining the error on the UltraVISTA photometric redshifts. Here the sample size refers to the number of galaxies with secure spectroscopic redshifts and $K_S \leqslant 24$. The median redshift (z_{med}) and I-band magnitude (I_{med}) of each sample are given, together with the error on photometric redshifts derived from the samples ($\sigma_{\Delta z/(1+z)}$) and the percentage of catastrophic outliers (η). Table from Ilbert et al. (2013).

of the errors reported in Table 6, weighted by the number of galaxies in each spectroscopic sample used to determine the error. These samples are rather small, sometimes only tens of galaxies, and may therefore overestimate the photometric redshift uncertainty if used independently.

Nevertheless, the uncertainty value that has been assumed may still underestimate the photometric redshift uncertainty at $z \geqslant 1.5$ and for faint galaxies. For this reason, a test using a larger photometric redshift error for galaxies at $z > 1.5$ has been performed (see Section 4.3.4), finding that the results exposed here are not significantly affected by larger photometric redshift uncertainties.

4.2 MEASURING LOCAL ENVIRONMENT AND THE GSMF

Local environment in the UltraVISTA survey has been determined using a fixed aperture method (see Section 2.1.1). The performance of this method with photometric redshifts has been extensively tested using mock galaxy catalogues (see Chapter 3). Following what has been found in this previous chapter, it has been used a cylinder, centred on each galaxy, with radius $R = 0.3, 0.6, 1$, and 2 Mpc and with height h equal to the 3σ photometric redshift error, according to Equation (8), with $n = 1.5$ and $\sigma_{\Delta z/(1+z)} = 0.01$.

All the galaxies in the sample were used both as targets and tracers for the density field estimation. Volume densities have been computed according to Equation (3), with $N(\vec{r})$ from Equation (6) and $V(\vec{r}) = \pi R^2 h$. Volume densities have been used because they take

SAMPLE	$z = 0.5$	$z = 1$	$z = 2$
All	$10^{8.5} M_\odot$	$10^{9.2} M_\odot$	$10^{9.7} M_\odot$
Quiescent	$10^{9} M_\odot$	$10^{9.5} M_\odot$	$10^{10.2} M_\odot$
Star-forming	$10^{8.5} M_\odot$	$10^{9.2} M_\odot$	$10^{9.7} M_\odot$

Table 7: Example of mass completeness limits at three different redshifts for all galaxies in the UltraVISTA survey and for only quiescent and star-forming ones.

the variations of the cylinder volume (due to the variation of the volume height inside the same redshift bin) into account on a galaxy by galaxy basis. In fact, as the cylinder length in the radial direction has been chosen proportional to the photometric redshift error of each galaxy, galaxies at different redshifts, even inside the same redshift bin, will have different volume sizes. This can create differences in their environment if not properly accounted for. By using volume densities the problem is solved in a self consistent fashion and environmental densities can be better compared (see Malavasi et al., 2016a, for details).

The UltraVISTA-COSMOS field has a complex shape, due to many holes left in the field by saturated stars. Galaxies close to edges or holes in the field can have their environmental measurement biased. In order to limit this effect a correction to the measured environments for galaxies too close to the edges has been applied. All galaxies for which the fraction of the area outside the survey edges (including holes in the field) was greater than 50% have been rejected while the measurement of the density field for all other galaxies has been corrected by dividing for the fraction of the aperture area inside the edges. Moreover, galaxies with RA (deg) > 150.55 and dec (deg) < 1.8 were not used in the measurement of the environment, as they lie in a small sky area far from the main field and they would have been too dominated by edge effects (see Figure 33). The sample sizes after the correction for the edge effects are reduced to 208 624, 208 446, 208 138, and 207 183 in the $R = 0.3, 0.6, 1$, and 2 Mpc case, respectively. Mass completeness limits for these samples have been calculated as in Pozzetti et al. (2010) and are in very good agreement with those of Ilbert et al. (2013). An example of the mass completeness limits for the total, passive, and star-forming populations is reported in Table 7.

High-density and low-density environments were selected as those above the 75th percentile or below the 25th percentile of the volume density distribution of galaxies with $M^* \geqslant 10^{10} M_\odot$, with the quartiles of the distribution computed at each redshift bin. This mass threshold has been chosen because the increase in the mass completeness limit of the sample with redshift can influence the density value

of the percentiles used to define high- and low-density environments. In fact, at low redshifts the sample is complete at lower masses, the dynamic range of the density measurement is large and the environmental density thresholds used to define high- and low-density environments are lower. Conversely at high redshifts the dynamic range is reduced, because the sample is complete only at higher masses compared to the low redshift case, and the threshold for the definition of high- and low-density environments is higher. The volume density distribution is sensitive to the mass completeness limit of the sample, because of the mass-density relation. In this way it would not be possible to compare the same kind of environments at low and high redshifts. By choosing a mass cut close to the mass completeness limit of the highest redshift bin it has been possible to compare the galaxy population at low and high redshift in a consistent way. High- and low-density environments have been defined for both quiescent and star-forming galaxies using the quartiles of the total galaxy population. In the following, the notation D_{75} and D_{25} will be used to indicate high- and low-density environments, respectively.

GSMFs have been calculated with the non parametric $1/V_{max}$ estimator (see Section 2.3). They have been calculated separately for quiescent and star-forming galaxies, both in high-density and low-density environments. A comparison of the GSMFs of this work and those of Ilbert et al. (2013) shows perfect agreement. As the thresholds for defining high- and low-density environments have been calculated using only galaxies more massive than $10^{10} M_{\odot}$, the GSMF result roughly normalised at high masses, due to the fact that

$$\int_{M \geqslant 10^{10} M_{\odot}} \Phi_{D75}(M) dM = \int_{M \geqslant 10^{10} M_{\odot}} \Phi_{D25}(M) dM \qquad (24)$$

When calculating mass functions, if the number of galaxies in a given mass bin is lower than two, the prescription for small counts Poisson statistics of Gehrels (1986) has been applied, in the form of 1σ upper and lower limits of his Tables 1 and 2.

4.3 RESULTS

Although the analysis presented in this chapter is based on photometric redshifts, the method that has been used to reconstruct local density and that has been tested on mock galaxy catalogues (see Chapter 3) is able to provide a robust measurement of galaxy environments up to $z = 3$ and on various physical scales. The high-precision photometric redshifts of the UltraVISTA sample allow us to trace environmental effects on galaxy properties throughout cosmic history, contributing in the creation of a consistent picture of galaxy evolution.

Figure 33 shows an example of the excellent performance of the fixed aperture method in estimating the density field. This figure

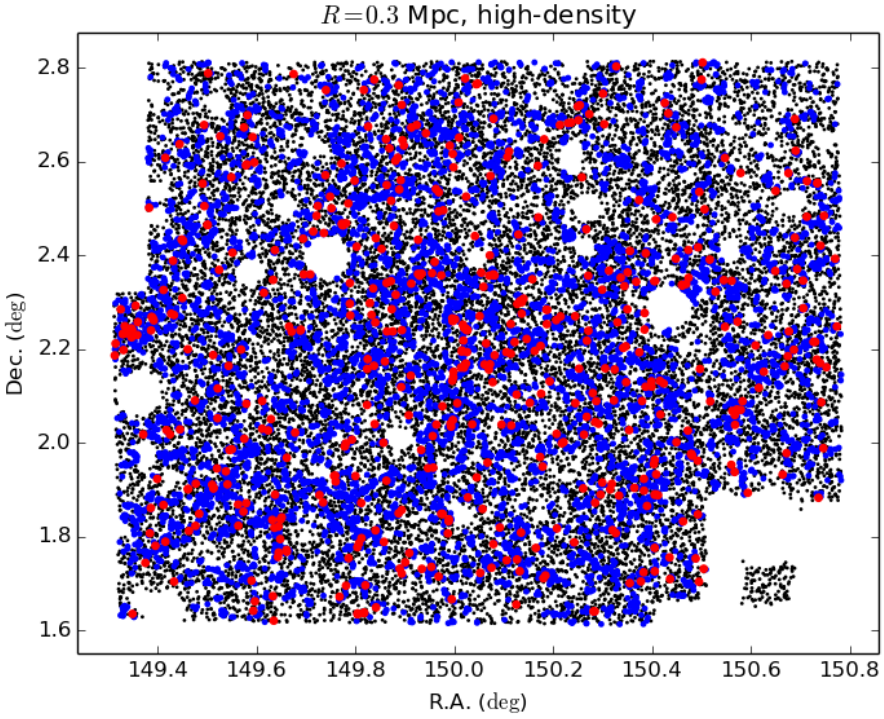
shows the UltraVISTA sky field in a representative high-redshift bin ($1.5 \leqslant z \leqslant 2.0$) for both high-density and low-density environments, distinguishing between quiescent and star-forming galaxies. Only the fixed aperture radius $R = 0.3$ Mpc is shown, as an example. It can be seen that the fixed aperture method implemented is able to identify galaxies in different environments. Galaxies belonging to high-density environments tend to be more clustered, while low-density galaxies appear spatially distributed in a more uniform fashion.

Interestingly, it can be seen how, although rare at this redshift, quiescent galaxies tend to be slightly more numerous in the high-density regions compared to the low-density ones. This trend can be expressed quantitatively by looking at the fraction of quiescent galaxies as a function of environment, redshift and mass (shown in Figure 34 for two representative radii, namely $R = 0.3$ Mpc and $R = 2$ Mpc). As expected the fraction of quiescent galaxies increases with cosmic time in both environments. Nevertheless, these fractions show how quiescent galaxies are more numerous in high-density environments compared to low-density ones as a function of mass. Although the difference between D_{75} and D_{25} is a function of mass and redshift, it remains well visible up to $z \sim 2$ for both the $R = 0.3$ Mpc and the $R = 2$ Mpc case. For masses $\sim 10^{11} M_{\odot}$, at $z \sim 0.5$ 60% of the galaxies in high-density environments are quiescent, while only 40% in low-density environments. At $z \sim 1$ the difference is reduced to $\lesssim 10\%$, but it is still visible. In the $R = 2$ Mpc case, differences of $\sim 10\%$ at $z \sim 0.5$ are reduced to $\sim 5\%$ at $z \sim 1$. In the same figure also the fractions of quiescent galaxies in different environments for the zCOSMOS Survey (Bolzonella et al., 2010, see their figure 8) are reported. A comparison between this work and Bolzonella et al. (2010) will be performed in Section 4.5.

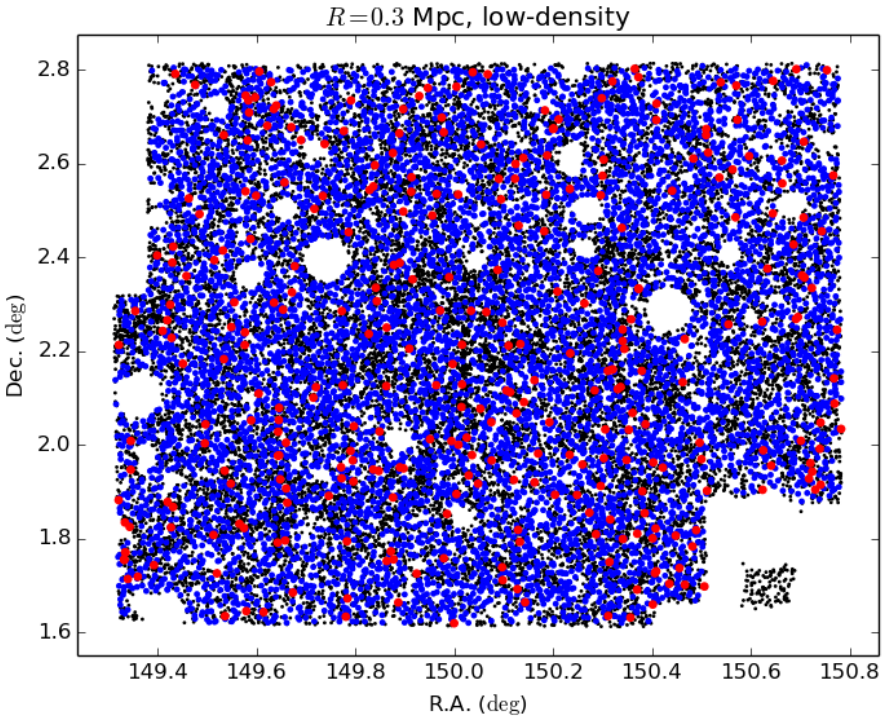
4.3.1 *The GSMF of the UltraVISTA sample*

Figure 35 shows the GSMF for all UltraVISTA galaxies. In the same figure also the GSMF for high-density and low-density environments are shown, in the case of $R = 0.3$ Mpc. GSMFs for other radii are not reported for the sake of clarity, although when performing the analysis on the shape of the GSMF in different environments all apertures will be considered. It can be seen how the GSMFs of high-density and low-density environments are different. The high-mass end of the GSMF (above $M = 10^{10.5-11} M_{\odot}$) is enhanced in the case of high-density environments, in comparison to low-density ones, while the low mass end is depleted. This difference can be appreciated up to $z \sim 2$, where no more differences can be seen between high-density and low-density environments.

Dividing the galaxy population into quiescent and star-forming galaxies, it is possible to see how the difference between high- and

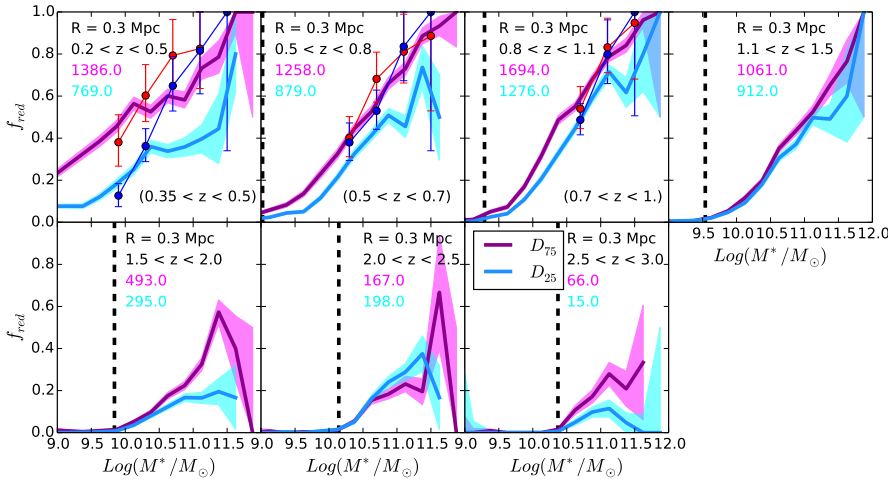


(a)

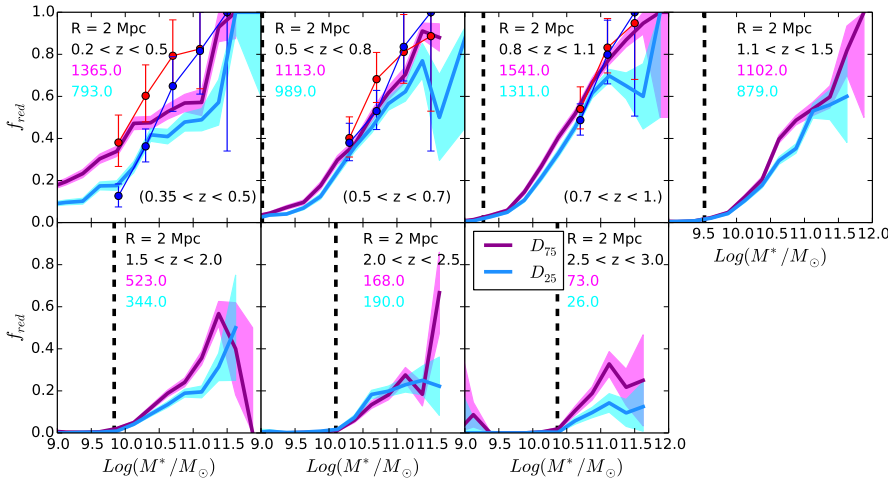


(b)

Figure 33: *Sky maps of the UltraVISTA field*. Black dots represent the total UltraVISTA sample, red points are quiescent galaxies, blue points are star-forming galaxies. (a): high-density environments, (b): low-density environments. Only the case for $R = 0.3$ Mpc, $z \in [1.5, 2.0]$ is shown.



(a)



(b)

Figure 34: *Fraction of quiescent galaxies*. The magenta line refers to high-density environments, the cyan line to low-density environments. The shaded regions correspond to the propagated errors on the fraction. For each redshift bin, the number of galaxies above the mass limit in the two environments (magenta for high-density environments and cyan for low-density environments) are reported. The vertical black dashed line corresponds to the mass completeness limit. (a): R = 0.3 Mpc. (b): R = 2 Mpc. In the first three redshift bins, the fractions of quiescent galaxies in high- and low-density environments from the work by Bolzonella et al. (2010, see their figure 8) are reported for comparison above the mass completeness limit. Red points and lines refer to high-density environments, blue points and lines to low-density environments. The redshift bins in which the fractions of Bolzonella et al. (2010) have been calculated are reported in parentheses in the bottom right corners of the plots.

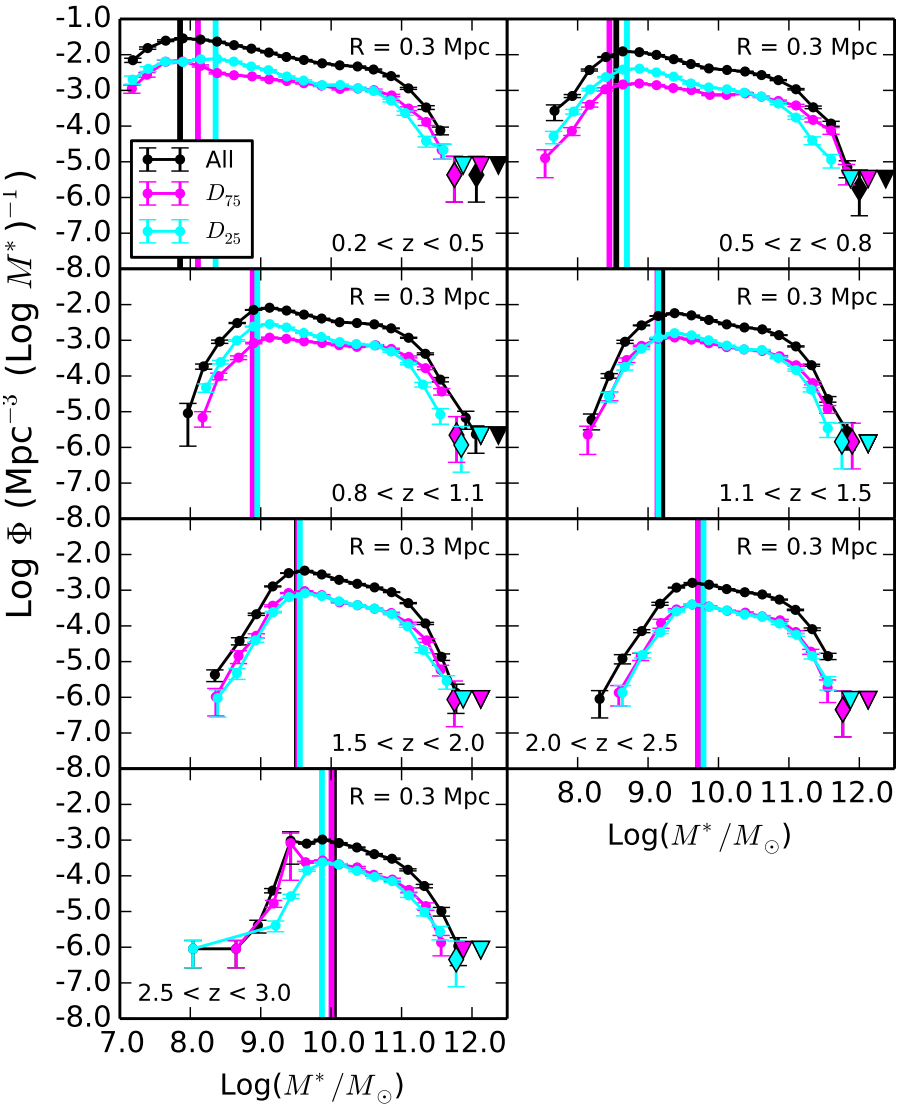


Figure 35: *GSMF of UltraVISTA galaxies - all galaxies*. The black curve refers to the total GSMF for all UltraVISTA galaxies, the magenta curve refers to high-density environments and the cyan curve refers to low-density environments. Vertical lines are the mass completeness limits, colour-coded as the corresponding GSMF. Diamonds represent mass bins with only one galaxy, triangles are upper limits for mass bins with zero galaxies. Error bars represent Poissonian errors.

low-density environments affects different parts of the GSMF in the case of the quiescent galaxy population (Figure 36) and in the case of the star-forming galaxy population (Figure 37). For quiescent galaxies, the enhancement of the high-mass end in high-density environments is visible in comparison to low-density environments up to $z \sim 2$. For the star-forming population, instead, the difference is mainly present at low masses (below $10^{11} M_{\odot}$) and at lower redshifts (below $z \sim 1.5$).

A more quantitative analysis on differences between high- and low-density GSMFs for the different galaxy populations can be performed by taking the ratios of the high-density to the low-density GSMF for the total galaxy population, quiescent and star-forming galaxies as a function of redshift (Figure 38). As stated in the previous section, in the quiescent and star-forming case the ratios are calculated using the quiescent and star-forming component of the total GSMF in high- and low-density environments. For this reason, the ratio can be greater than 1. It can be seen how the ratio of the high-density to low-density GSMF is typically higher in the case of quiescent galaxies compared to star-forming ones, at least up to $z \sim 2$ for both the $R = 0.3$ Mpc and the $R = 2$ Mpc case. The ratio of high-density to low-density GSMF is generally $\gtrsim 1$ for quiescent galaxies and it is generally $\lesssim 1$ for star-forming galaxies. This can be interpreted as quiescent galaxies being more represented in high-density environments and star-forming galaxies being more present in low-density environments. These ratios also show a trend with mass, both for quiescent and star-forming galaxies. High-density environments are dominated by a more massive galaxy population, and this is generally true for both quiescent and star-forming galaxies. Instead the ratio of high-density to low-density GSMFs for the total galaxy population follows the same ratio of star-forming galaxies at low masses and the one of quiescent galaxies at high masses, as expected.

4.3.2 *The shape of the GSMF in different environments*

Differences between the shape of the GSMF in high-density and low-density environments can be better analysed by taking the ratio of the high-mass end to the intermediate-mass range of the GSMF. In particular, the quantity

$$\log \frac{\int_{\log(M)\in[11,11.5]} \Phi(M) dM}{\int_{\log(M)\in[10,10.5]} \Phi(M) dM} \tag{25}$$

was calculated for both quiescent and star-forming galaxies in both high-density and low-density environments (shown in Figure 39). Upper limits due to mass bins with zero galaxies were not included in the computation, but mass bins with only one count have been considered.

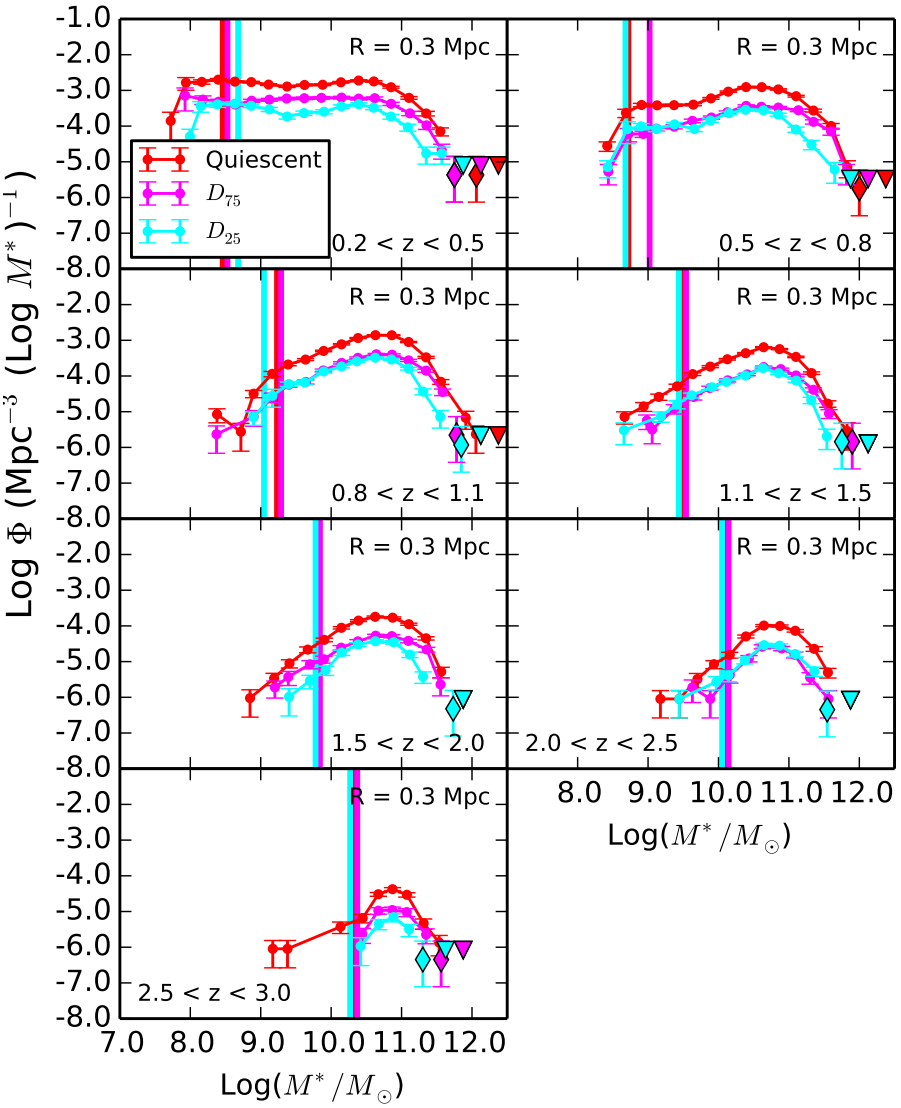


Figure 36: *GSMF of UltraVISTA galaxies - quiescent galaxies*. The red curve refers to the total GSMF for quiescent galaxies only, the magenta curve refers to high-density environments and the cyan curve refers to low-density environments. Vertical lines are the mass completeness limits, colour coded as the corresponding GSMF. Diamonds represent mass bins with only one galaxy, triangles are upper limits for mass bins with zero galaxies. Error bars represent Poissonian errors.

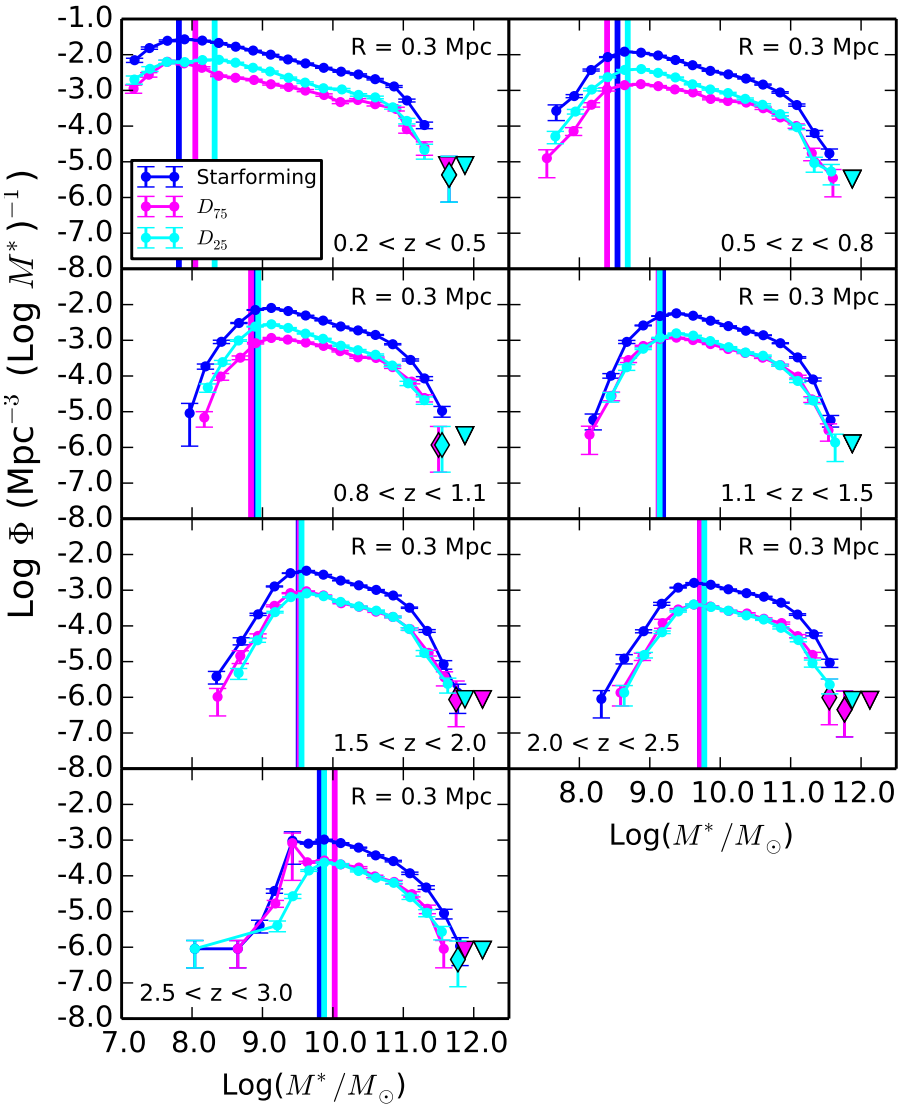
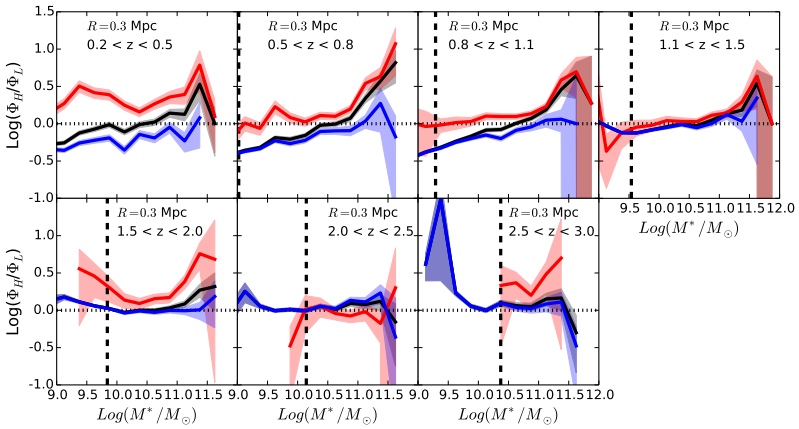
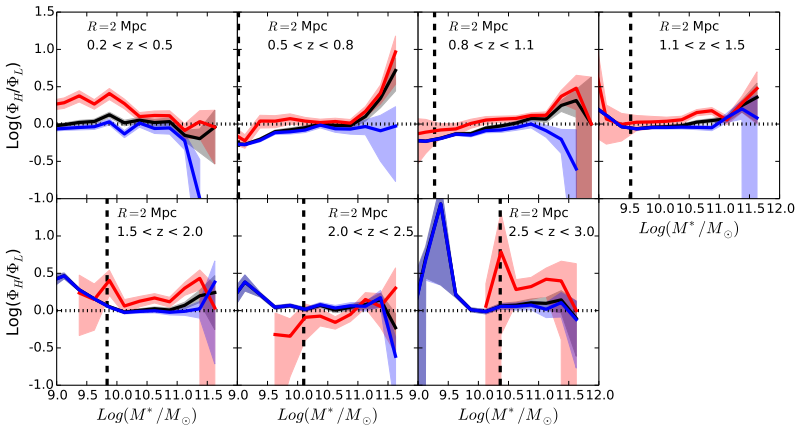


Figure 37: *GSMF of UltraVISTA galaxies - star-forming galaxies*. The blue curve refers to the total GSMF for star-forming galaxies only, magenta curve refers to high-density environments and cyan curve refers to low-density environments. Vertical lines are the mass completeness limits, colour-coded as the corresponding GSMF. Diamonds represent mass bins with only one galaxy, triangles are upper limits for mass bins with zero galaxies. Error bars represent Poissonian errors.



(a)



(b)

Figure 38: *Ratio of high- to low-density GSMFs.* Ratio of the high-density (Φ_H) to the low-density GSMF (Φ_L) as a function of mass and redshift. The black line refers to all galaxies, the red line to quiescent galaxies and the blue line to star-forming galaxies. The shaded regions correspond to the propagated errors on the ratio. The vertical black dashed line corresponds to the mass completeness limit. (a): R = 0.3 Mpc. (b): R = 2 Mpc.

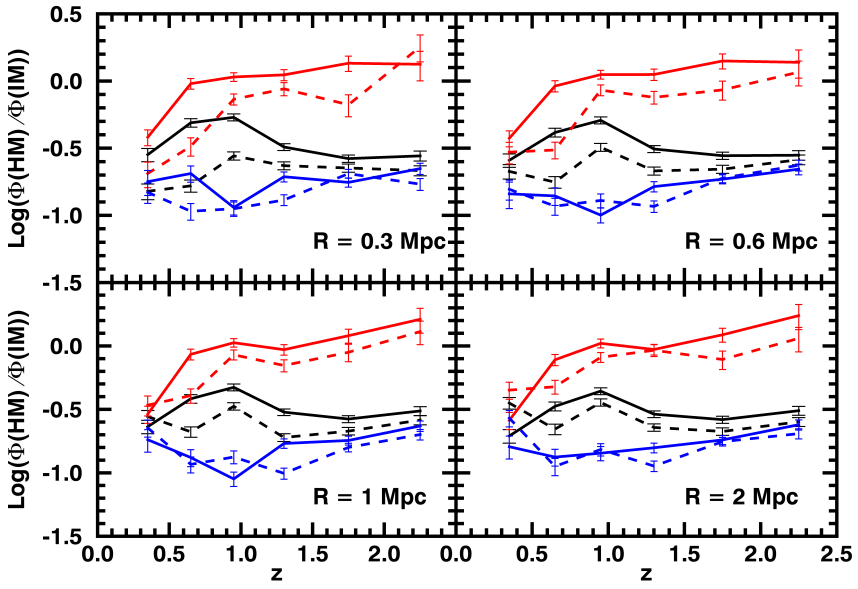


Figure 39: *Shape of the GSMF - high mass*. Ratio of the high-mass to the intermediate-mass end of the GSMF (see Equation (25)) as a function of redshift. Solid lines refer to high-density environments, dashed lines to low-density ones. Red lines represent quiescent galaxies and blue lines star-forming galaxies. In black the ratio for the total galaxy population is reported.

These ratios clearly show how the difference between high-density and low-density environments is present only for quiescent galaxies, as for star-forming galaxies no distinction and no trend in the ratios can be found at any redshift. For quiescent galaxies, the ratio of the high-mass to the intermediate-mass end of the GSMF is higher in high-density environments compared to low-density ones. This ratio also shows a trend with redshift, monotonically increasing up to $z \sim 2$ (for high-density, quiescent galaxies the trend with redshift is more evident for the $R = 1 - 2$ Mpc case). This reflects the gradual build-up of the intermediate mass part of the GSMF with cosmic time for the quenched galaxy population, and is in agreement with a scenario in which massive galaxies became passive at earlier times than lower mass galaxies (downsizing). The difference between high-density and low-density environments seems to be present for all radii, with no significant differences among them. It can be quantified to be roughly of $\sim 0.2 - 0.3$ dex.

Conversely, if the ratio of the intermediate-mass range to the low-mass end of the GSMF is taken, differences between D_{75} and D_{25} emerge only for star-forming galaxies at $z \leqslant 1.5$. In particular, it has been calculated the quantity

$$\log \frac{\int_{\log(M)\in[10.5,11]} \Phi(M)\,dM}{\int_{\log(M)\in[9.5,10]} \Phi(M)\,dM} \tag{26}$$

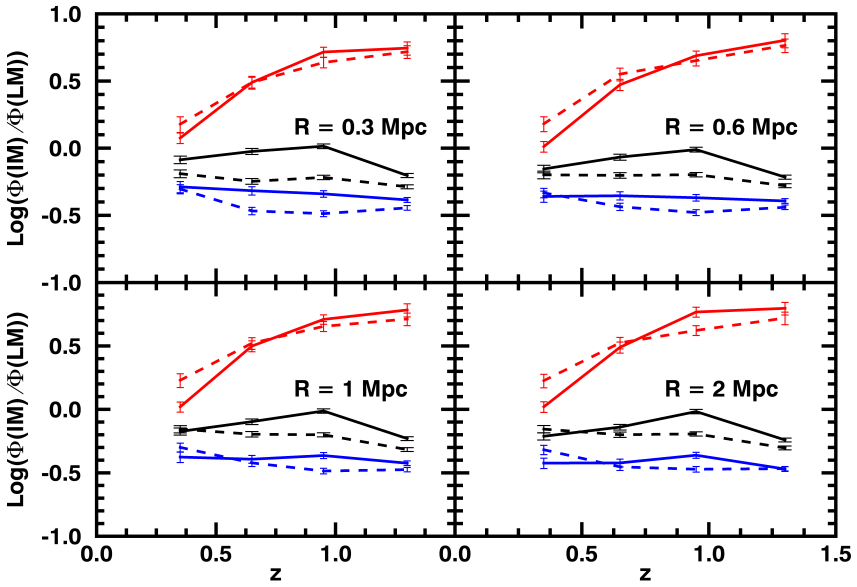


Figure 40: *Shape of the GSMF - low mass*. Ratio of the intermediate-mass to the low-mass end of the GSMF (see Equation (26)) as a function of redshift. Solid lines refer to high-density environments, dashed lines to low-density ones. Red lines represent quiescent galaxies and blue lines star-forming galaxies. In black the ratio for the total galaxy population is reported.

for both quiescent and star-forming galaxies in both high-density and low-density environments (shown in Figure 40). Again, upper limits due to mass bins with zero galaxies were not included in the computation, but mass bins with only one count have been considered. This figure clearly shows how a difference between high- and low-density environments is present only for star-forming galaxies up to $z \sim 1 - 1.5$ and not for quiescent galaxies. The ratio is < 1 for star-forming galaxies, and it is smaller in low-density environments by ~ 0.2 dex for environments measured with a fixed aperture radius of $R = 0.3$ Mpc. The difference between high- and low-density environments seems to get smaller increasing the fixed aperture radius. Nevertheless, an indication of low-mass star-forming galaxies being more present in low-density environments seems to be visible from the data. The monotonic trend with redshift of the intermediate-mass to low-mass ratio for the quiescent galaxies is an indication of a progressive steepening of the low-mass end of the quiescent GSMF with redshift (this seems to be at variance with what hinted by Davidzon et al., 2016, although the different redshift range explored, mass completeness limit and environmental definition do not allow any firm conclusion from the comparison). Nevertheless, this monotonic trend with cosmic time is not observed for the star-forming galaxy population, and this can again be related to the gradual build-up of galaxy mass with cosmic time, in a complementary way to what found be-

fore with the high- to intermediate-mass end ratios. The low-mass end of the quiescent GSMF is gradually enhanced as more low mass galaxies are quenched with cosmic time, while the low-mass end of the star-forming GSMF is continuously replenished by galaxies that increase their stellar mass through ongoing star-formation activity. This result is in agreement also with what found by Pozzetti et al. (2010, see their Figure 14) and Ilbert et al. (2013, see their Figure 6).

4.3.3 *The relative importance of quiescent and star-forming GSMF in different environments*

Figure 41 shows the mass at which the quiescent and the star-forming GSMFs intersect (M_{cross}) as a function of redshift and environment (i. e. the mass above which the GSMF is dominated by the quiescent population). It can be seen how M_{cross} is higher in low-density environments compared to high-density ones up to redshift $z \sim 1.5$, where the two curves become indistinguishable. This is in agreement with the current paradigm of galaxy evolution, which predicts that massive galaxies became quiescent at earlier times compared to less massive galaxies. Therefore, as redshift increases, the mass at which the quiescent GSMF starts to dominate over the star-forming GSMF increases as well. The fact that M_{cross} is higher in low-density environments compared to high-density ones is an evidence of the fact that the processes that lead to the quenching of the star-formation and to the transformation of star-forming galaxies into quiescent galaxies are more efficient in high-density environments, leading to less massive galaxies being already quenched, while at the same redshift, in low-density environments, they will still be star-forming. In high-density environments M_{cross} is a monotonically increasing function of redshift, increasing from $\sim 10^{10} M_{\odot}$ at $z \sim 0.5$ to $10^{11.5} M_{\odot}$ at $z \sim 2$. If upper limits to the value of M_{cross} derived when GSMFs do not intersect are considered, then an increase of M_{cross} as a function of redshift is roughly true also for low-density environments for redshifts $z \gtrsim 1$, while at lower redshifts Figure 41 shows an upturn in the value of M_{cross}. This upturn seems to become less evident going from $R = 0.3$ Mpc to $R = 2$ Mpc. The upturn at low redshifts of the M_{cross} in low-density environments is probably due to the lowest density environments probed by the fixed aperture method, especially on small scales (e. g. $R = 0.3$ Mpc). In such underdense environments quiescent galaxies are rare, as the fraction of quiescent galaxies is higher in high-density environments (see e. g. Figure 34). Therefore, the quiescent and star-forming GSMF will be more separated, especially at high masses, and the M_{cross} results higher.

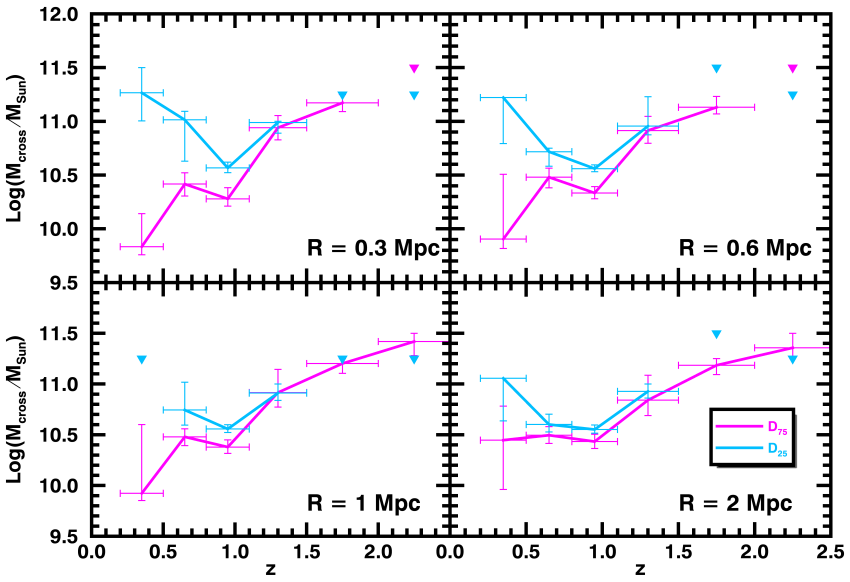


Figure 41: M_{cross}. Mass at which the star-forming and the quiescent GSMFs intersect. Magenta lines and points refer to high-density environments, while cyan lines and points refer to low-density environments. Triangles are upper limits to the value of M_{cross} in the case when the two GSMFs do not intersect, defined as the mass at which the two GSMFs are closer to each other.

4.3.4 *A test on the effect of photometric redshift uncertainties*

As discussed in Chapter 3, photometric redshift uncertainties are a major limitation in the reconstruction of the density field. Nevertheless, it is still possible to perform a study of the GSMF in different environments using photometric redshifts, provided that their uncertainty is small ($\sigma_{\Delta z/(1+z)} \lesssim 0.01$). In this case, differences between high- and low-density environments that are present in the GSMF calculated using each galaxy's true redshift up to $z \sim 2.5$ result damped when using photometric redshifts, but they will still be recovered. Following what shown in Chapter 3 and as explained in Section 4.1, for this work an uncertainty value for the photometric redshifts of $\sigma_{\Delta z/(1+z)} = 0.01$ has been chosen, which may be, nevertheless, underestimated at high redshift. Moreover, the photometric redshift uncertainty depends on K_S-band magnitude, as shown, for example, in Figure 2 of Scoville et al. (2013, reported in Figure 42 for reference). This figure shows the photometric redshift uncertainty as a function of K_S-band magnitude and redshift, together with the median K_S-band magnitude of a sample of galaxies extracted from the UltraVISTA survey and similar to the one used in this Chapter. It can be seen how $\sigma_{\Delta z/(1+z)}$ seems to evolve from a value of 1% at $z \leqslant 1.5$ with increasing redshift, reaching values of $\sim 3\%$ at $z > 1.5$. For this reason, the effect of a larger uncertainty has also been tested.

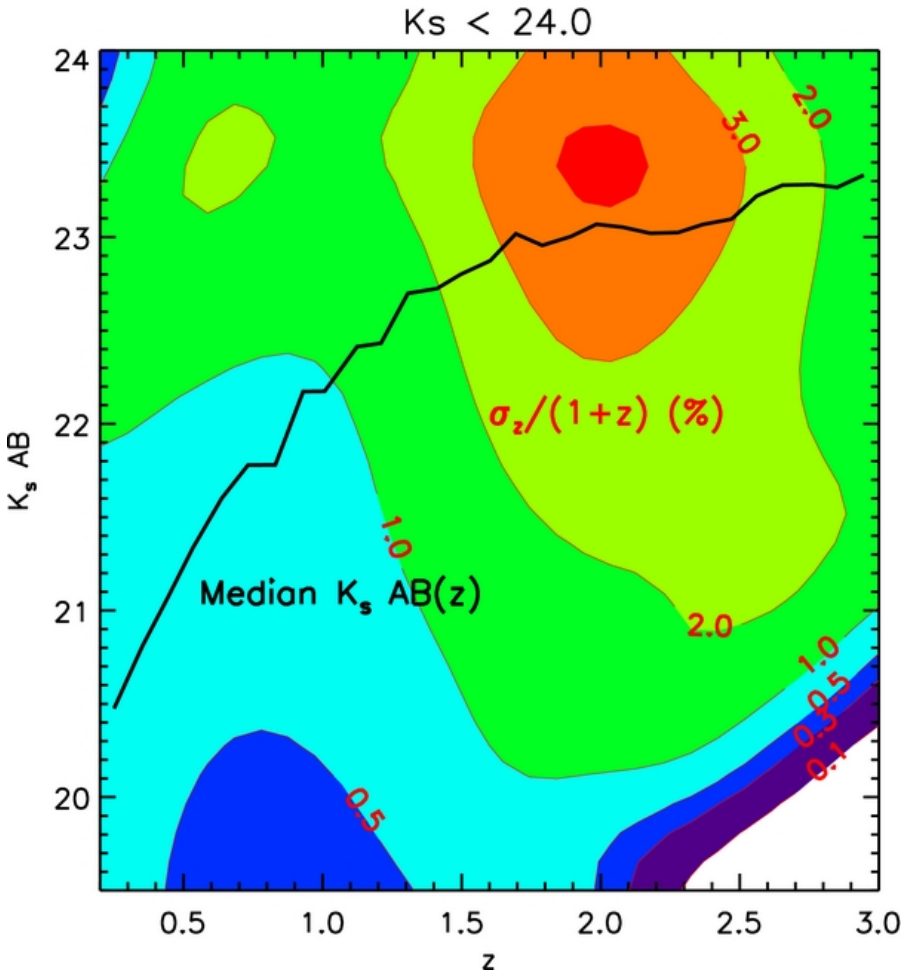


Figure 42: Map of the photometric redshift uncertainty for the UltraVISTA survey as a function of redshift and K_S-band magnitude. The black line is the median K_S-band magnitude for the UltraVISTA galaxies. Figure from Scoville et al. (2013).

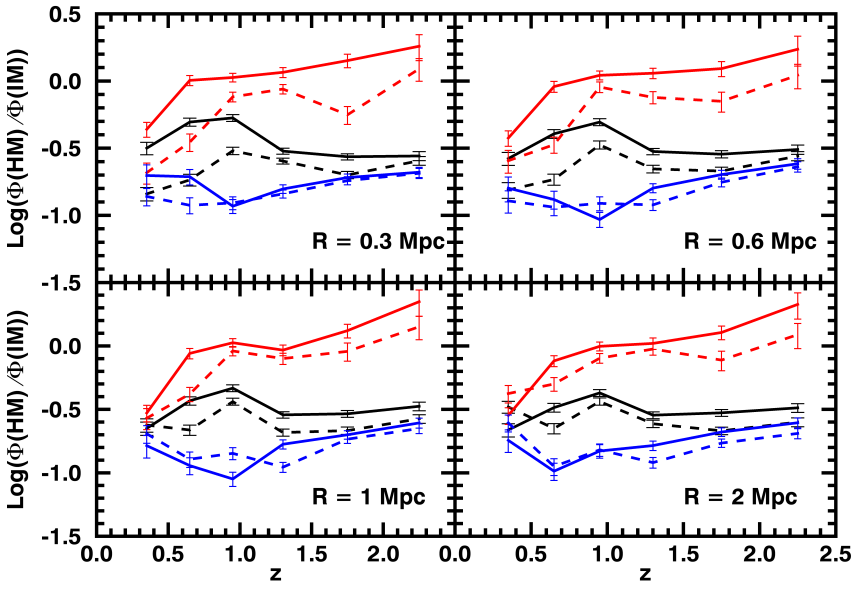


Figure 43: *Shape of the GSMF - increased errors*. Ratio of the high-mass to the intermediate-mass end of the GSMF (see Equation (25)) as a function of redshift. Solid lines refer to high-density environments, dashed lines to low-density ones. Red lines represent quiescent galaxies and blue lines star-forming galaxies. In black the ratio for the total galaxy population is reported.

Following Figure 42 the results have been rederived assuming a photometric redshift uncertainty of

$$\sigma_{\Delta z/(1+z)} = \begin{cases} 0.01 & \text{for } z \leqslant 1.5 \\ 0.03 & \text{for } z > 1.5 \end{cases} \tag{27}$$

As the main purpose of this work is to study the shape of the GSMF in different environments, it has been tested whether the differences that are seen between the GSMFs in high- and low-density environments (Figures 39 and 40) are maintained when considering larger errors at higher redshift. Because in the case of the ratio between the intermediate-mass and low-mass end of the GSMF the analysis is limited at $z \leqslant 1.5$, in Figure 43 it is reported for comparison only the ratio of the high-mass to the intermediate-mass end of the GSMF (see Figure 39), performed with the higher photometric redshift uncertainty value at high redshift. It can be seen how, even with larger photometric redshift errors, the trends are maintained, without significant differences. The increase in the redshift error affects only the analysis at $z \geqslant 1.5$ and in a negligible way.

4.4 DISCUSSION

Galaxies evolve in parallel with cosmic structures. As galaxies form, so do galaxy clusters, groups and the LSS and transformations in galaxy properties happen at the same time as changes in their local and global environment. It is therefore expected some correlation between galaxy environment and galaxy properties, as a function of redshift. The current understanding of the effect of the environment on galaxy evolution is that environment plays a role in determining the cease of star formation in galaxies and in causing their transformation from blue, actively star-forming, disc-like objects to red, quiescent spheroidal systems (see Chapter 1).

This picture of galaxy formation in relation to environmental effects is supported by evidence both on the theoretical and the observational sides. For example, many mechanisms correlated to galaxy environment have been proposed to end the star formation in a galaxy (see Section 1.3) and many correlations have been found between the main observables and the density field (see Section 1.2). In this Chapter observational evidence is presented of the presence of a complex interplay among galaxy mass, star-formation activity (or lack thereof) and local environment.

Although relying only on photometric redshifts, the work described in this Chapter is able to recover with good accuracy the environmental trends of the GSMF, by making use of a method that has been fully tested on mock galaxy catalogues. In Chapter 3 it has been found that photometric redshifts have the only effect of damping differences between the GSMFs in high- and low-density environments, not introducing any spurious effect.

It is found that the galaxy population is different in the various environments. High-density environments are populated by a higher fraction of quiescent galaxies and this distinction is particularly visible at high masses ($M \gtrsim 10^{11} M_{\odot}$) up to redshift $z \sim 2$. The lack of differences between the fraction of quiescent galaxies in different environments at higher redshifts is probably due to the fact that at higher redshifts structures are at an earlier stage of formation (see e. g. Chiang, Overzier, and Gebhardt, 2013) and quiescent galaxies, even the massive ones, are rarer (while star-forming galaxies still dominate). This evidence is complemented by the total GSMF divided according to local environment, which shows how in high-density environments massive galaxies are more represented up to $z \sim 2$. Therefore, peaks in the density field seem to constitute a particular kind of environment where galaxies are more massive and more quiescent. Environment, therefore, plays a role in shaping the galaxy population and is connected to the build up of galaxy mass and to the end of the star formation. Environmental effects are visible since $z \sim 2$ and on

scales of $R = 2$ Mpc, being effective for a long period of galaxy formation in a strong way.

A particularly interesting scenario for galaxy evolution is the one proposed by Gabor and Davé (2015). In their work, the authors used numerical simulations to investigate the new unified quenching model that they propose. In this model, a galaxy is quenched once the gas in its host halo becomes hot ($T \geqslant 10^{5.4}$ K) and this happens when the host halo reaches a mass of $10^{12} M_\odot$ (roughly corresponding to a stellar mass of $10^{10.5} M_\odot$). In this scenario, both "mass quenching" and "environmental quenching" (Peng et al., 2010) are seen as separate evidences of the same underlying quenching mechanism due to the presence of hot gas. This theoretical model can be used to give an interpretation of the results presented here. In particular, according to this model, the galaxies populating the high-mass end of the GSMF ($M \geqslant 10^{11} M_\odot$) are being quenched because they live in hot gas dominated haloes. Although they are found also in low-density environments (see Figure 6 of Gabor and Davé, 2015), massive haloes ($M \geqslant 10^{12} M_\odot$) are found preferentially in high-density environments. This is the cause of the difference between high and low-density environments seen in the GSMF of quiescent galaxies at masses $M \geqslant 10^{11} M_\odot$.

This difference is not seen in the GSMF of star-forming galaxies because massive galaxies in high-density environments are quenched, therefore they are not included in the star-forming GSMF. This goes in the direction of diluting the signal of potential differences in the high-mass end of the GSMF of star-forming galaxies as a function of environment. Instead, a difference is visible at low masses, with low-mass star-forming galaxies being more present in low-density environments. This is due to the fact that these galaxies live in too low-mass haloes to develop a hot gas environment and be quenched. Nevertheless, those living in high-density environments can still be quenched as satellites of more massive galaxies that live in hot gas dominated haloes. Therefore, the low-mass end of the star-forming GSMF is depleted in high-density environments compared to low-density ones. Interestingly, this trend should reflect on a difference in the low-mass end of the GSMF of quiescent galaxies in high-density environments, which seem to be absent in the data used for this work. This lack of a difference between the high- and low-density, low-mass end of the quiescent GSMF could be due to uncertainties in the photometric redshift calculation or in the distinction between quiescent and star-forming galaxies using the colour-colour diagram. Star-forming galaxies being the majority of the sample, a difference in the low-mass end of the GSMF can be recovered for them, but not for quiescent galaxies, which may suffer from residual contamination from star-forming galaxies at low masses. Nevertheless, a more accurate

analysis, with more precise redshifts and a larger data set has to be performed to solve the problem.

4.5 COMPARISON WITH PREVIOUS STUDIES

Thanks to its excellent combination of multiwavelength coverage, fairly large area, and availability of data sets with a high statistical power, the UltraVISTA-COSMOS is a perfect field where to perform studies of galaxy evolution. For this reason it has been deeply exploited in several works. As the GSMF is a very powerful tool to study the galaxy formation from a statistical point of view, many studies have investigated its relation with galaxy environment. In this section some of the main works that studied the GSMF in different environments are reviewed, from low redshifts (using spectroscopic samples) up to high redshift (using photometric redshift surveys).

The main spectroscopic survey that has been performed in the COSMOS field is the zCOSMOS Survey (see Lilly et al., 2007). Using the 10k spectroscopic sample of Lilly et al. (2009) in synergy with the COSMOS photometric sample (Capak et al., 2007; McCracken et al., 2010) and the environmental estimate by Kovač et al. (2010), Bolzonella et al. (2010) performed a thorough study of the GSMF in different environments up to $z = 1$. In their work, they found a difference between the GSMF of high- and low-density environments, with the massive end of the GSMF being more enhanced in high-density environments. This is qualitatively in agreement with the results of this Chapter in the common redshift range. In Figures 44 and 45 it is shown a comparison between the GSMF of this Chapter and those of Bolzonella et al. (2010, see their figure 5) for two common redshift bins. Quiescent and star-forming GSMFs both in high- and low-density environments have been compared. GSMFs have been normalised so to be equal at a given mass (that of the lowest mass bins considered for VIPERS GSMF, see below), in order to compare their shape. With the exception of quiescent galaxies in the range $0.8 \leqslant z \leqslant 1.1$ (where the GSMF of both high- and low- density environments are in good agreement), the GSMF of Bolzonella et al. (2010) and those of this work show a slightly different shape, with the GSMF of Bolzonella et al. (2010) displaying a steeper slope in the high-mass and/or low-mass end.

The steeper slope of the zCOSMOS mass functions could be due to the different environment estimator used in this Chapter (fixed aperture with $R = 0.3 - 2$ Mpc) and in Bolzonella et al. (2010) (distance to the 5th nearest neighbour), to the different definition of quiescent and star-forming galaxies (photometric type coming from SED fitting estimate for Bolzonella et al. (2010) *vs* restframe colour-colour classification for this thesis) or to the fact that Bolzonella et al. (2010) use spectroscopic redshift while photometric redshifts are used here.

Thanks to the higher statistics and lower mass limit compared to Bolzonella et al. (2010), a difference between high- and low-density environments is seen for both passive and star-forming galaxies.

A difference between this thesis and Bolzonella et al. (2010) is confirmed also by the fraction of quiescent galaxies in different environments, shown in Figure 34. In particular, while the fraction of quiescent galaxies in both environments is comparable to the one derived in this thesis for the lowest mass bins considered by Bolzonella et al. (2010), their fractions become quickly higher than those of this work with increasing mass and especially in low-density environments. This is likely due to the different environmental estimator used in this work and in Bolzonella et al. (2010) and it is likely the origin of the discrepancy observed in the values of M_{cross} between this chapter and Bolzonella et al. (2010, see their figure 7). In fact, while the value of M_{cross} in high-density environments is in agreement between this chapter and Bolzonella et al. (2010), here it is found an upturn in the values of M_{cross} in low-density environments which is totally absent in the work of Bolzonella et al. (2010).

Although a quantitative comparison is difficult, the range of densities corresponding to low-density environments explored in this chapter is much lower than that in Bolzonella et al. (2010). Therefore, in such environments the quiescent galaxy population will be more under-represented and the corresponding quiescent GSMF will result depressed with respect to the star-forming GSMF, with the corresponding value of M_{cross} increased, as confirmed by the different fraction of quiescent galaxies.

In Figures 44 and 45 a comparison between the GSMFs of this thesis and those of Davidzon et al. (2016, see their figure 4) is also reported. Although with a different spectroscopic data set (VIPERS Survey, Garilli et al., 2014; Guzzo et al., 2014), Davidzon et al. (2016) performed a thorough study of the GSMF in different environments at $z \lesssim 1$, finding consistent results with Bolzonella et al. (2010). Although their mass completeness limit allowed only a characterisation of the high-mass end of the GSMF, still their GSMFs for quiescent and star-forming galaxies in different environments are consistent within errorbars with those of this thesis in the overlapping redshift bins.

As the UltraVISTA sample offers high-quality photometric redshifts for a large statistical sample, a few works have been performed at redshift $z > 1$ in the COSMOS field. Two recent works have explored the dependence of the GSMF on the environment using the same sample as it has been done in this work. Both Scoville et al. (2013) and Darvish et al. (2015) used a 2D Voronoi tessellation performed in subsequent redshift slices to study how the galaxy population and the GSMF change in different environments. In particular, Darvish et al. (2015) found a strong evidence for massive ($M > 10^{11} M_{\odot}$), quiescent galaxies showing an increasingly important difference between

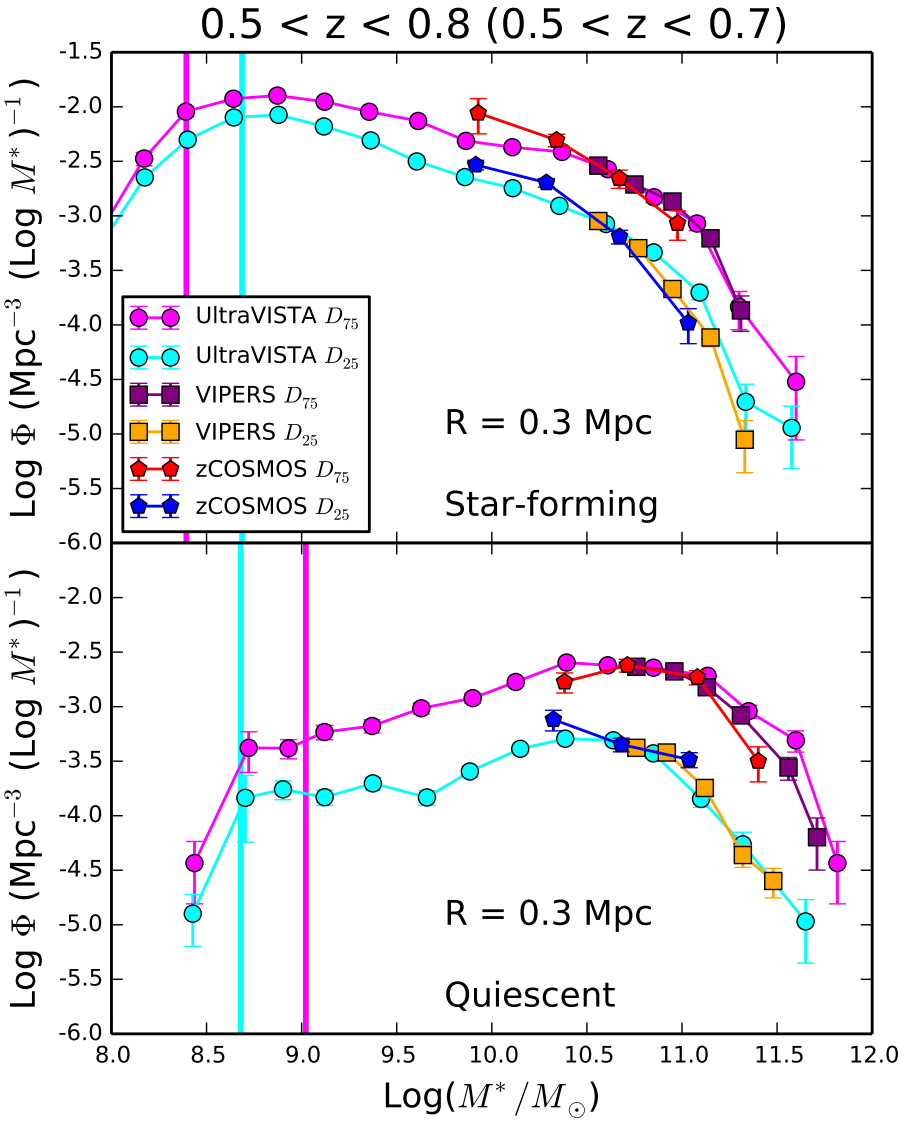


Figure 44: *Comparison with zCOSMOS and VIPERS GSMFs* - $0.5 \leqslant z \leqslant 0.8$. Magenta and cyan circles correspond to UltraVISTA GSMF (this thesis, R = 0.3 Mpc, high-density and low-density environments, respectively). Red and blue pentagons correspond to the zCOSMOS GSMF (Bolzonella et al., 2010, see their figure 5, red represents high-density environments and blue represents low density environments). Purple and orange squares correspond to the VIPERS GSMF (Davidzon et al., 2016, see their figure 4, purple represents high-density environments, orange represents low-density environments). For zCOSMOS and VIPERS GSMF only points above the respective mass completeness limits are shown. Vertical lines represent UltraVISTA mass completeness limits, colour-coded as the corresponding GSMF. Top panel refers to star-forming galaxies, bottom panel to quiescent galaxies. The redshift bin in which zCOSMOS GSMF have been calculated is reported in parentheses on top of the figure, the redshift bin in which VIPERS GSMF have been calculated is $0.65 \leqslant z \leqslant 0.8$.

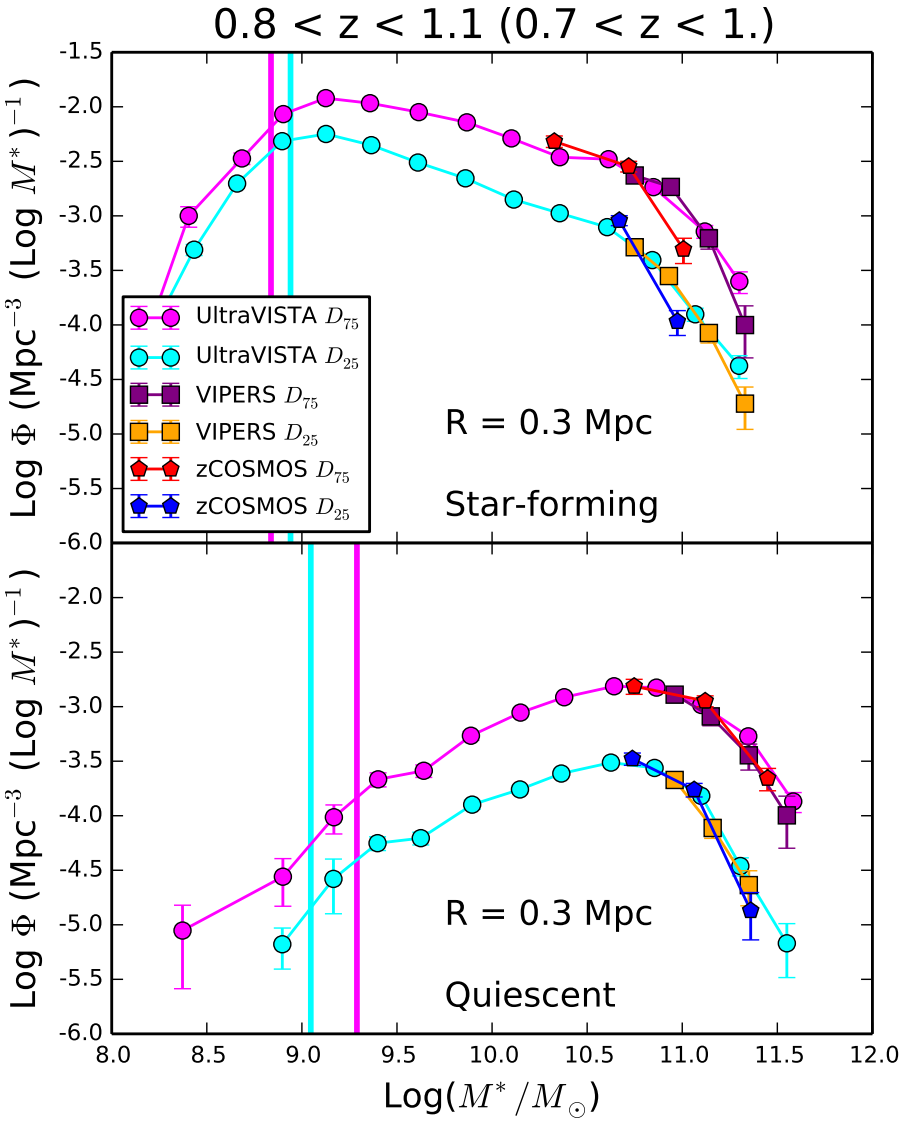


Figure 45: *Comparison with zCOSMOS and VIPERS GSMFs* - $0.8 \leqslant z \leqslant 1.1$. Magenta and cyan circles correspond to UltraVISTA GSMF (this thesis, R = 0.3 Mpc, high-density and low-density environments, respectively). Red and blue pentagons correspond to the zCOSMOS GSMF (Bolzonella et al., 2010, see their figure 5, red represents high-density environments and blue represents low density environments). Purple and orange squares correspond to the VIPERS GSMF (Davidzon et al., 2016, see their figure 4, purple represents high-density environments, orange represents low-density environments). For zCOSMOS and VIPERS GSMF only points above the respective mass completeness limits are shown. Vertical lines represent UltraVISTA mass completeness limits, colour-coded as the corresponding GSMF. Top panel refers to star-forming galaxies, bottom panel to quiescent galaxies. The redshift bin in which zCOSMOS GSMF have been calculated is reported in parentheses on top of the figure, the redshift bin in which VIPERS GSMF have been calculated is $0.8 \leqslant z \leqslant 0.9$.

high- and low-density environments at $z \lesssim 1.5$, in agreement with the results of this thesis, which see environmental effects disappear for quiescent galaxies at $z \sim 2$. Both the work by Darvish et al. (2015) and this work found no environmental effect for massive star-forming galaxies at any redshift. However, with this thesis it has also been possible to extend the analysis to low-mass star-forming galaxies, finding an environmental effect up to $z \lesssim 1.5$.

It is also important to mention that the work presented in this chapter is in agreement with what found in the UKIDSS UDS field (see Mortlock et al., 2015). By using the UKIDSS UDS data set and the CANDELS photometric redshifts (Galametz et al., 2013; Guo et al., 2013), Mortlock et al. (2015) found that the GSMF is different in high- and low-density environments up to $z \sim 1.5$. A comparison between this work and the work by Mortlock et al. (2015) in three high-redshift bins is shown in Figure 46. GSMFs are normalised to be equal at $M = 10^{11} M_{\odot}$, so to be able to compare their shape in a consistent way. It can be seen how the GSMFs of this work and those of Mortlock et al. (2015) are in good agreement except for the last redshift bin ($2.0 < z < 2.5$), where they show a different shape, with the GSMFs of Mortlock et al. (2015) characterised by a steeper slope. This difference in shape at high redshift could be due to the different high- and low-density environment definition (25th and 75th percentile of the volume density distribution in this thesis, 1σ deviation from the mean of the density distribution in Mortlock et al. 2015).

These comparisons show how the excellent UltraVISTA data set allows us to extend previous works done at low redshift with spectroscopic surveys (*e.g.* Bolzonella et al., 2010; Davidzon et al., 2016) and to complement other works performed at high redshift with photometric redshift surveys (*e.g.* Darvish et al., 2015; Mortlock et al., 2015).

4.6 CONCLUSIONS

In this chapter the GSMF and the high-precision photometric redshifts of the UltraVISTA survey (Ilbert et al., 2013; McCracken et al., 2012) have been used to outline a history of the role of local environment in galaxy evolution from $z = 3$ to $z = 0$. Although derived with photometric redshifts the results presented in this Chapter are robust and provide a reliable observational evidence to support theoretical scenarios of galaxy formation. The main findings of this Chapter can be summarised as follows:

1. The fraction of massive quiescent galaxies is higher in high-density environments compared to low density ones. The difference is visible at all explored scales ($R = 0.3 - 2$ Mpc), and it is present up to redshift $z \sim 2$. The fraction of quiescent galaxies increases with mass and decreases with redshift.

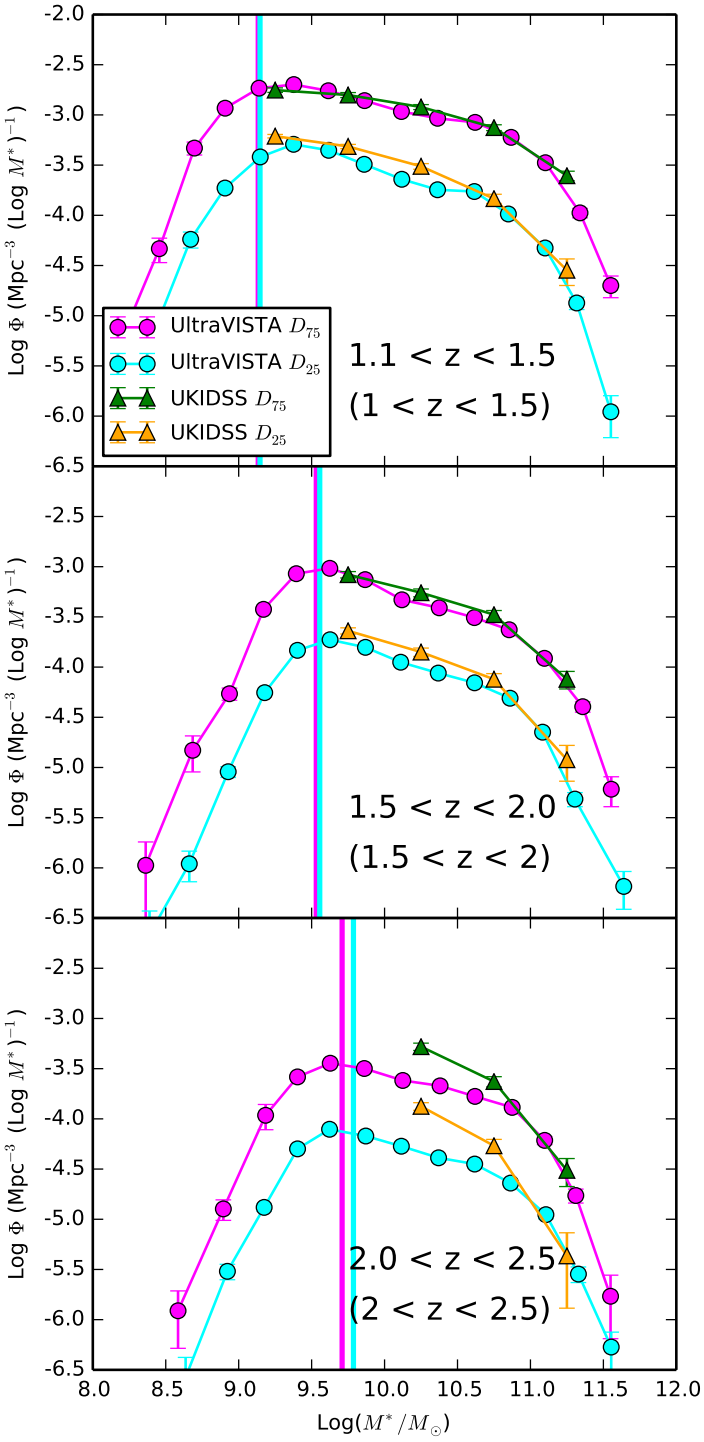


Figure 46: *Comparison with UKIDSS GSMFs.* Magenta and cyan circles refer to the UltraVISTA GSMFs (this thesis, R = 0.3 Mpc, high-density and low-density environments respectively). Green and orange triangles correspond to the UKIDSS-CANDELS GSMFs (Mortlock et al., 2015, see their figure 8, green refers to high-density environments, orange to low-density environments). Vertical lines correspond to the UltraVISTA mass completeness limits, colour-coded as the corresponding GSMFs. For the Mortlock et al. (2015) GSMFs, only points above the mass completeness limit have been considered. In each panel, the redshift bin in which the Mortlock et al. (2015) GSMF have been calculated is reported in parentheses.

2. The shape of the GSMF is different in high-density and low-density environments for the total galaxy population. The high-mass end of the GSMF ($\log(M^*/M_\odot) \in [11, 11.5]$) is enhanced with respect to the intermediate-mass ($\log(M^*/M_\odot) \in [10, 10.5]$ in high-density environments) up to $z \sim 2$.

3. The difference in the shape of the GSMF between high-density and low-density environments is visible for quiescent galaxies up to $z \sim 2$ and at masses $M > 10^{11} M_\odot$ and for star-forming galaxies up to $z \sim 1.5$ and at masses $M < 10^{11} M_\odot$.

4. No environmental effects seem to be present at $z \gtrsim 2$. This could be due to the fact that structures are at an earlier stage of formation and therefore environmental dependencies are not yet in place.

5. The mass at which galaxies become quiescent at a given redshift is lower in high-density environments compared to low-density ones. This effect is visible up to redshift $z \sim 1.5$. In high-density environments, the mass at which the quiescent GSMF starts to dominate over the star-forming GSMF is a monotonically increasing function of redshift.

It has been shown that local environment plays indeed a role in shaping galaxy evolution, in the redshift range $0 \leqslant z \leqslant 2$. High-density environments show an enhanced fraction of massive ($\sim 10^{11} M_\odot$) quiescent galaxies, compared to low-density ones. This work, with the large redshift range explored and the range of scales on which environment has been measured (from 0.3 to 2 Mpc) may help to shed new light on unsolved problems in galaxy evolution, while also representing an example of what could be achieved by future surveys such as Euclid and Wide Field Infra-red Survey Telescope (WFIRST).

5

GALAXY SEGREGATION INSIDE FILAMENTS AT $z \simeq 0.7$

Aside from the role of local environment in galaxy formation and evolution (extensively investigated by many works, see Chapter 1), in recent years global galaxy environment, the LSS, has gained an incremented attention, together with its relationship to the processes governing galaxy evolution, becoming the object of an increasingly systematic investigation. Different methods for recovering the LSS have been developed (see Chapter 2) and the analysis of the CW has transitioned from the study of single objects to a statistical survey of galaxy properties with respect to their position in the LSS.

So far, the impact of global environment on galaxy properties has been investigated only in numerical simulations and in the local Universe, ($z \leqslant 0.3$, e. g. using the SDSS and GAMA surveys). The study of global environment at high redshift is limited by the need of galaxy samples which grant a sufficient statistic and which are characterised by a sufficiently precise redshift information, so to use galaxies as tracers for structures. A characterisation of the CW and a study of the effect of global environment on galaxy evolution have so far been prevented by the lack of adequate data sets and the difficulty of performing large spectroscopic surveys at high redshift.

In this Chapter, the complete sample of the VIPERS survey (performed in the redshift range $0.4 \leqslant z \leqslant 1.2$) has been exploited to detect the filamentary structure of the CW at high redshift and to study the correlation between galaxy properties and their distance to the closest filament. The results reported in this Chapter are fully described in the paper Malavasi et al. (2016c).

This Chapter is structured as follows: the VIPERS survey is described in Section 5.1, while in Section 5.2 the method used to recover the LSS in the survey volume, already introduced in Section 2.2.2, is briefly summarised. Section 5.3 reports tests performed on simulations of the VIPERS survey to assess the method performance given the survey layout and observational properties. The application to the real VIPERS data is performed in Section 5.4, where a significative galaxy mass and type segregation within filaments is also reported. Conclusions are drawn in Section 5.5. The cosmology adopted for the analysis presented in this Chapter is the Planck Collaboration et al. (2015) cosmology, with $H_0 = 67.51$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.3121$, and $\Omega_\Lambda = 0.6879$.

5.1 DATA

The VIMOS Public Extragalactic Redshift Survey (VIPERS)[1] is a spectroscopic galaxy survey, magnitude-limited to $i_{AB} \leqslant 22.5$. It covers an overall area of about $16\,\mathrm{deg}^2$ and $8\,\mathrm{deg}^2$ in the W1 and W4 fields of the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS)-Wide imaging survey, respectively. VIPERS spectra were collected in low resolution mode, $R = 230$, leading to a radial velocity error of $\sigma_v = 175(1+z_{spec})$ km s^{-1}. The spectroscopic targets were pre-selected in a colour-colour space to remove galaxies below $z = 0.5$, which coupled with an optimised observing strategy, provides an average effective sampling rate of about 40%. More details can be found in the survey description papers by Guzzo et al. (2014), Garilli et al. (2014), and Scodeggio et al. (2016).

In this work the final galaxy sample is used, described in the latter paper (the so-called Public Data Release-2). Only the most secure redshifts, corresponding to quality flag $\geqslant 2$ in the VIPERS grading scheme (Confidence Level (CL) $> 97\%$) are considered. The mean number density of galaxies, $\overline{n}(z)$, varies significantly at the redshift boundaries of the survey, due to the magnitude limit, the target sampling rate and the colour selection (see Guzzo et al., 2014; de la Torre et al., 2013, for details). For this reason this analysis is limited to the 50 980 galaxies in the range $0.5 \leqslant z \leqslant 0.85$, where the typical spatial resolution in terms of mean inter-galaxy separation, $\langle D_z \rangle \sim \overline{n}(z)^{-1/3}$, is the highest ($7.7 < \langle D_z/\mathrm{Mpc} \rangle < 10$). These values are comparable with those of the GAMA survey (Driver et al., 2011, with $4.6 < \langle D_z/\mathrm{Mpc} \rangle < 8.8$ for $0.1 < z < 0.3$ and $r \leqslant 19.8$), and make VIPERS the first galaxy redshift survey well suited for studying the CW at high redshift. The stellar masses for the objects in the adopted sample and the classification between active and passive populations were derived according to Moutard et al. (2016b), on the basis of the SED-fitting analysis of the multi-wavelength data collected in the VIPERS regions[2] (Moutard et al., 2016a).

5.2 DETECTING FILAMENTS IN THE VIPERS SURVEY

In order to trace the CW in VIPERS, the Discrete Persistent Structure Extractor (DisPerSE) algorithm (see Sousbie, 2011; Sousbie, Pichon, and Kawahara, 2011, and Chapter 2 for a complete description) was used. DisPerSE identifies filaments as ridges in the density field (calculated using the DTFE). DisPerSE uses the discrete Morse theory to extract critical points, where the gradient of the density field vanishes (e. g. maxima and saddle points), and the field lines connecting them. It then pairs the critical points in topological features, called "critical pairs", us-

1 http://www.vipers.inaf.it

2 The VIPERS-MLS: http://cesam.lam.fr/vipers-mls/

ing the persistent homology theory. The robustness of each feature (including the filaments) is assessed by the relative density contrast of its critical pair, the so-called persistence, which is chosen to pass a certain signal-to-noise (S/N) threshold. The noise level is defined relative to the variance of persistence values obtained from random sets of points. Because DisPerSE is based on a topologically-motivated algorithm, it is both very robust and flexible through the choice of the persistence threshold. Since it filters out the sampling noise, it enables an unsmoothed density field, more noisy but less biased, to be analysed. By construction it is also multi-scale: it builds a network which adapts naturally to the uneven sampling of observed catalogues. The persistence threshold is calibrated on mocks to account for the specific design of VIPERS.

5.3 TESTS ON VIPERS MOCK GALAXY CATALOGUES.

The performance of DisPerSE has been tested on an updated version of the VIPERS mock galaxy catalogues described in de la Torre et al. (2013, 2016) matching the VIPERS final geometry. The parent catalogues include all the galaxies down to the magnitude limit $i_{AB} = 22.5$ together with the selection function at $0.4 < z < 0.6$ due to the VIPERS colour pre-selection. The VIPERS-like catalogues are built from the parent ones with all the observational effects applied (i. e. pointing strategy, Target Sampling Rate (TSR), photometric and spectroscopic masks including gaps between VIsible Multi-Object Spectrograph (VIMOS) quadrants, and random errors on redshift) and perfectly reproduce all the final observational properties of the VIPERS survey. A set of mock galaxy catalogues intermediate between the full parent catalogue and the final VIPERS mocks has also been created, by introducing only one or a limited selection of observational issues at a time (e. g. only the random undersampling due to the TSR, or only the gaps due to the VIMOS quadrants) in order to test specific effects.

The impact of observational biases on the skeleton reconstruction can be assessed by comparing the skeletons obtained from the various mock catalogues. To quantify the differences between two skeletons a pseudo-distance between the two skeletons to be compared is defined. In practice, a skeleton S_a is composed by N_a short straight segments, s_a^i. The pseudo-distance from a skeleton S_a to a skeleton S_b, $D(S_a, S_b)$, is defined as the PDF of the distances between each segment of S_a, s_a^i, and its closest segment in S_b, s_b^j (Sousbie, Colombi, and Pichon, 2009). The distributions $D(S_a, S_b)$ and $D(S_b, S_a)$ are composed by N_a and N_b distances, respectively. There is no reason for the pseudo-distance $D(S_a, S_b)$ to be identical to $D(S_b, S_a)$. Indeed the discrepancy between the two PDFs is related to the differences between the two skeletons.

5.3.1 *Preliminary tests: peculiar velocities, undersampling, gaps*

The first performed test regarded the effect of peculiar velocities on the skeleton reconstruction. Top panel of Figure 47 shows the PDFs of the pseudo-distances obtained by comparing the skeletons S_{rs} and S_{zs}, both measured on the parent mock galaxy catalogue. S_{rs} considers each galaxy's cosmological redshift and S_{zs} introduces errors on the redshift measurement due to the peculiar velocities of galaxies. Both skeletons were measured with a 3σ persistence threshold and the comparison is performed in the redshift range $0.5 < z < 0.85$. It can be seen that both skeletons are reconstructed with the same amount of detail (the two distributions have the same shape). Moreover, the average lengths of the skeletons, defined as the total skeleton length divided by the survey volume (expressed in Mpc/Mpc^3), are also reported in the figure and their values are comparable. An estimate of the uncertainty in the location of the filaments is given by the modes of the PDFs which do not peak at distance $D \sim 0$ (corresponding to a perfect match between the segments of the two skeletons) but $D \sim 1.5$ Mpc. This means that the filaments are slightly displaced from their original position in a systematic way, which is due to the uncertainty in the galaxy positions due to peculiar velocities. Nevertheless, peculiar velocities are not affecting the amount of detail with which the skeleton can be reconstructed and they only introduce a small uncertainty (of the order of $\sim 1-2$ Mpc) in their exact location. As uncertainties in the redshift measurement due to peculiar velocities are intrinsic of redshift surveys, for the rest of this Chapter only comparisons between mock galaxy catalogues with peculiar velocities included in the redshift measurement will be considered.

The middle panel of Figure 47 shows the PDFs of the pseudo-distances obtained by comparing the skeletons S_{zs} and S_u, both measured on the parent mock galaxy catalogue. S_{zs} is defined as above (in the following also referred to as S_{Par} for simplicity, it being measured on the full parent mock catalogue) and S_u introduces in the parent mock galaxy catalogue both peculiar velocities and a further random undersampling by an amount corresponding to the mean TSR ($\sim 50\%$). Two effects can be seen in the plot: first of all, the modes of the distribution are shifted to smaller values than before (the two distributions now peak at $D \sim 1$ Mpc). This means that the uncertainty in the skeleton position introduced by the random undersampling is lower than that introduced by peculiar velocities. Secondly, the two PDFs now have a different shape. The asymmetry between $D(S_{zs}, S_u)$ and $D(S_u, S_{zs})$ reflects the fact that S_{zs} (full sampling) has much more details which have no counterpart in S_u. The random undersampling introduces a loss of detail in the skeleton measurement, and this effect will be discussed in Section 5.3.2 through the use of the persistence threshold.

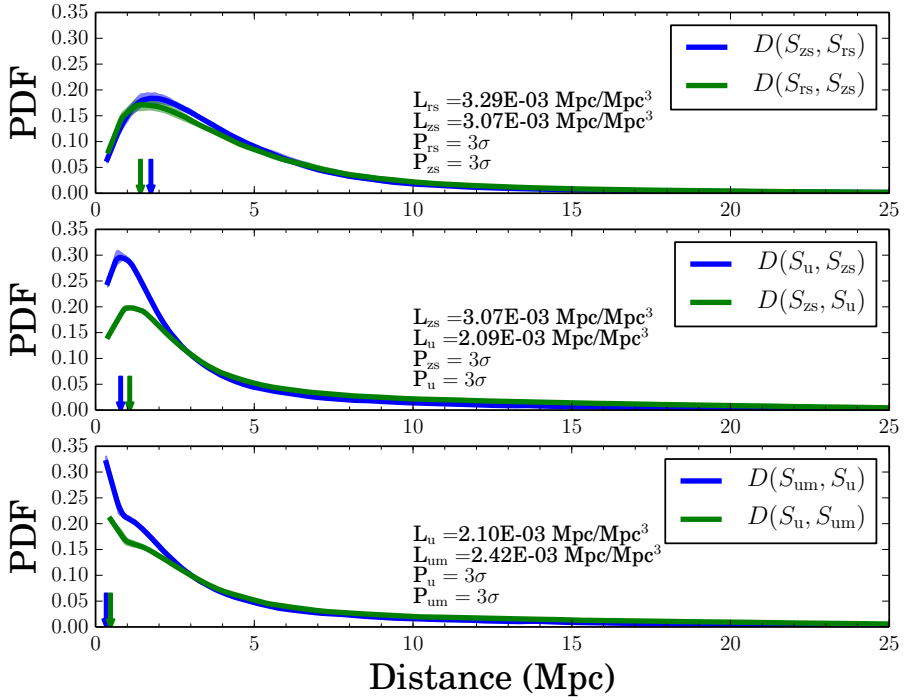


Figure 47: PDFs of the pseudo-distances between the parent skeleton measured using each galaxy's cosmological redshift (S_{rs}), the parent skeleton measured including errors on the redshift measurement due to peculiar velocities (S_{zs}), the parent skeleton with further undersampling by the mean TSR (S_u), and the parent skeleton undersampled by the mean TSR and with the VIPERS photometric and spectroscopic masks applied (S_{um}). Blue lines refer to S_{zs} projected onto S_{rs} (top panel), S_u projected onto S_{zs} (middle panel), S_{um} projected onto S_u (bottom panel). Green lines are the reverse. Solid lines correspond to the mean of 10 mocks and the shaded areas enclose the 1σ variation (please note that uncertainties are negligible). Vertical arrows show the modes of the distributions. All skeletons have been derived with a 3σ persistence threshold.

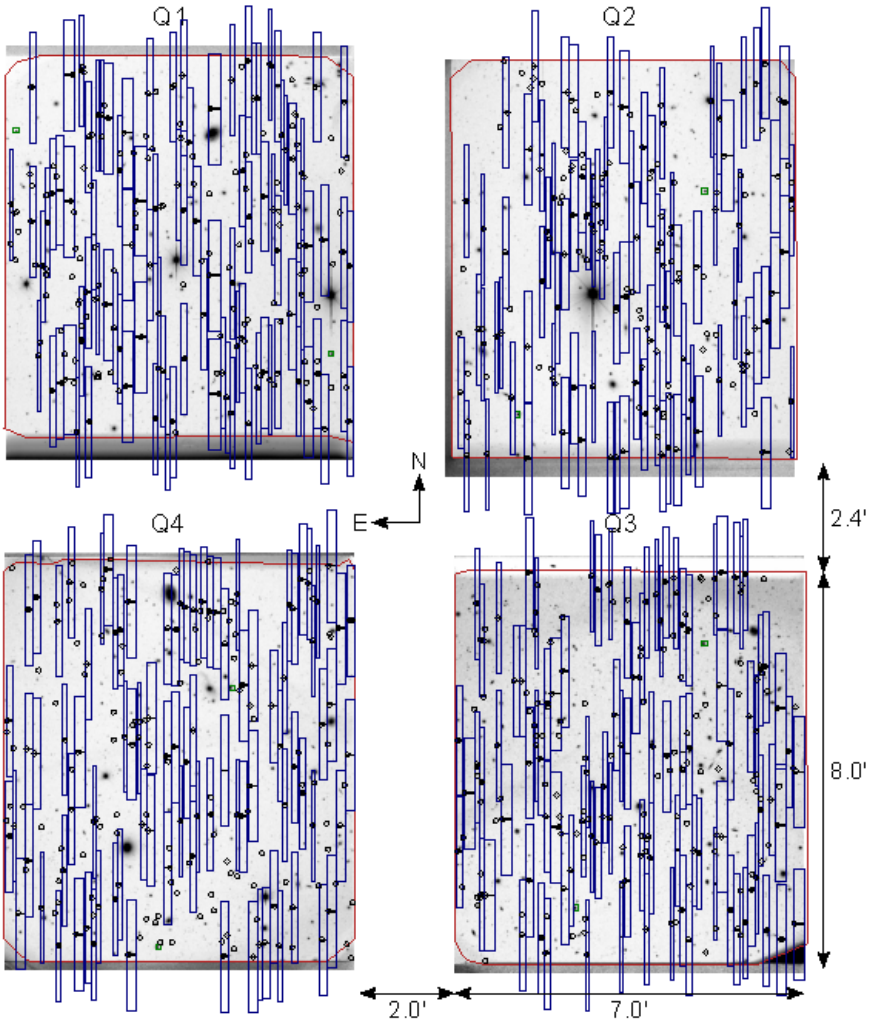


Figure 48: *Example of a VIMOS pointing*. Example of a VIMOS pointing in the VIPERS survey. The dimensions of the gaps between the VIMOS quadrants are reported, together with an example of the slit-positioning (blue rectangles). Figure from Guzzo et al. (2014).

Aside from the random undersampling, the VIPERS survey is also characterised by gaps in the galaxy spatial layout introduced by the physical disposition of the VIMOS detectors. An example can be seen in Figure 48, where it is shown how the distance between the detectors is of $\sim$ 2.4 arcminutes. These bands, where no galaxies have a spectroscopic measurement, together with the slit-positioning procedure, which systematically undersamples denser regions (an example of slit-positioning is also shown in the figure) can introduce biases in the detection of LSS.

The effect of gaps in the skeleton reconstruction is shown in the bottom panel of Figure 47. This plot shows the PDFs of the pseudo-distances obtained by comparing the skeletons S_u and S_{um}, both measured on the parent mock galaxy catalogue, S_u being defined as above

and S_{um} being the skeleton measured on the undersampled parent mock catalogue to which the VIPERS photometric and spectroscopic masks have been applied (and which therefore includes the gaps in the field, but not the systematic undersampling of denser regions due to the slit-positioning procedure). As it can be seen, the asymmetry between the two PDFs is low, meaning that there is only a slight loss of detail introduced by the VIMOS gaps. Moreover, the modes of the PDFs indicate that they peak at $D \lesssim 0.5$ Mpc, meaning that the uncertainty in the skeleton positions introduced by the gaps in the field is negligible.

In fact, the DTFE provides by construction a simple way to reconnect LSS features with smoothed variations in density (Aragón-Calvo et al., 2015), such as filaments across large gaps. For this reason, and based on the results on simulations discussed here, no correction for the gaps in the VIPERS survey is applied. An example of how the DTFE can cross gaps and recover information on the density field where galaxies are missing is shown in Figure 49. This figure shows the effect on the measured density field (bottom row, middle column) of introducing a gap in the starting galaxy distribution (grey shaded area, top row, middle column). It can be seen that the DTFE is able to recover some information on the density field inside the gap by means of the galaxy distribution at the gap borders. This can be enhanced by using the DTFE to provide some form of interpolation of the density field inside the gap (bottom row, right column).

5.3.2 *A comparison of the parent and VIPERS mock galaxy catalogues*

As a final test, the skeletons measured from the parent mock catalogue (S_{Par}) and VIPERS-like mock catalogue, complete of all the characteristics and systematics of the VIPERS survey (S_{VIP}), have been compared. S_{VIP} also includes the gaps in the field due to VIMOS quadrants and the undersampling due to the slit-positioning algorithm. The PDFs of the pseudo-distances obtained from the comparison in the redshift range $0.5 < z < 0.85$ are shown in Figure 50. In the upper panel, DisPerSE is run with a persistence threshold of 3σ in both catalogues. This threshold guarantees that less than 1% of critical pairs are spurious, as tested on random field simulations (Sousbie, 2011). S_{VIP} is reconstructed with less accuracy and detail due to the lower sampling, the effect of gaps, and the undersampling due to the slit-positioning. An estimate of the uncertainty in the location of the filaments is given by the modes of the PDFs which do not peak at distance $D \sim 0$ (corresponding to a perfect match between the segments of the two skeletons) but $D \sim 1.5 - 2$ Mpc. The asymmetry between $D(S_{Par}, S_{VIP})$ and $D(S_{VIP}, S_{Par})$ reflects the fact that S_{Par} (full sampling) has much more details which have no counterpart in S_{VIP}. On the other hand 90% of the segments of S_{VIP} have a counterpart in

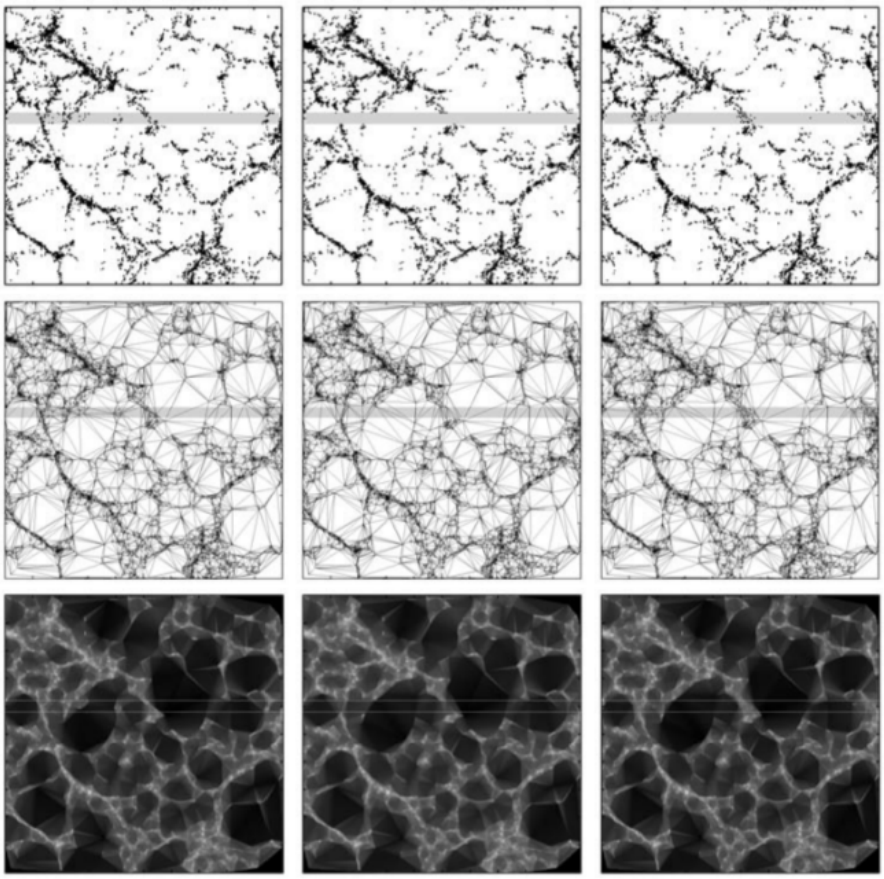


Figure 49: This figure shows an example of the spatial distribution of galaxies in the SDSS (York et al., 2000, top), the tessellation of the galaxy distribution obtained with the DTFE (middle), and the resulting density field (bottom). Left column shows the complete galaxy sample, while middle column shows the effect of a gap introduced in the galaxy distribution (grey shaded area) on the density field reconstruction. Right column shows a possible reconstruction of the density field in the gap based on DTFI (the original figure and a description of the DTFI can be found in Aragón-Calvo, 2007).

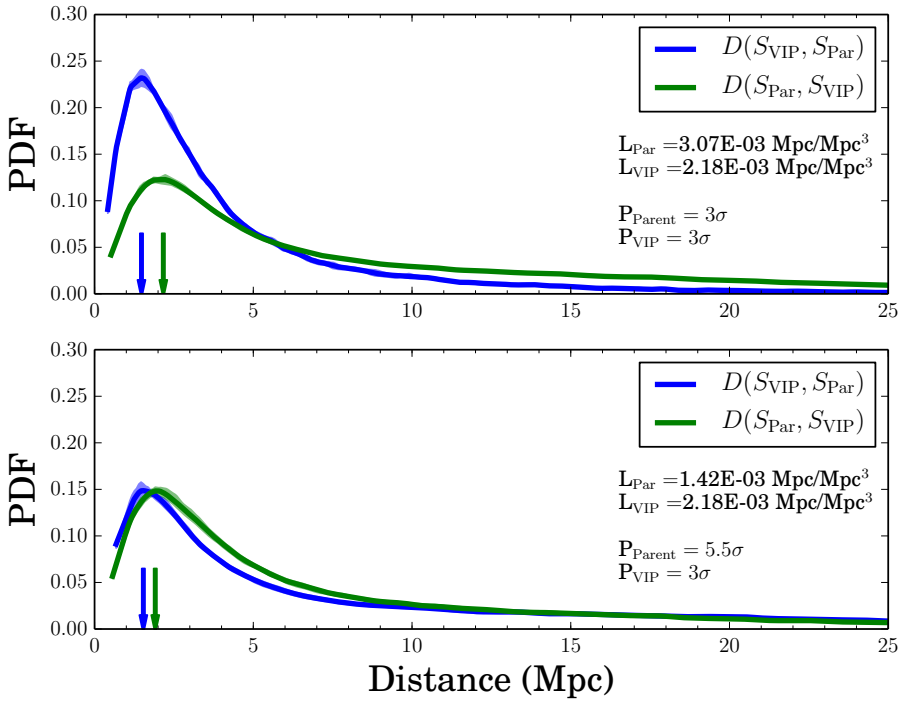


Figure 50: PDFs of the pseudo-distances between the parent and VIPERS-like skeletons. Blue lines refer to $S_{\rm VIP}$ projected onto $S_{\rm Par}$ while green lines are the reverse. Solid lines correspond to the mean of 10 mocks and the shaded areas enclose the 1σ variation (please note that uncertainties are negligible). Vertical arrows show the modes of the distributions. Top panel: the two skeletons are extracted with a 3σ persistence threshold. Bottom panel $S_{\rm Par}$ is extracted with a 5.5σ persistence threshold.

$S_{\rm Par}$ with distances $D \lesssim 10$ Mpc, illustrating the small fraction of spurious filaments. The average length of the VIPERS-like ($L_{\rm VIP}$) and the parent ($L_{\rm Par}$) skeletons, defined as the total skeleton length divided by the survey volume (expressed in Mpc/Mpc^3) are also reported in the figure. $L_{\rm VIP}$ is shorter than $L_{\rm Par}$, as expected for a skeleton with fewer details. In the lower panel, the persistence threshold on the parent skeleton is increased to 5.5σ. The length, $L_{\rm Par}$, is shortened with only the most significant filaments still present. The two PDFs for $D(S_{\rm Par}, S_{\rm VIP})$ and $D(S_{\rm VIP}, S_{\rm Par})$ are now much closer in amplitude and shape. Most of the segments in $S_{\rm Par}$ (75%) have a counterpart in $S_{\rm VIP}$ with $D \leqslant 10$ Mpc. The modes of the distributions are almost unchanged. Even if the two skeletons are more similar, the skeleton $S_{\rm VIP}$ tends to oscillate around its true location (as reconstructed by $S_{\rm Par}$), making the length $L_{\rm VIP}$ longer than $L_{\rm Par}$.

In conclusion, the close match between the two PDFs indicates that the skeleton reconstructed at 3σ for the VIPERS-like catalogue is able to detect the most robust filaments (corresponding to a 5.5σ persistence

threshold in a fully sampled data set) and contains a small fraction of spurious filaments.

5.4 RESULTS

DisPerSE is run on the VIPERS fields with a 3σ persistence threshold. Figures 51 and 52 show the filamentary network for the W1 and W4 fields, overplotted on a map of the density contrast $\delta = n_{DTFE}/\overline{n}(z) - 1$, where n_{DTFE} is the local DTFE density estimate. Even in the 2D projections, it is possible to see that filaments reveal the ridges of the 3D density field which, by construction, connect the density peaks between them via saddle points. The averaged length of the skeletons are similar with $L \sim 0.0013$ and $L \sim 0.0016$ Mpc/Mpc^3 in the W1 and W4 fields respectively. At low ($z \leqslant 0.5$) and high redshift ($z \geqslant 0.85$), the number of filaments drops as a consequence of the lower sampling and only the most secure filaments are detected, as expected with DisPerSE. Thanks to the large contiguous volume probed by VIPERS, large voids, partly delineated by the filaments, are visible in both fields with radii as large as $R \sim 30$ Mpc (see Hawken et al., 2016; Micheletti et al., 2014, for void analysis in VIPERS).

As an example of the effect of increasing the persistence threshold on a skeleton measurement, Figures 53 and 54 show the same map of the CW as Figures 51 and 52 with overlaid also the skeletons measured with persistence thresholds of 4σ and 5σ. It can be seen how increasing the persistence threshold there is a systematic loss of filaments from the more underdense regions, as only the most secure filaments that are located in higher density regions are maintained.

5.4.1 *Mass segregation inside filaments*

It is now investigated how different galaxy populations are distributed within this filamentary network in the redshift range $0.5 \leqslant z \leqslant 0.85$. The distance D_{skel} of each galaxy to the nearest filament is measured. In the figures, D_{skel} is normalised to $\langle D_z \rangle$ to take care of the variations of the mean inter-galaxy separation across the redshift range considered.

A first analysis of the relations among galaxy mass, SFR and distance to the closest filament can be seen in Figure 55. In this figure a 2D distribution of $D_{skel}/\langle D_z \rangle$ as a function of mass and SFR is shown. A relation is clearly visible, with more massive galaxies being located closer to filaments, in particular in the range $M^* \gtrsim 10^{11} M_\odot$. Although less well defined, a relation is also visible with respect to SFR: galaxies with low values of SFR ($\log(SFR/(M_\odot yr^{-1})) \lesssim -3$) are also closer to filaments.

A more quantitative analysis can be performed by looking at the normalised PDFs of the distances. These PDFs, calculated in the W1

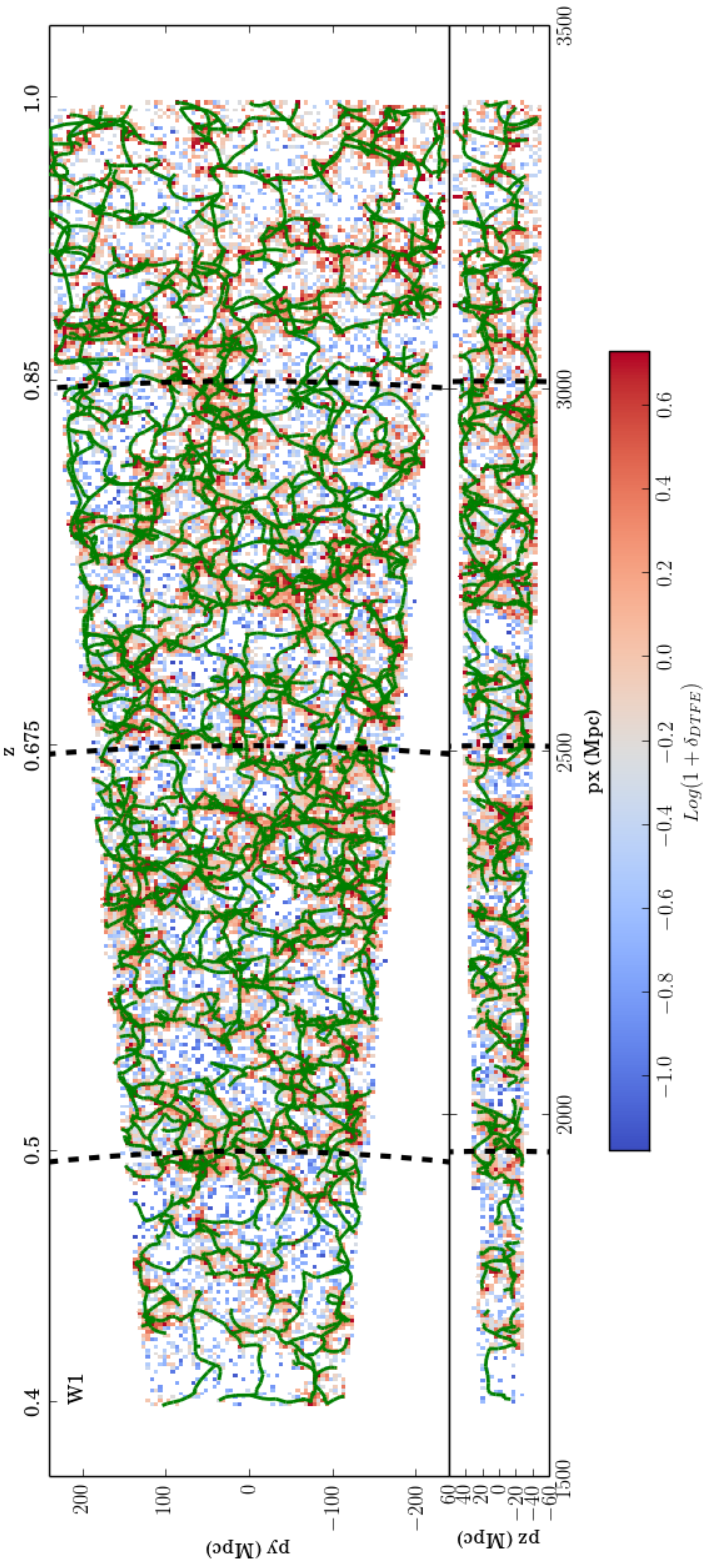


Figure 51: Projected distribution of the filaments reconstructed with DisPerSE (in dark green) in the VIPERS W1 field between $0.4 \leqslant z \leqslant 1$. The density contrast, $\log(1+\delta_{\mathrm{DTFE}})$, is averaged on cells of $5 \times 5\,\mathrm{Mpc}^2$ and colour-coded as indicated (white for empty cells). Top row is the projected distribution along the dec direction ($\Delta\delta = 2$). Bottom row is the projected distribution along the RA direction (in the central regions with $\Delta\alpha = 2$).

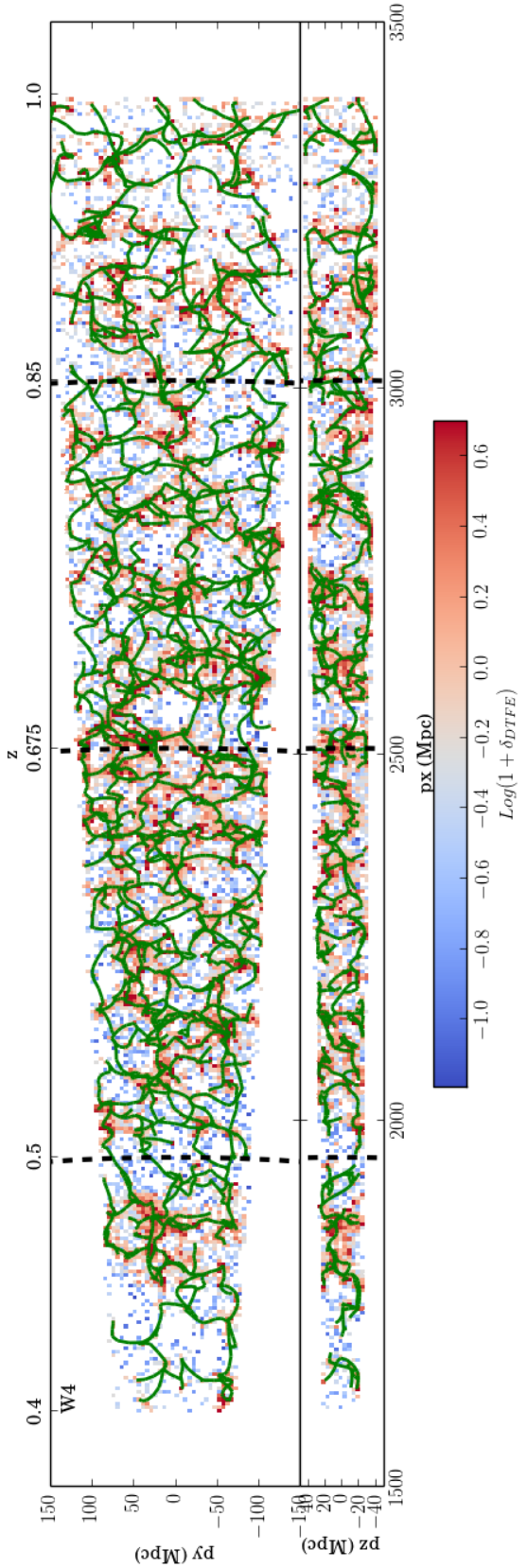


Figure 52: Projected distribution of the filaments reconstructed with DisPerSE (in dark green) in the VIPERS W4 field between $0.4 \leqslant z \leqslant 1$. The density contrast, $\log(1+\delta_{\mathrm{DTFE}})$, is averaged on cells of 5×5 Mpc^2 and colour-coded as indicated (white for empty cells). Top row is the projected distribution along the dec direction ($\Delta\delta = 2$). Bottom row is the projected distribution along the RA direction (in the central regions with $\Delta\alpha = 2$).

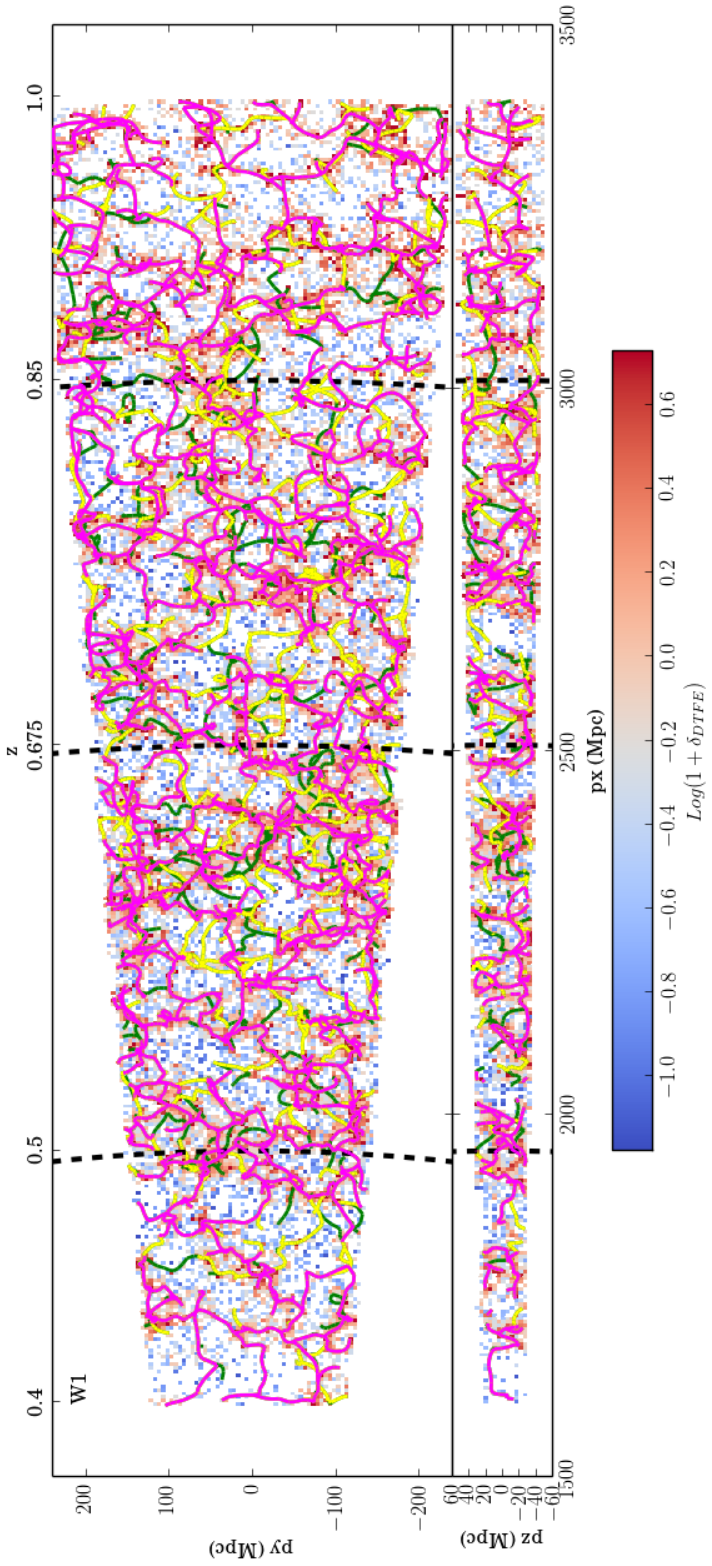


Figure 53: Same as Figure 51, but with skeletons reconstructed with higher persistence thresholds also shown. The persistence threshold varies from 3σ (dark green skeleton), to 4σ (yellow skeleton), and 5σ (magenta skeleton).

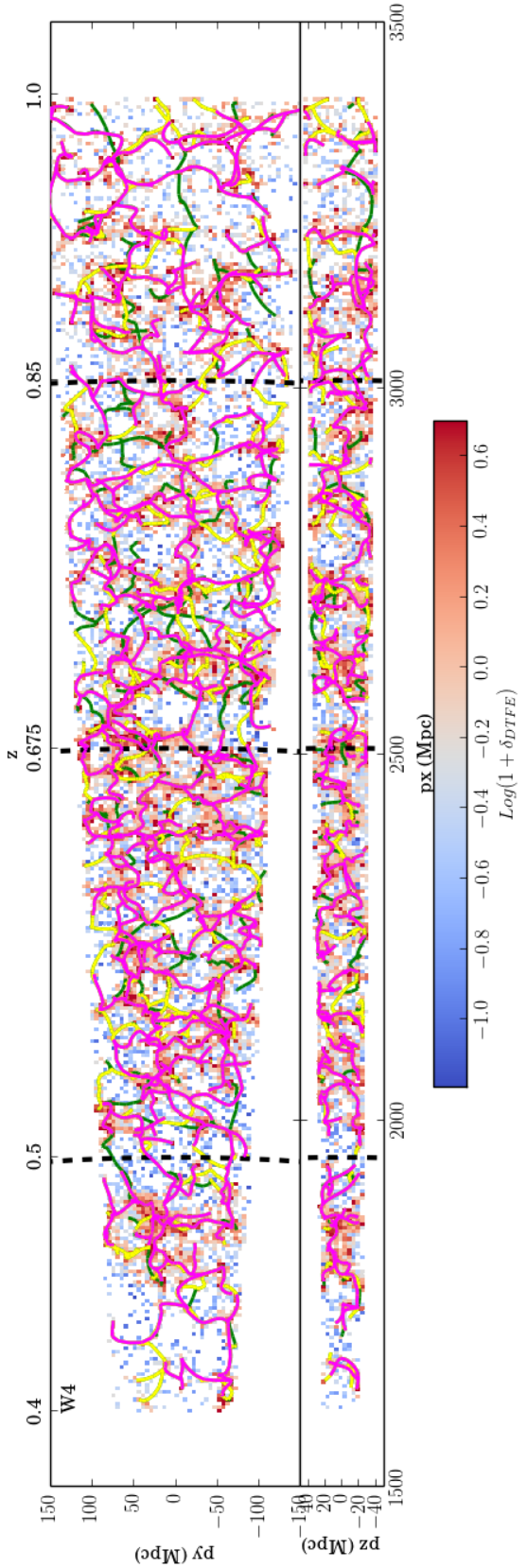


Figure 54: Same as Figure 52, but with skeletons reconstructed with higher persistence thresholds also shown. The persistence threshold varies from 3σ (dark green skeleton), to 4σ (yellow skeleton), and 5σ (magenta skeleton).

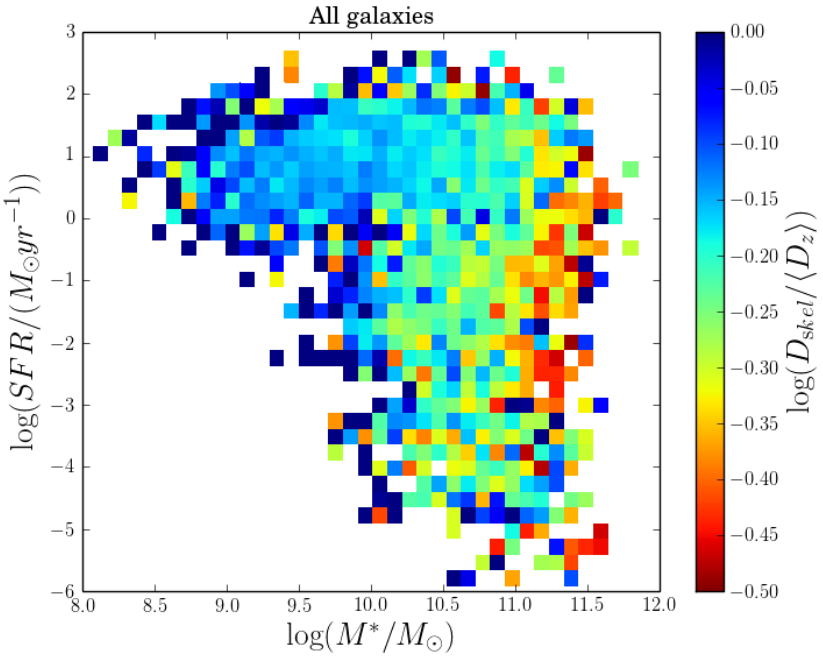


Figure 55: *Distance to filaments as a function of* SFR *and stellar mass*. Map of the galaxy distance to filaments, normalised to the mean inter-galaxy separation ($D_{skel}/\langle D_z \rangle$, colour coded as indicated, white for empty bins) as a function of SFR and stellar mass. Both W1 and W4 fields are considered. Note that in order to be complete at all values of SFR a limit in mass at $M^* \geqslant 10^{10.5} M_\odot$ has to be adopted.

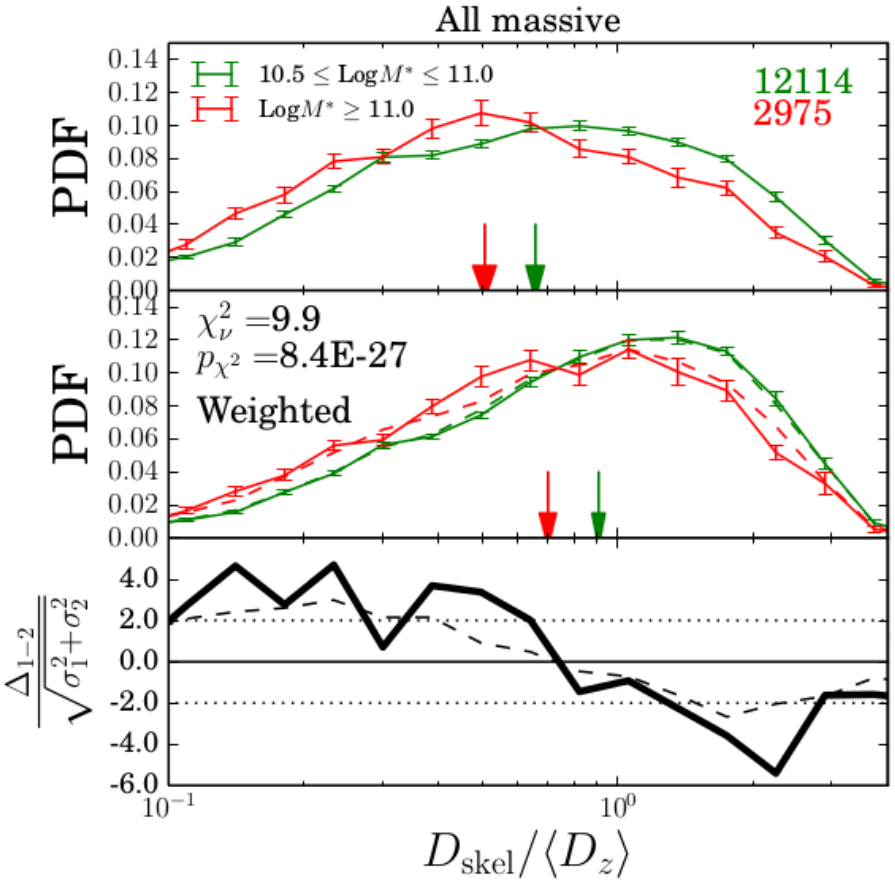


Figure 56: Distributions of the unweighted (top row) and density-weighted (middle row) normalised distances ($D_{skel}/\langle D_z \rangle$), for low *vs* high mass galaxies. Vertical arrows indicate the medians of the PDFs and errorbars represent the dispersions computed with block-wise jackknife resampling. The PDFs after reshuffling of the samples (see text) are shown with dashed lines. The differences between the weighted distributions are shown in the bottom row. The values resulting from the χ^2 test of independence are reported in the middle panel.

and W4 fields combined, are shown in Figure 56 (low *vs* high mass galaxies), Figure 57 (star-forming *vs* quiescent galaxies), and Figure 58 (star-forming galaxies in three mass bins). The errorbars are estimated with a block-wise (1 deg^2) jackknife resampling. The samples in Figures 56 and 57 are limited to $M^* \geqslant 10^{10.5} M_\odot$ to be complete in stellar mass for both quiescent and star-forming galaxies (see Davidzon et al., 2016) while a mass cut of $M^* \sim 10^{10} M_\odot$ is used in Figure 58, when only star-forming galaxies are considered.

A relation between distance and stellar mass is observed (Figure 56), with more massive galaxies being closer to filaments as indicated by the shift in the median values of the two PDFs. Passive galaxies are also found to be closer to filaments (Figure 57). While a large fraction (47%) of massive ($M^* \geqslant 10^{10.5} M_\odot$) galaxies are also passive, by

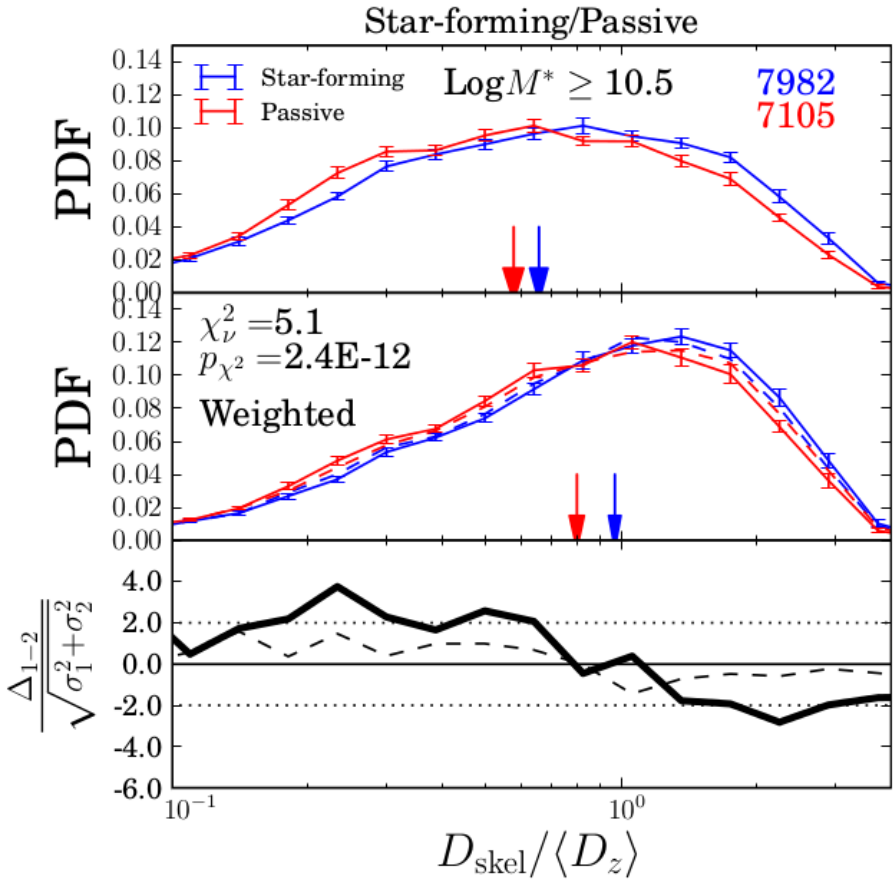


Figure 57: Distributions of the unweighted (top row) and density-weighted (middle row) normalised distances ($D_{skel}/\langle D_z \rangle$), for quiescent *vs* star-forming galaxies. Vertical arrows indicate the medians of the PDFs and errorbars represent the dispersions computed with block-wise jackknife resampling. The PDFs after reshuffling of the samples (see text) are shown with dashed lines. The differences between the weighted distributions are shown in the bottom row. The values resulting from the χ^2 test of independence are reported in the middle panels.

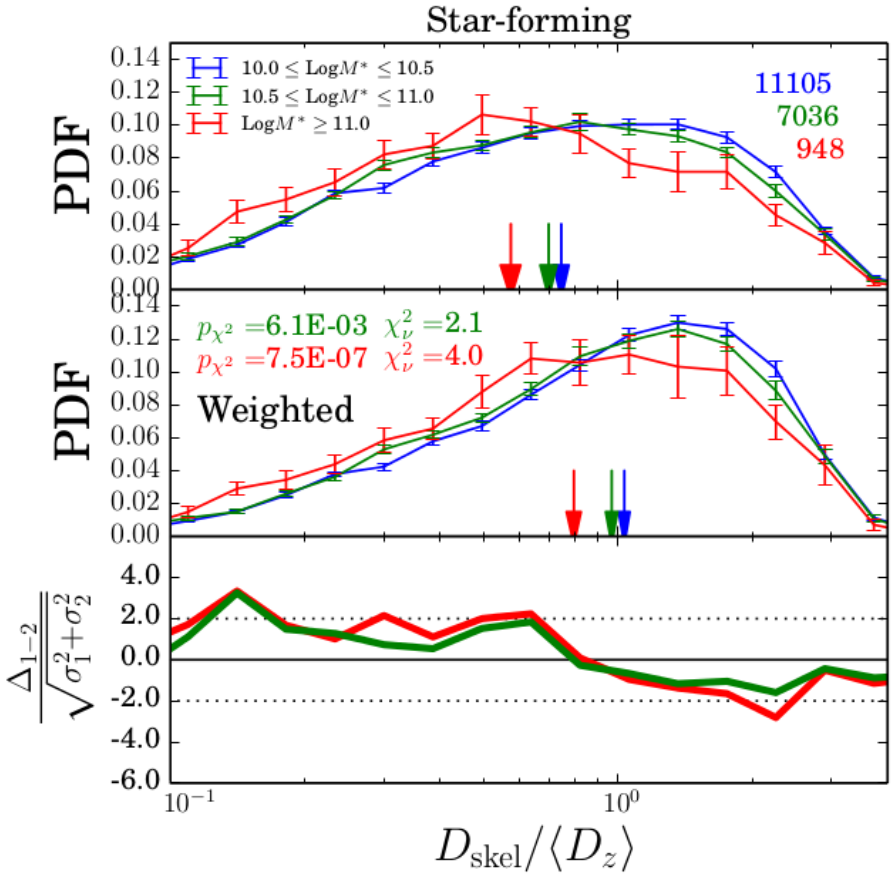


Figure 58: Distributions of the unweighted (top row) and density-weighted (middle row) normalised distances ($D_{skel}/\langle D_z \rangle$), for star-forming galaxies in three mass bins. Vertical arrows indicate the medians of the PDFs and errorbars represent the dispersions computed with block-wise jackknife resampling. The differences between the weighted distributions are shown in the bottom row. Only the differences between high-intermediate (green) and high-low (red) mass bins are shown. The values resulting from the χ^2 test of independence are reported in the middle panels.

looking at the star-forming population alone it is possible to observe a similar trend, albeit weaker, with the most massive star-forming galaxies being closer to filaments (Figure 58). However, since the aim of this Chapter is to evaluate the impact of the filaments on galaxy properties, it is necessary to take into account the contribution of the nodes of the density field, usually related to galaxy groups and clusters, which are at the intersections of filaments and are known to be privileged regions where quenching is more efficient. This task is not easy, as there are partial overlaps between the local density and the CW environment (Aragón-Calvo et al., 2010). In fact, local density will increase close to LSS features such as clusters and filaments. As proposed by Gay et al. (2010), to minimise the node contributions each galaxy is weighted by the inverse of the density field smoothed using a Gaussian filter with $\sigma = 3$ Mpc (see Chapter 2). The weighted PDFs are shown in the middle rows of Figures 56, 57, and 58. A shift in the medians of the PDF to larger distances is observed but the trends remain. Alternative approaches are also adopted by rejecting galaxies in high density regions ($\delta \geqslant 4$), located within groups, (according to Iovino et al., 2016), or by keeping only the filaments with a higher persistence threshold. They do not qualitatively change the results discussed in this section.

The significance of the observed trends is illustrated by the residuals between the weighted distributions expressed in units of σ (bottom rows of Figures 56, 57, and 58). The deviations exceed 2σ in most of the bins except for star-forming samples alone (due to shot noise in the most massive bin). It is also possible to quantify the differences with the χ^2 test of independence and the probabilities of observing such a difference by chance which are negligible (listed in the middle panels of the same Figures). This confirms the existence of a weak but statistically significant segregation effect inside the filaments and suggests that galaxy processing (i. e. the effect on galaxies of processes that induce quenching of the star-formation) happens also during the drift of galaxies towards the nodes of the CW.

As local density increases with decreasing distance to structures of the CW, and tight correlations exist among galaxy mass, local density, and galaxy type, it is necessary to test whether the relations described in the paragraphs above are due to the anisotropic environments of the LSS or are simply another way to recover known relation with local density. To this aim, in Figure 56 it is also analysed how the mass-density relation is hidden in the observed mass segregation. The sample is split in local density bins and the stellar masses are reshuffled between the galaxies in each bin. The mean PDF distributions for 10 random reshufflings are shown as dashed lines on the middle panel. The PDFs for the low and high mass bins are close to the original ones, which shows that the mass segregation exists even after reshuffling the masses, if the mass-density relation is preserved. Therefore

the mass segregation inside the filaments emerges naturally from the mass-density relation and the anisotropic distribution of the density in the CW.

A similar approach is adopted for the galaxy type segregation (Figure 57). A galaxy type (passive/star-forming) is randomly attributed to galaxies by conserving the type fraction observed in different stellar mass bins. The mean PDF for 10 random reshufflings are shown as dashed lines (middle panel). In this case the segregation almost vanishes, with a difference between the two PDFs of less than 1σ. The observed type segregation therefore does not arise just from the mass-type fraction relation but could have its origin in the dynamics of the large scale anisotropic structures of the CW.

5.5 DISCUSSION AND CONCLUSIONS

In this Chapter the first characterisation of large scale filamentary structures at $z \sim 0.7$ has been reported, carried out in the cosmological volume probed by the VIPERS spectroscopic survey. The reconstruction is based on the DisPerSE code and the capability of VIPERS to capture such a CW's filamentary network is tested on simulations. A small but significant trend for galaxies with different stellar masses and stellar activity to segregate near the filaments is observed, with the most massive and/or passive galaxies being closer to filaments. The signal persists even after down-weighting the contribution of nodes and high density regions.

The galaxy segregation quantified in this Chapter is a first step in support of a new paradigm in galaxy formation, where large scale cosmic flows play a role in shaping galaxy properties. Beyond the observed anisotropy of the mass distribution (which follows naturally from the mass-density relation and the anisotropy of density in the CW), other physical parameters (e. g. stellar activity controlled in part by gas accretion or morphology and size controlled in part by angular momentum) are expected to be affected by this large scale environment. Indeed, the results presented here are fully consistent with the ingredients of the spin alignment theory presented in Codis, Pichon, and Pogosyan (2015) which relies on these large-scale cosmic flows. The stellar activity segregation observed here can be interpreted inside this theory. Low mass or star-forming galaxies are preferentially located at the outskirts of filaments, a vorticity rich environment (Laigle et al., 2015), where galaxies acquire both their angular momentum (leading to a spin parallel to the filaments) and their stellar mass essentially via smooth accretion (Welker et al., 2015). The stellar-mass segregation observed here is the next step, where at higher mass, a transition is predicted in simulations, when more massive post-mergers drifting along filaments convert the orbital momentum of their progenitors into spin perpendicular to the axis of the

filament, with increased efficiency for higher merger rate (Dubois et al., 2014; Welker et al., 2015). The most massive galaxies, dominated by the quiescent population, should preferentially complete their stellar mass assembly in the core of filaments by merging.

6

GALAXY OVERDENSITIES AROUND RADIO SOURCES AT $z \lesssim 2$

Active Galactic Nuclei (AGNs) are a class of astrophysical sources resulting from the large amount of radiation emitted by material accreting onto the SMBHs at the centre of galaxies. As the output of energy from these sources is very large, they are commonly regarded as a possible mechanism to end the star-formation in galaxies. In particular, the energy input from AGNs was first considered as a viable quenching mechanism as it allowed to reconcile cosmological simulations with observations (see e. g. Croton et al., 2006; Granato et al., 2004).

Increasing observational evidence supporting the idea that galaxies and their AGNs co-evolve has also been gathered. For example, a tight correlation exists between SMBH and bulge mass (see e. g. Ferrarese and Merritt, 2000; Gebhardt et al., 2000; Magorrian et al., 1998), while SFR and AGN activity have a similar redshift evolution (see e. g. Madau et al., 1996; Ueda et al., 2003). By noting that X-ray and radio-AGNs have different clustering properties, Hickox et al. (2009) suggested a possible scenario for the SMBH-galaxy co-evolution: while the host galaxy transforms from a star-forming to a quiescent one, the AGN passes from a quasar, X-ray emitter phase to a radio-galaxy one. These transformations happen at earlier epochs for haloes of higher mass. This evolution of AGN type with galaxy transformation could be the reason behind the estimate that radio AGNs constitute $\sim 10\%$ of the whole AGN population (see e. g. Hickox et al., 2009), although projection effects and the limiting flux of current radio surveys may also play a role. The reason for such a small fraction of radio-to-total AGNs has still not been fully understood. Since the analysis conducted in this Chapter focuses on radio-AGN samples, in later text it will explicitly be referred to this kind of sources.

As stated in Chapter 1, it was already known that many radio-AGNs reside in early-type galaxies and that the probability that a galaxy hosts an AGN is increasing with stellar mass and local density. In this Chapter the relation between radio-AGN presence and local environment up to $z \sim 2$ is explored by making use of the large statistical data set offered by the Very Large Array (VLA)-COSMOS survey at 1.4 GHz and by the COSMOS photometric redshift survey. This work is fully described in the paper Malavasi et al. (2015)

This Chapter is structured as follows: the data and the method used to define the radio-AGN sample are described in sections 6.1 and 6.2,

respectively. In section 6.3 the method for comparing the overdensity richness distributions is presented, while in sections 6.4 and 6.5 the results for various sub-samples are discussed. In section 6.6 a correlation with known clusters and groups catalogues is attempted and in section 6.7 the integrated mass functions of radio AGNs and normal galaxies are analysed. Conclusions are drawn in section 6.8. Throughout the text, masses are expressed in units of solar masses, while radio luminosities are expressed in SI units. The assumed cosmology is $\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$, and $H_0 = 70$ km s^{-1} Mpc^{-1}. For the SED fitting procedure, in order to derive stellar masses and SFRs, Ilbert et al. (2010) used SED templates produced with a Chabrier (2003) Initial Mass Function (IMF).

6.1 DATA

In this Chapter, two main data sets have been used. Optical data (with photometry in 30 bands from UV to mid-IR) were used to estimate photometric redshifts for all the sources, to determine the environment around every object in every sample, and to extract control samples as explained later in the text. Radio data at 1.4 GHz, instead, were used to create the radio source and AGN samples around which the environment has been studied.

6.1.1 *Optical data*

The data at optical wavelengths are constituted of the photometric sample of Capak et al. (2007), used in the construction of the version 1.8 of the photometric redshift catalogue of Ilbert et al. (2009). This sample is a compilation of photometric data taken from the COSMOS survey (Scoville et al., 2007), which covers a field with an area of about $1.4^\circ \times 1.4^\circ$ that is centred at RA (J2000.0) = $10^h00^m28.6^s$ and dec (J2000.0) = $+02^\circ12'21.0''$.

In Ilbert et al. (2009), photometric redshifts have been calculated through a SED fitting procedure, using 30 broad-, intermediate-, and narrow-band filters from UV to mid-IR frequencies (see their Table 1). The fit was performed with a χ^2 minimisation algorithm on a template library using the *Le Phare* code (S. Arnouts & O. Ilbert). The redshifts obtained with the described procedure have been compared with those obtained using the zCOSMOS spectroscopic survey (Lilly et al., 2007). The accuracy of the photometric redshifts (z_p) is estimated to be

$$\sigma_{\Delta z/(1+z)} = 0.06 \tag{28}$$

This value will be assumed throughout this Chapter. The error has been conservatively selected as the maximum value of those obtained

as a function of both magnitude and redshift in Ilbert et al. (2009, see their Figure 9), taking the depth of the optical catalogue used for the analysis into account. To check the consistency of such a choice, the effect of degrading the value to $\sigma_{\Delta z/(1+z)} = 0.1 - 0.2$ for high-redshift sources has been tested. As explained later, it was found that it does not change significantly the results described here.

Other sources of uncertainties in the photometric redshifts determination that must be kept in mind are the so-called catastrophic errors. The definition of catastrophic errors is applied to those sources for which the redshift calculation fails in the form of

$$\frac{|z_p - z_s|}{1 + z_s} > 0.15 \tag{29}$$

where z_s is the spectroscopic redshift. It is estimated that the fraction η of catastrophic errors rises from 1% to 20% going from $i^+ < 22.5$ to $i^+ \sim 24$.

For this analysis only objects that are not in masked areas and that have magnitude $i^+ < 26.5$ have been considered. This magnitude limit corresponds approximately to the point where the magnitude-number-count histogram stops increasing. At this limit the average number of filters used for the photometric redshift determination is 18. Although the effect of such a deep cut in limiting magnitude is to increase the uncertainties in the physical quantities derived from the optical photometry, it also has the advantage of greatly increasing the statistics for the environment estimate. For this reason, this cut will be assumed for all samples in this Chapter. As explained extensively in section 6.3, tests were performed by introducing brighter magnitude cuts from $i^+ < 25.5$ to $i^+ < 24$ to check the problematics related to also including in the analysis very faint objects. These tests showed that the magnitude cut has no major effect on the results that have been found.

From the optical catalogue, a sample of 823 939 sources with optical data and measured photometric redshift has been extracted (hereafter "O sample"). SFRs and stellar masses (M^*) are both derived through SED fitting with population synthesis models (see Ilbert et al., 2010), together with other physical parameters.

6.1.2 *Radio data*

The radio data are taken from the VLA-COSMOS Large Project version 2.0 catalogue (Schinnerer et al., 2007, 2010), whose observations were carried out with the VLA in its A configuration. This grants a resolution of about 1.5″ at Full Width Half Maximum (FWHM) in the observation wavelength of 1.4 GHz. A total of 23 pointings was necessary to cover the full COSMOS field, for 240 hours of observation, performed between 2004 and 2005. The observations reached a sensitivity of about 11 μJy r.m.s. and the 5σ detection limit catalogue

BIN	z	R SAMPLE	MR SAMPLE	AGN SAMPLE
Bin 1	0 - 0.7	532	425	119
Bin 2	0.7 - 1	320	290	100
Bin 3	1 - 2	575	503	53

Table 8: Size of every sample and redshift range of every bin used in the data analysis.

contains 2417 radio sources, 78 of which have been classified as multi-component radio sources.

The search for the optical counterparts used the photometric redshift sample (version 1.8) by Ilbert et al. (2009), with no limitation on the optical magnitude, while for the radio catalogue, the VLA-COSMOS Large Project version 2.0 was used, but without the 78 multi-component radio sources (2339 radio sources).

The optical identification used the likelihood ratio method, described in Sutherland and Saunders (1992), Ciliegi et al. (2005), and Brusa et al. (2007). For the likelihood ratio analysis a Q value (the probability that the optical counterpart of the radio sources is brighter than the magnitude limit of the optical catalogue) of 0.9 was adopted and a likelihood ratio cut-off of 0.1 ($1 - Q$). Using a likelihood ratio cut-off equal to $1 - Q$ ensures that all the optical counterparts of the radio sources with only one identification and a likelihood ratio greater than the cut-off value have a reliability greater than 0.5. With these thresholds, 2069 radio sources with an optical counterpart have been found, $\sim$ 26 of which ($\sim$ 1.3%) could be spurious positional coincidences. Applying the same cuts as used to define the O sample (objects not in masked areas and with magnitude $i^+ < 26.5$) to the 2069 identified radio sources, a sample of 1806 optically identified radio sources was created (hereafter "R sample"). The data analysis has, nevertheless, been performed only out to $z \leqslant 2.0$, i. e. on a selection from the R sample of 1427 objects. The sizes of the various samples can be found in Table 8, while a summary of the sample definitions can be found in Table 9.

In Figure 59 the redshift distributions for the O, R, MR and AGN samples (the last two are introduced later in the text) are compared. It can be seen that the distributions referring to radio galaxies peak at lower redshift, while the distribution for the optical galaxies is wider in a more extended redshift range.

The greatest difference in the samples is highlighted in Figure 60, which shows the stellar mass distributions for the same samples as in the previous plot. It can be easily seen that the galaxies showing radio emission (R, MR, and AGN samples) are found almost only in high-mass galaxies (peak around $\log(M^*/M_\odot) \sim 11$), while objects

SAMPLE NAME	DEFINITION
O	Optical objects, $i^+ \leqslant 26.5$
R	Optically identified radio objects, $i^+ \leqslant 26.5$
MR	Optically identified radio objects, $i^+ \leqslant 26.5$, $\log(M^*) \geqslant 10$
AGN	AGNs, $i^+ \leqslant 26.5$, $\log(M^*) \geqslant 10$, $\log(\frac{SSFR}{yr^{-1}}) \leqslant -11$
RO	Control sample of R, extracted from O
MO	Control sample of MR, extracted from O, same M^* distribution
QO	Control sample of AGN, extracted from O, same M^* distribution, same SSFR range

Table 9: Samples definition.

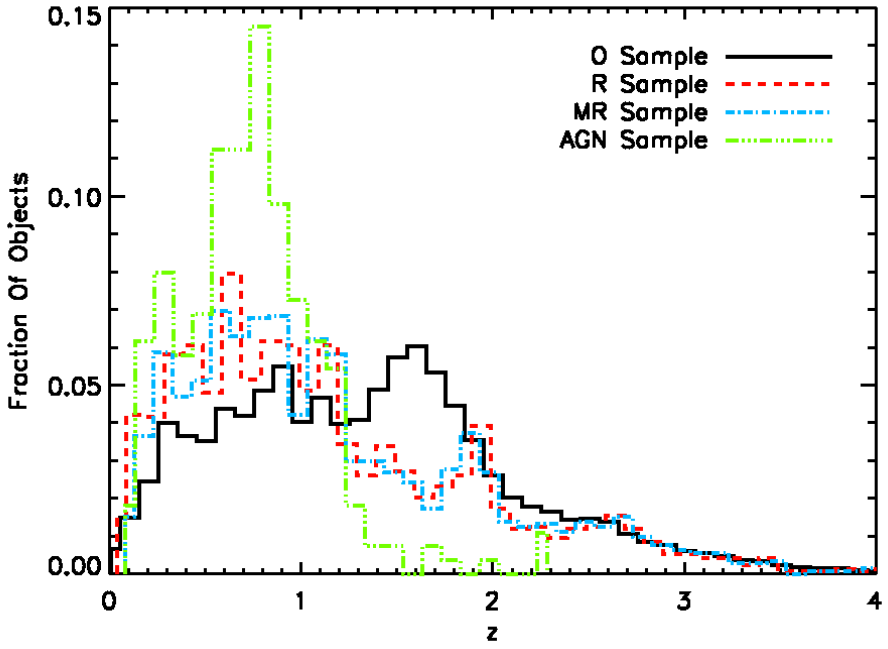


Figure 59: *Redshift distributions*. The black solid line refers to the O sample, the red dashed one to the R sample, the cyan dot-dashed line refers to the MR sample and the green triple dot-dashed one refers to the AGN sample.

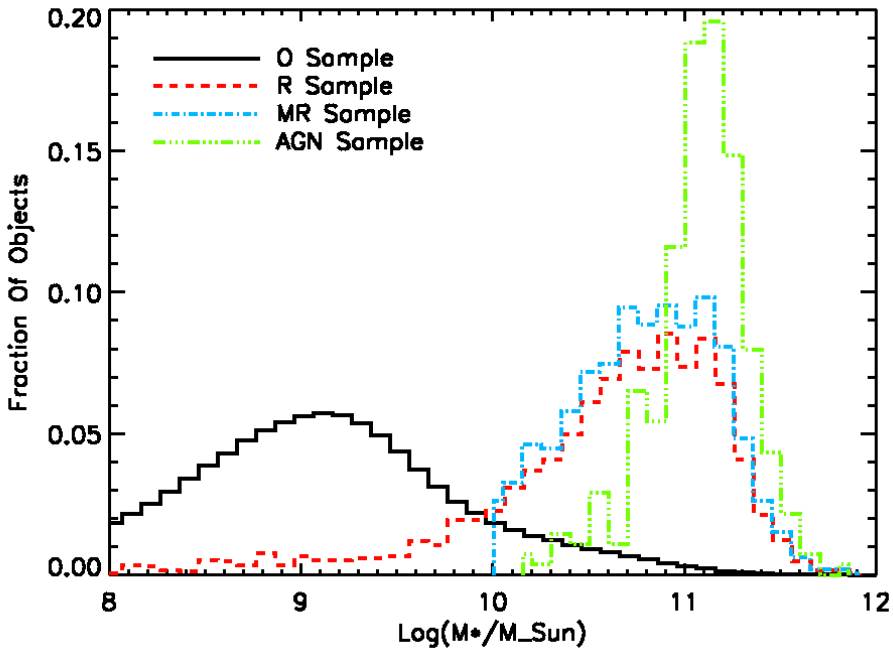


Figure 60: *Stellar mass distributions*. The black solid line refers to the O sample, the red dashed one to the R sample, the cyan dot-dashed line refers to the MR sample and the green triple dot-dashed one refers to the AGN sample.

from the O sample span a much wider range of masses, with a peak at $\log(M^*/M_\odot) \sim 9$.

This could be due both to an evolutionary effect, which sees the onset of radio emission mainly in high-mass galaxies, or to the limiting flux of the VLA-COSMOS, which could limit the kind of galaxy detected to only the most massive and luminous ones. Aside from the physical process responsible for the shape of the stellar mass distributions, the fact that the radio galaxy samples are located at higher masses with respect to normal galaxies must be considered, in order not to introduce a bias in the conclusions drawn from these samples. For this reason a cut in stellar mass has been introduced by selecting only the most massive radio sources. When considering only radio sources at $\log(M^*/M_\odot) \geqslant 10$, the sample size becomes of 1448 objects, ("MR sample"). Again, the data analysis has only been conducted on a selection of 1218 with $z \leqslant 2.0$. The cut to $\log(M^*/M_\odot) \geqslant 10$ also accounts for the possibility that the radio sources without optical identification are faint galaxies with low stellar mass. Therefore selecting only high-mass radio sources should avoid the insurgence of any bias in the results from missing optical identifications of radio sources.

6.2 THE AGN SAMPLE

It is well known that radio emission could be due both to star-formation and to AGN phenomena. Several methods have been suggested for dividing the two populations where each of the two phenomena is dominant. There are several recipes for segregating the radio population into sources with emission from an AGN or from star-formation (see e. g. Bardelli et al., 2009; Best et al., 2005; Smolčić et al., 2008, and references therein). In this Chapter two different ways have been adopted, which, nevertheless, share a common hypothesis.

The first way of defining a sample of radio AGNs is through a cut in sSFR. For each object, the sSFR is defined as the ratio of the SFR and the stellar mass, with both quantities derived from the SED fitting procedure. Quiescent galaxies are selected, and it is assumed that for this kind of sources all the radio flux comes from AGN activity. For this purpose, the sSFR *vs* stellar mass (sSFR-M^*) plane is analysed (Figure 61). From this plot, the locus of star-forming galaxies is clearly visible as a horizontal region within a narrow range of sSFR ($-11 \leqslant \log(SSFR/yr^{-1}) \leqslant -7$). To the right of the figure, a vertical band located at $10 \leqslant \log(M^*/M_\odot) \leqslant 12$ represents the galaxies that are switching their star-formation off ("the dead line"). As AGN sample it has been chosen to select radio sources with $\log(M^*/M_\odot) \geqslant 10$ and $\log(SSFR/yr^{-1}) \leqslant -11$. With these cuts, a sample of 276 radio active galactic nuclei has been defined (272 at $z \leqslant 2$).

Although efficient, this method adopts an extreme hypothesis, since it assumes that all radio flux is coming from radio AGN activity in quiescent galaxies. The cuts applied to the sample may actually result inaccurate as more reliable methods of distinguishing radio activity from AGN phenomena and from star formation can be developed.

One of these is described in Bardelli et al. (2010). It consists in the comparison between SFRs as resulted from the SED fitting of Ilbert et al. (2010) and as estimated from the radio power as done in Bell (2003). The hypothesis is that all radio flux exceeding that given by the SFR_{SED} is due to AGN activity.

The SFR_{Radio} has been computed using the Bell (2003) formula (which is derived for a Salpeter IMF), with a correction of ~ 2 in the normalisation introduced to be consistent with Herschel-PACS Evolutionary Probe (PEP) data (Bardelli et al. in prep.). This normalisation correction also considers the difference between IMFs used in the SFR derivation from optical SED fitting and in the formula from Bell (2003). In this particular work, the normalisation is less important, since only the identification of the locus of star-forming galaxies matters. The comparison between the two SFRs is reported in Figure 62. In this figure it is clearly visible an elongated region due to the star-forming galaxies with a slight zero point shift and tilt. To fit this relation, only those objects that have $\log(SFR_{radio}) \leqslant \log(SFR_{SED}) + 2$ have

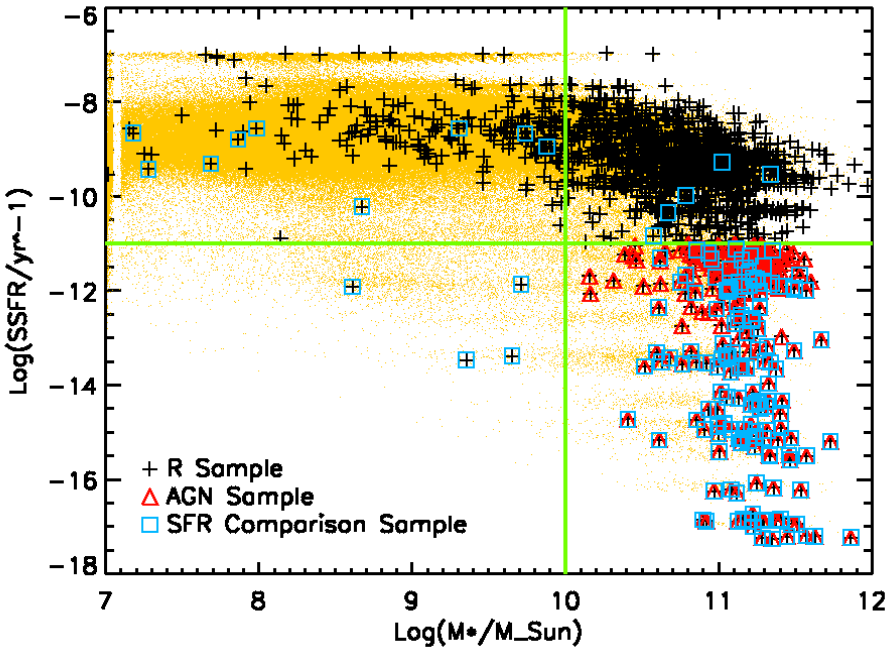


Figure 61: *sSFR-M* plane*. The black crosses refer to the R sample, the red triangles to the AGN sample defined through the sSFR cut and the cyan squares to the AGN sample defined through SFR comparison. The green lines correspond to the cuts in stellar mass and sSFR described in text. In yellow, the O sample is reported for comparison.

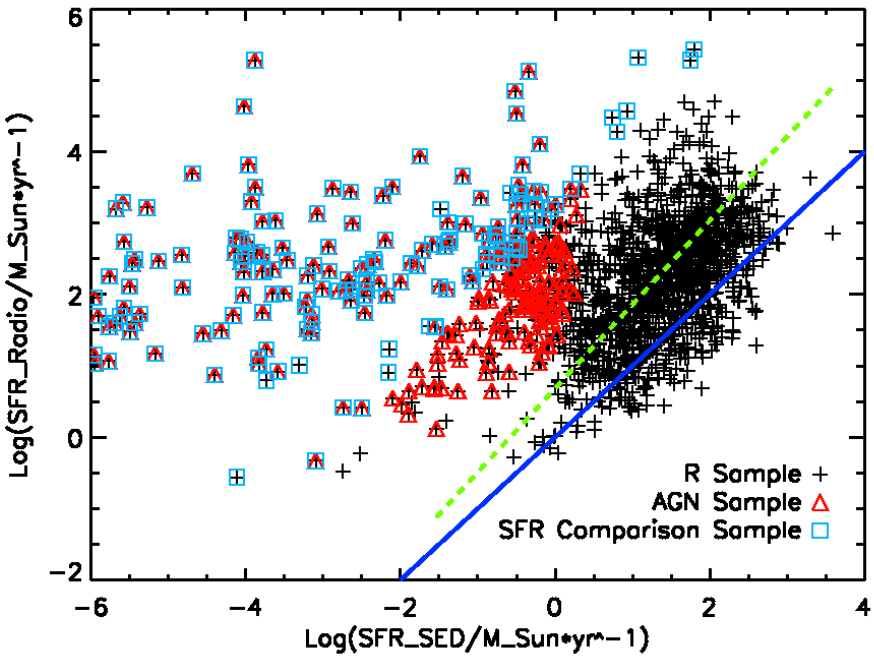


Figure 62: *AGN sample definition*. The black crosses represent the R sample, the red triangles represent the AGNs selected through the cut in sSFR, while the cyan squares represent the AGN sample defined through SFR comparison. The dashed green line represents the linear fit relation described in text, while the solid blue line shows the $SFR_{radio} = SFR_{SED}$ relation.

been considered. The linear fit resulted in $\log(SFR_{radio}) = 1.176 \times \log(SFR_{SED}) + 0.685$. The distribution of the distances from the best-fit line was then fitted with a Gaussian distribution, obtaining a dispersion of 1.40. To be very conservative, all the objects that lie at a distance greater than 3σ from the fit line are defined as AGNs. With this method, 164 sources are selected (154 at $z \leqslant 2$).

It has to be noted that this second method is more conservative, since the assumptions are less extreme: it is now possible for every galaxy, even for the quiescent ones, that part of the radio emission could be due to residual star-formation. The discrimination is set to decide which of the two phenomena (radio AGN activity or the formation of stars) prevails on the other. With this method, the probability of correctly selecting a radio AGN source (purity) is increased, in spite of completeness, since some radio AGN sources may not be individuated by the selection criteria. Figures 61 and 62 both show that the radio AGN samples derived with the two methods overlap.

Since the infrared emission is a better tracer of the SFR, a more rigorous approach would be to apply the method described above using the SFR derived from IR data, rather than from optical ones. For this reason the R sample has also been matched to the DR1 cata-

logue of the Herschel-PEP[1]. A description of the survey, observational strategies, and data reduction techniques may be found in Lutz et al. (2011), while SFR and physical quantities were obtained through a SED fitting procedure performed with a modified version of the Multi-wavelength Analysis of Galaxy Physical Properties (MAGPHYS) code (see Berta et al., 2013; Delvecchio et al., 2014, and references therein for a detailed description). A total of 923 sources have been found by matching the R sample, 45 of which are in the AGN sample and are all situated at lower SFR values compared with other Herschel sources.

A sample of 83 AGNs was created from comparing the SFR_{Radio} with the SFR from the IR data, which was derived through a procedure of SED fitting that allows for correction of the AGN contribution to galaxy emission (SFR_{IR}). This method is equal to the one applied in previous paragraphs: the SFR_{Radio} and SFR_{IR} are compared and the bulk of the relation is linearly fitted to isolate sources with an excess of SFR_{Radio} with respect to SFR_{IR}. The distribution of the distances of each object from the best-fit line is fitted with a Gaussian distribution and only sources residing farther from the best-fit line than 3σ are defined as AGN.

The environmental analysis was performed on these sources in the same way as for the other AGN samples. The environmental segregation effect is recovered, but unfortunately the sample is too small to make the results significant.

Extensive tests were performed to check whether the various methods of defining AGNs yield differences in the results found in this Chapter. No significant discrepancies emerge from the tests. The AGN sample defined through the cuts in sSFR and stellar mass was used, since it permits slightly more complete statistics, but all the conclusions derived in the following also remain valid if the radio AGNs are selected through the SFR_{Radio} and SFR_{SED} (or SFR_{IR}) comparison.

6.3 METHOD

The aim of the work described in this Chapter was to investigate the presence and properties of galaxy overdensities in the COSMOS field, especially at high redshift. The environmental estimator adopted for this work has been a fixed aperture method (see Section 2.1.1), with an aperture shaped as a parallelepiped with a base side of 1 Mpc (comoving), centred on the considered source, and height twice $\Delta z = 3 \times \sigma_{\Delta z/(1+z)} \times (1 + z_p)$. In the following, three kinds of samples will be defined, two of which already introduced in Section 2.1.

TARGET SAMPLE: this kind of sample is composed of the objects around which the environment is to be estimated and which

1 http://www.mpe.mpg.de/ir/Research/PEP/DR1

are placed at the centre of each parallelepiped. Target samples are the R, MR, and AGN samples.

TRACER SAMPLE: these objects are those that are counted inside each parallelepiped and that are used to estimate the environment around every galaxy of each Target sample. The O sample is used as Tracer sample.

CONTROL SAMPLE the absolute values and richness distributions of overdensities do not have any direct meaning if not compared with those obtained with a Control sample. To create it, the redshift has been maintained for each galaxy of a Target sample, but coordinates have been randomly extracted. The way this is achieved is different for each sample and is explained thoroughly in section 6.3.1. The environment was then estimated again for each Control sample using the same Tracer sample as for the corresponding real ones. The richness distributions between each sample and its corresponding Control one have been compared with a Kolmogorov-Smirnov (KS) test. To not be influenced by only one extraction, this process was repeated 100 times.

The environmental analysis was performed in three different redshift bins: $z \in [0.0 - 0.7[$, $[0.7 - 1.0[$, and $[1.0 - 2.0]$. It is important to mention that all the environmental analysis was done both on a 1 Mpc and on a 2 Mpc scale. It was found that all the results discussed in the following sections have their significance greatly increased in the 1 Mpc compared to the 2 Mpc case. This could be because when adopting a smaller scale, only the central regions of the bound structures were considered, especially at high redshift, where clusters and groups were in an earlier stage of formation. It is known that in such regions (i. e. at the centre of overdensities), galaxy formation happens on shorter timescales. Therefore any existing relation between processes involved in galaxy evolution and environment (such as the presence of radio-AGNs) is likely to be shown more on a 1 Mpc scale. Larger scales (such as the 2 Mpc case) may have the signal diluted by the inclusion of regions in structures that are not yet well formed and by an increase in the number of interlopers owing to the photometric redshift error. This is confirmed by the fact that the results of this work are much more significant when using the 1 Mpc scale compared to the 2 Mpc case. For this reason, in the following, only the results obtained with an estimate of the environment on a scale of 1 Mpc are discussed.

As already mentioned in previous sections, several tests have been performed to check whether the choice of the photometric redshift error and of the magnitude cut performed on the sample do in fact affect the results of this Chapter. In particular, the environmental analysis was performed by also using a cut on the O sample at brighter

magnitudes: $i^+ < 25.5$, $i^+ < 25$ and $i^+ < 24$. Only the magnitude cut on the O sample has been changed, since no AGN sources have $25 < i^+ < 26.5$. In this way, while the magnitude distribution for the AGN sample is maintained, those of the Control samples and of the sources used to trace the environment have been changed. In the first two magnitude cuts, the assumption of the photometric redshift error has also been degraded to $\sigma_{\Delta z/(1+z)} = 0.1 \div 0.2$ in the third redshift bin ($z \in [1.0 - 2.0]$). The significance of the results is maintained for all the combinations, except for the AGNs in the highest redshift bin when using a magnitude cut for the O sample of $i^+ < 24$. This is not unexpected, though, since tracing the environment using only the most luminous sources and comparing AGN with luminous control galaxies has the effect of comparing objects residing (especially at high redshift) in similar environments, therefore reducing the significance of the signal of environmental segregation. It is therefore possible to conclude that neither the magnitude cut used nor the value of the photometric redshift error assumed have a systematic effect on the results presented in this Chapter.

6.3.1 *Control sample definition*

As briefly explained above, the Control samples for the R, MR, and AGN sub-samples were created with optical data. For every galaxy in the R, MR, and AGN samples, coordinates have been randomly extracted from the O sample in the same redshift bin for every Target sample. In addition, the following prescriptions were used separately for each sub-sample.

R SAMPLE: The Control sample RO was extracted from O, so as to have the same number of sources as the R sample in every redshift bin.

MR SAMPLE: The Control sample MO has been extracted from O, so as to have the same number of sources as the MR sample in every redshift bin. Moreover, every galaxy of the Control sample has been extracted among those having stellar mass in an interval of ± 0.5 dex from the mass of each galaxy composing the MR sample. In this way, the MO sample was extracted with the same stellar mass distribution as the MR one.

AGN SAMPLE: The Control sample QO has been extracted from O, so to have the same number of sources as the AGN sample in every redshift bin. It has also been extracted having the same stellar mass distribution as the AGN sample, with the same method as used to extract the MO Control sample. Also, the galaxies used as a pool for the QO sample extraction were selected to be quiescent ($\log(SSFR/yr^{-1}) \leqslant -11$), therefore the QO sample is

SAMPLE	$0 \leqslant z < 0.7$	$0.7 \leqslant z < 1$	$1 \leqslant z \leqslant 2$
R *vs* RO	100 ($6.6 \cdot 10^{-8}$)	100 ($7.5 \cdot 10^{-9}$)	100 ($6.5 \cdot 10^{-7}$)
MR *vs* MO	100 ($6.5 \cdot 10^{-7}$)	100 ($6.9 \cdot 10^{-7}$)	100 ($9.2 \cdot 10^{-6}$)
AGN *vs* QO	100 ($8.6 \cdot 10^{-5}$)	100 ($6.0 \cdot 10^{-6}$)	87 (0.006)

Table 10: *Number of extractions of the various Control samples that are below the significance threshold of 0.05*. In brackets, the median value of the KS probability value distributions is given. The probability resulting from the KS test is the probability that the two distributions analysed were extracted from the same parent population.

located in the same lower right-hand region of Figure 61 as the AGN sample.

6.4 RESULTS

In this Section the results of the analysis of the environment performed on the various Target samples are analysed. As explained above, to make the differences in the environment around radio sources and around galaxies with no radio emission (represented by the extracted Control samples) more apparent, the overdensity richness distributions were compared using a KS test, repeated with 100 independent extractions of the various Control samples. When applied, the KS test gives the probability that two distributions come from the same parent population. To quantify the difference in the environments recovered, only richness distributions that have a KS probability value lower than 0.05 of being extracted from the same parent population will be considered as significant. In Table 10 the number of Control sample extractions (out of 100) that are below the significance threshold of 0.05 are reported, together with the median values of the KS probability value distributions (in brackets).

6.4.1 *R Sample*

At first, the environment is estimated around sources from Target samples R and RO. It has to be kept in mind that the R sample is the whole radio catalogue. This sample does not distinguish whether the radio emission comes from star formation or AGN activity. The Tracer sample used is represented by the whole O sample. Results of the KS test between R and the 100 RO extractions are listed in the first line of Table 10. From these values it can already be seen that the environment around radio sources is significantly different from the environment around sources with no sign of radio emission, in every redshift bin. The total of the Control sample extractions is always

below the significance threshold with very low median values. The way in which these environments differ is exemplified by Figure 63.

In this Figure, the overdensity richness distributions are shown for the R sample and for all the 100 extractions of the RO Control sample. The richness distribution of the Control sample extraction that corresponds to the median of the KS probability value distribution is in evidence. It can be seen that in every redshift bin the radio sources are systematically distributed at higher overdensity richness than normal galaxies.

Although important, this result could be due to other effects, such as the different mass distributions or evolutionary status of R and RO galaxies. A deeper investigation of the phenomenon is thus needed.

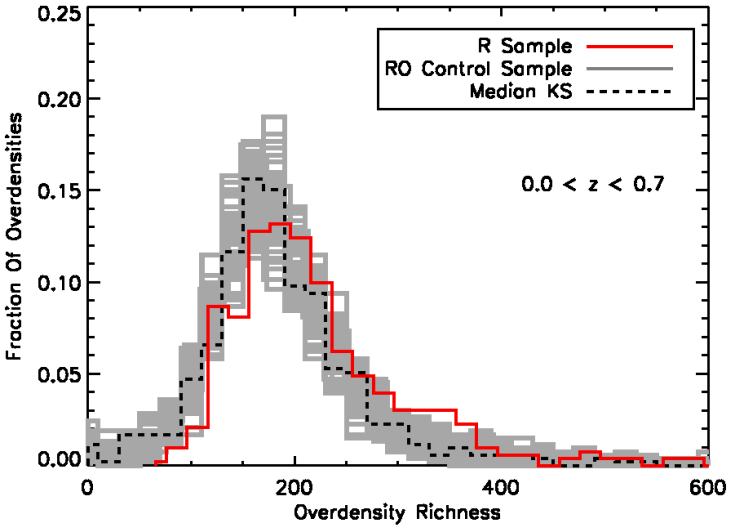
6.4.2 *M Sample*

It has been shown (see Figure 60) that optical galaxies and radio sources have different mass distributions. In particular, radio sources are primarily found in high-mass galaxies, which in turn are known to reside primarily in high-density environments. This could be the actual reason behind the difference in environmental density found between sources from the R and RO samples described in the previous section.

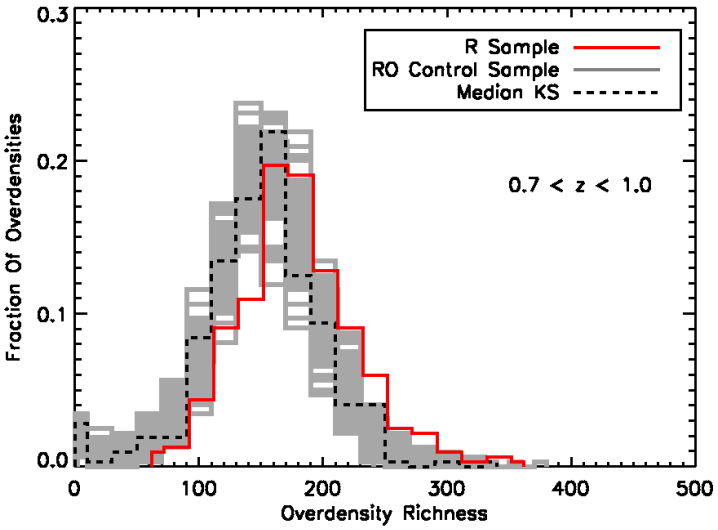
The sample MR and its corresponding Control sample MO have been created to rule out this effect. These samples are designed to include only massive galaxies ($\log(M^*) \geqslant 10$) and to have the same stellar mass distribution. In this way, because the MR sample is composed of radio sources and its corresponding Control sample MO of optical galaxies, any difference in the environment between the two can be ascribed to the actual presence of radio emission.

The second line of Table 10 confirms the trend already found for the R and RO sample. The environment around radio sources and the environment around normal galaxies are significantly different, as demonstrated by the fact that all 100 extractions of the MO Control sample are below the significance threshold, with very low median values. Again a visual inspection of the overdensity richness distributions, such as those reported in Figure 64 for the MR sample and all of the 100 extractions of the MO Control sample, confirms that overdensities found around massive radio sources are distributed at higher richness values than those around massive galaxies without radio emission.

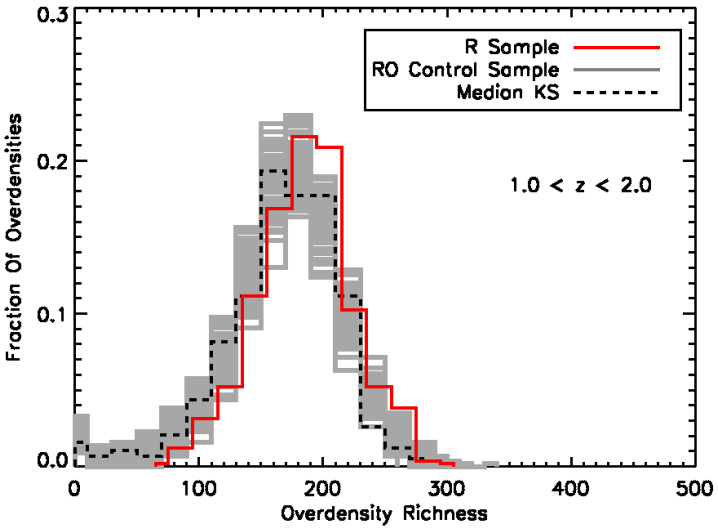
It is therefore possible to exclude a stellar mass effect behind the differences in the environment around normal galaxies and radio-emitting ones. Instead these differences are to be ascribed to the physical process behind the radio emission and its correlation with environmental density.



(a)

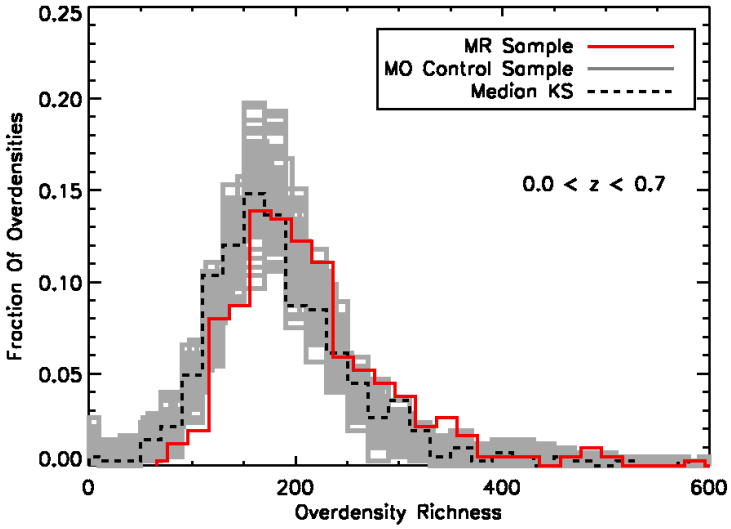


(b)

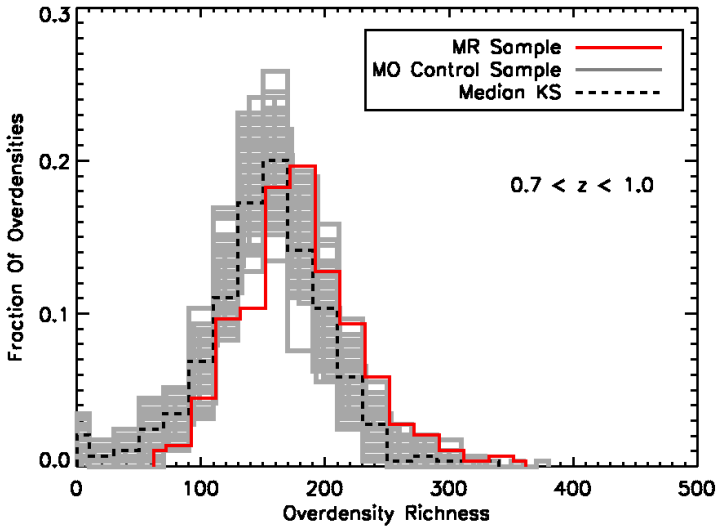


(c)

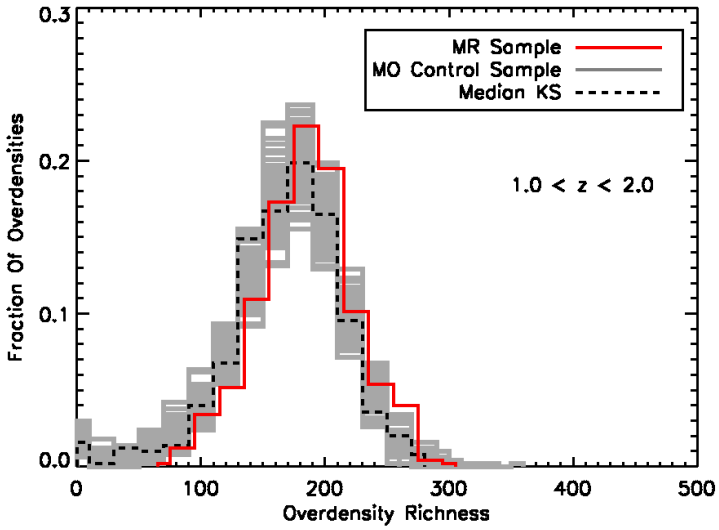
Figure 63: *Galaxy overdensity richness distributions, R and RO sample*. The top panel refers to $0 \leqslant z < 0.7$, middle panel to $0.7 \leqslant z < 1$, and the bottom panel to $1 \leqslant z \leqslant 2$. The red solid line refers to the R sample, and the grey lines correspond to the 100 extractions of the RO Control sample. The dashed black line is the richness distribution of the Control sample extraction corresponding to the median value of the KS probability value distribution.



(a)



(b)



(c)

Figure 64: *Galaxy overdensity richness distributions, samples MR and MO.* The top panel refers to $0 \leqslant z < 0.7$, middle panel to $0.7 \leqslant z < 1$, and the bottom panel to $1 \leqslant z \leqslant 2$. The solid red line refers to the MR sample, the grey lines correspond to the 100 extractions of the MO Control sample. The dashed black line is the richness distribution of the Control sample extraction corresponding to the median value of the KS probability value distribution.

6.5 THE ENVIRONMENT OF RADIO AGNS

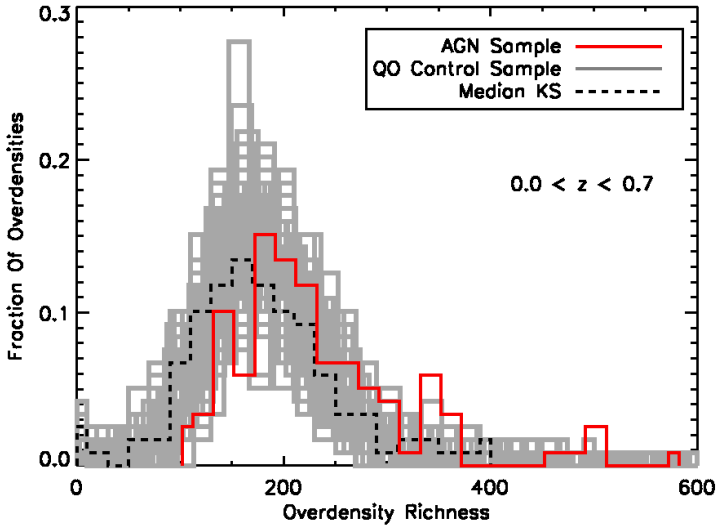
The main focus of this study is the environment around radio AGN sources in the VLA-COSMOS survey field. This task has also been performed in a similar way in other works (Best, 2000; Hatch et al., 2014, see e. g.). In this thesis, a general approach is adopted and the results are in good agreement with the literature.

The radio-AGN sample for environment definition has been extracted from the whole population of radio sources in the way described in section 6.2. By using a cut in sSFR, only those sources whose radio emission is due to AGN activity were selected and their environment determined using the O sample as Tracer sample.

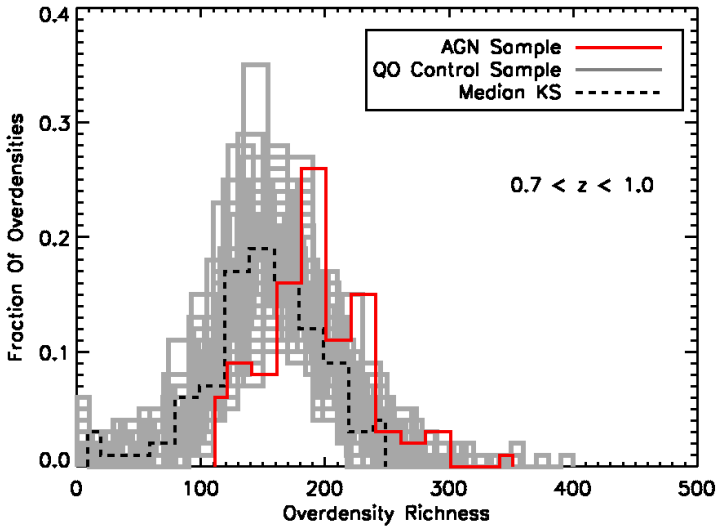
Nevertheless, the method used to isolate radio-AGNs has the effect of also selecting radio sources hosted by massive and quiescent galaxies. These two kinds of objects are both naturally found in high-density environments, therefore leading to a possible bias in the conclusions drawn. In fact, any difference in the environment around radio-AGNs compared to optical galaxies could actually be due to quiescent galaxies usually being found inside overdensities or to massive galaxies being more clustered. To not be affected by these problems, the QO Control sample has been selected carefully to avoid this bias. Galaxies in the QO sample were extracted in the same sSFR range and with the same stellar mass distribution of the AGN sample. In this way, every difference in the environmental properties can be ascribed to the presence of radio-AGN phenomena.

The results of the comparison of the overdensity richness distributions are shown in the third line of Table 10. It can be seen that the environment around radio-AGN sources is significantly different from the environment around normal galaxies, with the vast majority of the extractions of the QO sample being under the significance threshold. Also the median values confirm this trend, since they are always under 5%. This is the most important result of this Chapter, by showing a definite correlation between environment and the presence of radio-AGN phenomena.

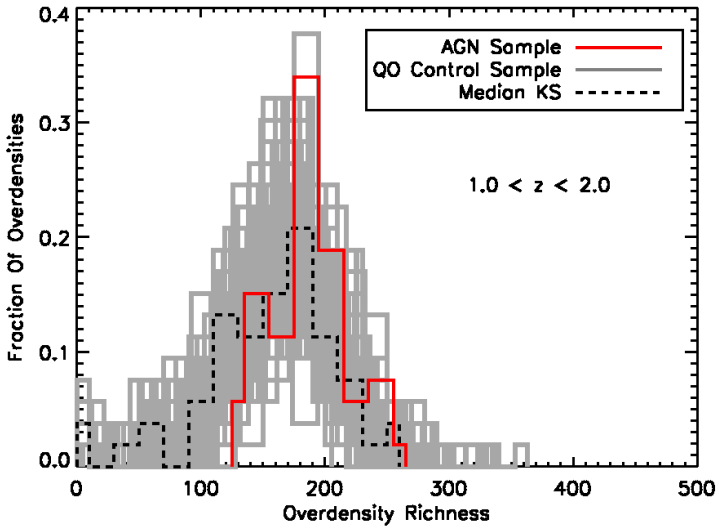
Figure 65 shows the overdensity richness distributions for the AGN sample and for the 100 extractions of the QO Control sample. It can be clearly seen that overdensities around radio-AGNs are distributed at higher richness values than those around normal galaxies. It is therefore possible to conclude that the environment around radio-AGNs is significantly richer than around normal galaxies, with this effect being due to radio-AGN presence rather than to radio-AGNs being found primarily in massive or quiescent galaxies (see e. g. Hickox et al., 2009; Ledlow and Owen, 1996). In other words, galaxies in denser environments show an enhanced probability of hosting a radio-AGN than galaxies in less dense environments.



(a)



(b)



(c)

Figure 65: *Galaxy overdensity richness distributions, samples AGN and QO.* The top panel refers to $0 \leqslant z < 0.7$, middle panel to $0.7 \leqslant z < 1$, and the bottom panel to $1 \leqslant z \leqslant 2$. The solid red line refers to the AGN sample, the grey lines correspond to the 100 extractions of the QO Control sample. The dashed black line is the richness distribution of the Control sample extraction corresponding to the median value of the KS probability value distribution.

As a check on the results, the environment around a sample of radio-emitting star-forming galaxies was also estimated, as selected in the upper right-hand quadrant of Figure 61. These sources, characterised by $\log(M^*) \geqslant 10$ and $\log(SSFR/yr^{-1}) \geqslant -11$, were compared with a Control sample extracted in the same sSFR region and with the same stellar mass distribution. An environmental effect has been found in this sample, too, at all redshifts, with star-forming radio sources in higher density environments. The median values of the KS probability value distribution, though, are higher by an order of magnitude than those in the last line of Table 10, except in the highest redshift bin, even if the sample of star-forming radio sources is ~ 10 times greater than the AGN one (450 star-forming sources against 53 AGNs). Although the great inequality in statistics between radio-AGNs and star-forming sources surely plays a major role in determining the significance of the environmental effect for the latter kind of objects, the cause of this is also probably residual differences in the SFR between the star-forming radio sources and the Control sample ones. This is proven by the fact that, although they are situated in the same SFR range, the star-forming radio sources are detected by the VLA instrument while the Control sample galaxies are not. Therefore the SFR is higher in the radio-detected sample compared to the Control one, or there could still be some residual contamination from AGN emission, which induces the environmental effect on these sources.

6.5.1 *High-power and low-power radio sources*

It has been suggested (seee. g. Chiaberge et al., 2010; Miley and De Breuck, 2008) that low-power radio-AGNs are often found in high-density regions (see also Bardelli et al., 2009; Castignani et al., 2014b; Donoso et al., 2010). In fact, low-power radio-AGNs are used to detect protoclusters at high redshifts (see e. g. Cooke, Pettini, and Murphy, 2012; Wylezalek et al., 2013). In this section we explore whether the environmental density actually depends on the radio power. To do this, each sample was divided between a high-power sub-sample ($\log(L_{1.4GHz}) \geqslant 24.5$) and a low-power one ($24 \leqslant \log(L_{1.4GHz}) < 24.5$). The reason for the lower limit at $\log(L_{1.4GHz}) = 24$ is to avoid the effect of the flux limit of the VLA-COSMOS survey. Such an effect is visible in Figure 66, which shows the dependence on the redshift of the radio luminosity for the R sample.

It can be seen that only the most luminous sources will be detected at high redshifts, owing to the survey limiting flux (effect known as the *Malmquist Bias*). This could lead to a bias in the results based on the sample, since more luminous objects could represent a different population of radio sources than do less luminous ones. This problem can be solved by considering only those sources that are more

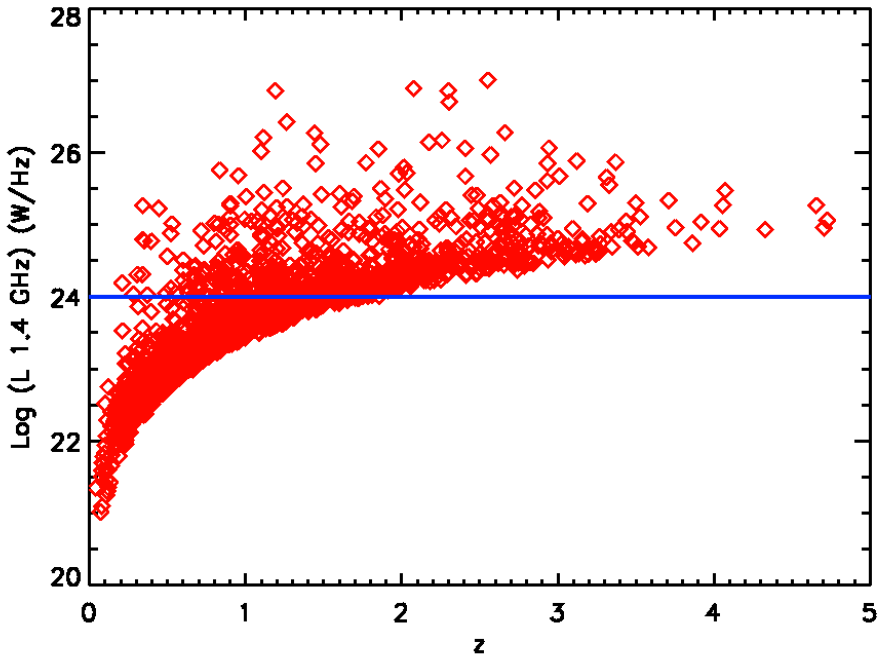


Figure 66: *Redshift dependence of radio luminosity*. Only the R sample is represented.

SAMPLE	$0 \leqslant z < 0.7$	$0.7 \leqslant z < 1$	$1 \leqslant z \leqslant 2$
AGNH and QOH	6	16	19
AGNL and QOL	16	30	17

Table 11: Number of galaxies in each sample for every redshift bin. High-power and low-power distinction.

luminous than $\log(L_{1.4\,GHz}) = 24$. In this way, a complete sample of radio sources up to $z = 2$ can be created.

It is worth noting that the limit at $\log(L_{1.4\,GHz}) = 24.5$, set to divide the high-power and low-power radio sub-samples, roughly corresponds to the canonical division between FRI and FRII objects (Fanaroff and Riley, 1974).

Therefore the samples AGNH and QOH (where H in the sample name stands for high power) and the samples AGNL and QOL (where L stands for low power) have been created. The number of galaxies in each sample is shown in Table 11.

The results of the analysis of the environment in relation to the distinction in radio power are now discussed. In particular, in Table 12 the number of Control sample extractions below the significativity level of 0.05 are reported, together with the median values of the KS probability value distribution.

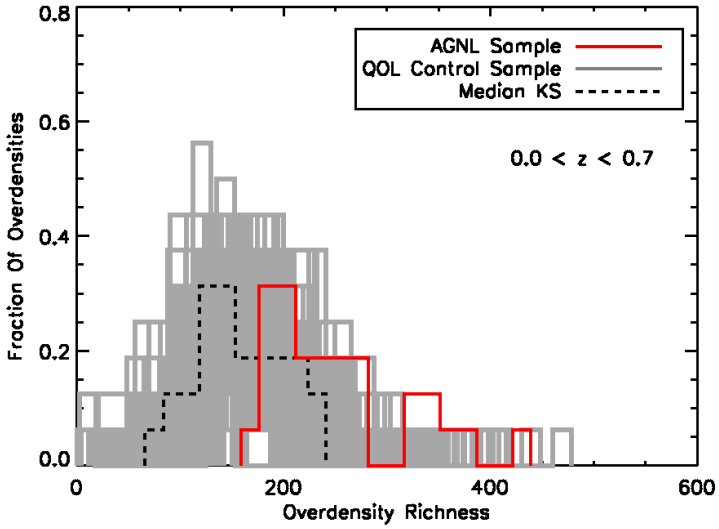
It is possible to see that the high-power and low-power samples have different KS probabilities: the environmental effect remains present

SAMPLE	$0 \leqslant z < 0.7$	$0.7 \leqslant z < 1$	$1 \leqslant z \leqslant 2$
AGNH *vs* QOH	9 (0.08)	43 (0.07)	17 (0.25)
AGNL *vs* QOL	93 (0.002)	99 (0.0006)	37 (0.08)

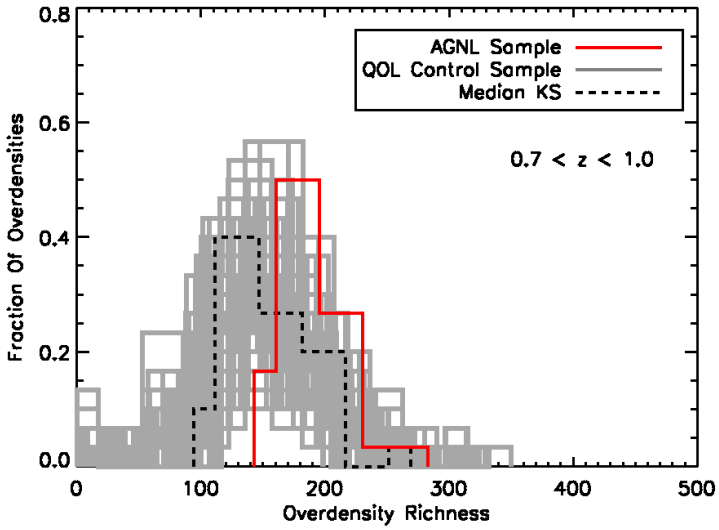
Table 12: Number of extractions of the various Control samples that are below the significance threshold of 0.05. In brackets, the median value of the KS probability value distributions is given. Low-power and high-power radio source distinction.

only for low-power radio sources. In fact, in Table 12 it can be seen that the number of extractions below the significance level is fairly low for the-high power radio-AGNs, never reaching even 50% of the total, and the median KS probability value is always greater than 0.05. On the other hand, for low-power radio-AGNs the number of extractions below the significance threshold is close to 100%. Only in the highest redshift bin the value is very low (37%), but it is nevertheless higher than the corresponding value in the same redshift bin for the high-power AGNs. In Figure 67 overdensity richness distributions are shown for the case of the AGNL sample and for the extractions of the corresponding QOL Control sample. That the distributions of the AGNL sample are systematically shifted towards higher richness values confirms the trend.

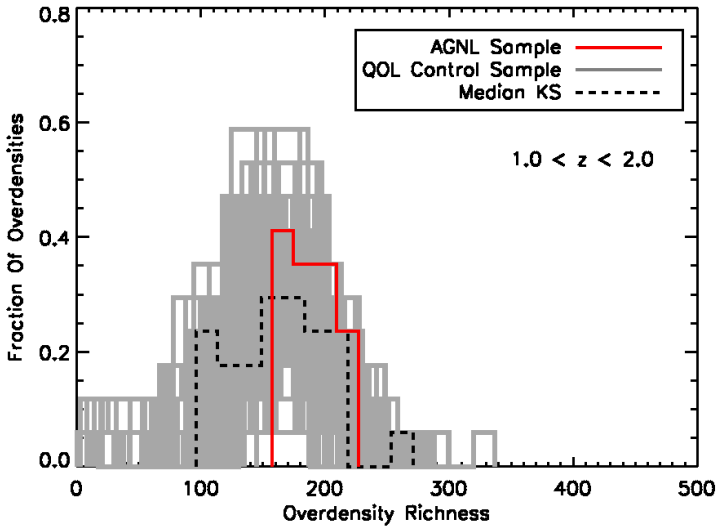
This is another important result of this Chapter: when considering the environment of radio sources, low-power AGNs are usually found in high-density environments. Therefore, higher overdensity richness values enhance the probability that a galaxy hosts a low-power radio-AGN. This result is not unexpected. In fact, it could arise as a direct consequence of different accretion modes for low-power and high-power AGNs. In particular, theory predicts that low-power radio-AGNs are fuelled by hot gas (such as the one found at the centre of clusters where cooling flows are present), whereas high-power AGNs are fuelled by cold gas (see for example Hardcastle, Evans, and Croston, 2007, and references therein). If this were true, then low-power AGNs (fuelled by hot gas) would be expected to reside at the centre of rich overdensities, such as clusters and groups (where hot gas is present). On the other hand, high-power AGNs would not be found in high-density environments, both for the lack of cold gas necessary for accretion and for the inefficiency of wet mergers (due to the high velocity dispersions of galaxies at the centre of clusters and rich overdensities), which are the trigger mechanism that has been proposed for the ignition of high-power radio-AGNs. The results presented in this Chapter seem to support this view of AGN and galaxy formation.



(a)



(b)



(c)

Figure 67: *Galaxy overdensity richness distributions, samples AGNL and QOL.* The top panel refers to $0 \leqslant z < 0.7$, the middle panel to $0.7 \leqslant z < 1$, and the bottom panel to $1 \leqslant z \leqslant 2$. The solid red line refers to the AGNL sample, the grey lines correspond to the 100 extractions of the QOL Control sample. The dashed black line is the richness distribution of the Control sample extraction corresponding to the median value of the KS probability value distribution.

6.6 CORRELATION WITH KNOWN CLUSTER AND GROUP CATALOGUES

In the previous sections it has been found that radio-AGNs live in environments that are on average denser than those of normal galaxies. It has therefore been explored whether these overdensities are related to virialised or bound objects. Two compilations of clusters and groups from the literature have been used: that of Finoguenov et al. (2007), found in the COSMOS field analysing deep Chandra and XMM observations (see references in the paper), and the zCOSMOS group catalogue (Knobel et al., 2012). For this comparison, only the two most distant redshift bins of the AGN sample have been considered ($z \in [0.7-1.0[$ and $z \in [1.0-2.0]$). At those redshifts, a 0.5 Mpc value corresponds approximately to 71 ($z = 0.7$), 63 ($z = 1.0$) and 61 ($z = 2.0$) arcsec. However, the largest source of uncertainties comes from the large error in the photometric redshifts.

To be conservative, all objects that are within a radius of ~ 60 arcsec and $|\Delta z| < 0.12$ for $0.7 \leqslant z < 1.0$ and ~ 60 arcsec and $|\Delta z| < 0.2$ for $1.0 \leqslant z \leqslant 2.0$ of the cluster locations reported in the two catalogues have been considered, in order to take the varying error in the photometric redshifts estimate into account.

In Table 13, the 26 associations with the Finoguenov et al. (2007) sample are listed, while in Table 14 those with the zCOSMOS group catalogue of Knobel et al. (2012) are reported. It has been found that $\sim 10-20\%$ of the AGN sample is in groups and clusters. Furthermore, no significant difference between the radio power distributions of the radio AGNs in clusters and groups and of the radio sources presented in this analysis is detected. The only noticeable source is the one with $\log(L_{1.4\,GHz}) = 26.86$ in a cluster in the most distant redshift bin, which is also the brightest object in the AGN sample.

For the X-ray luminosities, there is a slight tendency of clusters hosting a radio-AGN at $z \in [0.7-1.0[$ to be brighter, with a KS probability of being extracted by the same population of ~ 0.09. However the statistics is not sufficient to say anything for the highest redshift bin.

In the case of zCOSMOS groups, the richness distributions are significantly different, with groups hosting a radio-AGN being richer (KS probability of $\sim 10^{-3}$). This also holds when considering only groups with more than three elements.

6.7 MASS FUNCTIONS

It has also been investigated whether there is a relative evolution of the GSMF of radio-AGNs and those of normal galaxies. To perform this task, stellar masses as determined in the UltraVISTA Survey (Ilbert et al., 2013; McCracken et al., 2012) are used. A limiting magnitude for

RA_{cl}	dec_{cl}	z_{cl}	D	Δz	L_X	$\log(L_{1.4GHz})$
			$0.7 \leqslant z < 1$			
150.41386	1.84759	0.969	40.176	0.015	$2.9883 \cdot 10^{43}$	25.68
150.27736	2.05303	0.908	41.112	-0.0521	$9.6268 \cdot 10^{42}$	23.76
150.02382	2.20323	0.942	7.1856	0.0988	$1.9325 \cdot 10^{43}$	23.35
149.64966	2.20925	0.954	0.120744	-0.0165	$1.7111 \cdot 10^{43}$	24.34
150.21454	2.28010	0.881	26.3124	-0.0419	$1.3734 \cdot 10^{43}$	24.59
149.95262	2.34188	0.942	0.0144	0.0156	$1.5213 \cdot 10^{43}$	23.47
149.95262	2.34188	0.942	41.76	-0.0002	$1.5213 \cdot 10^{43}$	24.16
149.92926	2.40902	0.874	0.0072	-0.0044	$8.4425 \cdot 10^{42}$	23.58
149.63988	2.34912	0.951	41.112	-0.0369	$3.2975 \cdot 10^{43}$	24.06
149.62355	2.39918	0.845	0.108072	0.0051	$1.9522 \cdot 10^{43}$	24.31
150.15298	2.39447	0.899	26.01	0.0193	$9.4442 \cdot 10^{42}$	23.64
149.66927	2.47365	0.957	11.4408	-0.0044	$1.8343 \cdot 10^{43}$	23.92
149.56174	2.42195	0.846	0.0144	0.0332	$1.7645 \cdot 10^{43}$	24.20
150.00713	2.45343	0.731	0.1476	0.0229	$5.2795 \cdot 10^{42}$	24.15
149.92343	2.52499	0.729	30.8016	0.0016	$1.1257 \cdot 10^{44}$	23.74
150.10533	2.72392	0.727	2.39364	-0.0023	$6.9659 \cdot 10^{42}$	23.89
149.91772	2.70088	0.889	3.00132	-0.0355	$2.1160 \cdot 10^{43}$	24.28
150.58397	2.32155	0.720	46.116	0.0182	$4.2323 \cdot 10^{42}$	23.39
150.05057	2.13923	0.959	1.09008	0.1189	$5.7184 \cdot 10^{42}$	24.19
			$1 \leqslant z \leqslant 2$			
150.76245	1.79362	1.258	10.4004	-0.0051	$5.5689 \cdot 10^{43}$	26.43
149.51855	2.09959	1.382	48.096	0.1839	$7.8229 \cdot 10^{43}$	24.00
150.59309	2.53890	1.045	7.056	-0.1063	$3.1937 \cdot 10^{43}$	23.63
149.59763	2.44004	1.168	10.4076	-0.023	$2.6946 \cdot 10^{43}$	26.86
150.57024	2.49864	1.146	17.586	-0.0083	$2.5386 \cdot 10^{43}$	24.13
150.12646	1.99926	1.019	13.644	-0.1744	$1.2897 \cdot 10^{43}$	23.65

Table 13: *Candidate associations of radio AGNs and X-ray clusters from the catalogue of Finoguenov et al. (2007).* RA_{cl}, dec_{cl} and z_{cl} are the coordinates of the cluster centre. D is the distance between the cluster centre and the source of the AGN sample considered on the plane of the sky (in arcsec), while Δz is the distance in redshift between the two. L_X and $\log(L_{1.4GHz})$ are the X-ray (in the band $[0.1 - 2.4]$ keV, expressed in $\frac{erg}{s \times cm^2}$) and 1.4 GHz luminosities (in $W \times Hz^{-1}$).

RA_{gr}	dec_{gr}	S/P	z_{gr}	D	Δz	$\log(L_{1.4GHz})$
			$0.7 \leqslant z < 1$			
149.920567	2.521800	12/82	0.7297	17.6328	0.0023	23.74
150.005475	2.451841	12/49	0.7311	8.1792	0.023	24.15
150.212666	2.281762	7/14	0.8812	18.7812	-0.0417	24.59
149.552005	2.423054	5/25	0.8446	35.2692	0.0318	24.2
149.914194	2.694681	5/54	0.8898	55.008	-0.0669	24.05
149.914194	2.694681	5/54	0.8898	28.5552	-0.0347	24.28
150.101070	2.268664	3/4	0.6855	47.16	-0.0547	23.36
150.094158	2.063360	3/7	0.725	53.064	0.01	23.61
150.580055	2.328653	3/6	0.7274	28.3644	0.0256	23.39
150.023066	2.516794	3/2	0.7473	31.554	0.0314	24.92
150.508547	2.653296	3/7	0.8102	34.056	0.0103	25.02
149.569460	2.418630	3/107	0.8529	30.258	0.0401	24.2
150.515963	2.005564	3/11	0.8797	15.822	0.0089	24.29
149.788439	2.757904	3/10	0.9093	24.6348	-0.0272	24.16
150.475201	1.626689	3/1	0.9713	32.6952	-0.0236	23.64
150.098295	2.056180	2/3	0.6375	47.34	-0.0775	23.61
150.017350	2.442392	2/5	0.6681	54.288	-0.04	24.15
149.725800	2.770682	2/7	0.7032	42.48	-0.0508	23.98
149.838840	1.683059	2/7	0.747	0.126828	0.0067	23.3
149.731350	2.759319	2/6	0.7655	18.3024	0.0115	23.98
150.231175	2.072644	2/1	0.7995	40.212	-0.0594	23.61
150.465750	2.423217	2/4	0.8266	9.9792	-0.0006	24.4
150.410760	1.801465	2/5	0.8456	13.6152	0.0372	24.77
149.985860	2.325898	2/3	0.849	32.7384	-0.0764	23.94
150.091666	2.599807	2/4	0.8929	20.2572	-0.0961	24.07
			$1 \leqslant z \leqslant 2$			
150.216932	2.273971	2/5	0.893	44.424	-0.0299	24.59
150.445160	1.845406	2/8	0.8943	20.6964	0.0232	24.6
150.023318	2.205914	2/12	0.9409	5.7276	0.0977	23.35
149.963878	2.363113	2/10	0.9454	50.04	0.0032	24.16
149.649660	2.209250	2/7	0.9539	0.120744	-0.0166	24.34
149.843245	2.573333	2/8	0.9657	42.408	-0.0285	25.08
149.633185	2.457927	3/2	1.1688	40.716	-0.0832	25.06
150.128820	1.922544	2/1	1.0083	53.172	-0.0832	23.64

Table 14: *Candidate association between radio-AGNs and groups from the catalogue of Knobel et al. (2012).* RA_{gr}, dec_{gr} and z_{gr} are the coordinates of the group centre. S/P is the number of sources with spectroscopic and photometric redshift respectively, D and Δz are the distances from the group centre to the position of the radioAGN source on the plane of the sky and along the redshift direction respectively. $\log(L_{1.4GHz})$ is the 1.4 GHz luminosity (in $W \times Hz^{-1}$).

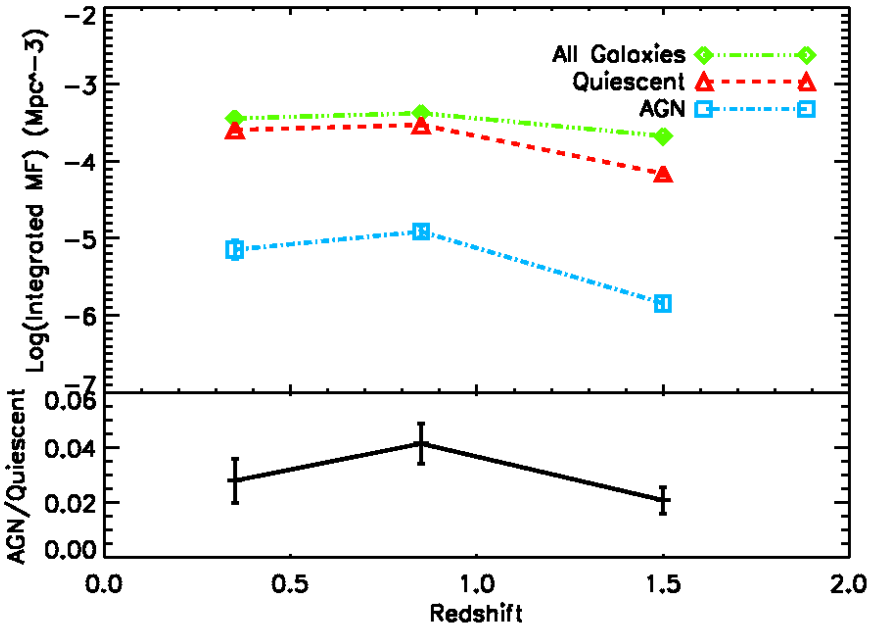


Figure 68: *Redshift evolution of the integrated GSMFs*. Upper panel: value of the integrated GSMFs. The curves refer to all UltraVISTA galaxies (green diamonds) and to a selection of only the quiescent ones (red triangles) and the radio-AGN sample (cyan squares). Lower panel: ratio of the integrated GSMF values for AGN hosting and quiescent galaxies. Points have been placed at values of redshift corresponding to the mean redshift of each redshift bin.

the UltraVISTA Survey of $K_S = 24$ is assumed in the data. Moreover, the radio power limit of $\log(L_{1.4\,GHz}) \geqslant 24$ previously introduced has also been applied to the AGN sample (leading to a total of 88 sources), since lack of completeness in the radio data could bias the estimate of the redshift evolution of the GSMFs.

GSMFs are computed with the non-parametric $1/V_{max}$ estimator (Avni and Bahcall 1980, see also Bolzonella, Pelló, and Maccagni 2002 and references therein for further information). For radio objects, the V_{max} is calculated by taking the different radio and K_S limiting magnitudes of the surveys into account. In particular, the minimum between the V_{max} set by radio and optical data was used. GSMFs were calculated for all UltraVISTA galaxies, a sub-sample of quiescent ones (selected through a cut to $\log(SSFR/yr^{-1}) \leqslant -11$) and the AGN sample. The size of the AGN sample is too small to draw conclusions as a function of mass, therefore only the integral of the stellar mass functions for galaxies more massive than $\log(M^*) = 11$ is considered. (This new cut further reduces the sample to 71 sources.) These results are shown in Figure 68. Error bars were computed by simply summing all the $1/V_{max}^2$ terms due to Poisson statistics, as explained in Bolzonella, Pelló, and Maccagni (2002).

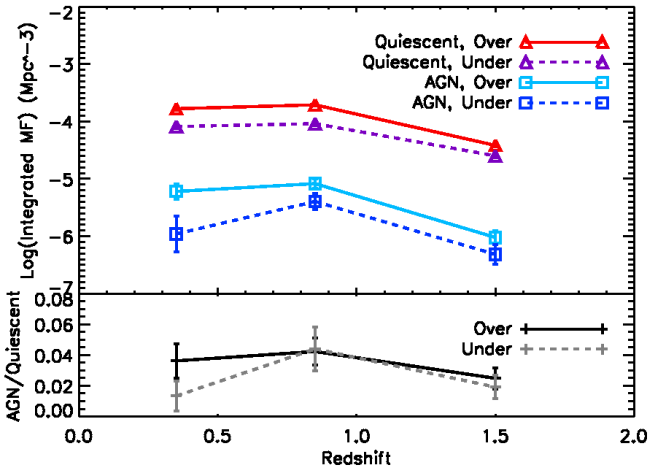


Figure 69: *Environment effect on the integrated* GSMFs. All samples have been split in two, as stated by the legend, with the terms "Over" and "Under" referring to high-density environments (over the median value) and low-density ones (under the median value). Upper panel: red and purple triangles refer to quiescent UltraVISTA galaxies, cyan and blue squares to the AGN sample. Lower panel: black and grey curves represent the ratio of AGN hosting to quiescent galaxies. Solid lines are for high-density environments, dashed lines for the low-density ones. Points have been placed at values of redshift corresponding to the mean redshift of each redshift bin.

Although the integrated GSMF for all galaxies varies, within a factor ~ 2, between the redshift bins $[1.0-2.0]$ and $[0.7-1.0[$ the corresponding variation for the integrated GSMF of quiescent and radio-AGNs is even stronger (a factor ~ 5 and almost an order of magnitude for the two samples respectively). In these same redshift bins the ratio of radio-AGN hosting to quiescent galaxies increases from ~ 0.02 to ~ 0.04. This increase could confirm that the presence of a radio-AGN is a phenomenon that evolves with quiescent galaxies, but at a different rate.

To explore whether the ratio of radio-AGNs to quiescent galaxies also evolves as a function of the local density, all samples have been divided in two, considering objects above and below the median of the overdensity richness distributions of all galaxies as obtained with the fixed aperture method. With these cuts, the number of objects in the AGN sample in the high/low density regions are, respectively, 13/3, 33/11,17/10 for the low, middle, and high redshift bins. For the quiescent sample, the statistics are, instead, always in the 2000-3000 range in every redshift bin.

As expected, the objects from the AGN sample are systematically more numerous above the median value than below. This could indicate the effect previously discovered in a more significant way of radio-AGNs being found in higher density environments. The effect

of the environment on the integrated mass functions and on the ratio of AGN-hosting to quiescent galaxies is shown in Figure 69.

The ratio of AGN-hosting galaxies to quiescent ones shows that the environmental segregation is present. In particular the ratio is higher in high-density environments than in low-density ones at $z \leqslant 0.7$. In the first redshift bin there is a factor ~ 4 difference between high- and low-density environments. This seems to indicate that the environmental segregation is somewhat more important for radio AGNs. Also, while the evolution is similar, both the integrated mass function of the AGN sample and the ratio of AGN to quiescent galaxies show an enhanced probability that a galaxy hosts a radio AGN in high-density environments in the lowest redshift bin. This result is in very good agreement with what has been found in previous sections with other methods.

6.8 CONCLUSIONS AND SUMMARY

In this Chapter, the role of the environment in determining the presence and the properties of radio-AGN phenomena has been investigated. AGN phenomena are predicted and needed by the theory of galaxy formation, therefore understanding the way environment influences their development could lead to a better comprehension of the processes through which galaxies are formed. The deep photometric redshift sample extracted from the COSMOS survey (Ilbert et al., 2009) was used, together with the VLA-COSMOS survey (Schinnerer et al., 2007) for the radio data.

From these surveys, various samples were extracted, with particular focus on AGN samples. Results may be summarised as follows:

1. Analysing the R sample and the RO Control sample, it was found that the environment where radio sources reside is significantly denser than the environment around galaxies without radio emission.

2. These results are not due to radio sources being hosted by high-mass galaxies (which in turn are known to reside in denser environments). In fact, by repeating the environment estimate on samples designed to have the same stellar mass distributions (MR and MO samples) the effect of radio sources residing in systematically and significantly denser environments still holds.

3. A sample of radio-AGNs was extracted from the whole catalogue of radio sources. The environment around radio-AGNs is significantly denser than the environment around normal galaxies in the same stellar mass and specific star formation rate range (QO sample). This shows that the environmental segregation that was found is actually due to an enhanced probability that

galaxies host a radio-AGN in denser environments. This led to concluding that environment plays a role in determining the onset of AGN phenomena in galaxies and that these phenomena are important in the process of galaxy formation.

4. The environmental effect in enhancing the probability of a galaxy hosting a radio-AGN is dominated by low-power radio-AGNs. In fact, the difference in the environment between low-power radio-AGNs and normal galaxies is significant, while for high-power radio-AGNs it is not. This was found by splitting all the previously analysed samples according to a cut in their radio luminosity. This effect of low-power radio-AGNs showing the only signal of environmental segregation could be related to the thermodynamic properties of the gas fuelled to the AGN (see e. g. Hardcastle, Evans, and Croston, 2007).

5. After analysing the integrated GSMFs, calculated for the whole population of the UltraVISTA galaxies and sub-samples of only quiescent and radio AGN-hosting ones, hints of a redshift evolution can be found. This can be seen both in the decrease in the value of the integrated GSMFs with redshift for the samples of quiescent galaxies and radio-AGNs, as well as in the same decrease with redshift in their ratio.

6. Exploring the dependence from the environment of the integrated GSMFs, it is found that the fraction of radio-AGNs is always enhanced in high-density environments, in agreement with the stronger results found using richness distributions. Moreover, radio-AGNs are also the population with the greatest difference in the values of integrated mass functions between high- and low-density environments.

The picture that emerges from this work is that a high-density environment is indeed related to the presence and the properties of radio-AGNs. This is in good agreement with the current theory of galaxy formation, which predicts that galaxies in high density environments undergo star-formation quenching faster and in a more efficient way. Future studies and the increase in the statistics for the mass functions could open the way to more thorough analysis of the existing correlations.

7 CONCLUSIONS

The work presented in this thesis constitutes an analysis, based on observations, of the complex interplay between galaxy environment, stellar mass, SFR, and radio-AGN presence. Galaxy environment (both at the scale of local galaxy overdensities and in terms of LSS features such as filaments) is thought to deeply affect galaxy evolution and in particular to be related to star-formation quenching and the build-up of stellar mass. In this thesis, the relation between environment and galaxy evolution has been studied through the use of both spectroscopic redshift data sets and of photometric redshift surveys in a large redshift range ($0.2 \leqslant z \leqslant 3$), carefully taking into account the effect of observational biases.

The analysis presented in this thesis explores the relation between environment and galaxy evolution under many points of view. Nevertheless, the common thread among all these studies is the quenching of the star-formation in galaxies as a function of their mass and environment. This has been investigated both in terms of a specific physical process (e. g. radio-AGN activity) or of a specific type of environment (e. g. the filaments of the LSS) as well as in terms of a statistical analysis of how the galaxy population is distributed in different galaxy types, mass ranges, and local densities (by studying the GSMFs in different environments). The main results of this thesis are summarised in the following.

LOCAL ENVIRONMENT WITH PHOTOMETRIC REDSHIFTS

In Chapter 3 mock galaxy catalogues have been used to check whether it is still possible to measure the GSMF in different environments if the density field is estimated using photometric redshifts instead of spectroscopic ones. The results of this work are presented in the paper Malavasi et al. (2016a). In particular, it was found that:

- It is possible to accurately measure local environment only with high-precision photometric redshifts ($\sigma_{\Delta z/(1+z)} \lesssim 0.01$).
- In case a fixed aperture method is adopted for the measurement of the local density field, the parameters defining the volume (i. e. the radius of the cylinder and the length of the cylinder in the radial direction) have to be carefully tuned to lower the uncertainty on the environment measurement. A length of the cylinder in the radial direction of the same size as the 3σ photo-

metric redshift uncertainty and a radius for the cylinder corresponding to the physical scale explored are the best option.

- Environments at smaller scales ($R = 0.3$ Mpc) are reconstructed better than environments at larger scales ($R = 2$ Mpc).
- If these prescriptions are applied to high-precision photometric redshifts ($\sigma_{\Delta z/(1+z)} = 0.01$), any difference that is present between GSMFs in high- and low-density environments (when the environment is measured using each galaxy's true redshift) will be damped when switching to photometric redshifts, but it will still be recoverable. In addition, no spurious differences between the GSMFs in high- and low-density environments seem to be created.

THE SHAPE OF THE GSMF IN DIFFERENT ENVIRONMENTS

In Chapter 4 the results of the previous Chapter have been applied to the high-precision photometric redshift data set of the UltraVISTA survey in the COSMOS field. The GSMFs for quiescent and star-forming galaxies up to $z = 3$ have been measured and their shape in different local environments has been compared. The results of this work are presented in the paper Malavasi et al. (2016b). In particular, it was found that:

- High-density environments are dominated by a population of massive ($M^* \gtrsim 10^{11} M_\odot$), quiescent galaxies.
- Less massive ($M^* \lesssim 10^{11} M_\odot$), star-forming galaxies are more frequent in low-density environments.
- The mass at which galaxies become quiescent at a given redshift is lower in high-density environments compared to low-density ones. In high-density environments, the mass at which the quiescent GSMF starts to dominate over the star-forming GSMF is a monotonically increasing function of redshift.
- The effects of local environment on the GSMFs are visible up to $z \sim 2$ and for physical scales in the range $0.3 - 2$ Mpc.

GALAXY PROPERTIES IN RELATION TO THEIR POSITION WITHIN FILAMENTS

In Chapter 5 the DisPerSE algorithm has been applied to the VIPERS survey, to recover the network of filamentary structures of the CW in the survey volume at $0.5 \leqslant z \leqslant 0.85$. The performance of the method given the survey layout and observational properties has been extensively tested on simulations of the survey volume. The distribution

of galaxy distances from filaments has been measured for galaxies in various mass bins (for masses $M^* \geqslant 10^{10-10.5} M_\odot$) and for star-forming and quiescent galaxies separately. The results of this work are presented in Malavasi et al. (2016c). In particular, it was found that:

- More massive galaxies ($M^* \geqslant 10^{11} M_\odot$) are closer to the core of filaments with respect to less massive galaxies.
- Passive galaxies are closer to the filament centres with respect to star-forming galaxies. This effect is visible also at fixed stellar mass, therefore suggesting that filaments directly affect the SFR of galaxies.
- These results hold even after down-weighting the contribution of peaks of the density field and are due to galaxies being located within the anisotropic environment of the CW.

THE ENVIRONMENT OF RADIO SOURCES AND RADIO-AGNS

In Chapter 6 the environment around a sample of radio sources and radio-AGNs extracted from the VLA-COSMOS survey has been analysed and a comparison with the environment measured around control samples of galaxies with no radio emission has been performed. The results of this work are presented in the paper Malavasi et al. (2015). In particular, it has been found that:

- Radio sources and radio-AGNs are located in environments that are significantly denser than those around galaxies having the same mass and SFR, but no sign of radio emission.
- The signal of environmental segregation is present only for low-power radio-AGNs ($24 \leqslant \log(L_{1.4GHz}) \leqslant 24.5$) and not for high-power ones.
- Environmental effects are visible up to $z \sim 2$.

7.1 A CONSISTENT VIEW OF GALAXY EVOLUTION

Throughout their lifetime, galaxies experience changes in their environment, which reflect on transformations in their properties. This work provides observational evidence in support of a picture in which galaxy evolution proceeds in parallel with the evolution of cosmic structures.

As already detailed in Chapter 1, it is widely accepted that galaxies form from the gravitational collapse of baryonic matter, that infalls and cools in the potential wells set by the DM distribution. Cold

flows can penetrate deep into the hot gas-filled DM haloes and sustain intense star-formation in primordial disks. Although the relative importance of the infall of gas cooling from the halo, the accretion of gas directly from the CW through cold flows, and the contribution of wet merger episodes is still to assess, star-formation in protogalaxies is regulated from the availability of a cold gas reservoir. This same cold gas is thought to possibly trigger AGN activity, if it is allowed to reach the central SMBH. As a galaxy evolves, so does the environment around it, with the formation of the LSS features and of local high-density regions. Both these kind of environments can then affect the SFR of a galaxy and bring it to quiescence.

As suggested by the different shape of the GSMFs in high- and low-density environments and the mass at which the GSMF of quiescent galaxies starts to dominate over the GSMF of star-forming ones, galaxies that inhabit high-density regions experience an accelerated evolution: they undergo more intense star-formation at high redshift (thus becoming more massive), and their star-formation is quenched at earlier times. These environmental effects start to be visible already at $z \sim 2$, meaning that they have been affecting galaxies for a long time, since an epoch when many structures had still to completely form and many galaxies had still to become passive. Moreover, the signature of local environment is not completely diluted when considering scales as large as 2 Mpc, meaning that the effect of the local density field in shaping the galaxy population is strong enough to be visible even at scales larger than the typical scale of clusters/groups ($\sim$ 1 Mpc).

Radio-AGN phenomena are thought to be caused by the accretion of hot gas on to the SMBH (the so-called "radio mode") and they are regarded as the possible subsequent phase of SMBHs that have been accreting cold gas (the so-called "quasar mode"). AGNs experiencing "quasar mode" accretion are thought to be a possible quenching mechanism in massive galaxies. When the cold gas reservoir has been completely removed by "quasar mode" accretion, AGNs begin the "radio mode" phase, which impedes new infall of fresh gas and the onset of new star-formation episodes. Although these processes are commonly regarded as internal quenching mechanisms, radio-AGNs (connected to the "radio mode" feedback) are located in denser environments than galaxies with the same stellar mass and SFR but no sign of radio emission. This is an example of the relation between possible quenching mechanisms and local environment, which supports the scenario in which galaxies in high-density environments have a higher probability of ending their star-formation at earlier times.

Aside from local environment, galaxies are embedded in a network of much larger structures, the Cosmic Web, composed of vast empty regions (voids), surrounded by flattened structures (walls), framed by elongated, linear matter overdensities (filaments) at whose inter-

sections galaxy clusters are found. These various kinds of matter distributions are the result of the evolution of primordial, Gaussian density fluctuations and represent an evolving landscape in which galaxies are formed. In particular, matter flows from underdense regions, reaching walls. Walls wind up forming filaments, through which matter flows to reach galaxy clusters. As galaxies form in this evolving network of structures, they follow the same path, entering filaments in the perpendicular direction, then flowing through filaments and finally reaching galaxy clusters. As galaxy properties change with the distance from filament centres, the CW can also represent a likely source of galaxy pre-processing, thus turning a fraction of galaxies quiescent before they reach the the group or cluster environment.

7.2 FUTURE PERSPECTIVES

The galaxy evolution scenario depicted above has been deduced through a complex analysis of extensive data sets. At low redshift ($z \lesssim 1.5$), our understanding of the Universe mainly derives from the use of wide-area spectroscopic redshift surveys, such as e. g. the SDSS (York et al., 2000), GAMA (Driver et al., 2011), zCOSMOS (Lilly et al., 2007), and VIPERS (Guzzo et al., 2014). These surveys have allowed us to map galaxy environment in detail, thanks to large statistical samples and precise position (redshift) information. The study of galaxy environment (both local and LSS) necessarily relies on these two requirements. As the general picture of galaxy evolution and its relation to environment is more or less in place in the local Universe, it is now required to extend it at further distances, to investigate the earliest stages of galaxy (and structure) formation. Studying galaxy environment (both local and global) at high redshift is not a simple task.

In order to obtain large statistical data sets with precise redshift information, the spectra of large numbers of galaxies on wide sky areas have to be taken. To achieve this goal, spectroscopic surveys that make extensive use of Multi-Object Spectrographs (MOSs) have been already undertaken, while many more have been planned. Aside from the already mentioned SDSS, GAMA, zCOSMOS, and VIPERS, the Baryon Oscillation Spectroscopic Survey (BOSS)[1] (Dawson et al., 2013) has just been completed, providing the spectra at optical wavelengths of $\sim$ 1.5 million luminous galaxies up to $z \sim 0.7$ over an area of $10^4\,\mathrm{deg}^2$. In the next couple of years, also the Prime Focus Spectrograph (PFS)[2] instrument (Takada et al., 2014) at the Subaru telescope will become available, measuring the optical to NIR spectra of $\sim$ 2400 galaxies at once on a $1.3\,\mathrm{deg}^2$ field of view in the redshift range $1 \leqslant z \leqslant 2$, while the Multi-Object Optical and Near-infrared Spectrograph (MOONS) instrument at the Very Large Telescope (VLT) will mea-

1 https://www.sdss3.org/surveys/boss.php
2 http://pfs.ipmu.jp/index.html

sure ~ 1000 spectra simultaneously in the optical and NIR bands over a 500 sqarcmin field of view. Later, also the Hobby-Eberly Telescope Dark Energy Experiment (HETDEX)[3] (Hill et al., 2008) will become operational, measuring optical spectra of ~ 34 000 objects at the same time, obtaining redshifts for LAEs in the redshift range $1.9 \leqslant z \leqslant 3.5$. In the meanwhile, the Euclid survey (Laureijs et al., 2011) and WFIRST[4] (Content et al., 2013) will operate from space obtaining both photometry and slitless spectroscopy. In particular, Euclid will obtain spectroscopy of ~ 50 million star-forming galaxies (H_α emitters) in the redshift range $0.9 \leqslant z \leqslant 1.8$ over an area of $15\,000\,\mathrm{deg}^2$ (see e. g. Pozzetti et al., 2016).

On the other hand, instruments that make use of Integral Field Unit Spectrographs (IFUs) such as the Palomar Cosmic Web Imager (CWI)[5] (Matuszewski et al., 2010) are focused on the measurement of the faint emission from the IGM and can provide a view of the gaseous phase of the CW in the redshift range $1.5 \leqslant z \leqslant 4$, offering a complementary approach to existing techniques such as the Lyman-α tomography (see e. g. Lee et al., 2014).

Complementary to spectroscopic information, photometric surveys can be used to measure photometric redshifts of galaxies, which can still be employed to study galaxy evolution in relation to environment (although they are characterised by a larger uncertainty). For example, as observations of the same sky field reach deeper magnitudes and the number of filters in which galaxies are observed is increased, improved photometric redshifts can be derived for a given data set. For example, photometric redshifts in the COSMOS field have improved their precision from $\sigma_{\Delta z/(1+z)} = 0.06$ (Ilbert et al., 2009) to $\sigma_{\Delta z/(1+z)} = 0.01$ (Ilbert et al., 2013) and lower (Laigle et al., 2016), opening the possibility to perform more accurate studies of galaxy evolution and local environment as well as detections of LSS features. Moreover, new surveys such as the Advanced Large Homogeneous Area Medium Band Redshift Astronomical (ALHAMBRA) survey[6] (Moles et al., 2008) and J-PAS Benítez et al. (2015), by using several (up to more than 50) narrow-band filters, are able to provide very high-precision photometric redshifts (with $\sigma_{\Delta z/(1+z)} \sim 0.003 - 0.005$) only one-two orders of magnitude larger than redshifts derived from spectroscopic information, but for much larger samples (covering ~ $8000\,\mathrm{deg}^2$ in the redshift range $0.1 \leqslant z \leqslant 2$).

Although for other surveys (e. g. the photometric part of the Euclid and WFIRST surveys) the expected uncertainty for the measurement of photometric redshifts is larger ($\sigma_{\Delta z/(1+z)} \sim 0.05$), still the possibility to couple photometric redshift information with redshifts measured from the corresponding spectroscopic part and the large area covered

3 http://hetdex.org/
4 http://wfirst.gsfc.nasa.gov/index.html
5 http://www.srl.caltech.edu/sal/cosmic-web-imager.html
6 http://alhambrasurvey.com/

(of the order of 15 000 $\deg^2$ for Euclid) could make them suitable for the study of local environment and LSS, granting sufficiently large statistical samples for the study of the CW. In turn, the LSS measured from spectroscopic surveys could be used to improve the photometric redshift measurement for these surveys (e. g. with the "photo-web" method, see Aragón-Calvo et al., 2015). This could also be applied to surveys such as the combined Kilo-Degree Survey (KiDS)-VISTA Kilo-degree Infrared Galaxy survey (VIKING) data set[7] ($\sim$ 1500 $\deg^2$ observed in the optical and NIR bands, see Edge et al. 2013; de Jong et al. 2015) and to some extent to the VISTA Deep Extragalactic Observations (VIDEO) survey ($\sim$ 10 $\deg^2$ observed in the NIR bands, Jarvis et al. 2013) although there is the possibility that, due to the large photometric redshift uncertainty expected for all these surveys, only studies of local and global environment in 2D (in projection on the plane of the sky for sources in a large redshift bin) are possible. Nevertheless, these data sets have the advantage that they could be used in synergy with surveys performed in the radio domain with instruments such as the Square Kilometer Array (SKA)[8] (see e. g. Ciliegi and Bardelli, 2015).

In the near future, the number and size of both spectroscopic and photometric surveys will dramatically increase and vast data sets with precise positions for a large number of galaxies will allow us to obtain a detailed map of the Large Scale Structure of the Universe. As these data sets are coupled with an increased knowledge of the processes that regulate the onset and the end of star-formation in galaxies coming from the theoretical side, the problem of determining the role that environment plays in galaxy evolution will likely become close to a more complete solution.

7 KiDS: http://kids.strw.leidenuniv.nl/index.php
VIKING: http://www.astro-wise.org/projects/VIKING/index.shtml
8 https://www.skatelescope.org/

"Così tra questa
immensità s'annega il pensier mio:
e il naufragar m'è dolce in questo mare."
— Giacomo Leopardi,
L'Infinito

ACKNOWLEDGEMENTS

This thesis is the result of more than three years of work, during which I have learnt many things, some of which related to astronomy. I like to look at this manuscript as a milestone of the goals achieved in the past three years, both in science and in my personal life. The path has been long, hard, and devious, but luckily for me (and probably less for them) I had people that helped me walk it all the way to the end, and with which now I have the opportunity to share this moment and the great happiness and satisfaction that it brings.

As an exhaustive list of all the people I should thank for having reached this goal is beyond the possibility of this Section (:-)), I hope that, if somebody reading this looks for his name and does not find it, he/she will not be too upset. If for any reason you think I should be grateful to you for having got this far, then trust me: I surely am.

I thank my supervisors Andrea Cimatti, Lucia Pozzetti, Sandro Bardelli, and Olga Cucciati for their valuable help in obtaining the results described in this thesis. It has been a pleasure to work with you and you have taught me many things.

While staying in Marseille I had the opportunity to work with many people who gave me a precious help to greatly improve the work described here. I am particularly thankful to Olivier Ilbert, Sylvain de la Torre, Didier Vibert, Marie Treyer, Iary Davidzon, and Thibaud Moutard for their advices, their teachings, or simply the chats that we had. Special thanks go to Stéphane Arnouts, thanks to which I have learnt many of the things related to astronomy which I mentioned at the beginning of this Section.

Regarding the other things that I have learnt, those not related to science, I have to thank my family and Milena. They have been a constant support and presence. Mamma, Papà, Cecilia, grazie di esserci e di esserci stati in questi tre anni. Siete stati, siete e sempre sarete importanti per me. Grazie di cuore.

Milena, i motivi per dirti grazie sono e saranno sempre talmente tanti che non riuscirò mai a elencarli tutti. Grazie della tua costante e instancabile presenza, del tuo assiduo e immutabile sostegno. Ma soprattutto, grazie di nuotare al mio fianco in questo grande mare.

Last, but not least, I am thankful to the Lu Granola cookies “au gros éclats de chocolat” , the Coca-Cola Company, and a lot of beer breweries from all around the world for providing part of the biochemical fuel that allowed me to reach this important goal.

BIBLIOGRAPHY

Abadi, M. G., B. Moore, and R. G. Bower (1999). In: *MNRAS* 308, pp. 947–954.

Abbas, U. and R. K. Sheth (2006). In: *MNRAS* 372, pp. 1749–1754.

Alpaslan, M. et al. (2015). In: *MNRAS* 451, pp. 3249–3268.

Alpaslan, M. et al. (2016). In: *MNRAS* 457, pp. 2287–2300.

Annunziatella, M. et al. (2016). In: *A&A* 585, A160.

Aragón-Calvo, M. A. (2007). PhD thesis. Rijksuniversiteit Groningen.

Aragon-Calvo, M. A., M. C. Neyrinck, and J. Silk (2016). In: *ArXiv* 1607.07881.

Aragón-Calvo, M. A. et al. (2007). In: *A&A* 474, pp. 315–338.

Aragón-Calvo, M. A. et al. (2010). In: *ApJ* 723, pp. 364–382.

Aragón-Calvo, M. A. et al. (2015). In: *MNRAS* 454, pp. 463–477.

Arnouts, S. et al. (2007). In: *A&A* 476, pp. 137–150.

Avni, Y. and J. N. Bahcall (1980). In: *ApJ* 235, pp. 694–716.

Baldry, I. K., K. Glazebrook, and S. P. Driver (2008). In: *MNRAS* 388, pp. 945–959.

Baldry, I. K. et al. (2004). In: *ApJ* 600, pp. 681–694.

Baldry, I. K. et al. (2012). In: *MNRAS* 421, pp. 621–634.

Balogh, M. L., J. F. Navarro, and S. L. Morris (2000). In: *ApJ* 540, pp. 113–121.

Balogh, M. L. et al. (1998). In: *ApJL* 504, pp. L75–L78.

Balogh, M. L. et al. (2001). In: *ApJ* 557, pp. 117–125.

Balogh, M. L. et al. (2004). In: *ApJL* 615, pp. L101–L104.

Bardelli, S. et al. (2009). In: *A&A* 495, pp. 431–446.

Bardelli, S. et al. (2010). In: *A&A* 511, A1.

Barnes, J. E. and L. E. Hernquist (1991). In: *ApJL* 370, pp. L65–L68.

Barnes, J. E. and L. Hernquist (1992). In: *Nature* 360, pp. 715–717.

Barnes, J. E. and L. Hernquist (1996). In: *ApJ* 471, p. 115.

Bekki, K. (1998). In: *ApJL* 502, pp. L133–L137.

Bell, E. F. (2003). In: *ApJ* 586, pp. 794–813.

Bell, E. F. et al. (2003). In: *ApJS* 149, pp. 289–312.

Benítez, N. et al. (2015). In: *Highlights of Spanish Astrophysics VIII*. Ed. by A. J. Cenarro et al. SEA - Sociedad Española de Astronomía, pp. 148–153.

Berta, S. et al. (2013). In: *A&A* 551, A100.

Best, P. N. (2000). In: *MNRAS* 317, pp. 720–736.

Best, P. N. et al. (2005). In: *MNRAS* 362, pp. 25–40.

Beucher, S. and C. Lantuejoul (1979). In: *International Workshop on Image Processing*. Ed. by H. Blum. CCETT/IRISA, Rennes, France, p. 2.1.

Beucher, S. and F. Meyer (1993). In: *Mathematical Morphology in Image Processing*. Ed. by E. R. Dougherty. Marcel Dekker, Inc., New York, p. 433.
Beygu, B. et al. (2013). In: *AJ* 145, p. 120.
Beygu, B. et al. (2016). In: *MNRAS* 458, pp. 394–409.
Blanton, M. R. and J. Moustakas (2009). In: *ARA&A* 47, pp. 159–210.
Blanton, M. R. et al. (2003). In: *ApJ* 594, pp. 186–207.
Blanton, M. R. et al. (2005). In: *ApJ* 629, pp. 143–157.
Bolzonella, M., R. Pelló, and D. Maccagni (2002). In: *A&A* 395, pp. 443–463.
Bolzonella, M. et al. (2010). In: *A&A* 524, A76.
Bond, J. R., L. Kofman, and D. Pogosyan (1996). In: *Nature* 380, pp. 603–606.
Boselli, A. and G. Gavazzi (2006). In: *PASP* 118, pp. 517–559.
Boselli, A. and G. Gavazzi (2014). In: *A&ARv* 22, p. 74.
Bower, R. G. et al. (2006). In: *MNRAS* 370, pp. 645–655.
Brough, S. et al. (2006). In: *MNRAS* 369, pp. 1351–1374.
Brouwer, M. M. et al. (2016). In: *MNRAS* 462, pp. 4451–4463.
Brusa, M. et al. (2007). In: *ApJS* 172, pp. 353–367.
Bundy, K. et al. (2006). In: *ApJ* 651, pp. 120–141.
Byrd, G. and M. Valtonen (1990). In: *ApJ* 350, pp. 89–94.
Capak, P. et al. (2007). In: *ApJS* 172, pp. 99–116.
Castignani, G. et al. (2014a). In: *ApJ* 792, p. 113.
Castignani, G. et al. (2014b). In: *ApJ* 792, p. 114.
Cayatte, V. et al. (1994). In: *AJ* 107, pp. 1003–1017.
Chabrier, G. (2003). In: *PASP* 115, pp. 763–795.
Chiaberge, M. et al. (2010). In: *ApJL* 710, pp. L107–L110.
Chiang, Y.-K., R. Overzier, and K. Gebhardt (2013). In: *ApJ* 779, p. 127.
Chincarini, G. (2015). In: *Thirteenth Marcel Grossmann Meeting: On Recent Developments in Theoretical and Experimental General Relativity, Astrophysics and Relativistic Field Theories*. Ed. by K. Rosquist. World Scientific Publishing Co. Pte. Ltd., pp. 781–822.
Christlein, D. and A. I. Zabludoff (2005). In: *ApJ* 621, pp. 201–214.
Chung, A. et al. (2008). In: *Frontiers of Astrophysics: A Celebration of NRAO's 50th Anniversary*. Ed. by A. H. Bridle, J. J. Condon, and G. C. Hunt. Vol. 395. ASPCS, p. 364.
Ciliegi, P. and S. Bardelli (2015). In: *AASKA* 14, 150, p. 150.
Ciliegi, P. et al. (2005). In: *A&A* 441, pp. 879–891.
Cimatti, A., E. Daddi, and A. Renzini (2006). In: *A&A* 453, pp. L29–L33.
Cimatti, A. et al. (2002). In: *A&A* 392, pp. 395–406.
Codis, S., C. Pichon, and D. Pogosyan (2015). In: *MNRAS* 452, pp. 3369–3393.
Codis, S. et al. (2012). In: *MNRAS* 427, pp. 3320–3336.
Coil, A. L. et al. (2011). In: *ApJ* 741, p. 8.
Colberg, J. M. et al. (2008). In: *MNRAS* 387, pp. 933–944.

Cole, S. et al. (2000). In: *MNRAS* 319, pp. 168–204.
Cole, S. et al. (2001). In: *MNRAS* 326, pp. 255–273.
Coles, P. and F. Lucchin (2002). *Cosmology: The Origin and Evolution of Cosmic Structure, Second Edition*. Wiley-VCH.
Colless, M. et al. (2001). In: *MNRAS* 328, pp. 1039–1063.
Comparat, J. et al. (2015). In: *A&A* 575, A40.
Content, D. et al. (2013). In: *UV/Optical/IR Space Telescopes and Instruments: Innovative Technologies and Concepts VI*. Ed. by H. A. MacEwen and J. A. Breckinridge. Vol. 8860. Proc. SPIE, 88600E.
Cooke, R., M. Pettini, and M. T. Murphy (2012). In: *MNRAS* 425, pp. 347–354.
Cooper, M. C. et al. (2005). In: *ApJ* 634, pp. 833–848.
Cooper, M. C. et al. (2006). In: *MNRAS* 370, pp. 198–212.
Cooper, M. C. et al. (2010). In: *MNRAS* 409, pp. 337–345.
Cowie, L. L. and A. Songaila (1977). In: *Nature* 266, pp. 501–503.
Croston, J. H. et al. (2008). In: *MNRAS* 386, pp. 1709–1728.
Croton, D. J. et al. (2005). In: *MNRAS* 356, pp. 1155–1167.
Croton, D. J. et al. (2006). In: *MNRAS* 365, pp. 11–28.
Cucciati, O. et al. (2006). In: *A&A* 458, pp. 39–52.
Cucciati, O. et al. (2010). In: *A&A* 520, A42.
Cucciati, O. et al. (2012). In: *A&A* 548, A108.
Cucciati, O. et al. (2014). In: *A&A* 565, A67.
Cucciati, O. et al. (2016). In: *MNRAS* 462, pp. 1786–1801.
Darvish, B. et al. (2014). In: *ApJ* 796, p. 51.
Darvish, B. et al. (2015). In: *ApJ* 805, p. 121.
Davidzon, I. et al. (2013). In: *A&A* 558, A23.
Davidzon, I. et al. (2016). In: *A&A* 586, A23.
Davis, M. and M. J. Geller (1976). In: *ApJ* 208, pp. 13–19.
Davis, M. et al. (1982). In: *ApJ* 253, pp. 423–445.
Davis, M. et al. (2003). In: *Discoveries and Research Prospects from 6- to 10-Meter-Class Telescopes II*. Ed. by P. Guhathakurta. Vol. 4834. Proc. SPIE, pp. 161–172.
Dawson, K. S. et al. (2013). In: *AJ* 145, p. 10.
De Lucia, G. (2011). In: *A&SSP* 27, p. 203.
De Lucia, G. et al. (2012). In: *MNRAS* 423, pp. 1277–1292.
Dekel, A. and Y. Birnboim (2006). In: *MNRAS* 368, pp. 2–20.
Dekel, A. et al. (2009). In: *Nature* 457, pp. 451–454.
Delaunay, B. (1934). In: *Bull. Acad. Sci. USSR* 7, p. 793.
Delvecchio, I. et al. (2014). In: *MNRAS* 439, pp. 2736–2754.
Dey, A. et al. (2016). In: *ApJ* 823, p. 11.
Dickinson, M., M. Giavalisco, and GOODS Team (2003). In: *The Mass of Galaxies at Low and High Redshift*. Ed. by R. Bender and A. Renzini. ESO, Springer-Verlag, Germany, p. 324.
Donoso, E. et al. (2010). In: *MNRAS* 407, pp. 1078–1089.
Dressler, A. (1980). In: *ApJ* 236, pp. 351–365.
Driver, S. P. et al. (2011). In: *MNRAS* 413, pp. 971–995.

Dubois, Y. et al. (2014). In: *MNRAS* 444, pp. 1453–1468.
Eardley, E. et al. (2015). In: *MNRAS* 448, pp. 3665–3678.
Edelsbrunner, H., D. Letscher, and A. Zomorodian (2002). In: *Discrete Comput. Geom.* 28, p. 511.
Edge, A. et al. (2013). In: *The Messenger* 154, pp. 32–34.
Efstathiou, G., R. S. Ellis, and B. A. Peterson (1988). In: *MNRAS* 232, pp. 431–461.
Elbaz, D. et al. (2007). In: *A&A* 468, pp. 33–48.
Etherington, J. and D. Thomas (2015). In: *MNRAS* 451, pp. 660–679.
Fadda, D. et al. (2008). In: *ApJL* 672, p. L9.
Fanaroff, B. L. and J. M. Riley (1974). In: *MNRAS* 167, 31P–36P.
Farouki, R. and S. L. Shapiro (1981). In: *ApJ* 243, pp. 32–41.
Ferrarese, L. and D. Merritt (2000). In: *ApJL* 539, pp. L9–L12.
Finoguenov, A. et al. (2007). In: *ApJS* 172, pp. 182–195.
Florack, L. M. J. et al. (1992). In: *Image and Vision Computing* 10, p. 376.
Fontana, A. et al. (2004). In: *A&A* 424, pp. 23–42.
Forman, R. (1998). In: *Adv. Math.* 134, p. 90.
Forman, R. (2002). In: *Sém. Lothar. Combin.* 48, B48c.
Fossati, M. et al. (2015). In: *MNRAS* 446, pp. 2582–2598.
Fujita, Y. (1998). In: *ApJ* 509, pp. 587–594.
Fujita, Y. (2001). In: *ApJ* 550, pp. 612–621.
Fujita, Y. and M. Nagashima (1999). In: *ApJ* 516, pp. 619–625.
Gabor, J. M. and R. Davé (2015). In: *MNRAS* 447, pp. 374–391.
Galametz, A. et al. (2013). In: *ApJS* 206, p. 10.
Gallazzi, A. et al. (2009). In: *ApJ* 690, pp. 1883–1900.
Garilli, B. et al. (2014). In: *A&A* 562, A23.
Gay, C. et al. (2010). In: *MNRAS* 404, pp. 1801–1816.
Gebhardt, K. et al. (2000). In: *ApJL* 539, pp. L13–L16.
Gehrels, N. (1986). In: *ApJ* 303, pp. 336–346.
Geller, M. J. and J. P. Huchra (1989). In: *Science* 246, pp. 897–903.
Ghigna, S. et al. (1998). In: *MNRAS* 300, pp. 146–162.
Giodini, S. et al. (2012). In: *A&A* 538, A104.
Giovanelli, R. and M. P. Haynes (1983). In: *AJ* 88, pp. 881–908.
Giuricin, G. et al. (2001). In: *ApJ* 554, pp. 857–872.
Gómez, P. L. et al. (2003). In: *ApJ* 584, pp. 210–227.
Granato, G. L. et al. (2004). In: *ApJ* 600, pp. 580–594.
Grützbauch, R. et al. (2011). In: *MNRAS* 411, pp. 929–946.
Gunn, J. E. and J. R. Gott III (1972). In: *ApJ* 176, p. 1.
Guo, Y. et al. (2013). In: *ApJS* 207, p. 24.
Guzzo, L. et al. (1997). In: *ApJ* 489, pp. 37–48.
Guzzo, L. et al. (2014). In: *A&A* 566, A108.
Hahn, C. et al. (2015). In: *ApJ* 806, p. 162.
Hahn, O. et al. (2007). In: *MNRAS* 375, pp. 489–499.
Hardcastle, M. J., D. A. Evans, and J. H. Croston (2007). In: *MNRAS* 376, pp. 1849–1856.
Hashimoto, Y. and A. Oemler Jr. (1999). In: *ApJ* 510, pp. 609–613.

Hatch, N. A. et al. (2011). In: *MNRAS* 410, pp. 1537–1549.
Hatch, N. A. et al. (2014). In: *MNRAS* 445, pp. 280–289.
Hawken, A. J. et al. (2016). In: *in prep.*
Haynes, M. P., R. Giovanelli, and G. L. Chincarini (1984). In: *ARA&A* 22, pp. 445–470.
Henriksen, M. and G. Byrd (1996). In: *ApJ* 459, p. 82.
Hickox, R. C. et al. (2009). In: *ApJ* 696, pp. 891–919.
Hill, G. J. et al. (2008). In: *Panoramic Views of Galaxy Formation and Evolution*. Ed. by T. Kodama, T. Yamada, and K. Aoki. Vol. 399. ASPCS, p. 115.
Hubble, E. P. (1925). In: *The Observatory* 48, pp. 139–142.
Hubble, E. P. (1926). In: *ApJ* 64.
Hubble, E. P. (1936). *Realm of the Nebulae*. Yale University Press, New Haven.
Hubble, E. (1934). In: *ApJ* 79, p. 8.
Icke, V. (1985). In: *A&A* 144, pp. 115–123.
Ilbert, O. et al. (2009). In: *ApJ* 690, pp. 1236–1249.
Ilbert, O. et al. (2010). In: *ApJ* 709, pp. 644–663.
Ilbert, O. et al. (2013). In: *A&A* 556, A55.
Iovino, A. et al. (2016). In: *in prep.*
Jarrett, T. H. et al. (2000). In: *AJ* 119, pp. 2498–2531.
Jarvis, M. J. et al. (2013). In: *MNRAS* 428, pp. 1281–1295.
Kartaltepe, J. S. et al. (2015). In: *ApJS* 221, p. 11.
Kauffmann, G. and S. D. M. White (1993). In: *MNRAS* 261.
Kauffmann, G. et al. (2004). In: *MNRAS* 353, pp. 713–731.
Kereš, D. et al. (2005). In: *MNRAS* 363, pp. 2–28.
Knobel, C. et al. (2012). In: *ApJ* 753, p. 121.
Kochanek, C. S. et al. (2001). In: *ApJ* 560, pp. 566–579.
Kodama, T. and R. Bower (2003). In: *MNRAS* 346, pp. 1–12.
Kovač, K. et al. (2010). In: *ApJ* 708, pp. 505–533.
Krogager, J.-K. et al. (2014). In: *ApJ* 797, p. 17.
Lacey, C. and S. Cole (1993). In: *MNRAS* 262, pp. 627–649.
Lagos, C. d. P. et al. (2012). In: *MNRAS* 426, pp. 2142–2165.
Lai, C.-C. et al. (2016). In: *ApJ* 825, p. 40.
Laigle, C. et al. (2015). In: *MNRAS* 446, pp. 2744–2759.
Laigle, C. et al. (2016). In: *ApJS* 224, p. 24.
Larson, R. B., B. M. Tinsley, and C. N. Caldwell (1980). In: *ApJ* 237, pp. 692–707.
Laureijs, R. et al. (2011). In: *ArXiv* 1110.3193.
Le Fèvre, O. et al. (2005). In: *A&A* 439, pp. 845–862.
Le Fèvre, O. et al. (2015). In: *A&A* 576, A79.
Ledlow, M. J. and F. N. Owen (1996). In: *AJ* 112, p. 9.
Lee, K.-G. et al. (2014). In: *ApJL* 795, p. L12.
Levy, L. et al. (2007). In: *AJ* 133, pp. 1104–1124.
Lewis, I. et al. (2002). In: *MNRAS* 334, pp. 673–683.
Li, I. H. et al. (2011). In: *MNRAS* 411, pp. 1869–1879.

Lilly, S. J. et al. (2007). In: *ApJS* 172, pp. 70–85.
Lilly, S. J. et al. (2009). In: *ApJS* 184, pp. 218–229.
Lindeberg, T. (1998). In: *Int. J. Computer Vision* 30, p. 48.
Lutz, D. et al. (2011). In: *A&A* 532, A90.
Lynden-Bell, D. (1971). In: *MNRAS* 155, p. 95.
Madau, P. and M. Dickinson (2014). In: *ARA&A* 52, pp. 415–486.
Madau, P. et al. (1996). In: *MNRAS* 283, pp. 1388–1404.
Magorrian, J. et al. (1998). In: *AJ* 115, pp. 2285–2305.
Mahajan, S., S. Raychaudhury, and K. A. Pimbblet (2012). In: *MNRAS* 427, pp. 1252–1265.
Malavasi, N. et al. (2015). In: *A&A* 576, A101.
Malavasi, N. et al. (2016b). In: *submitted to MNRAS*.
Malavasi, N. et al. (2016c). In: *accepted for publication in MNRAS*.
Malavasi, N. et al. (2016a). In: *A&A* 585, A116.
Mandelbaum, R. et al. (2009). In: *MNRAS* 393, pp. 377–392.
Marinoni, C. et al. (2002). In: *ApJ* 580, pp. 122–143.
Martin, D. C. et al. (2007). In: *ApJS* 173, pp. 342–356.
Martin, D. C. et al. (2016). In: *ApJL* 824, p. L5.
Martínez, H. J., H. Muriel, and V. Coenda (2016). In: *MNRAS* 455, pp. 127–135.
Martínez, V. J. and E. Saar (2002). *Statistics of the Galaxy Distribution*. Chapman & Hall/CRC, Boca Raton.
Matthews, T. A., W. W. Morgan, and M. Schmidt (1964). In: *ApJ* 140, p. 35.
Matuszewski, M. et al. (2010). In: *Ground-based and Airborne Instrumentation for Astronomy III*. Ed. by I. S. McLean, S. K. Ramsay, and H. Takami. Vol. 7735. Proc. SPIE, 77350P.
McCracken, H. J. et al. (2010). In: *ApJ* 708, pp. 202–217.
McCracken, H. J. et al. (2012). In: *A&A* 544, A156.
McNaught-Roberts, T. et al. (2014). In: *MNRAS* 445, pp. 2125–2145.
Merritt, D. (1983). In: *ApJ* 264, pp. 24–48.
Merritt, D. (1984). In: *ApJ* 276, pp. 26–37.
Merson, A. I. et al. (2013). In: *MNRAS* 429, pp. 556–578.
Micheletti, D. et al. (2014). In: *A&A* 570, A106.
Mihos, J. C. (1995). In: *ApJL* 438, pp. L75–L78.
Mihos, J. C. and L. Hernquist (1994). In: *ApJL* 425, pp. L13–L16.
Mihos, J. C. and L. Hernquist (1996). In: *ApJ* 464, p. 641.
Miley, G. and C. De Breuck (2008). In: *A&ARv* 15, pp. 67–144.
Moles, M. et al. (2008). In: *AJ* 136, pp. 1325–1339.
Moore, B., G. Lake, and N. Katz (1998). In: *ApJ* 495, pp. 139–151.
Moore, B. et al. (1996). In: *Nature* 379, pp. 613–616.
Moore, B. et al. (1999). In: *MNRAS* 304, pp. 465–474.
Mortlock, A. et al. (2015). In: *MNRAS* 447, pp. 2–24.
Moustakas, J. et al. (2013). In: *ApJ* 767, p. 50.
Moutard, T. et al. (2016a). In: *A&A* 590, A102.
Moutard, T. et al. (2016b). In: *A&A* 590, A103.

Muldrew, S. I. et al. (2012). In: *MNRAS* 419, pp. 2670–2682.
Muzzin, A. et al. (2013). In: *ApJ* 777, p. 18.
Newman, J. A. et al. (2013). In: *ApJS* 208, p. 5.
Norberg, P. et al. (2002). In: *MNRAS* 332, pp. 827–838.
Nulsen, P. E. J. (1982). In: *MNRAS* 198, pp. 1007–1016.
Oemler Jr., A. (1974). In: *ApJ* 194, pp. 1–20.
Oesch, P. A. et al. (2016). In: *ApJ* 819, p. 129.
Onodera, M. et al. (2012). In: *ApJ* 755, p. 26.
Peebles, P. J. E. (1975). In: *ApJ* 196, pp. 647–651.
Peng, Y.-j. et al. (2010). In: *ApJ* 721, pp. 193–221.
Penzias, A. A. and R. W. Wilson (1965). In: *ApJ* 142, pp. 419–421.
Pichon, C. et al. (2011). In: *MNRAS* 418, pp. 2493–2507.
Planck Collaboration et al. (2015). In: *ArXiv* 1502.01589.
Planck Collaboration et al. (2016). In: *A&A* 594, A1.
Platen, E., R. van de Weygaert, and B. J. T. Jones (2007). In: *MNRAS* 380, pp. 551–570.
Porter, S. C. et al. (2008). In: *MNRAS* 388, pp. 1152–1160.
Postman, M. and M. J. Geller (1984). In: *ApJ* 281, pp. 95–99.
Pozzetti, L. et al. (2007). In: *A&A* 474, pp. 443–459.
Pozzetti, L. et al. (2010). In: *A&A* 523, A13.
Pozzetti, L. et al. (2016). In: *A&A* 590, A3.
Press, W. H. and P. Schechter (1974). In: *ApJ* 187, pp. 425–438.
Quilis, V., B. Moore, and R. Bower (2000). In: *Science* 288, pp. 1617–1620.
Renzini, A. (2006). In: *ARA&A* 44, pp. 141–192.
Ricciardelli, E. et al. (2014). In: *MNRAS* 445, pp. 4045–4054.
Richstone, D. O. (1976). In: *ApJ* 204, pp. 642–648.
Robins, V. (2000). PhD thesis. University of Colorado.
Rojas, R. R. et al. (2004). In: *ApJ* 617, pp. 50–63.
Rojas, R. R. et al. (2005). In: *ApJ* 624, pp. 571–585.
Sadler, E. M., C. R. Jenkins, and C. G. Kotanyi (1989). In: *MNRAS* 240, pp. 591–635.
Sadler, E. M. et al. (2002). In: *MNRAS* 329, pp. 227–245.
Sandage, A. (1961). *The Hubble atlas of galaxies*. Carnegie Institution, Washington.
Sandage, A., G. A. Tammann, and A. Yahil (1979). In: *ApJ* 232, pp. 352–364.
Sandage, A. and N. Visvanathan (1978). In: *ApJ* 225, pp. 742–750.
Schaap, W. E. (2007). PhD thesis. Rijksuniversiteit Groningen.
Schaap, W. E. and R. van de Weygaert (2000). In: *A&A* 363, pp. L29–L32.
Schechter, P. (1976). In: *ApJ* 203, pp. 297–306.
Schinnerer, E. et al. (2007). In: *ApJS* 172, pp. 46–69.
Schinnerer, E. et al. (2010). In: *ApJS* 188, pp. 384–404.
Scodeggio, M. et al. (2016). In: *in prep.*
Scoville, N. et al. (2007). In: *ApJS* 172, pp. 1–8.

Scoville, N. et al. (2013). In: *ApJS* 206, p. 3.
Seldner, M. et al. (1977). In: *AJ* 82, pp. 249–256.
Shapley, H. and A. Ames (1932). In: *Annals of Harvard College Observatory* 88, pp. 41–76.
Shectman, S. A. et al. (1996). In: *ApJ* 470, p. 172.
Silverman, J. D. et al. (2015). In: *ApJS* 220, p. 12.
Smolčić, V. et al. (2008). In: *ApJS* 177, pp. 14–38.
Somerville, R. S. and R. Davé (2015). In: *ARA&A* 53, pp. 51–113.
Sousbie, T. (2011). In: *MNRAS* 414, pp. 350–383.
Sousbie, T., S. Colombi, and C. Pichon (2009). In: *MNRAS* 393, pp. 457–477.
Sousbie, T., C. Pichon, and H. Kawahara (2011). In: *MNRAS* 414, pp. 384–403.
Spitzer Jr., L. and W. Baade (1951). In: *ApJ* 113, p. 413.
Springel, V. et al. (2005). In: *Nature* 435, pp. 629–636.
Strateva, I. et al. (2001). In: *AJ* 122, pp. 1861–1874.
Strazzullo, V. et al. (2015). In: *A&A* 576, p. L6.
Sutherland, W. and W. Saunders (1992). In: *MNRAS* 259, pp. 413–420.
Takada, M. et al. (2014). In: *PASJ* 66, R1.
Tempel, E., R. S. Stoica, and E. Saar (2013). In: *MNRAS* 428, pp. 1827–1836.
Thomas, D. et al. (2005). In: *ApJ* 621, pp. 673–694.
Toniazzo, T. and S. Schindler (2001). In: *MNRAS* 325, pp. 509–530.
Toshikawa, J. et al. (2016). In: *ApJ* 826, p. 114.
Trenti, M. and M. Stiavelli (2008). In: *ApJ* 676, pp. 767–780.
Treu, T. et al. (2003). In: *ApJ* 591, pp. 53–78.
Trujillo, I., C. Carretero, and S. G. Patiri (2006). In: *ApJL* 640, pp. L111–L114.
Ueda, Y. et al. (2003). In: *ApJ* 598, pp. 886–908.
Vogt, N. P. et al. (2004). In: *AJ* 127, pp. 3300–3324.
Voronoi, G. F. (1908). In: *J. Reine Angew. Math.* 134, p. 198.
Vulcani, B. et al. (2012). In: *MNRAS* 420, pp. 1481–1494.
Wake, D. A. et al. (2008). In: *MNRAS* 391, pp. 1674–1684.
Warmels, R. H. (1988). In: *A&AS* 72, pp. 19–87.
Welker, C. et al. (2015). In: *ArXiv* 1502.05053.
White, S. D. M. and M. J. Rees (1978). In: *MNRAS* 183, pp. 341–358.
Wylezalek, D. et al. (2013). In: *ApJ* 769, p. 79.
Yee, H. K. C. and R. F. Green (1987). In: *ApJ* 319, pp. 28–43.
York, D. G. et al. (2000). In: *AJ* 120, pp. 1579–1587.
Zel'dovich, Y. B. (1970). In: *A&A* 5, pp. 84–89.
Zhang, Y. et al. (2013). In: *ApJ* 779, p. 160.
de Jong, J. T. A. et al. (2015). In: *A&A* 582, A62.
de Lapparent, V., M. J. Geller, and J. P. Huchra (1986). In: *ApJL* 302, pp. L1–L5.
de la Torre, S. et al. (2013). In: *A&A* 557, A54.
de la Torre, S. et al. (2016). In: *in prep.*

van de Weygaert, R. and E. Platen (2011). In: *Int. J. of Mod. Phys. Conf. Series* 1, pp. 41–66.
van der Burg, R. F. J. et al. (2013). In: *A&A* 557, A15.