The environmental effects of the LSS : characterization of the baryonic components

Iris del Carmen Santiago Bautista

Université Fédérale
Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

Cotutelle internationale : Universidad de Guanajuato

Présentée et soutenue par

Iris SANTIAGO BAUTISTA

Le 25 février 2020

Étude des propriétés des galaxies dans les structures filamentaires

---

Ecole doctorale : **SDU2E - Sciences de l'Univers, de l'Environnement et de l'Espace**

Spécialité : **Astrophysique, Sciences de l'Espace, Planétologie**

Unité de recherche :
IRAP - Institut de Recherche en Astrophysique et Planetologie

Thèse dirigée par
Etienne POINTECOUTEAU et Cesar CARETTA

Jury

M. Johan VAN HOREBEEK, Rapporteur
M. Laerte SODRE JUNIOR, Rapporteur
Mme Fatima ROBLES, Examinatrice
M. Etienne POINTECOUTEAU, Directeur de thèse
M. Cesar CARETTA, Co-directeur de thèse

Universidad de Guanajuato
Departamento de Astronomía
DCNyE

The environmental effects of the LSS: characterization of the baryonic components

by
Iris Santiago-Bautista

a dissertation submitted in fulfilment
of the requirements for the degree of

Doctor of Philosophy

with specialization in astronomy

Directed by:

Cesar Caretta
Etienne Pointecouteau
Hector Bravo-Alfaro

Guanajuato, Mexico

February, 2020

Abstract

The baryonic component of the Large Scale Structure (LSS) of the Universe is composed by concentrations of gas and galaxies forming groups, clusters, elongated filaments and widely spread sheets which probably underline the distribution of dark matter. Nevertheless, according to the current cosmological models, most of the baryonic material in the Universe has not yet been directly observed. Numerical simulations suggest that from one-half to two-thirds of all baryons may be located out of clusters of galaxies, pervading the structures between them.

The most concentrated structures, which we call systems of galaxies (i.e., groups and clusters) usually contain high density hot gas ($1-10$ keV) that cools radiatively, emits at X-rays wavelengths and interacts with the cosmic microwave background at millimeter wavelengths (Sunyaev Zel'dovich effect, SZ). For the less dense structures, filaments and sheets, the baryons are probably in moderately hot gas phase ($0.01-1$ keV), commonly named as warm hot intergalactic medium (WHIM). In this PhD Thesis, we study the environmental effects associated to the different components of the LSS. For the galaxy systems, we aim to characterize the intra cluster medium (ICM) through the analysis of the S-Z effect. We employ the ACT and Planck data to analyze the gas pressure profiles of a sample of low mass galaxy clusters. For the least dense structures, we assembled a sample of filament candidates composed by chains of clusters that are located inside superclusters of galaxies. We aim to probe the filament structure skeletons and characterize their components (galaxies, groups/clusters and gas).

Resumen

La componente bariónica de la estructura a gran escala del Universo está compuesta por concentraciones de gas y galaxias formando grupos, cúmulos, filamentos elongados y amplias paredes. Dichas estructuras probablemente reflejan la distribución de materia oscura en el Universo. Sin embargo, de acuerdo al modelo cosmológico actual, la mayor parte de la materia bariónica en el Universo no ha sido observada aún. No obstante, las simulaciones numéricas nos sugieren que entre la mitad y dos tercios de los bariones quizá se encuentran entre cúmulos de galaxias, poblando las estructuras que los conectan. Las estructuras más concentradas, generalizadas por nosotros como sistemas (i.e. grupos y cúmulos), usualmente contienen gas a altas densidades y temperaturas ($1-10$ keV) que se enfría radiativamente emitiendo fotones observables en rayos X e interactúa con la radiación cósmica de fondo por efecto Sunyaev–Zel'dovich (SZ) observado a longitudes de onda milimétricas. En las estructuras menos densas, los filamentos y paredes, los bariones se encuentran probablemente en un estado menos denso y a una temperatura moderada ($0.01-1$ keV). Este gas es comúnmente llamado medio intergaláctico templado. En esta Tesis de Doctorado estudiamos los efectos ambientales asociados a las diferentes componentes de la estructura a gran escala. En el caso de los sistemas nuestro objetivo es caracterizar el medio intracumular de cúmulos utilizando el efecto SZ. Para esto hacemos uso de observaciones de satélite Planck y del ACT (Atacama Cosmological Telescope) para analizar el perfil de presión del gas contenido para una muestra de cúmulos de baja masa. Por otro lado, para el estudio de las estructuras de baja densidad, los filamentos, construímos una muestra de candidatos a filamentos que consiste en cadenas de cumulos dentro de supercumulos de galaxias. Nuestro objetivo es probar la naturaleza filamentosa de estas estructuras así como caracterizar sus componentes (galaxias, cúmulos y gas).

Résumé

La composante baryonique de la structure à grande échelle de l'Univers est composée de concentration de gaz et de galaxies, donnant lieu à des groupes, à des amas, à des filaments allongés et à des murs étendus. Ces structures peuvent suivre la distribution de matière noire dans l'Univers. Néanmoins, selon le modèle cosmologique actuel, l'ensemble des matières baryoniques dans l'Univers n'a pas encore été observé. Cependant, les simulations numériques nous suggèrent qu'entre la moitié et deux tiers des parties des baryons sont localisées entre les amas de galaxies et peuplent les structures qui les relient. Les structures les plus concentrées, que nous appelons ici $\ll$ des systèmes $\gg$ (i.e. groupes et amas), ont généralement des gaz à haute concentration et une température élevée ($1-10$ keV). Cette température se refroidit en émettant des photons qui sont observables en rayons X. De plus les gaz interagissent avec les photons du fond diffuse cosmologique par l'effet Sunyaev–Zel'dovich (SZ) , observable à longueur d'onde millimétrique. Dans les filamentaires et murs qui sont des structures moins denses, les baryons sont probablement dans un état moins dense et à une température modérée ($0.01-1$ keV). Ces gaz tièdes sont appelés WHIM (Warm Hot Intergalactic Medium). Pendant cette Thèse de doctorat nous étudions les effets environnementaux associés aux différents composants de la structure à grande échelle de l'Univers. Pour les systèmes, l'objectif est la caractérisation du milieu intra amas en utilisant l'effet SZ. Pour cela nous utilisons les observations du satellite Planck et de l'Atacama Cosmological Télescope (ACT) afin d'analyser les profils de pression pour un échantillon d'amas de faible masse. D'autre part, pour l'étude des structures à faible densité (structures filamentaires). Nous avons construit un échantillon de candidats à filaments, cet échantillon se compose des chaînes d'amas reliées en une structure de super amas de galaxies. Notre objectif est de prouver leur nature filamentaire et de caractériser ses composants (galaxies, amas et gaz).

Acknowledgements

First of all, I would like to express my gratitude to my Thesis advisors, Dr. Cesar Caretta, Dr. Etienne Pointecouteau and Dr. Hector Bravo Alfaro for giving me the opportunity of carrying out this PhD, for their patience and for encouraging me to develop my ideas. Their knowledge and advice motivates me to continuously improve myself as a researcher.

I want to give a special thanks to my rapporteurs Dr. Gabriel Pratt, Dr. Laerte Sodre Junior for reading this Thesis, and to Johan VanHoerebeek, whose meeting sessions and comments have enriched this work. I would like to thank the jury members: Dr. Juan Pablo Torres Papaqui, Dr. Fatima Robles and Dr. Josep M. Masque and especially to Dr. Heinz Andernach for his comments and suggestions which have increased the value and quality of this research.

Also, I would like to give thanks to Genevieve Soulcail, director of the SDU2E Doctoral school, for all her help, effort and support provided during my enrollment to the University Paul Sabatier in a joint program with the University of Guanajuato. My thanks to Mrs. Cathala and Mrs. Laura Salas for all the help given during all the administrative processes for both universities.

I'm grateful to DAIP and CONACYT, this PhD research would not have been possible without the funding granted by them. I would also like to thank my comrades from IRAP and UG for numerous conversations and support in several moments during the PhD. To my friends in Mexico who opened their homes to me during my time in Guanajuato, and to my friends in France who have helped me in numerous ways.

Finally, I would also like to express my immense gratitude to my parents, for all their support and encouragement to pursue my objectives, regardless the difficulties of distance.

Contents

Introduction

With the availability of all sky extragalactic surveys, it has been observed that, at large scale, the galaxies (observable baryonic matter) in the Universe are organized in a web-like pattern. This pattern has been observed in several galaxy redshift surveys (e.g. Joeveer & Einasto, 1978; Davis et al., 1982; Shectman et al., 1996; Colless et al., 2001; Huchra et al., 2012). Such observations have shown that the large scale structure (LSS) was composed of elongated filaments, widely spread sheets and higher density knots. The latter are usually located at the intersection of filaments, where clusters and groups are hosted (e.g. Peebles, 1980). According to the current standard cosmological model (e.g. Riess et al., 1998) and observations of the cosmic microwave background (CMB) (e.g. Planck Collaboration et al., 2016b), the Universe is composed of 70% of dark energy and 30% of matter. Of the latter, about 85% is present in the form of cold dark matter (CDM) and only about 15% in the form of baryonic matter.

This model establishes that such structures of matter were formed through gravitational collapse, in a hierarchical scenario. In other words, the smaller halos (galaxies) formed first, then, they grew by fusion and accretion with other halos. Since the baryonic matter follows, to first order, the distribution of the dark matter, the galaxies and gas populate these structures accordingly (e.g. Eisenstein et al., 2005). The groups and clusters of galaxies are the next step of this hierarchical process of forming the densest regions in the LSS. About 85% of the baryonic matter in these structures is present in the form of hot gas (T $\sim 10^7 - 10^8$ K, or kT ~ 1–10 keV). This gas can be directly observed by its X-ray emission or through the SZ (Sunyaev–Zel'dovich) effect in the millimeter range. However, theoretical studies (e.g. Cen & Ostriker, 1999) and CMB measurements of the primordial nucleosynthesis (e.g. Planck Collaboration et al., 2016b) suggest that, in the local Universe, between one half and two thirds of the baryonic matter has not been detected yet at any wavelength. Nowadays, the results from several numerical

N-body simulations have been used to unveil where these baryons might be located, e.g. the Millennium (Springel et al., 2005), the Magneticum (Dolag et al., 2016) and Illustris (Vogelsberger et al., 2014). Their results regarding this subject suggest that these baryons might be in a warm and low density gas phase (T $\sim 10^5 - 10^7$ K or kT ~ 0.01–1 keV) filling filaments and sheets between galaxy clusters, constituting the so-called *warm hot intergalactic medium* (WHIM). The characterization of the WHIM thought observables such as X-rays or SZ effect is very challenging. However, deeper observations and the increase in detection sensitivity at millimetric, optical and X-rays wavelengths open the possibility to better study the dispersed component of the LSS while allowing the following subjects of study:

1. to detect and characterize the LSS components (e.g. topology, density, temperature, dynamical state, matter distribution and its evolution with time),
2. to better constrain the environment role in galaxy evolution,
3. to understand the mechanisms that shape the LSS matter distribution and
4. to evaluate and improve the current standard cosmological model.

In this framework, this Thesis is focused on studying the first and second subjects, characterizing the LSS components and investigating their effects on the galaxies evolution. By this, we aim to understand the mechanisms driving the LSS distribution, the third subject. Firstly, we present the characterization of LSS components (filaments, groups and clusters) along with the study of their environmental effects on the galaxy properties. Secondly, we present the characterization of galaxy clusters through the extraction of intra-cluster medium (ICM) pressure profiles from a high resolution SZ map.

In Chapter 1, we describe briefly the cosmological context that is used for our work and we introduce the current state of the art regarding the LSS components: galaxy clusters and filaments.

In Chapter 2, we introduce a general view of the current advances in detection of the disperse component of the LSS. Then, we describe in detail the methodology we implemented for the detection of clusters and filamentary structures. Our method aims the detection of chain-like structures inside a sample of galaxy superclusters (from Chow-Martinez et al., 2014) in the Local Universe (z $\sim$ 0.15).

As shown in Tanaka et al. (2007), the possibility to find elongated chain-like structures increases in superclusters.

In Chapter 3 we present the analyses we carried out to search for correlations between the galaxy properties (mass, morphology, activity, color and metallicity) and the environment in which these galaxies reside. For this analysis we used the Sloan Digital Sky Survey (SDSS) spectroscopic redshifts and sub-products (Albareti et al., 2017).

Chapter 4 is orientated towards the study of the gas component of galaxy clusters. We describe the methodology used for reconstructing the 3D radial profile of the thermal pressure for a sample of clusters, and the statistical analysis of the pressure distribution across the sample. These analyses were achieved using a composed SZ map from Planck and ACT (Atacama Cosmology Telescope) surveys, PACTmap (Aghanim et al., 2019a).

The final Chapter of this Thesis discusses the results of this work regarding the detection of LSS structures, their classification and characterization. We review the trends and correlations observed for galaxies inhabiting different components of the LSS detected by our algorithms, and discuss the possible environmental effects over their evolution. In addition, we discuss different approaches to carry out further analyses of the characterization of the gas component of clusters and filaments, in a similar way as the study carried out in this Thesis, using the SZ effect. We also discuss the achievements of, and our contribution to the LSS structure detection algorithms. Finally, the directions for further characterization of the LSS components using optical and gas components are presented. We also describe the future applications of the GSyF and GFiF algorithms. For this work we adopt (H_0, Ω_M, Ω_Λ, Ω_k, Ω_r) = (0.7, 0.3, 0.7, 0.0, 0.0).

Chapter 1

A Cosmology context

The different theories formulated in the history of cosmology aim to bring understanding on the mechanisms that govern the formation and distribution of matter and energy in the Universe. As the observational facilities provide deeper and more sensitive observations, in different wavelength regimes, the cosmological theories need to be improved in order to model the Universe as observed today. The study and characterization of the LSS distribution and its content through observables can be used to evaluate the cosmological models and see if they reproduce what is observed.

In this Chapter we present the principal aspects of the cosmology formalism that are behind the most accepted cosmological model and introduce the current state of art regarding the study of the main components of the LSS.

1.1 The standard cosmological model

1.1.1 The establishment of the ΛCDM model

The standard cosmological model, ΛCDM (e.g. Bond & Szalay, 1983; Doroshkevich & Khlopov, 1984), sometimes referred to as the concordance model, was motivated by observational studies of the LSS and the observations of the CMB. As summarized by Battaner & Florido (2000), Einasto (2009) and Einasto (2014), the cosmological model has changed to adapt to the observational facts along

with the new theories about structure formation; here we introduce the most important contributions to this cosmological model. At the beginning the model suggested the hypothesis that the first galaxies were formed from primordial density fluctuations of the baryonic medium (Harrison, 1970). However, there were problems with physical and kinematic properties of galaxies and clusters such as the potential energy needed to keep these structures bounded (Oort, 1940; Roberts, 1966; Rubin & Ford, 1970). Then, to explain these discrepancies, a dark matter in the form of dark stellar population and gaseous halos was proposed (e.g Ostriker & Peebles, 1973; Bahcall & Casertano, 1985). With the first observations of the CMB the model was discarded since the observed density fluctuations were smaller than predicted. As an alternative, the Hot Dark Matter model was proposed. This model suggested a non-baryonic relativistic neutrino-like particle as candidate for the dark matter, which explained the small fluctuations observed in the CMB (Cowsik & McClelland, 1973; Szalay & Marx, 1976; Tremaine & Gunn, 1979; Doroshkevich et al., 1980; Chernin, 1981). However, in this scenario, the structures observed today would not have had enough time to build up and the model was also abandoned. This gave place to a new hypothetical dark matter particle which moves slowly leading place for structure formation. This model was called Cold Dark Matter (CDM) (Blumenthal et al., 1982; Bond et al., 1982; Peebles, 1982; Pagels & Primack, 1982; Doroshkevich & Khlopov, 1984). It was originally proposed as a theory opposed to the Hot Dark matter (HDM) neutrino model. Finally, supernovae observations showed that the Universe was in a state of accelerated expansion. This brought up the need for a return of the cosmological constant Λ, see Section 1.1.3, to open space for a supposed vacuum energy, leading to the current ΛCDM model.

1.1.2 The fundamental parameters

Considering the expansion of the Universe is uniform, the distance and velocity of an object can be expressed as:

$$r(t) = a(t) \times D_C(t), \tag{1.1}$$

$$\dot{r} = H(t)r(t) + a(t)\dot{r}, \tag{1.2}$$

where $D_C(t)$ is the comoving distance, $H(t)$ the Hubble parameter and $a(t)$ the scale factor. The expansion rate is dictated by the term $= H(t)r(t)$ which relates the separation and recession velocity of distant galaxies. Then, the Hubble-Lemaître law is approximated to:

$$v(t) = H(t)D_C(t), \tag{1.3}$$

for the Local Universe. By convention, the sub-index 0 denotes the current value for the Hubble parameter at $z = 0$, usually written as $H_0 = 100\,h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, scaled in terms of h. For instance, Planck Collaboration et al. (2016b) measured h as 0.6772 ± 0.0046 using CMB measurements, while the Gaia mission measured a value of 0.7348 ± 0.0166 using a distance scale derived from Cepheid parallaxes (Riess et al., 2016). Another approach by Cuceu et al. (2019) measured $h = 0.676 \pm 0.011$ using Baryon Acoustic Oscillations, while the H0LiCOW collaboration (e.g. Suyu et al., 2017) measured $h = 0.733 \pm 0.0175$ using the time delay of lensed quasars. Another interesting result was the measurement by Hotokezaka et al. (2019) using the radio counterpart of GW170817, combined with gravitational wave and electromagnetic data, $h = 0.703 \pm 0.0515$. Considering that the Universe is isotropic at large scale and has a constant curvature, then, the space-time can be expressed by the Friedman-Robertson-Walker metric which, in spherical coordinates, is written as:

$$ds^2 = -c^2dt^2 + a^2(t)\left(\frac{dr^2}{1-kr^2} + r^2d\Omega^2\right), \tag{1.4}$$

with $a^2 > 0$ and $\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$. k determines the space curvature as:

$$k \begin{cases} > 0 \text{ the space-time is closed (spherical-like topology)}, \\ = 0 \text{ the space-time is flat}, \\ > 0 \text{ the space-time is open (hyperbolical-like topology)}. \end{cases}$$

Then, a useful observational measurement to estimate an object's distance is the redshift. The redshift is the result of the Doppler effect seen in the the galaxies optical spectra or magnitude, consequence of the expansion of space in all directions. Measuring the redshift of the spectral lines allow us to estimate when the light photons were emitted. The time between an object emitting photons and

the observation can be defined as:

$$\int_{t_{source}}^{t_{obs}} = c\frac{dt}{a(t)} = \int_{t_{source}+\lambda_{source}/c}^{t_{obs}+\lambda_{obs}/c} c\frac{dt}{a(t)}, \tag{1.5}$$

Then, the redshift is deduced as:

$$\frac{v_{source}}{v_{obs}} = \frac{a_{obs}}{a_{source}} = 1 + z, \tag{1.6}$$

with v_{source} the velocity of the object and v_{obs} observed velocity. For small redshifts, i.e. small distances, we can approximate $z \sim \frac{v_{source}}{c}$.

1.1.3 Dynamics of a Universe in expansion

Considering that the Universe is homogeneous and isotropic, its evolution can be described, in General Relativity, by the Friedmann-Lemaître Equations:

$$(\frac{\dot{a}}{a})^2 = \frac{8\pi\ G}{3}\rho_{tot} - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3} \tag{1.7}$$

$$\frac{\dot{a}}{a} = \frac{4\pi G}{3}\left(\rho_{tot} + \frac{3p_{tot}}{c^2}\right) \text{ with} \tag{1.8}$$

$$\rho_{tot} = \rho_\Lambda + \rho_m + \rho_r, \tag{1.9}$$

where $\rho_\Lambda = \Lambda/8\pi G$ corresponds to the dark energy contribution. Here Λ is the cosmological constant[1], $\rho_m = \rho_c + \rho_b$ corresponds to the matter density (where ρ_b is the baryonic density and ρ_c is the dark matter density); the density of radiation is ρ_r while k is the Universe curvature. A detailed derivation of Equation 1.7 from Einstein's equation can be consulted in Padmanabhan (2003).

Then, the density and pressure can be expressed in terms of the cosmological constant as:

$$\rho \to \rho + \frac{\Lambda c^2}{8\pi G} \text{ and } p \to p - \frac{\Lambda c^2}{8\pi G}. \tag{1.10}$$

and the Friedmann equations can be re-written in terms of energy density of pressure:

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi\ G}{3}\rho_{tot} - \frac{kc^2}{a^2}, \tag{1.11}$$

[1] Introduced by Einstein in 1917 to compensate the natural dynamics (expansion or contraction) of the Universe and make it static. Currently it is used to represent the dark or vacuum energy effect which causes an acceleration in the expansion of the Universe.

Considering Equation 1.11 with $k = 0$, i.e. a flat Universe, we can derive the critical density of the Universe as $\rho_c = \frac{3H_0^2}{8\pi G}$. The density parameter can be written in terms of the critical density as: $\Omega = \rho(t)/\rho_c(t)$.

Then the ρ_Λ, ρ_m and ρ_r can be re-written in terms of the density parameter (Peebles, 1993):

$$\Omega_m = \frac{8\pi G\rho_{tot}}{3H_0^2}, \quad \Omega_k = \frac{-k}{a^2H_0^2}, \text{ and } \Omega_\Lambda = \frac{\Lambda}{3H_0^2} \tag{1.12}$$

The Friedmann equation can be re-written in terms of the density parameter as:

$$\Omega_m + \Omega_r + \Omega_\Lambda + \Omega_k = 1 \tag{1.13}$$

Then the curvature of the Universe can be defined as:

$$kc^2 = a_0^2H_0^2(\Omega_{tot} - 1) = a_0^2H_0^2(\Omega_m + \Omega_r + \Omega_\Lambda) \tag{1.14}$$

Considering Equations 1.7, the fluid equation of state can be expressed as:

$$\dot{\rho} + 3H\left(\rho_{tot} + \frac{p_{tot}}{c^2}\right) = 0, \tag{1.15}$$

In order to simplify, the pressure and energy can be related by $p = p(\rho)$. The Equation 1.15 can be solved for a fluid with no torsion considering $p = w\rho$, where $w = (\gamma - 1)$ is constant and $\rho \propto a^{-3(1+w)}$, leading to the solution of the Friedman Equation of the form $a \propto t^{2/(-3(1+w))}$.

This solution can be used to estimate the evolution of energy density:
The Radiation dominated Universe: called Tolman Universe; it is defined by $w = 1/3$ and a radiation density $\Omega_r \approx 1$. Then the density and scale factor evolve as:

$$\rho_r(t) = \rho_{r0}\left(\frac{a_0}{a}\right)^4 \text{ with } a(t) = a_0\left(\frac{t}{t_0}\right)^{1/2} \tag{1.16}$$

Matter dominated Universe: if $w = 0$ then $p = 0$ then matter is pressureless, corresponding to the Einstein-de Sitter Universe with the values of energy density and scale factor:

TABLE 1.1: Cosmological parameters.

h	0.6772,
$100\,\Omega_\Lambda$	69.11 ± 0.62,
$100\,\Omega_m$	30.89 ± 0.62,
$100\,\Omega_c h^2$	11.88 ± 0.10,
$100\,\Omega_b h^2$	2.23 ± 0.014,
$100\,\Omega_r$	5×10^{-3},
$100\,\Omega_k$	0.00 ± 0.50.

$$\rho_m(t) = \rho_{m0}\left(\frac{a_0}{a}\right)^3 \text{ with } a(t) = a_0\left(\frac{t}{t_0}\right)^{2/3} \tag{1.17}$$

Curvature dominated Universe: In the case of an open Universe with $k < 0$, the Equation 1.7 does not have a solution, then:

$$\left(\frac{a_0}{a}\right)^2 = -\frac{k}{a^2} \text{ with } a(t) \propto a_0\left(\frac{t}{t_0}\right). \tag{1.18}$$

Energy dominated Universe: this is called de Sitter Universe. With $w = -1$ the scale factor has to be expanded by a Taylor series.

$$\rho_\Lambda(t) = \rho_0 \text{ with } a(t) = a_0 e^{H(t-t_0)}. \tag{1.19}$$

Combining Equations 1.6, 1.14 and 1.13, the Friedmann equation now can be expressed in terms of redshift and densities as:

$$H^2 = H_0^2(\Omega_r(1+z)^4 + \Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda), \tag{1.20}$$

Values of these densities have been recently measured using observations of the Planck satellite (Planck Collaboration et al., 2016b) results:

From the latter, we can highlight that, according to the ΛCDM cosmological model, the Universe is predominately composed of dark energy while its matter content is mainly composed of cold dark matter, its curvature is flat and it is in a state of accelerating expansion.

1.1.4 Distance measurements in the Universe

Comoving distance in the line of sight Considering two objects moving with the Hubble flow, the measurement of the distance between them (proper distance) at different times needs to be scaled. The distance from us to an object in the line of sight, the comoving distance D_C, is defined as:

$$D_C = \frac{c}{H_0} \int_0^z \frac{dz'}{E(z')}, \tag{1.21}$$

$$D_C = D_H \int_0^z \frac{dz'}{E(z')}, \tag{1.22}$$

with $E(z)$, defined by Peebles (1993), as:

$$E(z) = \frac{H(z)}{H_0} = \sqrt{\Omega_r(1+z)^4 + \Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda}, \tag{1.23}$$

Angular diameter distance and luminosity distances The angular diameter distance D_A is used to measure the actual size of objects observed with small angular sizes ($\delta\theta$), that is, separations perpendicular to the line of sight, defined as $\delta\theta \times D_A$. The angular diameter distance is used to convert from angular distances in physical separations and has the form:

$$D_A = \frac{D_m}{1+z}. \tag{1.24}$$

where D_m is the comoving angular diameter distance which depends on the Universe curvature and is defined as:

$$D_m \begin{cases} D_H \frac{1}{\sqrt{\Omega_k}} \sinh\left(\sqrt{\Omega_k}\frac{D_C}{D_H}\right) & \text{for } \Omega_k > 0, \\ D_C & \text{for } \Omega_k = 0, \\ D_H \frac{1}{\sqrt{\Omega_k}} \sin\left(\sqrt{\Omega_k}\frac{D_C}{D_H}\right) & \text{for } \Omega_k < 0, \end{cases}$$

Another useful distance measure is the luminosity distance which, in an expanding space, can be defined as:

$$D_L^2 = \frac{L_s}{4\pi F}, \tag{1.25}$$

where F is the flux received from an emitting source with absolute luminosity L_s. The luminosity distance is related with the comoving angular diameter distance

as:

$$D_L = (1+z)D_m = (1+z)^2 D_A. \tag{1.26}$$

1.2 The Universe's thermal history

The combination of observations, the cosmological model and particles physics, allows us to study the early Universe by reconstructing the evolution of its components.

The Planck Era: $T \gtrsim 10^{32}$ **K** According to the current cosmological model, at the beginning the Universe was in an extremely hot, dense and in ionized state.

The study of the evolution of the Universe allows one to reconstruct its thermal history up to the so-called Planck's time, $t_p = 10^{-43}$ seconds after the Big Bang. Before this time, quantum corrections become significant due to the extreme temperature, and density. At this moment the gravity ional force separated from the other forces. This corresponds to the moment at which the de Broglie wavelength of the Universe equals to its Schwarzschild radius:

$$\frac{\hbar}{mc} = \frac{2Gm}{c^2} \tag{1.27}$$

The GUT Era: $T \sim 10^{32} - 10^{29}$ **K** At $z \sim 10^{32}$, $t \sim 10^{-43} - 10^{-36}$s the GUT theories propose a model in which the asymmetry between matter and antimatter produce a phase transition. This transition makes the Universe expand exponentially, giving rise to the inflation theory (Starobinsky, 1980; Guth & Pi, 1982; Linde, 1982).This model gives a solution to the so called homogeneity problem, allowing the formation of the structures observed today from primordial density inhomogeneities, at a Hubble radius.

The particle Era: $T \sim 10^{15} - 10^{9}$ **K** When the Universe cools to a temperature about $T \sim 10^{16}$ K, ($t \sim 10^{-12}$s$-$2 min) the electromagnetic and nuclear weak forces separate. At this point the quarks, electrons, photons and gluons form a plasma in constant interaction. Afterwards, as the temperature decreases, at

about $T \sim 10^{13}$ K, the strong nuclear force allow the transition between quarks and hadrons, forming the protons and neutrons.

The primordial nucleosynthesis: $T \sim 10^{16}$ **K** The neutrinos and anti-neutrinos stop their interactions with electrons. As the baryons cooled, at temperatures about $T \sim 10^{9}$ K, the first deuterium and helium nuclei were formed. This process is called primordial nucleosynthesis and depends on the expansion rate of the Universe. This processes also determines the amount of baryonic matter in the Universe, which is measured to be between 4% and 5% of the critical density of the Universe.

In the Big Bang Model, nucleosynthesis occurs in the radiation-dominated epoch. The density of the Universe can be estimated as function of the temperature as $\rho(T) = (\pi^2/30)N_{rel}T^4$, where N_{rel} is the effective number of bosonic degrees of freedom. Under this scenario: when $a(t) \to 0$, the density $\rho \to \infty$ and $T \to \infty$. However, the primordial densities can be calculated assuming that the reaction:

$$p + e^{-} \longleftrightarrow n + \nu_e, \tag{1.28}$$

occurs in thermal equilibrium, such that the interaction rate, $\Gamma(t) = n\langle\sigma_\tau v_{rms}\rangle$ (where σ_τ is the Thomson cross section and v_{rms} is the rms velocity), allows hydrogen to form.

Photon decoupling and recombination: $T \sim 10^{4}$ **K** At this temperature the Universe is dominated by relativistic particles and the matter and radiation densities are in equilibrium (t$\sim 47kyr$,T$\sim 10^4$K, z≈ 3600). Then, at temperatures of $\sim 3\,000$ K the electrons combine with nuclei to form stable neutral atoms. At this moment, the photons decouple from the matter in all directions, as evidenced by the cosmic microwave background (CMB), See Section 1.3. One can define the temperature of the primordial photons at the moment of decoupling from the matter, dependent on the expansion as:

$$T(t) \propto \frac{1}{a(t)}. \tag{1.29}$$

The dark era and the formation of the first stars: $T \sim 15$ **K** The CMB temperature decreases as the Universe expands. During this time, under the effect

of gravity the first dark matter halos formed the first galaxies which agglomerated giving plate to the first galaxy clusters following a hierarchical structure formation scenario. Moreover, clouds of gas formed under the effect of gravity inside these halos, giving place to the first stars giving place to the reionization Era (T$\sim$ 60-19 K, 200Myr, 20$>$z$>$6).

1.3 The cosmic microwave background

The cosmic microwave background (CMB) was first observed by accident by Penzias & Wilson (1965) while carrying out an observation for radio emission. Thanks to their discovery they were awarded with the Nobel prize in 1978. The CMB is observed as an isotropic emission in the microwave window, coming from all directions in the sky. Its power spectrum represents the Universe a little bit after the moment at which matter and radiation density were equal.

1.3.1 CMB electromagnetic spectrum

The CMB photons are uniformly distributed and have a temperature of 2.725 ± 0.00335 K (Fixsen, 2009) with a black body spectrum defined as:

$$B_\nu(T_{CMB}) = \frac{2h\nu}{c^2}\left(1 - \exp\left(\frac{h\nu}{k_B T_{CMB}}\right)\right)^{-1} \quad [\mathrm{Wm^{-2}sr^{-1}Hz^{-1}}], \tag{1.30}$$

where h and k_B are the Planck and Boltzmann constants, respectively. The CMB observations suggest that the photons were indeed very close to equilibrium at some point in the early Universe. However, the CMB temperature presents small fluctuations (anisotropies) of the order $\Delta T_{CMB} = (T - \langle T \rangle)/\langle T \rangle \sim 10^{-5}$.

1.3.2 CMB angular power spectra

In general, the anisotropies of the CMB are quantified by their power spectrum. These anisotropies are observed as a function of direction over the celestial sphere. Therefore, they are expressed in terms of spherical harmonics:

$$T(\hat{\mathbf{n}}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} T_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}) \tag{1.31}$$

where $\hat{\mathbf{n}}$ is the direction of observation. The multipole index ℓ quantifies the scale (long wavelength modes corresponding to low values of ℓ). The CMB power spectrum as a function of ℓ is defined as:

$$C_\ell = \langle |T_{\ell m}|^2 \rangle \tag{1.32}$$

The anisotropies of the CMB contain information of the primordial density fluctuations in the early epoch of the Universe. In addition, CMB photons, on their trajectory across the Universe, have been perturbed by gravitational potentials or scattered by inverse Compton effect with ionized media. Such interactions are called secondary perturbations. The Sachs-Wolfe effect is attributed to temporal variation of the gravitational potential along the photon trajectory (Sachs & Wolfe, 1967). Other anisotropies observed are attributed to gravitational lenses (Planck Collaboration et al., 2018). In this case, anisotropies are generated by the change on the trajectory of photons due to the gravitational potential of a very massive object, like galaxy clusters. Another gravitational effect, called the Rees-Sciama effect (Rees & Sciama, 1968), can induce anisotropies in the CMB due to evolving gravitational potentials of non linear or matter structures as a result of collapse or expansion. Moreover, along their journey through the Universe, the CMB photons also interact with the gas in galaxy clusters, via the Sunyaev-Zel'dovich effect (Zel'dovich, 1970; Sunyaev & Zeldovich, 1972). This interaction is of particular interest for the study of galaxy clusters and will be discussed with more detail in Section 1.4.4.

The peak in the CMB power spectrum observed at $\ell \approx 200$ in the power spectrum (corresponding to an angular size of about 1°) represents the largest scale that had time to collapse at recombination. At smaller values of $\ell(\leq 50)$, corresponding to very large scales, the spectrum is flat, which is just proportional to the nearly scale-invariant spectrum at the moment fluctuations cross the Hubble radius.

Figure 1.1 shows the angular power spectrum of the CMB as measured by the Planck satellite (Planck Collaboration et al., 2016b). The vertical axis of the CMB power spectrum correspond to $\ell(\ell+1)C_\ell$, this is proportional to the square of the

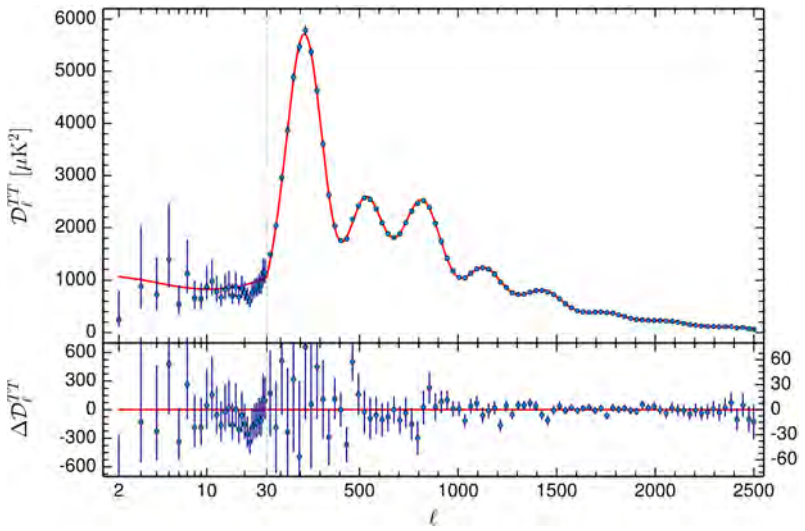

FIGURE 1.1: Power spectrum of the CMB as measured form Planck DR15 results. Figure extracted from Planck Collaboration et al. (2016b).

temperature fluctuations on the angular scale, $\theta \approx \pi/\ell$. In this representation, the CMB spectrum is about flat at large scales, $\ell < 50$.

1.4 The Large Scale Structure

1.4.1 Primordial overdensities and structure formation

In the ΛCDM model, the initial state of perturbations is assumed to be adiabatic and scalar. Such perturbations grow under the effects of gravity, leading to the formation of structures. These overdensities are measured in terms of the density contrast:

$$\delta = \frac{\rho(\hat{\mathbf{r}}, t) - \bar{\rho_{bg}}(t)}{\bar{\rho_{bg}}(t)}, \tag{1.33}$$

with $\bar{\rho_{bg}}(t)$ representing the mean background density measured at a time t. Considering only the contribution of gravity (Kaiser, 1986) in a Universe with $\Omega_{tot} = 1$, the spectrum of density fluctuations may be written as:

$$P(k) = k^n. \tag{1.34}$$

Considering $k \propto 1/r$ and $M \propto r^3$, the mass variance of fluctuations, σ^2, can be expressed as:

$$\sigma^2(k) \sim k^3 P(k) \propto r^{-(n+3)} \sim M^{-(n+3)/3} \tag{1.35}$$

Under these assumptions, by definition the fluctuations are self similar. However, as mentioned by Kaiser (1986), this property can only be applied at cluster scales, allowing to construct scaling relations, which are discussed in section 1.4.4.

Observations of the CMB power spectrum, as mentioned above, provide direct evidence of primordial perturbations. The initial density perturbations can be represented by Gaussian samples with mean zero, i.e. $\langle \delta \rangle = 0$, then the power spectrum is given by:

$$P(t,k) = \langle |\delta_k(t)|^2 \rangle \tag{1.36}$$

where δ_k is the Fourier transform of $\delta(t, \hat{\mathbf{r}})$. A complete characterization of the density perturbations at the time of decoupling is given by the CMB the power spectrum:

$$P(k) = \int \delta(\hat{\mathbf{r}}) \exp^{-ik\hat{\mathbf{r}}} \, d^3x = 4\pi \int \delta(R) \frac{\sin(kR)}{kR} r^2 \, dR \tag{1.37}$$

The spectrum is computed, when the density contrast is low, assuming a linear perturbation. However, when $\delta \sim 1$ analytical solutions are needed, some examples of such solutions can be consulted in Padmanabhan (2002). This model is called spherical top-hat approximation. The linear density contrast can be derived from solving Equation 1.33 in the regime of nonlinear perturbations, in a matter dominated Universe and it is defined as:

$$\delta_L = \frac{3}{5} \left(\frac{3}{4} \right)^{2/3} (\theta - \sin\theta)^{2/3}. \tag{1.38}$$

Figure 1.2 depicts the density as calculated from Equation 1.38. At the beginning, the perturbed region detaches from the cosmic expansion, starting to expand at lower rate. One can observe that at $\theta = 2\pi/3$ the solution is no longer linear. Afterwards, at $\theta = \pi$ the matter collapses after the so-called *turnaround* point. Finally, bounded structures are formed at θ larger than 2π.

Now, if the density fluctuations are very small, $\delta << 1$, then the solution to Equation 1.33 is non-linear and is then solved by an analytic model developed by Press & Schechter (1974), see Percival (2001) for details. Then the variance of

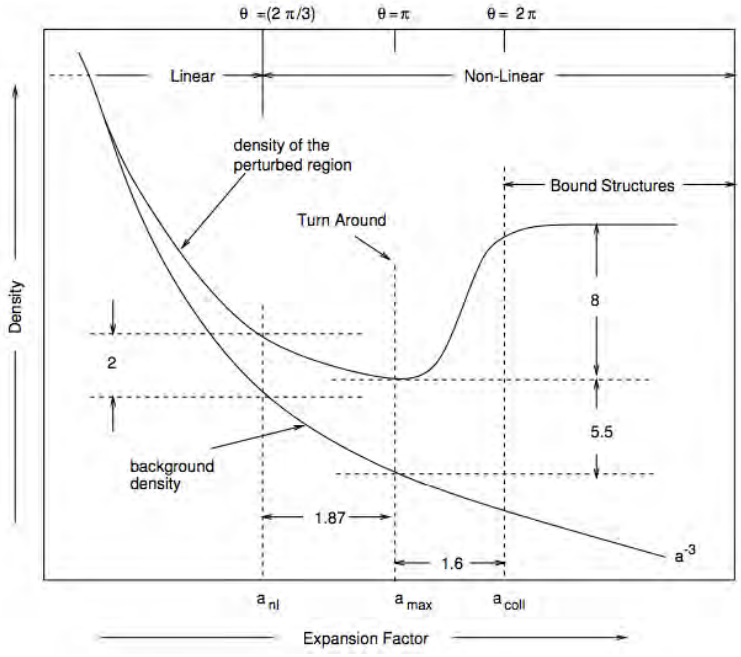

FIGURE 1.2: Evolution of an overdense region in a spherical top-hat approximation. Here $a_{\rm nl}$ refers to non-lineal and $a_{\rm coll}$ to collapse. Figure extracted from Padmanabhan (2002).

the mass fluctuations is given by:

$$\sigma(R)^2 = \int \frac{k^2}{(2\pi)^2} W_s^2(kR) P(k)\, dk, \tag{1.39}$$

where W_s is the top-hat filter in the Fourier space defining spheres of radius R. The amplitude of the power spectrum on the scale of 8 h_{70}^{-1} Mpc is called σ_8, defined as the r.m.s. density variation when a top-hat filter of 8 h_{70}^{-1} Mpc is applied. It can be expressed as:

$$\sigma_8 = \frac{1}{2\pi^2} \int W k^2 P(k)\, dk, \tag{1.40}$$

where W_s is the top-hat filter in the Fourier space:

$$W_s = \frac{3 j_1(kR_s)}{kR_s}, \tag{1.41}$$

with j_1 the first order of a spherical Bessel function.

1.4.2 Supercluster of galaxies

The Large Scale Structure (LSS) of the Universe, as observed today, shows a web-like pattern formed by groups and clusters of galaxies, elongated filaments, widely spread sheets, and voids (e.g. Peebles, 1980; Davis et al., 1982; Bond et al., 1996). The superclusters of galaxies are individual fragments of this LSS, forming sheets or just a weblike structure connecting filaments. The nodes where these filaments cross are the focus of clusters and groups of galaxies. Thus, superclusters of galaxies can be defined as the largest and most massive structures ongoing gravitational effects, although they are not virialized. They probably just passed the quasi non-linear regime described by the Zel'dovich' approximation (1970, see also the "sticking model" by Shandarin & Zel'dovich 1989), and may be close in time to the turn around point. As a consequence, the inter-cluster medium embedded in them (dark halos, gas and galaxies) dynamically interacts and organizes by falling in to the gravitational potential of the massive halos, forming bridges between pairs of clusters and groups.

Superclusters can be identified from groups of clusters of galaxies (e.g. Abell, 1961; Zucca et al., 1993; Einasto et al., 2001; Chow-Martinez et al., 2014) but also from the galaxies distribution, their local density and luminosity (e.g. Luparello et al., 2011; Costa-Duarte et al., 2011; Liivamägi et al., 2012). Recent numerical N-body simulations based on the ΛCDM cosmological model (e.g. *Millennium*, Springel et al., 2005; *Bolshoi*, Klypin et al., 2011; *Illustris*, Vogelsberger et al., 2014) reinforce that these structures are assembled under the effect of gravity and the process of gravitational collapse is still ongoing.

1.4.3 Galaxy clusters

Since 85% of the ICM is present in the form of hot gas, observations of such gas may be subject of study in order to characterize the ICM of clusters. There exist multiple studies of the ICM carried out through X-ray and SZ observations (e.g. Arnaud et al. (2001); Pointecouteau et al. (2005); Croston et al. (2008); Arnaud et al. (2010a); Planck Collaboration et al. (2011, 2013a); Bourdin et al. (2017)). These studies present analyses of the gas density, temperature and pressure of the ICM which confirm the strong similarity, in shape, of the clusters. There exist several models to estimate the mass distribution of galaxy clusters using

different observables. In the following paragraphs some examples used to model the clusters' density profiles are presented.

The isothermal sphere The distribution of galaxies in the central region of rich clusters can be approximated by the King distribution, (King, 1972):

$$n_{gal} = n^0_{gal}\left(1 + \frac{r}{R_c}\right)^{-3/2}, \tag{1.42}$$

with R_c the core radius and n^0_{gal} the central density. However, the mass estimated from this model diverges when $r \leftarrow 0$ and $r \leftarrow \infty$. Moreover, observations of galaxy clusters suggest that the velocity dispersion of galaxies is proportional to their distance from the cluster center. This leads to the analytical model of an isothermal sphere with a density distribution of the shape:

$$\rho(r) = \frac{\sigma_v^2}{2\pi G r^2}. \tag{1.43}$$

This model, originally proposed to estimate the dark matter profile for a self-gravitating isothermal sphere (as detailed by Binney & Tremaine, 1987), is used to estimate the cluster properties in analytical analyses.

Sérsic profile The Sérsic profile, (Sérsic, 1968) is a surface brightness 2D profile for galaxies. It is defined by:

$$\rho(r) = \rho_s \exp(-b_n[\frac{r}{r_s}]^{1/n} - 1). \tag{1.44}$$

where ρ_s is the surface brightness at radius r_s, measured at the half bright surface. The parameter b_n can be approximated as $2n - 1/3$, as shown by Ciotti & Bertin (1999), where n describes the shape of the light profile.

Beta profile Regarding the gas component of galaxy clusters, the density profile can be computed considering the beta profile (Cavaliere & Fusco-Femiano, 1976). This profile is a modified profile adapted for the distribution of gas. Here

the gas distribution is given as:

$$\rho_{gas}(r) = \rho_{go}\left[1 + \left(\frac{r}{r_c}\right)^2\right]^{-3\beta/2}, \tag{1.45}$$

with

$$\beta = \frac{\mu m_p \sigma^2}{kT_g}. \tag{1.46}$$

Navarro-Frenk and White cluster profile Numerical simulations of structure formation have provided a solid description of the gas behaviour under the influence of the key physical processes governing the intra-cluster medium (ICM Nagai et al., 2007; Battaglia et al., 2010). The dark matter density profile of galaxy clusters is approximated by a Navarro-Frenk and White profile, (Navarro et al., 1997, hereafter, NFW):

$$\rho(r) = \frac{\delta_c \rho_c}{(r/r_s)(1 + r/r_s)^2} \tag{1.47}$$

where ρ_c is the critical density at the cluster redshift, r_s is the scaled radius found by Navarro et al. (1997) and δ_c is the characteristic density (for a concentration parameter c) defined as:

$$\delta_c = \frac{200}{3}\frac{c^3}{\ln(1+c) - c/(1+c)} \tag{1.48}$$

Usually this profile is adjusted using an analytical solution of the NFW profile parametrized as:

$$\rho(r) = \frac{\delta_c \rho_{c0}}{x^\gamma (1 + c_{500x}^\alpha)^{(\beta-\gamma)/\alpha}}, \tag{1.49}$$

where α, β and γ are the best fitted slopes for the density profile.

The generalisation of the Navarro et al. (1997) profile (gNFW) for the distribution of dark matter and gas derived from early numerical simulations, as given by Nagai et al. (2007) is:

$$\mathbb{P}(x) = \frac{P_0}{(c_{500}\, x)^\gamma [1 + (c_{500}\, x)^\alpha]^{(\beta-\gamma)/\alpha}} \tag{1.50}$$

where $x = r/r_s$, $r_s = r_{500}/c_{500}$. r_{500} and c_{500} are the characteristic radius and the corresponding concentration encompassing 500 times the critical density of the Universe at the cluster redshift . α, β and γ are the inner, external and transition slopes (at r_s) of the profile. The gNFW profile thus provides a simple parametric description easily tested against observational constraints (e.g., Arnaud et al., 2010b; Planck Collaboration et al., 2013a; Eckert et al., 2013; Sayers et al., 2016; Romero et al., 2017; Bourdin et al., 2017; Ruppin et al., 2018). The afore-cited works have found a very good agreement between the gNFW predictions and the actually observed pressure distribution in X-ray or SZ, at least within the central part of the galaxy clusters.

1.4.3.1 The X-ray observations

As mentioned before, the ICM gas is observed to have temperatures of about 10^8 K, corresponding to 1-30 keV. This gas cools by thermal Bremsstrahlung emission which depends quadratically of the gas density n_e. Then, the X-ray surface brightness is used to estimate the gas density. The X-ray surface brightness by unit of solid angle is given by:

$$S_x = \frac{1}{4\pi(1+z)^4} \int n_e \Lambda(T_e, Z) dl, \tag{1.51}$$

where dl refers to the line of sight, $\Lambda(T_e, Z)$ is the cooling function which depends on the gas metallicity, Z and the temperature T_e. $\Lambda(T_e, Z)$ can be approximated as $\sqrt{T_e}$ Although the X-ray emission is suitable to characterize the ICM gas density and its metallicity, it depends on the cluster' density, then, at lower densities, photon detection requires larger observation times. Also, its redshift dependence makes it difficult to detect and characterize clusters at large redshift. Therefore, clusters can be studied using X-ray observations only in their densest region, i.e. close to the core. Analyses up to the cluster outskirts have as a consequence an increment in the observation time, also, the required instrumentation sensitivity increases, making such studies very challenging.

1.4.4 Scaling relations

In the hierarchical structure formation scenario, galaxy clusters are the largest gravitationally bounded structures. One can consider the hypothesis that clusters evolved purely by gravitational collapse, following a spherical collapse model, as proposed by Jeans (1902). Then, by assuming these systems are virialized, the matter contained in their gravitational potential will be in hydrostatic equilibrium. Also, one can assume the gas fraction in clusters, $f_{gas} = M_{gas}/M_{vir}$ is constant and representative of that of the Universe. Therefore, as a consequence of these approximations, galaxy clusters and groups halos have self-similar internal structure. This property allows to construct scaling relations between the cluster principal properties (Kaiser et al., 1995; Bertschinger, 1998).

This behaviour is observed for their global thermodynamical properties (e.g., Giodini et al., 2013a) as for their internal distribution (e.g., Pratt et al., 2019). The gas thermal pressure is a remarkable example of this self-similar behaviour. The integrated pressure over the volume of the cluster, i.e., the SZ Comptonization parameter, has proven to be an excellent proxy of the total gas content, thus of the total mass of the halo, as the thermal pressure is mildly affected by non-gravitational physics (AGN feedback, radiation cooling, etc) with respect to other proxies (e.g., X-ray total luminosity Pratt et al., 2019; Mroczkowski et al., 2019, for recent reviews).

Let us consider a spherical region of radius $R_{\Delta c}$ with a mean density of $\Delta c \rho_c(z)$ at redshift z. The total mass is given by $M_{\Delta c} = (4\pi/3)\Delta_c \rho_c(z) R^3_{\Delta c}$. The critical density is defined at any redshift as $\rho_c(z) = \rho_{c0} E(z)$, with $E(z)$ as defined in Equation 1.23. Then the cluster radius as function of redshift can be approximated as:

$$R_{\Delta_c} \propto M^{1/3}_{\Delta_c} F_z^{-(2/3)}. \tag{1.52}$$

with F_z the source flux. Now, considering the gas in clusters satisfies $k_B T \propto , M_{\Delta_c}/R_{\Delta_c}$, then we have:

$$M_{\Delta_c} \propto T^{3/2} F_z^{-1}. \tag{1.53}$$

This last equation can be used to derive scaling relations from X-ray observables. These relations generally link by a power law the cluster mass with a measured proxy (e.g. the luminosity, gas temperature or the galaxies' velocity dispersion).

Some of the most used scaling relations are between X-ray luminosity and temperature, mass and temperature and luminosity-mass relations (e.g. Ettori et al., 2004; Arnaud et al., 2005; Kotov & Vikhlinin, 2005; Pratt et al., 2009; Giodini et al., 2013b, the latter provides a review regarding scaling relations for clusters):

$$F_z M_{gas} \propto \quad T^{3/2}, \tag{1.54}$$

$$F_z^{-1} L_x \propto \quad T^2, \tag{1.55}$$

$$F_z^{-1} L_x \propto \quad (F_z M_{gas})^{4/3}, \tag{1.56}$$

$$F_z^{4/3} S_x \propto \quad T, \tag{1.57}$$

However, it is important to recall that scaling relations are a consequence of the assumption of hydrostatic equilibrium. In a more realistic scenario, the dynamical state of the clusters at radii larger than R_{500} ($\Delta_c = 500$) is perturbed by accretion of the infalling material from the filaments and dynamical interactions. Then, in this region one can no longer consider the gas to be in a virialized state. Thus, the characterization of the gas in the cluster outskirts is a crucial step to understand the formation and evolution of these structures.

The SZ effect observations The Sunyaev-Zel'dovich (SZ) effect is a result due to the inverse Compton effect between the hot electrons in the ICM and the photons of the CMB. As a result of this interaction, the spectrum of the CMB shows a deformation toward the regions of clusters.

The intensity of the SZ effect is characterized by the dimensionless Comptonisation parameter y. As defined in Planck Collaboration Int. V (2013), the Comptonisation parameter corresponds to the product of the average fractional energy transferred per collision, by an electron to a photon, and the average number of collisions, such that:

$$y = \frac{\sigma_T}{m_e c^2} \int P(l) dl, \tag{1.58}$$

Then the SZ flux expressed as the integrated Compton parameter, Y, within a given solid angle, Ω, is proportional to the thermal pressure of the ICM gas integrated over the line of sight:

$$Y(\Omega) = \frac{\sigma_{\rm T}}{m_{\rm e} c^2} \int_{\Omega} d\Omega \int_{los} P(l)\, dl \tag{1.59}$$

where $\sigma_{\rm T}$ is the Thomson cross-section, m_e is the mass of the electron, c is the speed of light and P is the pressure produced by the plasma of thermal electrons along the line of sight. Since the SZ effect is independent of the distance of the object to the observer, the SZ effect is related to the redshift only by the angular size at which the cluster is observed. The virial approximation for galaxy clusters is:

$$k_B T_e = \beta \mu m_p \frac{G M_{tot}}{R}, \tag{1.60}$$

where M_{tot} is the mas comprised in a sphere of radius R and β is a mass distribution factor and μ the molecular mass of ICM particles. Then, assuming the ICM gas can be approximated as an ideal gas, the $Y - M$ relation is:

$$Y(R_\Delta) \propto \frac{\sigma_\tau}{m_e c^2} \frac{\nu}{\nu_e} \left(\frac{\Delta G^2 H(z)^2}{16} \right)^{1/3} f_{gas} M_{tot,\Delta^{5/4}}, \tag{1.61}$$

where $_e$ the electron mass, f_{gas} is the ratio of the ICM gass mass to the total cluster mass, which is assumed to be constant.

As a complement to X-ray observations, SZ observations offer the possibility to study the integrated pressure over statistically significant samples of clusters (e.g. Planck Collaboration X, 2011; Planck Collaboration Int. III, 2013; Czakon et al., 2015; Bender et al., 2016; Dietrich et al., 2019) demonstrating a coherent view of their gas content between X-ray and millimetre measurements. The increasing coverage and improving resolution and sensitivity of SZ observations have also allowed to constrain the pressure distribution over the whole volume of clusters (Plagge et al., 2010; Planck Collaboration et al., 2013a; Sayers et al., 2013; Eckert et al., 2013).

Plagge et al. (2010) extracted the pressure profiles of 15 clusters up to 2 $\times$ R_{500} using the South Pole Telescope. Their analysis shows a consistency with the X-ray cluster parameters. Planck Collaboration et al. (2013a) presented an analysis of the extracted SZ signal from the Planck satellite for 62 clusters. Their results combine X-rays for r $<\mathrm{R}_{500}$ and SZ signal for r $>\mathrm{R}_{500}$ to extract the pressure profiles up to 3 $\times$ R_{500}. Another study carried out by Bourdin et al. (2017) used the Planck full mission data to extract the pressure profiles for two samples of galaxy clusters, 61 clusters at z~ 0.15 and 23 clusters at z~ 0.56. They combined X-ray for the inner part of the clusters and SZ for the outskirts.

1.4.5 The dispersed component of superclusters and the filaments

Filamentary structures in superclusters are observed between galaxy clusters forming bridges of galaxies or, at larger scales, composing chains of clusters. The complex morphology of superclusters has been studied previously by several authors. Basilakos et al. (2001) studied the supercluster morphology using a differential geometry definition of shape. Their results suggest that filamentary morphology was the dominant feature for their supercluster sample. Moreover, Costa-Duarte et al. (2011) found, using Minkowsky functionals, that half of their sample have a pancake morphology while the other half exhibited filamentary morphology.

Due to its relatively low density and temperature, the WHIM is very difficult to observe with the current observational facilities (in X-rays and SZ effect). However, there exist some observational evidence of a WHIM in filaments. Among the best studied cases are the pairs of clusters: A222-A223 (Werner et al., 2008; Dietrich et al., 2012), A3391-A3395 (Tittley & Henriksen, 2001) and A399-A401 (Sakelliou & Ponman, 2004), which seem to be connected, in each case, by a gas bridge detected in X-rays. In another study, Planck Collaboration et al. (2013b) report an study of the SZ signal detected between pairs of clusters at separations up to 10 Mpc. More recently, Tanimura et al. (2017) found evidence of gas in filaments by stacking the SZ signal between pairs of luminous red galaxies (with a separation of between 6 and 10 Mpc). They estimated a gas temperature of T$\sim 8.2 \times 10^7$ K (kT $= 7.1 \pm 0.9$ keV).

In this work, we define the material (galaxies, gas, dark matter) between two clusters or groups as a “bridge” if this material is, somehow, denser than the surroundings. Here we shall call a chain of three or more clusters/groups connected by bridges a “filament”.

Moreover, it has been observed that the different components of the LSS are frequently aligned with each other. Indications of such alignments in the optical bands have a long history (e.g. Sastry, 1968; Carter & Metcalfe, 1980; Binggeli, 1982; Lambas et al., 1988; West, 1994; Plionis & Basilakos, 2002; Lee & Evrard, 2007; Hao et al., 2011). There are many important results connecting the cosmological alignments, the filaments and the evolution of galaxies and their systems.

In what follows we mention only some examples. Plionis & Basilakos (2002) analysed the dynamical evolution of clusters in a hierarchical scenario under the hypothesis of clusters merging within large scale filaments. They concluded that clusters with traces of dynamical activity are significantly more aligned with their nearest neighbours. Altay et al. (2006) found that the shapes of nearby clusters are aligned if the clusters are connected by a filament. They conclude that matter infalling along filaments is an important factor in galaxy cluster intrinsic alignments a hypothesis proposed originally by West (1994). Godłowski & Flin (2010) observed alignments of groups of galaxies within the Local Supercluster on scales up to 20 Mpc scales. Also, correlations between two clusters have been detected on scales up to 30 Mpc using X-ray observations (Chambers et al., 2002; Wang et al., 2009; Paz et al., 2011). On the other hand, Chen et al. (2017a) performed an analysis of the color, magnitude, morphological type and activity type of galaxies in the neighborhood of filament candidates. Their results show that the red galaxies are located nearer to the skeleton of the filament than the blue galaxies. Another study carried out by Zhang et al. (2015) presents the results from observations that suggest that the major axis of elliptical galaxies tends to appear orientated parallel to the filament.

In other words, there are many evidences that the evolution of galaxies and their systems is deeply connected to their environment, especially with the filamentary structure that patterns the LSS.

1.5 The galaxies that populate the LSS

Galaxies are currently classified by their shape, activity, metallicity, among other properties. For example, galaxies can be classified by their Hubble morphological type in elliptical (E), spiral (S) or irregular (I). Galaxies can also be classified by their color, defined as the difference in brightness between two bands.

Moreover, galaxies can also be classified based on their emission lines using the galaxy spectrum. In the following paragraphs we introduce the galaxy classifications used on this work.

1.5.1 Galaxy morphological and spectral classification

One way to classify the galaxies is according to their spectrum, e.g., Kennicutt (1992) found a relation between the spectra of galaxies with their morphological Hubble type establishing one of the most used spectral classifications.

The integrated spectrum of galaxies is the integrated spectrum of the stars, dust and gas (Jones & Lambourne 2014). For example, the optical spectrum of HII regions present the so−called forbidden lines. These lines are produced only in regions of very low density ($n_e \lesssim 10^3$ cm^{-3}). Strong forbidden lines like [NII] λ6548 Å and [OIII] λ5007 Å are seen in HII regions. At higher densities collisional de-excitation begins to play a role (Osterbrock 1989). In spiral and irregular galaxies, the contribution of HII regions to the spectrum of the galaxy is significant while for the elliptical galaxies their contribution is not relevant since they do not have HII regions. Therefore, considering these emission lines, galaxies are classified in late (Sp,I) and early (E,S0) type.

The motion of the galaxy components, stars, dust and gas, within the galaxy, is observed as a Doppler shift of the spectral lines, and as a result absorption and emission lines become broader. Then, the systemic velocity of the galaxy can be measured from the wavelength shift of the spectral lines with respect to the rest frame wavelength.

1.5.2 Activity classification

Moreover, the spectrum of a galaxy can also provide information about the presence of an active galaxy nucleus (AGN) or star-formation (SF). AGNs come in a variety of types and using the optical spectra are classified as Seyfert 1/2, LINERs, QSOs and blazars (Jones & Lambourne 2014). The unified model for AGNs (Barthel 1989; Antonucci 1993; Urry & Padovani 1995) proposes that the principal component, which powers the AGN, is a supermassive black hole (SMBH). An active SMBH at the center of active galaxies is partially hidden by a torus of gas and dust. Under this model, the different types of AGNs can be explained by the orientation of the source with respect to our line of sight. Unlike normal galaxies, the integrated spectrum of a galaxy with an active nucleus, like the Seyfert 1/2 galaxies is identified by the presence of very-broad/broad Hα emission lines.

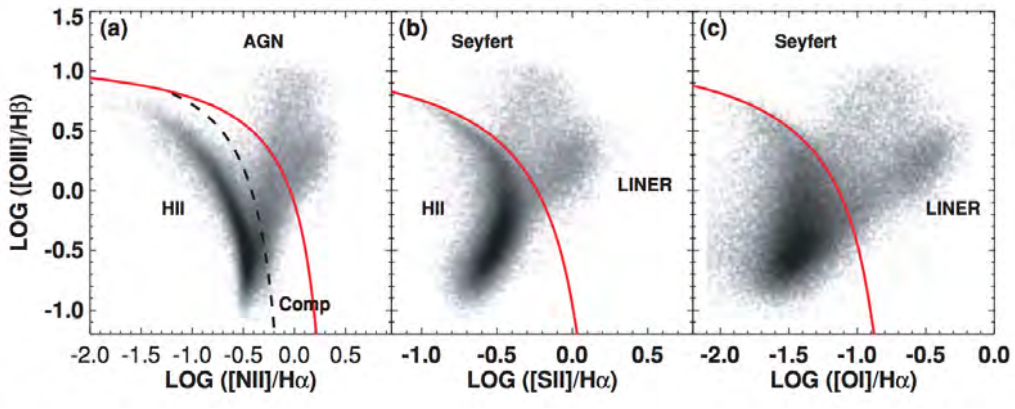

FIGURE 1.3: (a) The [NII]/Hα versus [O III]/Hβ diagnostic diagram for SDSS galaxies. The extreme starburst galaxies are located near the solid line and the AGN and HII–region–like galaxies are divided by the dashed line. (b) The [SII]/Hα versus [OIII]/Hβ diagnostic diagram. c) the [O I]/Hα versus [O III]/Hβ diagnostic diagram. Figure extracted from Kewley et al. (2006) .

The BPT (Baldwin, Phillips & Terlevich, Baldwin et al. 1981) diagnostic diagram allows to classify the galaxies as Seyfert 2, LINER or star forming galaxies based on the presence and ratio of certain emission lines in the spectrum. Figure 1.3 depicts the diagnostic diagrams for SDSS galaxies (Kewley et al., 2006). The galaxies that lie below the dashed line are classified as HII–region–like galaxies and those that lie above the dashed line are classified as AGNs. Galaxies that lie in between these two classifications lines are a mix of AGN and H II and are classified as composites or TOs (transition objects). Composite galaxies are likely to contain a metal–rich stellar population plus an AGN (Kewley et al. 2006).

1.5.3 Galaxies in clusters

The galaxy population inside galaxy clusters is predominantly composed of early type galaxies and about 35% is observed to correspond to late type galaxies (e.g. Roncarelli et al., 2010). Dressler (1980) carried out an analysis over a sample of 55 galaxy clusters. His results shown that the fraction of spiral galaxies in clusters correlates strongly with the environment in which they reside (a fraction of 80% of galaxies in the field, 60% in the cluster outskirts and 0% in the cluster center). Generally, the brightest cluster galaxy (BCG) is located near the center of the clusters. For rich clusters, the BCG galaxy is typically classified as an elliptical galaxy (Lauer & Postman, 1992). A fraction of these clusters presents

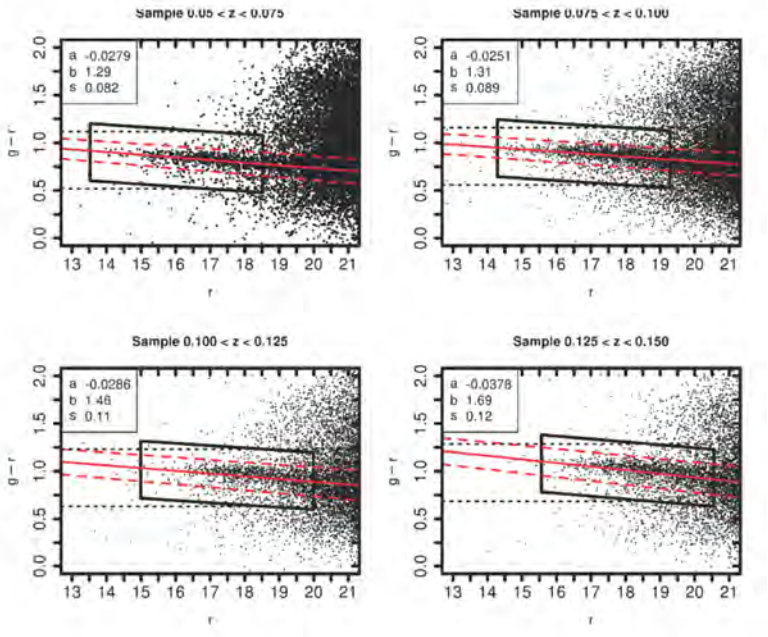

FIGURE 1.4: Color-magnitude diagrams for normal luminosity clusters at different redshift. The dotted lines are the first selection in colour and the black solid rectangles are the final red-sequence selections. The red solid and red dashed lines are the regression fit and errors, respectively. Figure extracted from Trejo-Alonso et al. (2014).

a very luminous elliptical galaxy with an extended low surface brightness envelope, called cD galaxy (Matthews et al., 1964). cD type galaxies are the result of several dynamical processes, Ostriker & Hausman (1977) suggest that the formation of these galaxies is the result of several mergers. Moreover, Dressler et al. (1997) found that the fraction of elliptical galaxies is higher in clusters at larger redshifts (z$\sim$0.5) compared to low redshift clusters. The observation and identification of the galaxy population of clusters have been used to detect clusters and determine the galaxy cluster membership (e.g Gladders & Yee, 2000, use the color-magnitude space to detect galaxy clusters).

The red sequence of galaxies in clusters

The color index of galaxies is observed to follow a straight relation with the galaxies' magnitude in clusters. When red galaxies are plotted in the color magnitude space, galaxies belonging to the cluster follow a linear distribution, called the red sequence. This configuration in the color magnitude diagram suggests that at a given moment galaxies were stick together, and evolved at the same time, $z \sim 2$

(e.g. Bower et al., 1992; Ellis et al., 1997; Gladders et al., 1998). The red sequence distribution can be adjusted by a line with a slope that suggest that the fainter galaxies are bluer than the brighter ones. Different analyses have shown that this sequence is related with the mass-metallicity relation, as shown by Kodama & Arimoto (1997) and Kauffmann & Charlot (1998) and not with the age of the cluster. As described by Kodama & Arimoto (1997), supernovae events in galaxy clusters would warm the interstellar medium and can evaporate the gas of low mass galaxies. This process results in an increment on the metallicity with the mass: the most massive and luminous galaxies have larger metallicities and are redder than low luminosity galaxies. Moreover, an analysis carried out by Trejo-Alonso et al. (2014), on a sample of 56 Abell cluster, found that X-ray underluminous clusters show a flatter red sequence slope than X-ray normal luminosity clusters. They suggest that underluminous clusters may be younger systems than clusters of normal X-ray luminosity.

Chapter 2

Detection of large scale structures: GSyF & GFiF algorithms

Recently, with the availability of large sky area databases, like the SDSS, the development of accurate structure detection algorithms has become a major concern in astronomy. Visually, the galaxy distribution shows filamentary chain-like structures which connect massive clusters and groups. However, the identification of these structures through a computational algorithm is not an easy task. A good algorithm should provide an identification which resembles the human visual identification. It also should provide quantitative results and should be founded in a robust and well defined numerical theory and all of this in an acceptable amount of time with reasonable computational resources.

Currently, several structure finding algorithms are available. For example, Cautun et al. (2013) proposed an automated algorithm which takes into account the density, tidal field, velocity divergence and velocity shear of the galaxies. Their results show a reliable identification of structures over a N-body simulation. Another approach, presented by Aragón-Calvo et al. (2010), makes use of segmentation techniques to trace the spines of the filaments. They applied their algorithm over a selected N-body simulation and compared their results against a heuristic Voronoi tessellation (VT).

Moreover, there have been several attempts to trace the distribution of the cosmic web using the SDSS database. For example, Platen et al. (2011) applied a Delaunay triangulation and VT over the galaxy positions to estimate the local

density. Then, they interpolated the density between neighbors to recover contiguous high density regions. Tempel et al. (2014) applied a Bissous model which adjusts cylinders to trace the filament spine. A different approach is presented by Chen et al. (2016): they apply kernel density estimators (KDE) in order to smooth the galaxy distribution with a Gaussian kernel. Along the same path, Bond et al. (2010) applied a smoother Hessian major axis filament finder for the SDSS. They extracted the filaments by computing the Hessian of the density distribution of the galaxies.

Also, combinations of different techniques have proven to be suitable for the identification of LSS structures. Smith et al. (2012) applied a multi-scale density estimator in combination with a Friends of Friends (FoF) algorithm over the SDSS galaxies. Their method allows them to recover structures from clusters up to filaments of $\sim 10 h^{-1}$Mpc length. Moreover, Alpaslan et al. (2014) found, using the Galaxy and Mass Assembly survey (GAMA), that there are fine filaments embedded inside the SDSS voids. These structures, called 'tendrils', have a lower density than the SDSS filaments and appear to be morphologically distinct, they are more isolated and span over shorter distances.

Since different methodologies identify relatively distinct structures, and every one of them has different advantages and drawbacks, we developed our own methodology, trying to improve the results over the available ones.

Therefore, in this Chapter we present in detail the algorithms we developed for the detection of internal structures of superclusters using the galaxy positions (RA$= \alpha$, Dec$= \delta$ and redshift$= z$). The *Galaxy System Finding algorithm* (GSyF) objective is to detect clusters and groups of galaxies, hereinafter called systems of galaxies, using VT and hierarchical clustering (HC) analyses. The *Galaxy Filaments Finding algorithm* (GFiF) seeks to detect the more elongated structures (bridges, filaments, and tendrils) using classical pattern recognition and machine learning methods. Since these algorithms use only the position of galaxies, they can be applied to various galaxy surveys and catalogs. Hereafter, the N galaxies inside a volume box (containing the supercluster volume) are considered as a set of points $x_1, x_2, ..., x_N \in X$, all being part of a sample X.

2.1 Machine learning methods applied to astronomy

2.1.1 Density estimators

The local number density is a useful measurement to characterize the environment in which a galaxy resides. For this work we implemented two techniques for measuring this local density: VT and KDE. Both methods have different advantages and limitations. For instance, VT is suitable to measure with accuracy the local density at the galaxy position while the strength of KDE is to measure the density on larger scales.

VT density estimator Voronoi tessellation (Voronoi, 1908) can be applied both over spatial distributions of galaxies to estimate the local 3D volume number density and over surface distributions to estimate the local 2D surface number density. VT has been widely used for the detection of overdensity regions in sky slices (e.g. Scoville et al., 2013; Darvish et al., 2015). Another approach is presented by Cybulski et al. (2014) which uses VT for the identification of filamentary structures in the Coma cluster region.

The VT of a sample X, Vor(X), can be defined as the subdivision of a 2D plane or a 3D space into cells with the property that the seed point $x_i \in X$ is located in the cell v_i if and only if the Euclidean distance $D_E(x_i, v_i) < D_E(x_i, v_j)$ for each $v_j \in X$ with $j \neq i$, i.e. VT partitions the space into optimal polygonal cells in a way that there is one cell for each galaxy position $x_i \in X$ (e.g. Platen et al., 2011).

Then, the density at x_i is determined as $d_i = 1/v_i$, with v_i being the volume (or area) of the cell around the galaxy x_i. One of the advantages of VT is that it provides a density estimation of high resolution, i.e. the density is estimated at the position of each galaxy.

Kernel density estimator In some cases VT fails on the identification of large overdensity regions, as mentioned by Cybulski et al. (2014). An alternative for the estimation of the number density is to apply KDE. In general, KDE methods work

by adjusting a kernel function over each observation in the sample. However, the choice of the correct kernel model and its intrinsic parameters are still under investigation within the community. Also, there are several attempts to apply adaptive Gaussian model kernels, in other words, to change the size of Gaussian model as a function of different parameters, e.g. the distance to the nearest neighbor (Chen et al., 2016) or a weighting function (Darvish et al., 2015).

For this work we set the kernels as Gaussian models inside VT cells. Therefore, VT method is first applied over the sample positions (see Section 2.1.1) and Gaussian ellipsoids are fitted inside each VT cell. Afterwards the Gaussian kernels $\phi_{\mathbf{\Sigma}}$ are calculated from the eigenvalues and eigenvectors of the fitted ellipsoids. Then, each n-dimensional kernel $\phi_{\mathbf{\Sigma}}$ centered in μ with covariance matrix $\mathbf{\Sigma}$ is represented as:

$$\phi_{\mathbf{\Sigma}}(\mathbf{x}-\mu) = (2\pi)^{d/2}|\mathbf{\Sigma}|^{-1/2}e^{-1/2(\mathbf{x}-\mu)^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu)}. \tag{2.1}$$

Then the KDE can be estimated as:

$$\hat{p}_{\mathrm{KDE}}(\mathbf{x}) = \sum_{i=1}^{N} \alpha_i \phi_{\mathbf{\Sigma}_i}(\mathbf{x}-\mathbf{x}_i), \tag{2.2}$$

where α_i is a weight factor calculated from the VT cell volume as v_i.

The identification of the overdensity regions is carried out through the projection of KDE kernels in 2D planes. This is done by superposing a regular rectangular grid to the data. Then, we estimate the density at each grid intersection by calculating the average density of all kernels that overlap at that point. Thus, observations closer to an evaluating point will contribute more to the density estimation than points that are far from it. Consequently, the density will be higher in areas with many observations than in areas with few observations. The identification of contiguous overdensity depends on the number of times the covariance of each kernel is being considered.

2.1.2 Hierarchical Cluster Analysis

Hierarchical clustering is a machine learning method whose objective is to group objects with similar properties. It has been used in different areas of science such as artificial intelligence, biology, medicine and business. In general, it can

be used to carry out pattern recognition analysis, allowing to regroup, segment and classify any kind of data. This method is equivalent to a reduction of the dimensionality of the data and decreases considerable the computing time. In astronomy, the most popular application of HC has been for the detection of substructures inside galaxy clusters, following the algorithm developed by Serna & Gerbal (1996). This algorithm considers the positions, redshifts and potential binding energy between pairs of galaxies to detect substructures (see also Guennou et al., 2014).

In general, hierarchical clustering methods are suitable for this kind of application since they allow the detection of structures on different scales and can be applied over different coordinate reference frames.

For this work, we chose an agglomerative hierarchical clustering method with a Ward's minimum variance clusterization criteria, described in detail by Murtagh & Legendre (2014). A detailed description of the HC algorithm can be found in Theodoridis & Koutroumbas (2009); Theodoridis et al. (2010) and Murtagh & Contreras (2011). In general, Ward's method works by merging the groups following the criterion:

$$\Delta D(c_1, c_2) = \frac{|c_1||c_2|}{|c_1| + |c_2|} ||c_1 - c_2||^2. \tag{2.3}$$

where ΔD is a term that measures the distance between two groups c_1 and c_2, respectively.

In our case, initially each point is considered as a group, sub-cluster or singleton, then each group can be agglomerated with a neighbor that has the minimum ΔD distance. The agglomeration continues until all points are grouped together into a single group.

The results of the HC clusterization can be represented by a dendrogram or hierarchical tree. A dendrogram represents, in a graphical form, the connections between elements and groups in different levels of agglomeration. The length of each connection line in the tree corresponds the distance between two elements or centroids connected. This representation also allows visualizing the principal branch structures where the singletons are the final leaves. The number of desired groups N_{cut} is, therefore, obtained by cutting the hierarchical tree at a certain level. The exact value of this level depends on the characteristics of the

sample or, more properly, on the underlying physics that is used to define the groups.

Each created group can be represented by a 2D/3D Gaussian model, $P_j(x)$. This allows us to classify the groups by their Gaussian properties, e.g. centroid (mean position, C_j), richness (number of members, N_j) and compactness (covariance, σ_j).

2.1.3 Graph definition

Graph theory-based algorithms have shown to be a suitable tool to analyze complex networks. Some of the most common subjects where these algorithms are applied successfully are social networks, computer vision, statistics, business and transportation networks.

A graph is a representation of the connections in a network. It is composed of "nodes" and "edges", where each node represents an object, and the edges represent the connections between each two nodes. Also, the edges can have weights that represent the strength of the connection. An undirected graph has edges that do not have direction. Generally, a graph is defined as $G = (U, E, W)$, with n nodes (or vertices) $u_i \in U$, m edges $e_{kl} \in E$ and a weight set W with a w_{kl} for each edge e_{kl}. The information of a graph can be represented by a square adjacency matrix. The values of the matrix elements indicate the weight of the connection between nodes. Hence, the adjacency matrix A of the graph G is defined as:

$$[A]_{kl} = \begin{cases} 1 & \text{if } (u_k, u_l) \in E \\ 0 & \text{otherwise} \end{cases}$$

where u_k and u_l are nodes in G.

2.1.4 Minimum Spanning Tree (MST)

A spanning tree connects nodes in a graph in a way that does not produce cycles. A graph can contain several unconnected spanning trees. Since the edges in a

graph can have weights, the minimum spanning tree algorithm (Graham & Hell, 1985) searches for a spanning tree that minimizes the total weight. This algorithm traces a tree-like continuous path for a group of edges and nodes in an optimal way. In particular, Kruskal's minimum spanning tree algorithm analyzes the edges in sequence, sorting them by weight. At the beginning, the shortest edge is analyzed and this is taken as the first tree branch. Then, the nodes are added to the tree under three conditions:

(i) only one node is added to the tree;

(ii) a node is added based on the number of connected edges;

(iii) their edges cannot be connected to another existing node in the tree.

The process continues with the following edges in the graph until all connected edges are analyzed. Finally, the tree is extracted from the graph and the process begins again with the remaining nodes until all are tested. As its name remarks, the result is a forest of minimized independent trees.

2.1.5 Dijkstra's shortest path

Dijkstra's algorithm (Dijkstra, 1959) is a classical method for searching the shortest path between two nodes in a graph. The path of length e_{kl} is defined between two nodes u_k and u_l as a sequence of nodes connected $u_1, u_2, ..., u_m$ if $k \neq l \; \forall \; k, l \in 1, ..., m$. In general Dijkstra's algorithm works as follows:

(i) The origin is selected by taking the node at the beginning of the path, u_0.

(ii) A distance value is assigned to all nodes: set to zero for the origin, $s(u_0)$, and to infinity for all the other nodes, $s(u_i) = \inf$.

(iii) Then, all nodes are marked as unvisited and u_0 is marked as current node a.

(iv) Subsequently, the algorithm calculates the distance from the current node a to all the unvisited nodes connected by the edges e_i as $s_{new} = s(e_{ai}) + w_{ai}$; here $s(e_{ai})$ is the distance from a to the node u_i and w_{ai} is the weight of the edge e_i.

(v) If $s(e_{ai}) + w_{ai} < s(e_i)$, then the distance is updated and the connected node label is updated as the current a.

(vi) After visiting all neighbors of the current node, they are marked as visited. A visited node will not be checked again; then the recorded distance $s(e_{ai})$ is final and minimal.

(vii) Finally, if all nodes have been visited, the algorithm stops. Otherwise, the algorithm sets the unvisited nodes with the smallest distance (from the initial node u_0, considering all nodes in the graph) as the next "current node" and continues from the second step.

A detailed description of the algorithm can be consulted in Santanu (2014).

2.2 Properties of systems

2.2.1 Coordinate transformation

In general, sky projected coordinates, α and δ, and redshift z can be converted to rectangular coordinates by applying the following transformation:

$$X = D_C \ \cos(\delta)\cos(\alpha)\,, \tag{2.4}$$

$$Y = D_C \ \cos(\delta)\sin(\alpha)\,, \tag{2.5}$$

$$Z = D_C \ \sin(\delta)\,, \tag{2.6}$$

where D_C is the co-moving distance obtained from Equations 1.23 and 1.22, by using the redshift and the cosmological parameters.

2.2.2 Velocity projection effects

In real astronomical observations one can observe that the use of radial velocity (redshift) is subject to projection effects in the line of sight called "Fingers of

God (FoG)". This effect is a result of the galaxies peculiar velocities added to the system velocity. As a consequence, the distribution of galaxies that are members of gravitationally bounded systems is elongated along the line of sight.

2.2.3 Virial mass and radius estimation

The virial mass and radius were estimated in order to characterize the systems of galaxies and to correct the positions of their members for the FoG effects.

This was achieved by applying a simplified version of the algorithm proposed by Biviano et al. (2006). This algorithm calculates iteratively the system's virial mass and radius, under the assumption of dynamical equilibrium, relative isolation of the systems and roughly spherical shapes. For this implementation, the algorithm does not use the surface pressure term correction based on the concentration parameter. This can lead to an overestimation of the virial radius, but, for this geometric analysis, the virial approximation is enough.

In summary, this algorithm works as follows:

(i) The galaxies are selected, in the $\alpha \times \delta$ projection, inside a cylinder of radius R_a = 1 h_{70}^{-1} Mpc, hereafter, aperture. In the line-of-sight (LOS) direction, the galaxies are selected within a range in velocity up to $S_a = \pm$3000 km s^{-1} with respect to the mean cluster velocity.

(ii) The mean velocity, v_{LOS}, and velocity dispersion, σ_v, are calculated using as a robust estimation Tukey's biweight (See eq. 9 in Beers et al., 1990) for the galaxies inside the cylinder.

(iii) The projected harmonic radius R_h of the galaxies inside the aperture within Δz is calculated as:

$$R_h = \frac{N(N-1)}{\sum_{i>j} R_{i\,j}^{-1}}, \tag{2.7}$$

with $R_{i\,j}$ as the distance between the galaxies projected positions (Girardi et al., 1998).

(iv) The mass M_a inside the cylinder of aperture R_a is estimated as:

$$M_a = \frac{3\pi}{2G}\,\sigma_v^2\,R_h, \tag{2.8}$$

where G is the gravitational constant and $3\pi/2$ is the deprojection factor.

(v) The virial radius is estimated as:

$$R_{vir}^3 = \frac{\sigma_v^2 \, R_h}{6\pi \, H^2(z)}, \tag{2.9}$$

by assuming a spherical model for nonlinear collapse, that is, by taking the virialization density as $\rho_{vir} = 18\pi^2[3H^2(z)]/[8\pi G]$, and approximating M_{vir} as M_a,

(vi) Then, the aperture R_a is updated to the calculated R_{vir} value, the mean velocity to v_{LOS}, and S_a to σ_v, defining a new cylinder.

The steps i-vi are repeated iteratively until the radius R_{vir} converges. M_{vir} is finally calculated at the end of the iteration process.

Then, the correction for the FoG effect is carried out by adjusting the position of the N_{mem} galaxies inside the final cylinder. This is done by scaling their comoving distances along the cylinder to the calculated virial radius.

2.3 Galaxy System-Finding algorithm (GSyF)

The GSyF algorithm was developed to detect the denser components of the LSS, systems of galaxies, and to correct the position of the system's galaxy members for the FoG effect. These are necessary steps to carry out before the application of GFiF to detect elongated LSS structures.

In this context, we implemented two strategies for the detection of systems. The first detects groups of galaxies above certain density value. The second, refines galaxy membership through a approximation. We need to point out that we start with the projected distribution of the galaxies since we do not have a priori the space distribution of the galaxies due to FoG effects.

The implementation of a new technique for the detection of galaxy systems allows the identification of new systems possibly not known before (especially the

galaxy groups). Also, this can lead to an improvement of the membership estimation of the superclusters themselves. In the following sections the GSyF algorithm is described in detail followed by a description of the optimization processes implemented[1].

2.3.1 Surface density baseline contrast

Since systems of galaxies are high to medium density galaxy concentrations, a measurement of the density contrast in terms of the density background is a convenient measure. However, since the distribution of galaxies in space is not isotropic, it is not possible to set directly a background density from the projected positions of the galaxies. Then, it is necessary to simulate a sample of equivalent isotropic distribution of the galaxies, in order to set the baseline value (see, e.g. Cybulski et al., 2014).

The methodology implemented by GSyF to measure the local surface density consists in:

(i) The baseline (background) density value, d_{bas} is estimated by simulating 1,000 random isotropic galaxy distributions each with the same number of galaxies over the same area.

(ii) The local surface density is calculated for each galaxy by applying the VT method (section 2.1.1) to the galaxies projected distribution in the plane of the sky. The VT individual area of the galaxy can be directly converted to a surface density estimation ($d_i = 1/a_i$), in this case in units of deg^{-2}.

$$d_{bas} = \frac{1}{m}\sum_{j=1}^{m}\frac{1}{n}\sum_{i=1}^{n} d'_{i,j}, \tag{2.10}$$

where $d'_{i,j} = 1/a'_{i,j}$ corresponds to the area of the point x'_i for the randomization j.

[1]The description of GSyF algorithm as well as the strategy to optimize its parameters using simulated mock volumes is presented in Santiago-Bautista et al. (2019b)

(iii) The density contrast, δ_i, is calculated from the VT densities with respect to a baseline (background) as:

$$\delta_i = \frac{d_i - d_{bas}}{d_{bas}}. \tag{2.11}$$

(iv) Only the N_{gal} galaxies with density contrasts above a certain value ($\delta_i \geq g$) or, equivalently, $d_i \geq (1+g)\, d_{bas}$ are selected;

here, the g parameter needed to be set by an optimization step, described in section 2.5.2. The density contrast was used to find contiguous cells with similar contrast value. Since the distribution of galaxies is not homogeneous in the sky at different redshifts, we need to calculate independent baselines values for each volume box under study. The selection of galaxies above the background density baseline is equivalent to separate galaxies from under-dense regions, i.e. field and void galaxies.

2.3.2 Grouping the galaxies using HC

GSyF applies the HC clustering method (section 2.1.2) over the galaxy projected positions (RA, Dec, $1000z$) (the 1000 factor corresponds to the weight needed for z values to be comparable to the sky coordinates values). The number of groups taken from the analysis is defined as a cut of the HC tree, fixed to $N_{cut} = N_{gal}/f$, with a segmentation parameter f, which is the expected mean number of elements per group. From the clusterization, only groups with $N_j \geq 3$ member galaxies were retained as system candidates. However, the selection of the optimal number of groups in the clusterization methods is still a topic under investigation in the pattern recognition community. Therefore, the value of f needed to be optimized for each volume under analysis (section 2.5.2).

2.3.3 Systems virial refinement

After identifying the galaxy system candidates, GSyF applies a refinement for galaxy membership and estimates the system's virial properties. This was achieved by applying the method described in section 2.2.3. In each iteration the virial estimation works as a filter for system membership. It also allows to re-group

duplicated/ over-clustered groups and permits to discard false detections of HC systems. For the first iteration of the virial refinement, GSyF uses the position of the brightest galaxy of the HC group as the system center. When the virial radius R_{vir} converges, the membership of the systems, N_{mem}, is established and the virial mass M_{vir} is calculated. Finally, the positions of the galaxies inside the final cylinder of radius R_{vir} are corrected for FoG effects. This is done by scaling their co-moving distances along the cylinder to the calculated virial radius. The pseudo-code of the GSyF methodology is presented in Algorithm 1 (Appendix A.1).

2.4 Galaxy Filament Skeleton-Finding Algorithm

In order to achieve one of the objectives of this work, namely to detect elongated and low-density structures, we implemented the GFiF algorithm over the galaxy positions. The GFiF methodology defined by us uses clusterization and graph methods in order to search for links, bridges and filaments connecting groups of galaxies. To have an accurate detection of elongated structures (filaments) we applied GFiF over the 3D distributions of galaxies whose positions were previously corrected for the FoG effect, i.e. after applying the GSyF algorithm.

2.4.1 Detection of low density regions

The GFiF algorithm applies a combined VT+HC method to the galaxy positions in the rectangular 3D space. The galaxy position in rectangular coordinates (X, Y, Z) is calculated according to equations 2.4, 2.5 and 2.6. Then, the VT local number densities corresponds to volume densities in units of Mpc^{-3}. We define the density baseline as the mean box volume number density, d_{bas}. However, for this analysis, the HC method is applied to all galaxies in the volume without density restrictions, that is, no baseline threshold is applied. Moreover, the application of HC is done in a way that allows to detect the more elongated and representative structures. This is achieved by using larger numbers on the segmentation parameter f, as compared with the GSyF f values. Afterwards, the GFiF algorithm performs as follows:

1. Once HC is applied, the Euclidean distance D_E between groups is measured, this is, the distance from all group centroids (nodes) to all its group neighbors. These connections (edges) can be represented by an undirected graph as described in section 2.1.3. The weights W of the edges are set by the Bhattacharyya coefficient, BC, defined as:

$$BC(P_1, P_2) = \sum_{x \in X} \sqrt{P_1(x) P_2(x)}. \tag{2.12}$$

   The Bhattacharyya coefficient quantifies the amount of similarity between two distributions $P_1(x)$ and $P_2(x)$. Thus, the orientation of the Gaussians of two groups weights the connection between them.

2. The mean lineal density of the edges is measured inside a cylinder of radius 1 h_{70}^{-1} Mpc defined along the edge. Then, the density is calculated as the number of galaxies inside the cylinder N_{edge} divided by the cylinder volume.

3. An edge is considered as a real *link* of galaxies connecting the two nodes based on the following:

   (a) The edge length D_E is smaller than a threshold, D_{max} (hereafter, linking length).

   (b) The mean lineal density of the edge is above $\bar{d} = N/V$.

   The ensemble of these links is considered as an undirected graph $G = (U, E)$, where U represents the nodes (the group's centroids), E the edges (connections between nodes) and W a matrix that gives weights to the connections. Each ensemble of connected links is a tree in the forest graph.

4. Then, Kruskal's minimum spanning tree algorithm (section 2.1.4) is used to identify independent trees and their dominant branches.

2.4.2 Chaining the filaments: GSyF + GFiF

In order to identify filaments connecting systems of galaxies, each GFiF tree is matched against the detected GSyF galaxy systems. However, richness of the detected systems depends on the redshift, because of the Malmquist effect. The number of galaxies in systems decreases with increasing redshift (see Figure

3.1). In other words, to have a comparable richness for two similar system, one at $z = 0.03$ and the other at $z = 0.13$, it is necessary to apply a correcting factor to the richness of the second one. Therefore, a lower limit for the richness of the systems at the supercluster redshift was set as:

$$\log_{10} N_{min} = a \log_{10} z + b, \tag{2.13}$$

with $a = -1.0$ and $b = -0.2$. This leads to a minimum richness of $N_{mem} = 30$ to 5 galaxies, from the nearest and farthest supercluster in our sample respectively.

The GFiF detected structures are classified, based on the number of GSyF galaxy systems embedded in them, as:

(i) The *nodes* are defined as the HC group's centroids.

(ii) An *edge* is defined as any connection between two nodes.

(iii) The real *links* between the systems are defined as the most promising edges, filtered according to their proximity and density.

(iv) *Spanning trees* are acyclic optimal connections of links. Some nodes inside a spanning tree may be detected as galaxy systems by the GSyF algorithm.

(v) A *bridge* is defined as a sequence of links and nodes between two systems.

(vi) A *filament* is identified if a spanning tree links three or more systems connected by bridges.

(vii) A *tendril* is a spanning tree that contains none or only one system.

(viii) The *skeleton* is the medial line of a filament. The method for finding it, which intends to reduce the dimensionality of the objects (in our case, galaxy filaments), is known as skeletonization.

Figure 2.1 provides a schematic of the definitions and Figure 2.2 shows schematically the steps of the GFiF algorithm.

The links forming filaments were refined using Dijkstra's algorithm (section 2.1.5). This refinement allows the identification of the filament skeleton, i.e. the principal

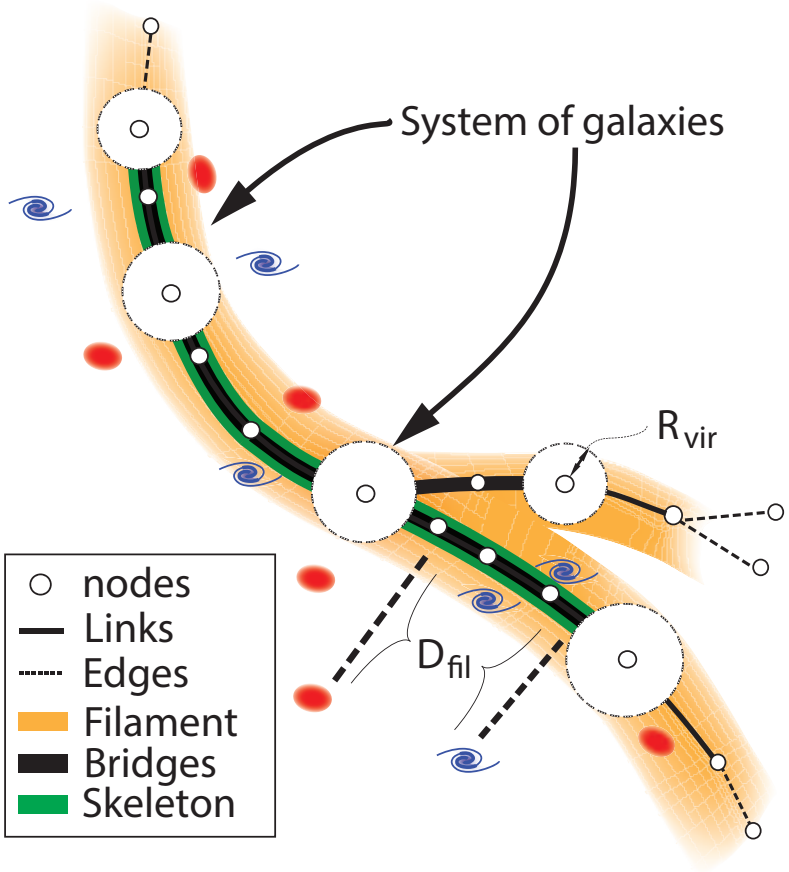

FIGURE 2.1: Representation of a filament. Graph nodes are represented by white circles and edges by dark lines. The three systems connected are represented by a dotted circle of radius R_{vir}. A bridge connecting two systems is represented as a bold black line. The distance from galaxies to the filament, dashed line, is measured in the perpendicular line to the edges.

branch connection. According to the pattern recognition literature, a skeleton represents the principal features of an object such as topology, geometry, orientation and scale.

2.5 Algorithm optimization for the SDSS

2.5.1 Mock maps modeling

As mentioned before, the detection efficiency of the GSyF method depends on the value of the segmentation parameter f and the contrast level parameter g. The strategy we implemented to find an optimized combination of these parameters was to simulate 30 mock galaxy maps for each supercluster of the sample.

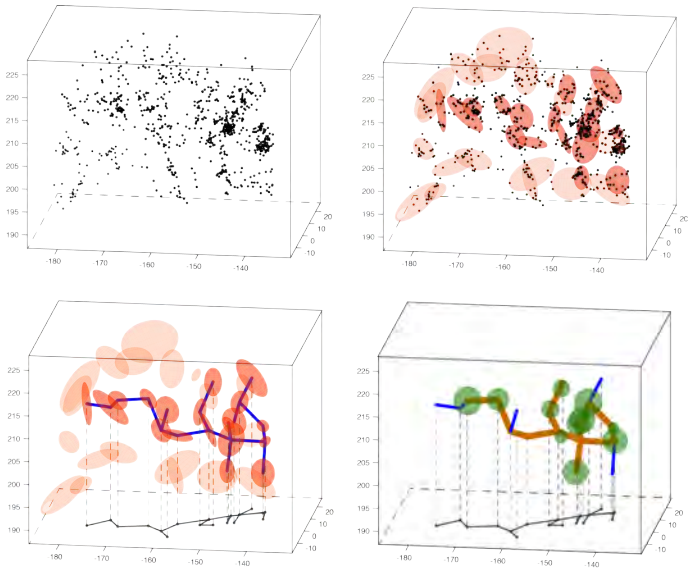

FIGURE 2.2: Illustration of the steps of the GFiF algorithm. In the first box (top-left) one can see the distribution of galaxies. In the second one (top-right) the HC groups are marked, with denser red colors representing the richer HC groups. The filtered edges (links) among the groups of the spanning tree are displayed in the third box (bottom-left). The last box (bottom-right) presents the systems (green circles), bridges (brown lines) and other links (blue lines) found among the groups of the preceding step.

The properties of the synthetic galaxy systems were estimated by the relations described by Pearson et al. (2015). These galaxy-based relations were calculated for the SDSS database and follow a power law of the form $\log_{10}(M_{500}) = \alpha \log_{10}(\frac{x}{x_0}) + \beta$, where x is the property under consideration.

Each mock map was constructed as follows: the simulated volume is filled with N_{synth} synthetic systems, randomly distributed. Each of them is filled with synthetic galaxies in the range $N_{elem} = 10 - 200$. The number of systems in the volume is set using the power function:

$$\log_{10}[N_{synth}(N_{elem})][h_{70}^3\ Mpc^{-3}] = m \log_{10}(N_{elem}) + b, \tag{2.14}$$

(hereafter, multiplicity function Berlind et al., 2006). The slope and intercept (m, b) of the multiplicity function are set to (-2.48, -2.1) for $z < 0.08$ while for higher z values are (-2.72, -2.4). The system proxies (M_{500} mass, R_{vir} radius and σ_v velocity dispersion) are calculated from:

$$\log_{10} M_{500} = 1.03 \log_{10}(N_{elem} - 2.63) + 0.34 \quad (2.15)$$

$$\log_{10} R_{vir} = 1.05 \ \log_{10}(M_{500} - 0.35) + 8.48 \quad (2.16)$$

$$\log_{10} \sigma_v^3 = 2.33 \ \log_{10}(M_{500} - 0.21) + 3.04 \quad (2.17)$$

Then, the systems are filled with galaxies following a normal distribution $N_{elem}(\mu_i, R_{vir})$ with $\mu_i = \{\alpha_i, \delta_i\}$. The FoG effect is incorporated to the system galaxies by adding a velocity dispersion $N_{elem}(z_i, \sigma_v)$. Finally, random galaxies are added to the box volume following the ratio: 60% of galaxies are distributed in the field, and 40% in the systems.

2.5.2 Optimization of GSyF parameters

The GSyF method was applied over the mock maps probing values for f in the range $f = \{3, 6, 9, ..., 36\}$ and for g in the range $\{-0.25, -0.15, 0.0, 0.15, 0.25, 0.5\}$. We observe that for larger values of f the HC grouping was not efficient in detecting individual systems. Moreover, the analysis on the g parameter suggests that the selection of this parameter does not impact the efficiency of GSyF detecting systems. Thus, we set the g parameter equal to 0.0 for all superclusters. The comparative of detection at different values of g is depicted in Figure 2.3 (panels a–d). The analysis on the f parameter (see Figure 2.3) (panels a–d) suggests that GSyF has a higher detection rate for synthetic systems (completeness) (above 85%, arriving close to 100% for $f < 5$) for f values below 10. However, at these f values the contamination rate (false detection) is between 5 and 35% decreasing for larger values of f. This analysis also reveals that the efficiency of the algorithm also depends on the final richness of the systems: success rates for richer systems ($N_{mem} \geq 20$) are much higher than success rates for poor ones ($N_{mem} \geq 10$), while failure rates do not change significantly, as shown in Figure 2.3 (panels e–f). The optimal value for f is chosen as the one that maximizes the completeness and minimizes the contamination, that is, the value that maximizes the function:

$$\Gamma = \frac{N_{detec}}{N_{synth}} + \left(1 - \frac{N_{fail}}{N_{synth}}\right). \quad (2.18)$$

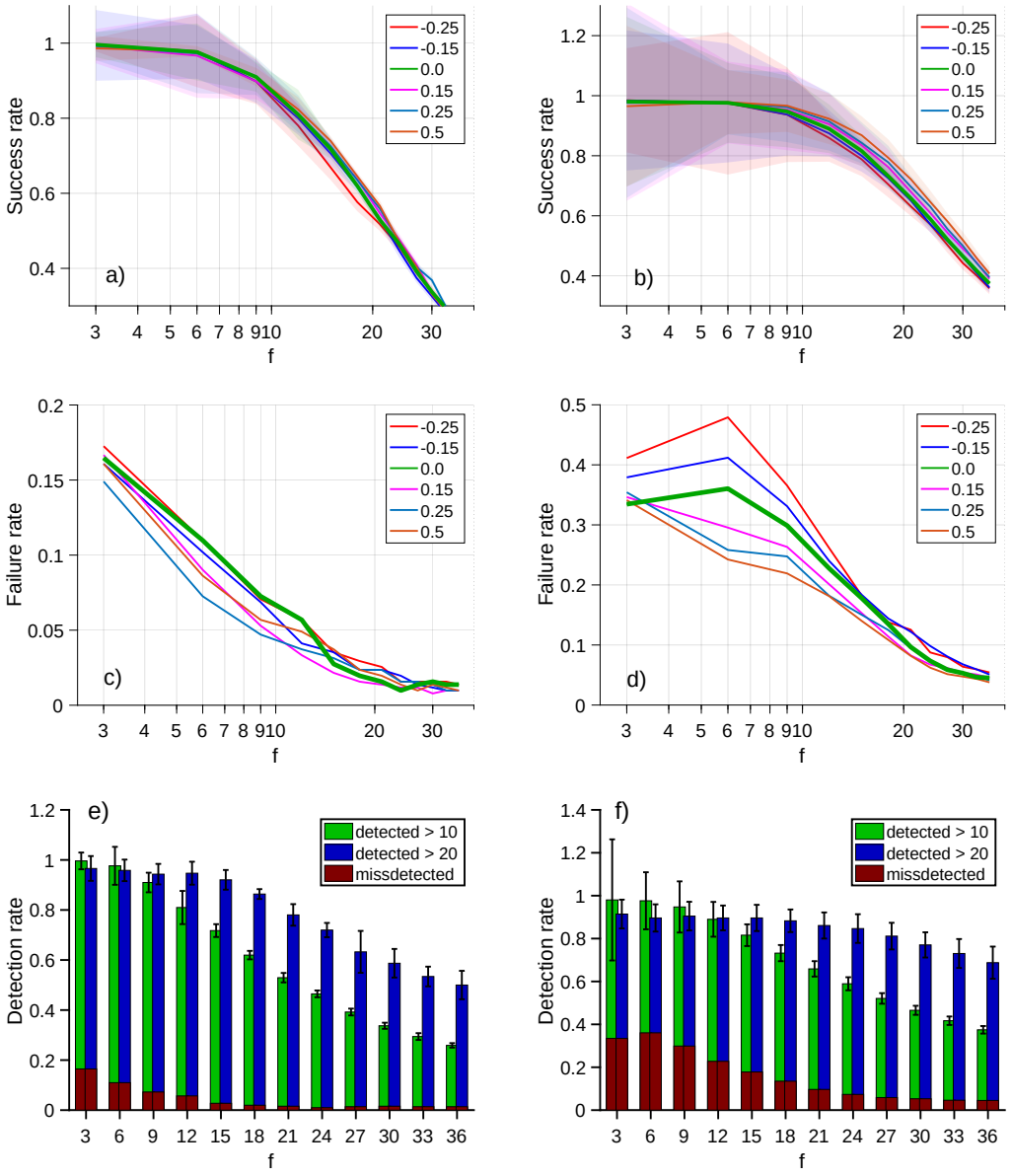

FIGURE 2.3: Optimization of the segmentation, f (x axis), and minimum contrast density, g (lines of different colors in panels a–d), for the supercluster volumes of MSCC-310 (z=0.06) and MSCC-454 (z=0.04), from the mock simulations. (a, b) Success rates; (c, d) failure rates; (e, f) detection rates for systems with $N_{mem} \geq 10$ (green) and $N_{mem} \geq 20$ (blue), with N_{mem} been the final number of members after the virial refinement. The failure rates are shown in brown.

2.5.3 Optimization of GFiF parameters

The results of the GFiF algorithm depend on several parameters, especially the number of HC groups, N_{cut} (or, equivalently, f), and the linking length D_{max} [2].

[2] The optimization for these parameters is described in detail in Santiago-Bautista et al. (2019b).

This optimization was carried out on the real galaxy positions. The GFiF algorithm is evaluated by counting the number of filaments detected setting the linking length D_{max} in the range of $\{D_1..., D_{end}\}$ and the segmentation parameter f in the range $\{f_1, ..., f_{end}\}$ h_{70}^{-1} Mpc. The number of filaments for each evaluation is accounted for in a combination matrix of these parameters $\Pi(f|D_{max})$. Finally, the optimal f and D_{min} parameters are set to:

$$N_{ske}(f|D_max) = max(\Pi(f|D_max)). \tag{2.19}$$

For the SDSS galaxy distribution, the linking length D_{max} was evaluated in the range $\{8, ..., 40\}$, and the segmentation f was evaluated in the range $\{8, ..., 40\}$. In practice, there is an optimal D_{max} for each evaluated value of f. The results of the optimization process for the supercluster MSCC-454 are shown in Figure 2.4. For this supercluster, the optimal parameter configuration is $f = 10$ and $D_{min} = 6$, with 9 filaments detected, as can be seen in panel a. Figure 2.5a depicts the projected distribution of the found filaments for the optimal parameter configuration, while Figure 2.5b depicts the results for the second-best configuration, $f = 20$ and $D_{min} = 8$.

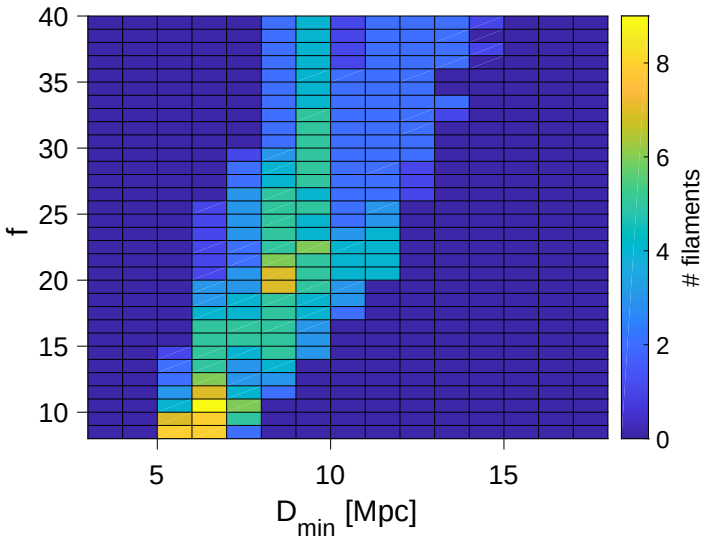

FIGURE 2.4: Optimization of the segmentation, f, and linking length, D_{min}, parameters for the supercluster MSCC-454. This optimization is based in the number of relatively dense and long filaments detected.

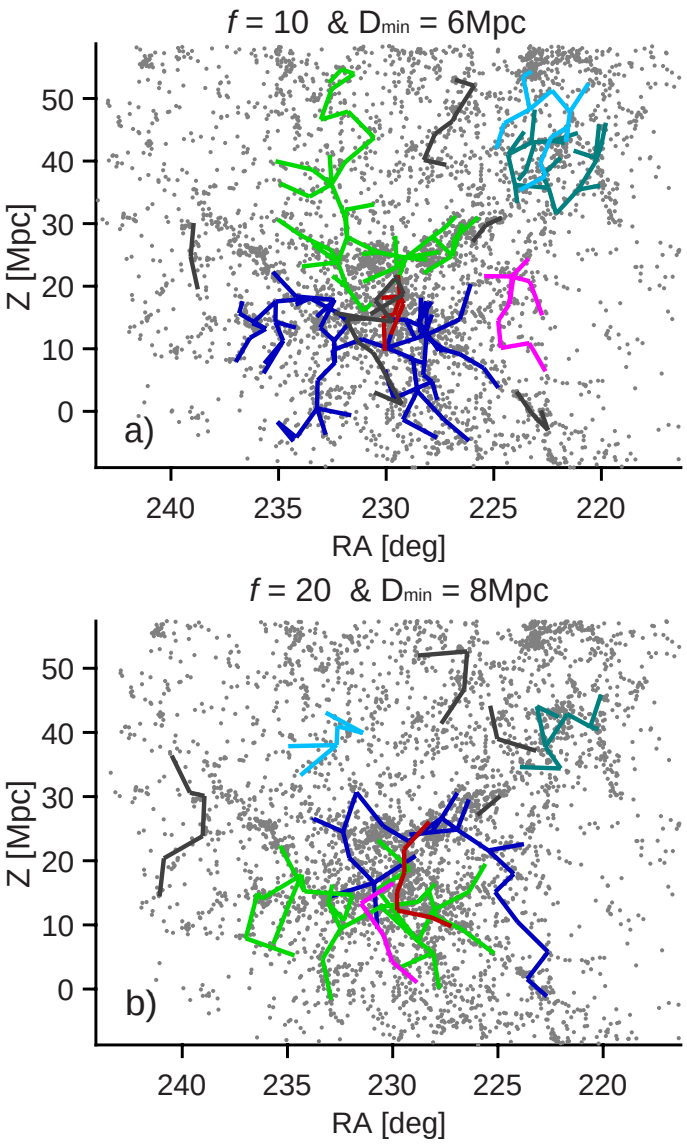

FIGURE 2.5: a) Filaments found by the GFiF algorithm for the solution parameters $f = 10\ h_{70}^{-1}$ Mpc and $D_{min} = 6\ h_{70}^{-1}$pc M. b) Filaments found by the GFiF algorithm for the solution parameters $f = 20\ h_{70}^{-1}$ Mpc and $D_{min} = 8\ h_{70}^{-1}$ Mpc.

2.5.4 Optimization results

The optimization for the GSyF algorithm f parameter was carried out for all supercluster in the present of study. Figure 2.6a shows the distribution of best f values for the superclusters in the sample with their mean redshift. Except for the first two points (the most nearby superclusters), one can see that there is no clear correlation between the two parameters, presenting a mean value around $f = 10$ with a considerable dispersion. In fact, for the two nearby superclusters the success rates are reduced because, as can be seen in Fig. 2.6b, for these

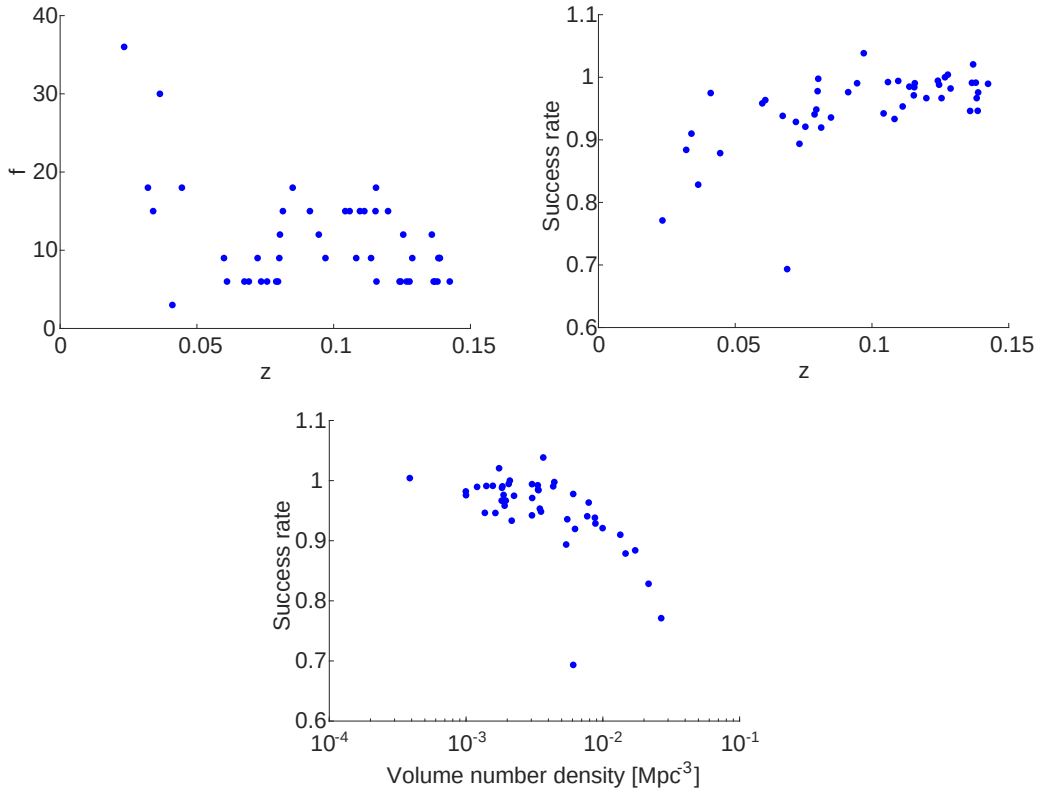

FIGURE 2.6: a) Distribution of best f values with redshift (z). b) Distribution of mean success rates with z. c) Distribution of mean success rates with volume number densities in each supercluster volume. Mean numbers are from the 30 mock simulations for each of the 42 superclusters.

superclusters the success rates are between 85%, while for all the others the success rates are around or above 90% (with a mean failure rate of about 10%). The low success rates for these two superclusters seem to be related to their higher number densities with respect to the others, as can be seen in Fig. 2.6c.

2.6 Discussion and conclusions of the Chapter

In this chapter we presented a new strategy for identifying clusters and groups of galaxies, the *Galaxy Systems Finding algorithm* (GSyF) and for identifying filaments, the *Galaxy Filaments Finding algorithm* (GFiF), inside superclusters of galaxies. We developed a methodology for GSyF that allows one to detect systems, their galaxy members, and to correct for FoG effects. This method uses the *Voronoi Tessellation* algorithm to measure local number densities and a *Hierarchical Clusterization* method to group the galaxies using their projected positions

(RA,Dec,z) and surface density. Afterwards, the groups are refined by calculating their virial radius and mass iteratively. Through this refinement, groups are either classified as detected systems or rejected as group mis-detections. As a result, GSyF provides, for the detected systems, galaxy membership, the virial mass and virial radius. In order to evaluate the methodology implemented by GSyF we used a set of mock maps. The analysis of mock maps was used to set the GSyF free parameters: the minimum density contrast for detecting the systems and the cut in the number of groups for HC method (segmentation parameter). The results of this evaluation shows that the GSyF algorithm has a completeness above 85% for 40 of our 46 superclusters, with a contamination (false positives) around 10%.

The GFiF algorithm uses the position of galaxies in 3D space to find links and bridges between concentrations of galaxies. For this algorithm, our methodology uses a combination of HC clustering, graphs and minimum spanning tree. Grouping the galaxies using HC is equivalent to connect them according to their distance. This allows to reduce the dimensionality of the graphs, facilitating the application of MST. Otherwise, the size of the generalized graph would make it difficult to search for optimal trees. It is worth to note that the grouping of galaxies allows to reduce considerably the computational time (consider that a supercluster has tens of thousands of galaxies).

GSyF and GFiF are probabilistic algorithms in the sense that they define systems and filaments as a function of the position and orientation of the Gaussian groups, which are the result of a *Hierarchical Clusterization* method. For GSyF, these measurements are refined by using a virial approximation, allowing the detection of gravitationally bounded systems of galaxies and discern mis-detections. For GFiF, the orientation and position of Gaussian groups are used to define a general tree from which independent structures are extracted.

However, we have shown that this methodology needs to be optimized for the number of HC groups and the linking length D_{max} in order to detect the most promising filaments. The optimization of these parameters is proposed to be implemented on the real data in the analysis, by testing a range of parameters. Then, the optimal combination of f and D_{min} will correspond to the one that maximizes the number of identified filaments.

Chapter 3

Characterization of systems and filaments through optical galaxies

One of the objectives of this work aims for the detection and study of filaments of galaxies, which can be defined as chains of clusters connected by bridges of galaxies (and probably by gas and dark matter). To achieve this goal, the approach implemented consists of applying the GSyF algorithm for the detection of systems of galaxies and GFiF for the detection of bridges and filaments. In this context, this Chapter describes the implementation of the algorithms described in the previous one. Both algorithms use only the position of the galaxies as measured from their projected positions and line of sight velocities to detect structures. This Chapter presents the results of our implementation of GSyF and GFiF to a sample of 46 boxes containing superclusters of galaxies. Then, we describe the samples of detected systems and filaments as well as the correlations between the principal features of galaxies and the structures they belong to.

3.1 The filament candidates sample

The supercluster sample used in this study consists of a selection of superclusters of galaxies from the Main SuperCluster Catalogue (MSCC) presented in Chow-Martinez et al. (2014). The full MSCC is an all-sky catalog that contains 601 superclusters, identified in a complete updated sample of rich Abell/ACO clusters, with redshifts from 0.02 to 0.15, by using a tunable FoF algorithm. Of

these, only those superclusters with five or more clusters and those for which their box volume (see below) was completely inside of the SDSS-DR13 surveyed area were selected for the present study.

The selection of superclusters took into account a sample of filament candidates by Chow-Martínez et al. (2019, in preparation). Roughly speaking, these filament candidates were identified as chains of at least three clusters, members of the superclusters, separated by less than 20 h_{70}^{-1} Mpc from each other. Therefore, one of the objectives of this work is to validate these filament candidates by searching for bridges of galaxies connecting the galaxy clusters. However, there exists the possibility that some of the filament candidates do not present bridges of galaxies connecting clusters. Also, some bridges may exist, but not in the direction that straightly connects the clusters.

The final sample consists on 46 superclusters of galaxies, which are listed in Table 3.1. The ID of the supercluster in MSCC is shown in column 1, with its proper name in column 2, when available. Column 3 presents sky coordinates, RA (α) and Dec (δ), of the supercluster mean position, while column 4 shows its mean redshift. Columns 5 and 6 list the richness (number of member clusters) and the number of filament candidates found previously in each supercluster (Chow-Martinez et. al. in prep). The IDs of the Abell/ACO member clusters are listed in column 7.

Both for the Abell/ACO clusters and for the galaxies in the superclusters box volumes (see section 3.1.1), the radial-angular coordinates were transformed to rectangular coordinates according to equations 2.4, 2.5 and 2.6.

3.1.1 The SDSS galaxies

The main galaxy sample of SDSS-DR13 is a suitable database to search for filamentary structures on the LSS since:

1. it covers a large sky area (14,555 square degrees), containing several MSCC superclusters;

2. it contains homogeneous photometric and spectroscopic data for galaxies with an astrometric precision of 0.1 arcsec rms and uncertainty in radial velocities of about 30 km s^{-1} (Bolton et al., 2012);

TABLE 3.1: Sample of MSCC superclusters.

SCl ID (MSCC)	Name	RA, Dec [deg, deg]	$\bar{z}$	N_{Cl}	N_{fil}	Abell/ACO clusters
(1)	(2)	(3)	(4)	(5)	(6)	(7)
55		17.75, 15.44	0.0614	5	1	A0150 A0152A A0154B A0158B A0160B
72		25.17, 0.64	0.0802	5	1	A0181A A0208A A0237A A0267B A0279A
75		28.09, -5.15	0.0937	7	1	A0256A A0256B A0266 A0269 A0274A A0274B A0277
76		28.35, -2.61	0.1299	16	3	A0211 A0233e A0255 A0256C A0261B A0265 A0267C A0268B A0271 A0274C A0279B A0281 A0285 A0295D A0303C A0308e
175		125.29, 17.07	0.0942	6	1	A0635A A0650B A0651A A0657A A0658A A0659
184		130.10, 30.24	0.1056	6	1	A0671B A0690C A0694 A0695B A0699B A0705A
211		147.87, 64.88	0.1191	8	1	A0764 A0802 A0804B A0845 A0871e A0906e A0975 A1014A
219		153.99, 19.14	0.1155	5	1	A0938B A0942A A0952A A0991B A0994A
222		155.14, 49.21	0.1382	10	2	A0915B A0927A A0950A A0965A A0990 A1002A A1003C A1003D A1004 A1040C
223		155.24, 62.94	0.1399	5	1	A0917 A0947A A0962A A1025A A1025B
229		156.14, 33.03	0.1423	7	2	A0924 A0951 A0982 A1007B A1036 A1045 A1053B
236		156.76, 10.38	0.0328	6	1	A0938A A0957A A0999A A1016A A1020A A1142A
238		156.98, 39.55	0.1068	21	2	A0967A A0971A A0971B A0972A A0995A A0997A A0997B A0997C A1010B A1021B A1021C A1021D A1026B A1028A A1031A A1031B A1033 A1040A A1050A A1054A A1055
248		159.49, 44.26	0.1246	5	1	A1040B A1050B A1054B A1056 A1074A
264		165.29, 12.20	0.1161	8	1	A1105C A1116A A1129A A1141A A1147A A1157 A1201B A1209A
266		165.91, 11.85	0.1273	8	1	A1131 A1137B A1141B A1147B A1152 A1159 A1183A A1209B
272		167.83, 41.33	0.0760	6	1	A1173 A1174A A1187 A1190 A1193A A1203
277		169.41, 49.67	0.1103	7	1	A1154 A1202B A1218B A1222 A1225 A1227A A1231A
278	Leo	169.37, 28.46	0.0333	6	1	A1177B A1179B A1185A A1228A A1257A A1267A
283		170.79, 20.34	0.1379	12	3	A1177C A1188 A1230B A1232B A1242A A1243B A1247e A1251 A1268 A1272 A1274 A1278
295	Com	173.63, 23.11	0.0223	5	1	A1100A A1177A A1179A A1367 A1656
310	UMa	175.91, 55.23	0.0639	21	1	A1212 A1270 A1291A A1291B A1291C A1318A A1318B A1324A A1324B A1349A A1349B A1377 A1383 A1396A A1396B A1400A A1400B A1400C A1436 A1452 A1457A
311		176.12, 9.93	0.0833	8	1	A1337A A1342A A1358A A1362B A1372A A1379 A1385A A1390
314		177.07, -2.01	0.0788	6	1	A1364A A1376A A1386A A1389A A1399A A1404A
317		177.42, -1.59	0.1278	13	1	A1373A A1373B A1376C A1386D A1386E A1386F A1389C A1389D A1392 A1399C A1407 A1411 A1419B
323		179.66, 27.26	0.1396	12	1	A1384A A1403A A1403B A1413B A1420C A1425B A1431B A1433C A1444C A1449B A1455C A1495
333		181.43, 29.34	0.0813	9	1	A1423A A1427 A1431A A1433A A1444B A1449A A1455B A1515A A1549A
335		182.42, 29.50	0.0732	6	1	A1444A A1455A A1478A A1480B A1486A A1519A
343		183.88, 14.31	0.0809	5	1	A1474 A1481A A1499A A1526C A1527A
360	Dra	190.94, 64.41	0.1055	11	1	A1518A A1539A A1544A A1559 A1566 A1579A A1621 A1640A A1646 A1674A A1718A
386		199.50, 38.33	0.0715	5	1	A1680A A1691 A1715A A1723B A1749B
407		208.55, 26.70	0.1364	6	2	A1797B A1817C A1817e A1818C A1819 A1824
414	Boo	211.31, 27.32	0.0709	24	1	A1775A A1775B A1781B A1795 A1797A A1800 A1817A A1818A A1831A A1831B A1832A A1863A A1869A A1869B A1873B A1873C A1874A A1886A A1898A A1903A A1908A A1909A A1912B A1921A
419		212.33, 7.17	0.1122	5	1	A1850 A1862 A1866A A1870 A1881
422		213.21, 28.95	0.1430	9	1	A1832B A1840B A1854 A1867A A1874B A1891B A1903C A1908B A1912E
430		216.72, 25.64	0.0982	6	1	A1909B A1910A A1912A A1912C A1926A A1927
440	BooA	223.17, 22.28	0.1170	9	1	A1939B A1972 A1976 A1980 A1986 A1988B A2001A A2006 A2021C
441		223.22, 28.40	0.1249	5	1	A1973A A1982D A1984 A1990A A2005B
454		228.28, 7.33	0.0456	6	1	A2020A A2028A A2033B A2040B A2055A A2063B
457		228.59, 6.98	0.0789	6	1	A2028B A2029 A2033C A2040C A2055B A2063C
460		229.70, 31.17	0.1142	9	1	A2025D A2034A A2049A A2056C A2059B A2062 A2067B A2069 A2083B
463	CrB	232.18, 30.42	0.0736	14	1	A2056A A2056B A2059A A2061A A2065 A2067A A2073A A2079A A2079B A2089 A2092A A2106A A2122A A2124
474	Her	241.56, 16.22	0.0363	5	1	A2147 A2151 A2152A A2153A A2159A
484		245.57, 42.39	0.1364	7	1	A2158B A2172 A2179 A2183 A2196 A2198D A2211A
579		351.82, 14.79	0.0427	5	1	A2572 A2589 A2593A A2593B A2657
586		354.20, 23.67	0.1274	5	1	A2611e A2619B A2627 A2647e A2650e

Notes. The superclusters have five or more Abell cluster members, with $z \leq 0.15$, and inside the SDSS-DR13 region. Superclusters with proper names are indicated in column 2.

3. it is roughly complete to the magnitude limit of the main galaxy sample ($r_{Pet} = 17.77$), which corresponds to an average $z \sim 0.1$, going (inhomogeneously) deeper for data releases after DR7 (Abazajian et al., 2009);

4. at the limit of our sample, $z = 0.15$, the SDSS spectra are complete for galaxies brighter than $M_r \sim -21$.

The SDSS-DR7 joins the SDSS-I/II spectra for one million galaxies and quasars. It has $\sim$6% incompleteness due to fiber collisions (Strauss et al., 2002) and another $\sim$7% incompleteness attributed to pipeline misclassification (Rines et al., 2007). These spectra are included in the final data release of the SDSS-III (Alam et al., 2015). The Barion Oscillation Spectroscopic Survey (BOSS) is part of the SDSS-III observations and has obtained spectra for another 1.4 million galaxies. The BOSS observations are divided in two main samples, LOWZ ($z < 0.4$) and CMASS ($0.4 < z < 0.7$). The SDSS-DR13 (Albareti et al., 2017) includes spectra for more than 2.6 million galaxies and quasars.

For the present analysis, only objects listed on the SpecObj sample with spectroscopic redshifts available denoting an extragalactic object (that is, galaxies and low-z quasars) were selected (downloaded from the SkyServer web service). The SpecObj table contains the best spectra (unique within a circle of 2 arcsec) called "sciencePrimary" objects. The galaxies were extracted within a redshift range from 0.01 to 0.15. Only galaxies with spectra quality flag "good" or "marginal" were used. A study of the 3D galaxy distribution is dependent on the galaxy distance measurement. Spectroscopic redshifts provide the accuracy needed for this kind of analysis. Nevertheless, several photometric redshift catalogs are available for the SDSS, which can be included to the sample in future analyses to test if their addition increases the signal of filament detection.

Moreover, there exist several sub-products that use the SDSS spectra to characterize the galaxy properties. In particular, the following catalogs were used to characterize the galaxy filaments: the MPA-JHU catalog (Brinchmann et al., 2004; Kauffmann et al., 2003; Tremonti et al., 2004) provides different galaxy properties (stellar mass, metallicity, activity type classification, star forming rate, among others). As explained by Tremonti et al. (2004), the galaxy properties in the MPA-JHU catalog are calculated by processing the galaxy spectrum in a way that even the weaker emission lines are detectable. The GRANADA group catalog is another SDSS-DR13 value added database that provides galaxy properties

such as mass and metallicities calculated from the galaxy magnitudes (Conroy et al., 2009). The SDSS images and magnitudes have also been used to estimate the galaxy properties using neuronal networks. One example of this approach is the morphological classification provided by Huertas-Company et al. (2011). These authors calculate a probabilistic morphological classification, for the SDSS-DR7 spectroscopic galaxies, by applying deep learning techniques that make use of their photometry. They also compare their automated classification with a sample of the Galaxy Zoo (Lintott et al., 2008, 2011) visual classification. They show that their classification into early and late types are in good agreement with the visual classification.

3.1.2 The superclusters' boxes

For this analysis the superclusters were confined to boxes defined in rectangular coordinates. The box walls were set at a minimum distance of 20 h_{70}^{-1} Mpc from the center of the farthest clusters in the supercluster. Then, all SDSS galaxies located inside its box volume were selected.

In Table 3.2 the box volumes (column 2) for each supercluster are listed as well as number of galaxies inside each volume (column 3), the mean volume and surface (sky projected) number densities (columns 4 and 5), the baseline density (column 6), the number of galaxies with surface density above the baseline density (column 7), the segmentation parameter f and the number of HC groups (see Section 2.3.2) (columns 8 and 9), FoG corrected groups of richness lower than 10 galaxies (column 10) and with richness higher than 10 galaxies (column 11), the ranges of radius and velocity dispersion for this last final list of groups (columns 12 and 13).

In particular, the superclusters MSCC 236, 314 and 317 lie close to the limits of the SDSS region: although all their member clusters are inside, their boxes were reduced to a margin of 10 h_{70}^{-1} Mpc in only one direction. For this reason, their analysis must be taken with caution due to this small data incompleteness.

Figure 3.1 shows the diminution of mean volume density of the boxes with redshift, due to Malmquist bias. The fitted function will be used as the selection function for the SDSS galaxies used in this work.

TABLE 3.2: Properties of the galaxies in the sampled superclusters.

SCl ID (MSCC) (1)	box volume [10^3Mpc^3] (2)	N (DR13) (3)	$d = 1/V$ [Mpc^{-3}] (4)	$d_{sup} = N/A$ [deg^{-2}] (5)	d_{bas} [deg^{-2}] (6)	N_{gal} $d_i > d_{bas}$ (7)	f (8)	N$_{HC}$ $N_j \geq 3$ (9)	N$_{FoG}$ $N_{mem} < 10$ (10)	N$_{FoG}$ $N_{mem} \geq 10$ (11)	R_{vir} [Mpc h_{70}^{-1} Mpc] (12)	σ_v [km s^{-1}] (13)
55	424.3	812	0.0019	77.6	8.3	468	27	57	11	5	1.1 - 2.4	245 - 806
72	549.7	1941	0.0035	232.0	22.1	1341	18	228	30	22	0.9 - 2.4	184 - 689
75	854.6	1607	0.0019	95.8	22.6	877	15	69	9	7	1.5 - 3.4	335 - 1144
76	2628.3	2617	0.0010	110.1	26.6	1536	27	204	14	11	1.3 - 3.6	214 - 1051
175	577.4	2504	0.0043	116.3	25.8	1315	6	172	22	10	1.2 - 2.7	220 - 735
184	692.9	2101	0.0030	70.5	20.7	1003	3	137	10	9	1.4 - 2.7	278 - 739
211	814.7	1484	0.0018	28.6	11.5	654	6	53	2	3	1.9 - 2.1	406 - 496
219	628.4	1913	0.0030	118.2	19.5	1273	6	175	13	10	1.4 - 3.7	292 - 1151
222	955.9	1865	0.0020	97.3	20.3	885	3	123	10	8	1.3 - 4.1	245 - 1207
223	777.0	776	0.0010	302.7	13.8	247	3	45	1	3	2.1 - 2.1	475 - 475
229	1352.4	1855	0.0014	45.2	22.3	745	3	106	6	2	2.2 - 2.2	498 - 498
236	643.6	8636	0.0134	52.9	10.3	4733	3	309	93	73	0.5 - 1.8	105 - 703
238	3861.8	8328	0.0022	74.4	20.3	4860	6	832	38	74	0.7 - 4.0	112 - 1293
248	690.1	1263	0.0018	72.4	17.4	564	3	75	5	2	1.7 - 2.8	366 - 730
264	923.7	1704	0.0018	59.6	24.5	626	3	105	11	12	1.3 - 3.5	245 - 1026
266	458.5	958	0.0021	49.5	28.8	318	3	55	6	3	1.6 - 3.1	320 - 823
272	138.4	1379	0.0100	135.3	28.0	654	3	87	10	5	1.1 - 2.4	219 - 699
277	905.3	2748	0.0030	76.0	20.2	1329	6	179	17	9	1.4 - 2.5	278 - 675
278	459.3	7920	0.0172	52.3	10.3	4116	6	222	80	35	0.5 - 1.9	112 - 711
283	1478.8	2320	0.0016	70.3	20.2	1379	12	239	10	17	1.5 - 3.4	295 - 907
295	535.5	14308	0.0267	48.5	7.2	7422	6	272	114	46	0.4 - 2.0	74 - 909
310	1558.8	12286	0.0079	76.9	15.7	7529	6	1015	116	139	0.8 - 3.0	140 - 1182
311	958.8	5270	0.0055	91.8	22.4	3050	6	416	48	40	0.8 - 2.4	131 - 704
314	91.9	558	0.0061	135.2	27.0	289	3	49	10	4	1.2 - 2.3	254 - 659
317	438.6	840	0.0019	104.2	38.2	433	6	76	10	5	1.8 - 3.3	366 - 929
323	1909.6	3330	0.0017	77.3	21.7	1764	6	304	17	21	1.5 - 3.7	295 - 1069
333	445.5	1968	0.0044	65.1	22.6	793	3	135	14	27	1.1 - 2.9	221 - 949
335	574.5	3099	0.0054	62.2	21.4	1285	3	211	29	38	0.8 - 3.0	144 - 973
343	427.9	2679	0.0063	105.8	19.2	1526	6	196	23	25	0.8 - 2.4	131 - 675
360	657.7	2199	0.0033	80.1	15.3	934	12	160	15	16	1.3 - 2.5	253 - 653
386	535.9	3256	0.0061	54.9	17.2	1600	9	257	33	40	1.0 - 2.7	211 - 852
407	800.0	1126	0.0014	48.9	22.8	481	12	79	5	5	1.5 - 4.0	280 - 1184
414	1245.9	10902	0.0088	93.0	23.1	6366	6	1066	144	161	0.8 - 3.2	140 - 1191
419	497.6	1723	0.0035	91.7	19.7	1103	6	196	25	20	1.2 - 3.3	211 - 976
422	884.6	1065	0.0012	41.9	24.2	382	3	62	2	6	2.1 - 2.3	474 - 526
430	437.0	1603	0.0037	88.6	22.9	871	3	121	20	9	1.3 - 2.4	281 - 647
440	1017.1	3442	0.0034	99.4	72.9	917	6	143	24	14	1.5 - 3.3	309 - 935
441	516.1	1058	0.0021	60.8	20.9	425	6	59	2	3	3.0 - 3.0	796 - 796
454	389.0	5704	0.0147	99.7	18.9	3231	6	524	84	106	0.7 - 1.9	142 - 610
457	529.6	4072	0.0077	129.0	22.7	2605	6	443	58	44	1.0 - 3.1	187 - 1038
460	1041.1	3499	0.0034	108.4	27.3	1925	3	335	35	23	1.3 - 3.6	238 - 1073
463	959.2	8466	0.0088	121.6	22.4	5278	3	898	113	113	0.6 - 3.1	103 - 1077
474	343.8	7424	0.0216	109.2	15.3	4506	9	166	64	26	0.6 - 2.6	122 - 1115
484	805.0	1319	0.0016	43.1	19.2	571	6	86	5	4	2.3 - 3.0	536 - 793
579	658.9	1477	0.0022	142.9	1.2	1234	3	149	19	23	0.6 - 1.9	128 - 659
586	962.1	373	0.0004	18.3	13.6	72	3	8	0	0	-99.0 - -99.0	-99 - -99

Notes. The value -99 is set when GSyF does not detect systems.

3.2 Implementation of GSyF and GFiF algorithms

In what follows the application of the detection algorithms presented in Chapter 2 to one of the superclusters in our sample, MSCC310, the *Ursa-Majoris* Supercluster (see Tables 3.1 and 3.2) is described in detail. This supercluster contains 21 Abell clusters, with redshifts in the range from 0.05 to 0.08 and its volume is the fourth largest in the sample: it occupies an area in the sky of about 1 700 deg^2, equivalent to a volume of $(116\ h_{70}^{-1}$ Mpc$)^3$ (including the 20 h_{70}^{-1} Mpc added to the box limits from the farthest clusters).

The volume is filled by $N = 12\,286$ SDSS galaxies with spectroscopic redshift. This corresponds to a mean surface density of 24 gal.deg^{-2} or 0.008 gal.h_{70}^3 Mpc^{-3},

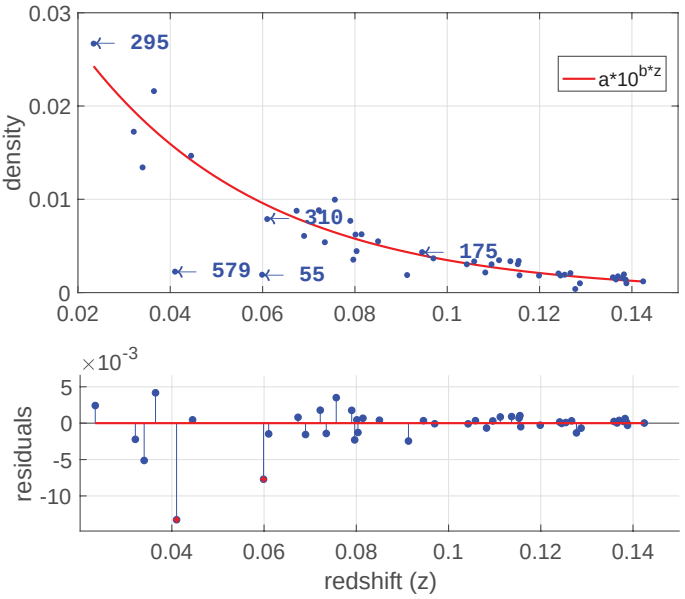

FIGURE 3.1: Distribution of mean volume densities, for the 46 superclusters in our sample, with redshift, blue points. The red line corresponds to the best fit of a power law function. Residuals of the fitting are showed in the bottom panel. MSCC579 and MSCC55 were excluded from the fitting.

see Table 3.2.

3.2.1 Application of GSyF to MSCC310

As mentioned before, the transformation from radial-angular coordinates to rectangular coordinates is more complicated for the galaxies. Their peculiar velocity may bias their redshift-space coordinate, especially when they are members of clusters and groups of galaxies, being subject to the “Finger of God” (FoG) effect. Therefore, in order to detect systems of galaxies and correct this effect the GSyF algorithm was implemented over the galaxy distribution, as described in section 2.3.

First we applied the VT algorithm over the projected distribution of MSCC-310 galaxies to calculate the local surface density d_i for each galaxy. The surface density baseline, d_{bas} (15.7 deg^{-2}) was calculated from 1 000 simulations of the galaxy distribution. Then the HC algorithm was applied to the $N_{gal} = 7\,529$ galaxies with $\delta_i > 0$.

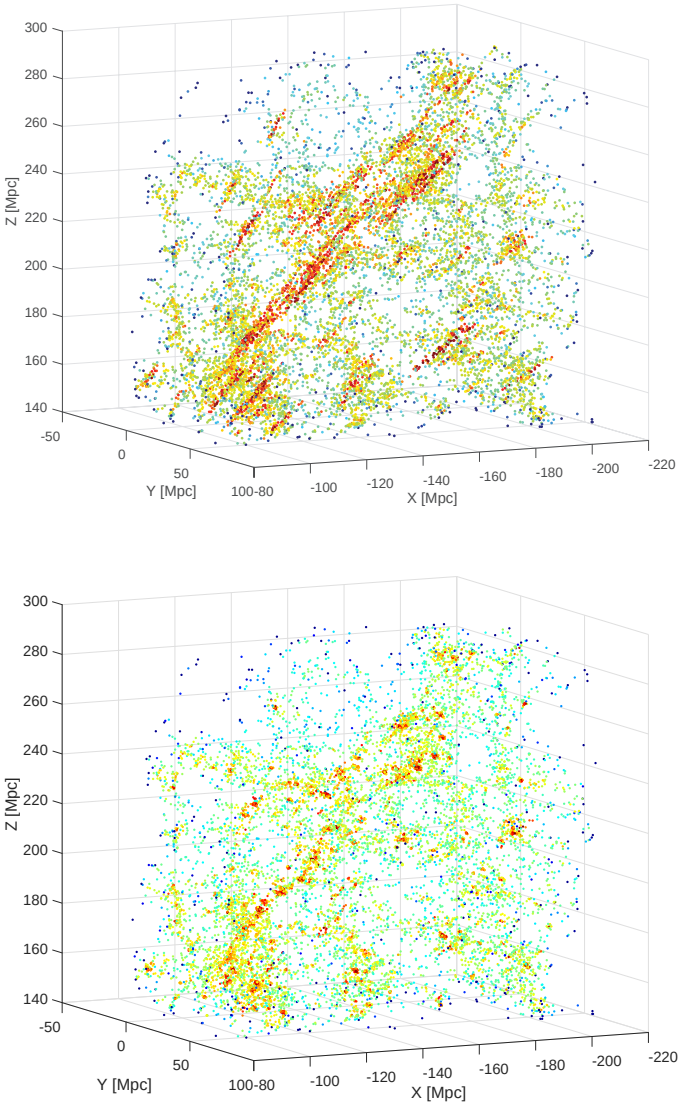

FIGURE 3.2: 3D distribution of galaxies for the MSCC-310 supercluster volume. Top, galaxy positions before the application of the FoG correction. Bottom, Galaxy positions after correction for FoG effects. The color represents density as calculated from 3D VT. The higher density is represented in red color while greener to blue color represent lower density.

The best f parameter for the implementation of the HC algorithm was calculated from 30 mock simulations ($f = 6$, in this case), using the optimization methodology described in section 2.5.2. The HC implementation grouped the galaxies in $N_{cut} = 1\,140$ groups. As expected, these groups have, on average, ~ 6 members. Of these, 1 015 groups with $N_j \geq 3$ were selected.

For the first iteration of the virial refinement the center of each HC group was set at the position of the galaxy member brightest r-band and closest to the geometric

HC centroid. For MSCC-310 groups, the mean difference between the geometrical center and their brightest galaxy position was found to be about 350 h_{70}^{-1} kpc. On average, the virial refinement needed six iterations to produce convergence of the virial radius. This refinement resulted in 139 systems with $N_j \geq 10$ for the MSCC-310 volume. The refinement also confirmed 116 smaller systems with $5 \leq N_j < 10$.

In Table 3.3 the properties of the first 25 richest systems for the MSCC-310 supercluster are listed. Column 1 assigns a sequential number to the systems, while column 2 presents their richness. The coordinates of the brightest member of the system are indicated in columns 3, 4 and 5, while columns 6, 7 and 8 show the coordinates of the final position of the centroid. The other calculated properties of the systems – velocity dispersion, harmonic and virial radius – are presented in columns 9, 10 and 11 respectively. Column 12 denotes the cross-reference with Abell clusters. The range of virial radius of the GSyF systems with $N_{mem} \geq 10$ in MSCC-310 is from 0.7 to $2.5\ h_{70}^{-1}$ Mpc. For groups with $5 \leq N_{mem} < 10$ the range of virial radius lies is from 0.4 to $0.9\ h_{70}^{-1}$ Mpc. After the refinement, the projected central position of the systems changed, on average, by 170 h_{70}^{-1} kpc, while the redshift was refined for some cases up to $\Delta z \sim 0.001$ or $\Delta\sigma_v \sim 300\,\mathrm{km\,s^{-1}}$.

As an example, the richest system in MSCC-310 is the cluster A1291A. Its HC initial centroid position (set as the position of the brightest galaxy in the HC group: $\alpha = 172.73$, $\delta = 56.49$ and $z = 0.0611$) changed by 13 h_{70}^{-1} Mpc after 17 iterations of the virial refinement (the final centroid position corresponds to $\alpha = 173.01$, $\delta = 56.09$, $z = 0.0535$). This position is 240 h_{70}^{-1} kpc from the brightest galaxy detected of A1291A which has coordinates ($\alpha = 173.05$, $\delta = 56.048$, $z = 0.0585$, Lauer et al., 2014). This means that the position was mostly adjusted along the line of sight coordinate.

Finally, the comoving distance D_C of the member galaxies in each system was corrected by re-scaling their dispersion range to the R_{vir} of the system prior to conversion to rectangular coordinates. An example of the MSCC-310 volume, before and after the correction, is shown in Figure 3.2.

TABLE 3.3: Main properties for the 25 richest systems identified in the volume of the supercluster MSCC-310.

System	N_{mem}		C_{BGM}			C_{FoG} centroid			σ_v	R_h	R_{vir}	cross-ref
No.	#	RA	Dec	z	RA	Dec	z_{LOS}	[km s^{-1}]	[h_{70}^{-1} Mpc]	[h_{70}^{-1} Mpc]	ACO Nr.
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
1	123	173.10	55.97	0.0515	173.01	56.09	0.0535	1182.20	1.09	2.54	A1291A
2	103	174.01	55.08	0.0571	174.18	55.20	0.0587	1103.00	1.28	2.56	A1318A
3	95	180.27	56.37	0.0648	180.07	56.20	0.0649	761.62	1.04	1.87	A1436
4	94	167.10	44.15	0.0587	167.11	44.07	0.0590	643.51	0.81	1.53	A1169
5	91	176.84	55.73	0.0515	176.81	55.69	0.0518	712.05	0.80	1.64	A1377
6	82	177.19	54.52	0.0601	177.06	54.64	0.0604	845.19	0.95	1.94	A1383
7	67	168.85	54.44	0.0695	168.91	54.51	0.0700	659.12	0.91	1.63	
8	61	175.28	55.19	0.0593	175.26	55.30	0.0609	1103.10	1.30	2.58	A1349A
9	61	163.40	54.87	0.0716	163.54	54.84	0.0722	639.99	0.89	1.58	
10	59	172.33	54.13	0.0689	172.45	54.08	0.0690	582.26	0.75	1.40	A1270
11	54	158.25	56.75	0.0448	158.33	56.82	0.0454	459.21	0.69	1.16	
12	52	180.23	51.42	0.0666	180.47	51.65	0.0649	1069.30	1.08	2.37	A1452
13	50	152.32	54.21	0.0465	152.41	54.42	0.0460	414.65	0.68	1.08	
14	46	183.70	59.91	0.0600	183.60	59.90	0.0599	443.38	0.75	1.17	A1507B
15	43	168.07	57.08	0.0471	168.13	57.05	0.0467	490.66	0.77	1.26	
16	42	178.38	52.69	0.0716	178.60	52.77	0.0695	760.89	0.84	1.74	
17	39	151.22	54.57	0.0470	151.00	54.66	0.0472	459.58	0.64	1.13	
18	39	163.28	56.33	0.0772	163.36	56.34	0.0745	1003.20	0.89	2.13	
19	36	172.43	55.38	0.0685	172.45	55.42	0.0684	534.44	0.54	1.19	
20	36	162.95	55.39	0.0739	162.90	55.35	0.0737	367.35	0.68	1.00	A1112A
21	34	182.19	53.33	0.0813	182.19	53.32	0.0821	572.89	0.66	1.33	
22	33	181.31	43.17	0.0529	181.41	43.20	0.0526	503.74	0.67	1.23	
23	31	177.05	52.85	0.0503	177.05	52.60	0.0505	555.54	0.67	1.31	
24	31	178.57	55.47	0.0508	178.68	55.20	0.0512	583.72	0.82	1.45	
25	31	151.31	53.15	0.0463	151.32	52.99	0.0451	430.60	0.68	1.11	

Notes. The complete version of this table and the tables of systems of the other superclusters can be found electronically at: www.astro.ugto.mx/catalog

3.2.2 Application of GFiF to MSCC-310

The galaxy positions and redshifts were transformed from sky coordinates to rectangular ones following the equations 2.4, 2.5 and 2.6. This transformation was carried out after correcting the galaxies co-moving distances D_C for the FoG effect. The N_{gal} was now taken to be the total number of galaxies in the box of MSCC-310, 12 286, to which the GFiF method was applied. The VT algorithm was then applied to calculate the volumetric numerical densities. The optimization of the f and D_E parameters was carried out as described in section 2.5.3.

With this, the HC method identified 768 low density groups using the optimal segmentation parameter $f = 16$. As expected, the application of the HC algorithm on all galaxies detected richer groups ($\sim$ 15 galaxies on average now) and more elongated, with a mean σ_j of 1.8 h_{70}^{-1} Mpc compared with the mean σ_j of 0.5 h_{70}^{-1} Mpc found with GSyF.

After, we calculated the Euclidean distance D_E between the group centroids while the weight was set as the BC coefficient between groups. The linking length estimated by the optimization process (see section 2.5.3) was $D_{max} = 8\,h_{70}^{-1}$ Mpc resulting on 334 edges. These edges were filtered by the minimum mean lineal density along the edge cylinders (in this case 0.008 gal.Mpc^{-3}), resulting on 273 links. This resulted in 34 unconnected trees, obtained from the application of MST. As stated before, and following the definition by Chow-Martínez et al. (2019, in preparation), the objective of GFiF is to search for the filaments which have at least 3 galaxy systems connected by bridges. Only 9 trees were found linking 3 or more systems of galaxies with a richness N_{mem} above 11 galaxies. The richness limit was calculated according to equation 2.13. The remaining trees correspond to isolated bridges (that is, connecting only one pair of system) and tendrils (connections between nodes with less than two systems embedded). Although these structures are also a sub-product of the algorithm, the analysis presented in the following sections concern only the filaments. The filament are the largest structures in the volume, this result is shown in the dendrogram depicted in Figure 3.3 (top panel) which shows 9 dominant filaments for the MSCC-310 supercluster.

Concerning the systems embedded in the structures, from the 359 HC groups (nodes) in the spanning trees, 116 matched with the systems with $N_{mem} \geq 11$ identified with GSyF. From these, 61 were found to be in filaments (53%), 26 (22%) in bridges between pairs of systems, and 29 (25%) not connected by bridges, that is, relatively isolated.

The filaments detected by GFiF algorithm in the MSCC-310 supercluster and their main properties are listed in Table 3.4. Column 1 assigns a sequential number to the filament; column 2 lists the number of systems detected by GSyF linked by the filament; column 3 shows the number of galaxies attributed to the filament; columns 4 to 6 are the mean, minimum and maximum redshift of the filament; column 7 correspond to the mean number density inside the filament; column 8 is the mean transversal radius of the filament skeleton measured as $3\times d_{bas}$; columns 9 and 10 show the number of nodes that constitute the filament and the number of central skeleton nodes, respectively; column 11 is the length of the filament skeleton.

The filaments inside MSCC-310 volume are depicted in Figure 3.3. This figure shows the skeletons plotted in RA [deg] $\times$ Z [Mpc] rectangular coordinates to

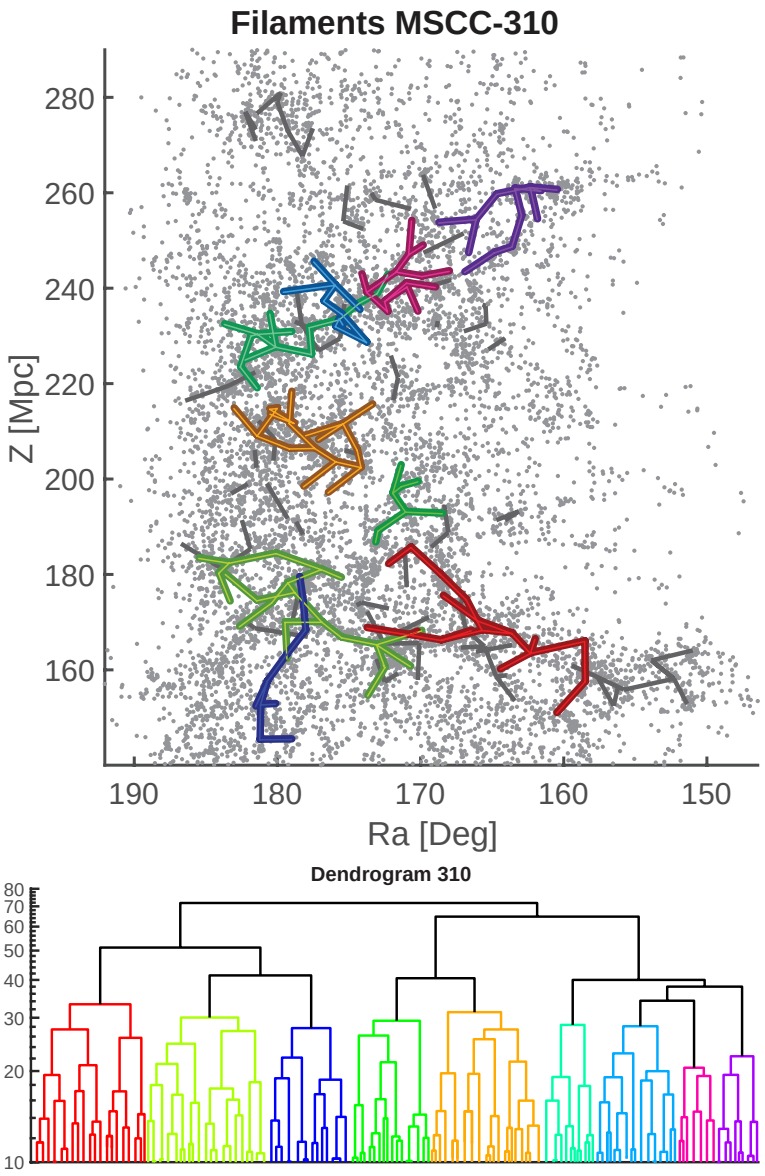

FIGURE 3.3: Results of GFiF algorithm for the supercluster volume MSCC-310. Top: The 9 detected filaments are shown in color over the SDSS galaxies distribution, gray points. The most prominent segments are depicted in dark-gray color. Bottom: Dendrogram color is set according to structures. The y axis on the dendrogram plot indicates the distance at each level of the tree.

facilitate the recognition of structures both in projection and depth. Filaments are depicted in different colors, both in the upper and lower panels. The black lines correspond to bridges that connect less than 3 systems. The longest path of the filaments detected in the MSCC-310 volume ranges from 30 to 70 h_{70}^{-1} Mpc and comprise up to 10 systems. Moreover, the paths between pairs of systems chained together by bridges have lengths from 5 to 24 h_{70}^{-1} Mpc.

TABLE 3.4: Main properties of the filaments extracted with GFiF for the supercluster MSCC-310.

Fil. Nr. (1)	N_{sys} systems (2)	N_{gfil} gals. (3)	redshift [mean, (4)	min, (5)	max] (6)	mean density [h_{70}^3 Mpc^{-3}] (7)	R_{fil} [h_{70}^{-1} Mpc] (8)	N_{nod} filament (9)	skeleton (10)	L_{fil} [h_{70}^{-1} Mpc] (11)
MSCC-310-F1	10	714	0.0609	0.0518	0.0689	0.4180	2.81	18	11	61.6
MSCC-310-F2	8	696	0.0502	0.0443	0.0588	0.5210	2.84	22	11	51.8
MSCC-310-F3	7	506	0.0481	0.0427	0.0528	0.3486	2.50	19	10	49.0
MSCC-310-F4	8	490	0.0656	0.0585	0.0710	0.4263	2.46	14	10	59.0
MSCC-310-F5	7	461	0.0700	0.0642	0.0774	0.5303	2.60	13	7	47.6
MSCC-310-F6	6	409	0.0725	0.0651	0.0791	0.2986	1.97	12	9	39.3
MSCC-310-F7	4	385	0.0551	0.0479	0.0619	0.4062	2.76	7	5	20.7
MSCC-310-F8	4	254	0.0546	0.0485	0.0617	0.2118	2.22	9	7	33.7
MSCC-310-F9	4	155	0.0464	0.0437	0.0528	0.1464	1.13	9	6	17.9

3.3 Validation of the methods

3.3.1 Checking the identified systems of galaxies

In order to validate the GSyF algorithm, the list of identified systems was compared with different cluster and group catalogs available for the SDSS region. For MSCC-310, GSyF detected 139 systems with ten or more galaxies and another 116 systems with $5 \leq N_{mem} < 10$. A match was considered positive if the projected positions of the system in the two compared catalogs were not farther than 1 h_{70}^{-1} Mpc, while for redshift space the difference was set as $\Delta z = 0.007$ which corresponds to ± 2100 km s^{-1}.

For the rich clusters, the GSyF results were compared with the original Abell/ACO catalog (Abell et al., 1989), based on the most recent parameter measurements for its clusters (e.g. Chow-Martinez et al., 2014). We also compared them with the galaxy positions provided by the *brightest cluster galaxy* catalog (BCG, Lauer et al., 2014). Regarding catalogs for the SDSS we compared with the C4 cluster catalog (Miller et al., 2005), which was computed for the SDSS-DR2. These comparisons were carried out using all systems detected by GSyF down to systems with a richness of 5 galaxies. By using a tolerance aperture of 1 h_{70}^{-1} Mpc, 19 of the 37 Abell/ACO clusters inside the MSCC-310 box were detected as systems of richness above 5 galaxies with our method (51%), while the equivalent number was 26 (76%) for the 34 clusters in C4. There are 11 BCG clusters embedded in the volume and 8 (73%) of them have GSyF counterparts. However, incrementing the aperture to 2 h_{70}^{-1} Mpc results in an increase in the detection of Abell clusters to 29/37 (78%), C4 clusters to 33/34 (97%) and BCG clusters to 100%, see Table 3.6.

TABLE 3.5: Glossary of parameters used by GSyS and GFiF algorithms

Param.	Description	MSCC-310
	Amount and density of galaxies	
N	Nr. of galaxies in the volume	12,286
$\bar{d} = \frac{N}{V}$	Mean volume number density (supercluster box)	0.008 Mpc^{-3}
N_{gal}	Nr. of galaxies above the baseline d_{bas}	6,842
$d_{sup} = \frac{N}{A}$	Mean surface number density (sky projection)	76.9 deg^{-2}
N_j	Nr. of galaxies in the HC detected group	each HC
N_{mem}	Nr. of member galaxies of detected FoG system	each system
N_{min}	Minimum nr. of galaxies for systems in filaments at different z	11
N_{gfil}	Nr. of galaxies per GFiF filament detected	each filament
N_{iter}	Number of randomizations	each SC
	Amount of structures	
$N_{cut} = \frac{N_{gal}}{f}$	Nr. of extracted HC groups	1140, 767
N_{HC}	Number of HC detected groups	1015,768
N_{FoG}	Nr. of systems (that survived the FoG filter)	235
N_{nodes}	Nr. of nodes for GFiF algorithm	359
N_{edges}	Nr. of edges for GFiF algorithm	334
N_{links}	Nr. of connections that survived filters 1 and 2	316
N_{trees}	Nr. of trees after MST	34
N_{brid}	Nr. of bridges between two systems	17
N_{ske}	Nr. of filaments in the supercluster box	9
N_{tren}	Nr. of tendrils in the supercluster box	18
N_{nod}	Nr. of nodes in filament	each filament
N_{tot}	Total number of systems above N_{min} in box which matched nodes	116
N_{sys}	Nr. of systems embedded in filaments	61
N_{pair}	Nr. of systems forming pairs connected by bridges	26
N_{isol}	Nr. of systems not forming filaments or pairs	29
	GSyF	
d_i	Local VT (surface or volume) density	each galaxy
d_{bas}	Density baseline	each SC
δ_i	Local density contrast	each galaxy
f	Segmentation parameter (OPTIMIZATION)	6,16
C_j, σ_j	Properties of HC detected group (centroid and mean sky position)	each HC
C_{BGM}	Position of Brightest HC Group Member	each HC
R_a, S_a, M_a R_h	Parameters of iterative process for FoG correction	each HC
σ_v, v_{LOS}, R_{vir}, M_{vir}	Properties of FoG systems	each system
C_{FoG}	Properties of FoG detected group (centroid and mean sky position)	each system
Δz	Difference in velocity	each system
	GFiF	
D_E	Euclidean distance between two nodes (edge size)	each edge
BC	Bhattacharyya coefficient (edge weight)	each edge
W	Weight of the connection	each edge
D_{max}	Linking length (first filter) (OPTIMIZATION)	8 Mpc
D_{cyl}	Cylinder density (second filter)	each edge
d_{fil}	Mean density of filament	each filament
R_{fil}	Mean radius of filament	each filament
L_{fil}	Longitude of filament skeleton	each filament
D_{fil}	Euclidean distance between galaxy and filament skeleton	each galaxy

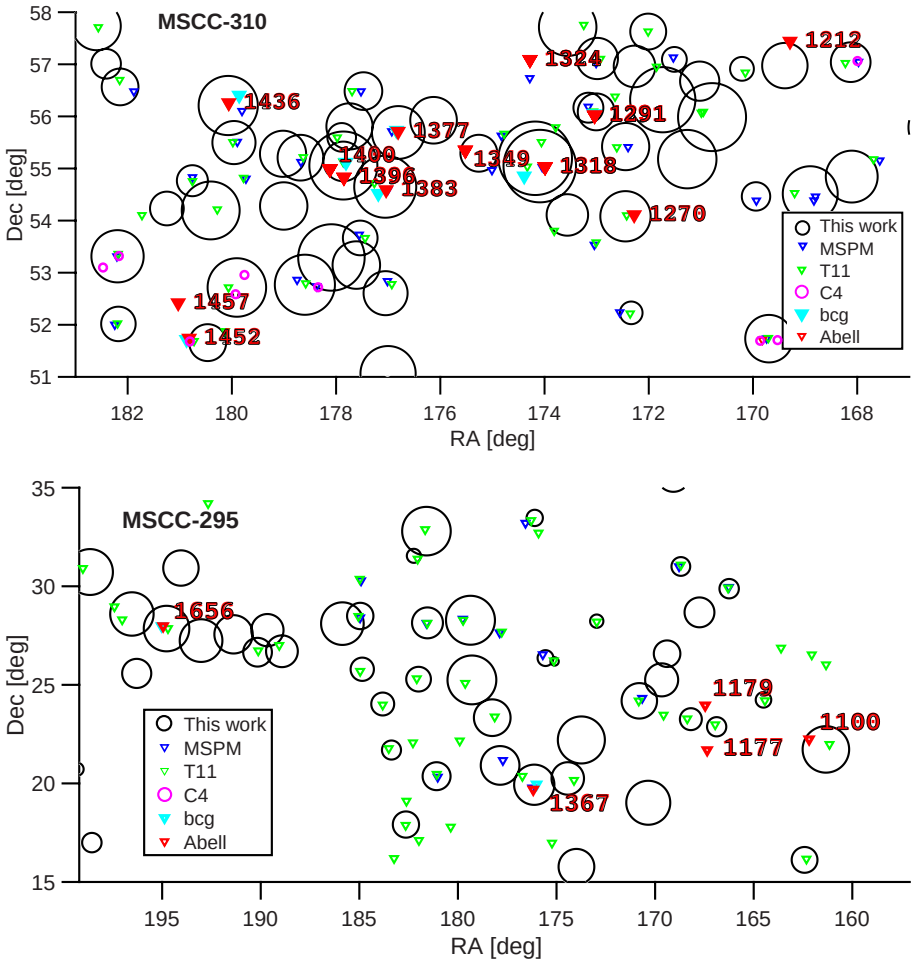

FIGURE 3.4: Projected distribution in the sky of systems detected by GSyF. Top: Systems detected for the MSCC-310. Bottom: Systems detected for the MSCC-295 supercluster. The system radii is shown as a circle of $r = R_{vir}$. For comparisons, the position of systems reported by Abell, C4, MSPM and T11 catalogs are depicted by color points: red, pink, blue and green, respectively.

The increase of $\sim 21\% - 27\%$ in all rich cluster matches when using a larger aperture size can be related to the fact that the mean separation of galaxy members increases for lower richness systems, and the determination of the cluster center then is subject to this separation, see Table 3.6. For example, A1452 and A1507B have a GSyF counterpart located at $\sim$1.5 h_{70}^{-1} Mpc distance and $\Delta\sigma_v$ of ~ 630 km s^{-1} and 120 km s^{-1} respectively (See Table 3.3 systems No. 12 and 14), while their C4 counterparts are 0.7 and 0.4 h_{70}^{-1} Mpc , respectively.

For less rich clusters and groups, GSyF systems were compared with the *Multi-scale Probability Mapping clusters/groups* catalog (MSPM, Smith et al., 2012) and the Tempel et al. (2011) catalog (here after T11), based on the SDSS-DR7 and -DR8 respectively. There are 105 clusters detected by T11 and 79 clusters

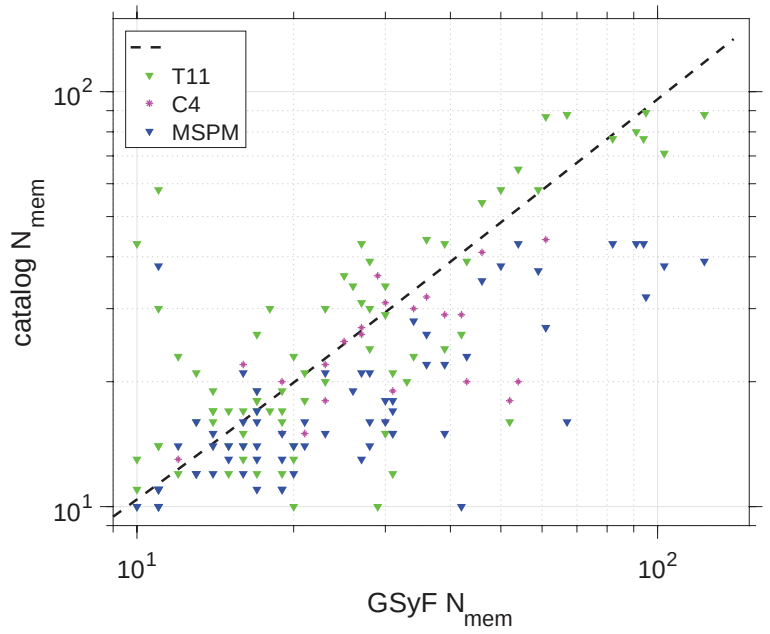

FIGURE 3.5: Comparison of MSCC-310 supercluster GSyF system richness against the richness measured by other catalogs for the matching systems. Symbol colors are the same as the ones in the top figure. The dashed line represents the identity.

found via MSPM with richness larger or equal to 10 galaxies for the MSCC-310 volume. Our algorithm detected systems that correspond to 70% of the T11 clusters (167 detections) and 80% of the MSPM clusters (249 detections) within an aperture of 1 h_{70}^{-1}Mpc, see Table 3.6. On the other hand, one can observe that both catalogs, T11 and MSPM, have a larger number of systems matched with GSyF systems of richness between 10 and 5 galaxies, causing the match ratio to decrease. This is acceptable for our purposes since GSyF has been constructed to find the clusters that present a FoG effect, although our algorithm can clearly find poorer systems.

It is interesting to note that, for rich clusters (Abell, BCGL and C4), the mean separation (Table 3.6) between different catalogs is about 0.4-0.5 h_{70}^{-1} Mpc and 100-150 km s^{-1}, with 1 h_{70}^{-1} Mpc of tolerance, while for groups (T14 and MSPM) these numbers are 0.35 h_{70}^{-1} Mpc and about 200 km s^{-1}.

The systems detected in the main portion of the MSCC-310 supercluster are depicted, on a sky projected distribution, in Figure 3.4, by black circles with radius equal to the measured virial radius. The system positions from Abell, C4, BCG, MSPM and T11 catalogs are depicted as red, pink, cyan, blue and green symbols.

TABLE 3.6: GSyF systems listed in other catalogs for the MSCC-310 superclus-ter.

Aperture=1 h_{70}^{-1} Mpc						
Other catalog	Number	Fraction	Number	Fraction	separation	
	$N_{gal} > 5$		$N_{gal} > 10$		h_{70}^{-1} Mpc	$\Delta\sigma_v$
C4	26/34	76%	24/34	71%	0.43	100
Abell	19/37	41%	17/37	46%	0.45	295
BCG	8/11	73%	8/11	73%	0.5	340
MSPM	142/213	67%	63/79	80%	0.34	145
T11	192/315	61%	73/105	70%	0.33	230
Aperture=2 h_{70}^{-1} Mpc						
	$N_{gal} > 5$		$N_{gal} > 10$			
C4	33/34	97%	32/34	94%	0.77	166
Abell	29/37	78%	24/37	65%	0.82	300
BCG	11/11	100%	11/11	100%	0.80	430
MSPM	167/213	78%	68/79	86%	0.55	193
T11	249/315	79%	85/105	81%	0.64	320

Similar analyses can be done for the other superclusters in our sample. For example, for the *Coma* supercluster (MSCC-295), the GSyF algorithm detected, in total, 160 systems. Among these, one can find A1656 cluster, which is composed of 579 galaxies. The estimated virial radius and mass for this system are respectively, $1.96\ h_{70}^{-1}$ Mpc and 7.7×10^{14} $M_{\odot}$. The second richest system is A1367, which has 243 galaxies and its radius and mass are, respectively, $1.73\ h_{70}^{-1}$ Mpc and 5.3×10^{14} $M_{\odot}$. These estimations are in good agreement with those measured by Rines et al. (2003).

3.3.2 Checking the filament skeletons

In order to check the efficiency of the GFiF algorithm, the filaments detected in MSCC-310 were compared against those presented by Tempel et al. (2014) (hereafter, T14) as extracted from their table 2. The T14 catalog is suitable to evaluate the GFiF filament detection since it lists filaments derivated from SDSS-DR7. The T14 filament position are given in *survey coordinates* (η and λ) with (η and λ) = (0.,0.), (RA, Dec) = (185., 32.5). The transformation to angular coordinates was carried out according to the following transformation:

$$Dec = \sin^{-1}\{\cos(\lambda) * \sin(\eta + 32.5)\}, \tag{3.1}$$

$$RA = \tan^{-1}\left\{\frac{\sin(\lambda)}{\cos(\lambda)*\cos(\eta+32.5)}\right\} + 185.0. \tag{3.2}$$

There are about 630 T14 filaments that lie in the sampled volume of MSCC-310 supercluster. These filaments have a mean length of 9 h_{70}^{-1} Mpc while the largest one has a length of 48 h_{70}^{-1} Mpc. As a comparison, the filament skeletons detected by GFiF have a mean length of 42 h_{70}^{-1} Mpc and the largest one has a length of 62 h_{70}^{-1} Mpc. The T14 filaments are a sequence of contiguous points forming a line. Then, the match rate was evaluated as the mean distance from the points to the skeleton of the GFiF filaments. The GFiF filaments show a match of 40% with the T14 filaments. Moreover, this number increases to 80% when comparing with GFiF isolated bridges and tendrils. The mean distance between the medial axis of the T14 filaments matching the nearest filament/tendril detected by GFiF is $\sim$ 1.5 h_{70}^{-1} Mpc.

The detected GFiF filaments are depicted over T14 filaments in Figure 3.6. As can be seen in this Figure, GFiF detects the most prominent (dense) filaments of these listed by T14.

3.3.3 Comparison with KDE density maps

For further validating the results from GSyF and GFiF algorithms a KDE method was applied to the galaxies in the MSCC-310 volume. As mentioned before, the KDE method makes use of the local density of the galaxies to estimate the density probability function that better describes de galaxy distribution. This validation was carried out over the 2D projections (density maps) of the 3D KDE (XY, XZ, YZ). For the KDE implementation, the kernel size was set as 1 Σ (see section 2.1.1). Since each kernel is created based on the VT cell, then the baseline density was scaled in multiples of d_{bas}. Then, we selected those regions for which $d_{kde} > 1\, d_{bas}$ in the RA$\times$Dec projected density map. A comparison of the position of the density peak of each region with the centroids of the GSyF systems revealed that 93 $N_j \geq 10$ GSyF systems (76%) match density peaks above $3\ d_{bas}$. The remaining 24% are identified with density peaks between $(1-2)\ d_{bas}$. Moreover, we observe that the filament edges are connecting these density peaks

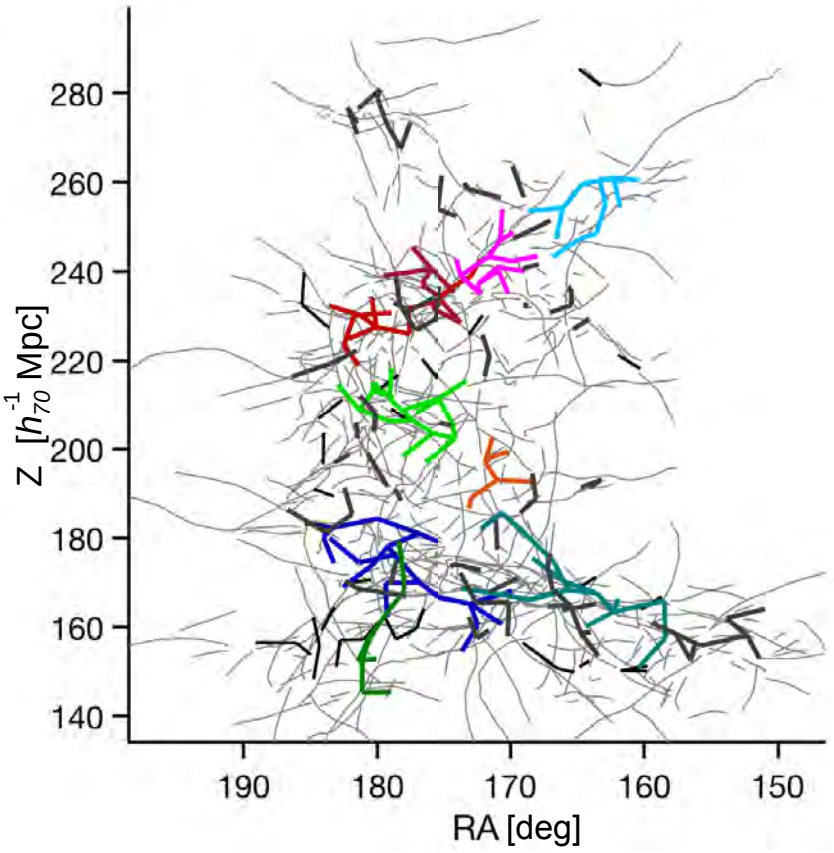

FIGURE 3.6: Comparison of GFiF filaments for the MSCC-310 supercluster against the T14 filaments for the SDSS DR8. Gray lines are T14 filaments. Colored lines depict filaments from the present work.

forming chains of overdensity regions, i.e. chains of systems. In Figure 3.7 the systems detected by GSyF are represented by circles of $r = R_{vir}$ over the galaxy density distribution as obtained from KDE in a RA$\times$Dec projection. In Figure 3.8 the filaments are shown overlaid to the KDE density map for the MSCC-310 volume. These density is expressed in units of the mean number density.

3.4 Filament Properties

3.4.1 Main properties of the filaments

The same methodology described for the MSCC-310 supercluster, was applied to the 46 superclusters of our sample, detecting about 1 500 systems of galaxies

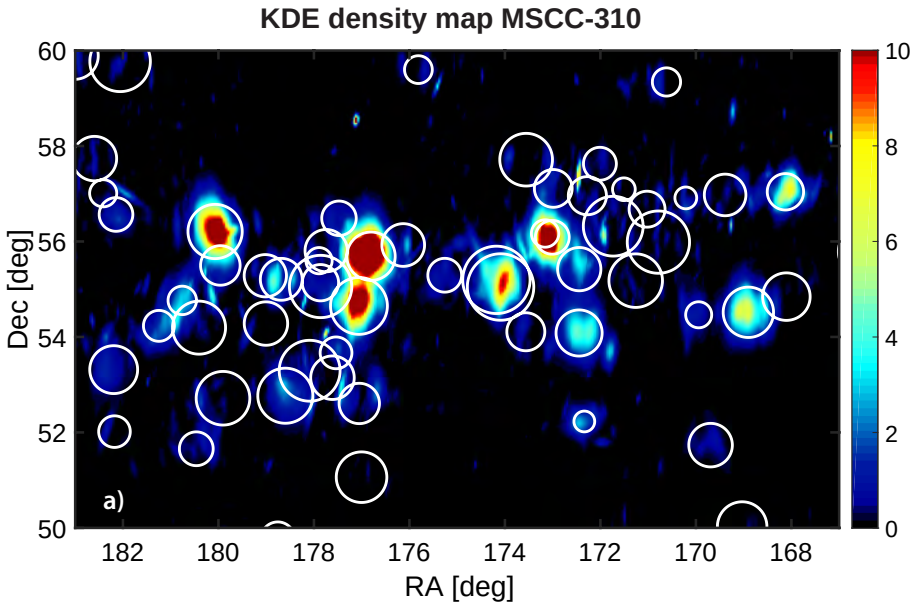

FIGURE 3.7: RA.Dec. projected density map as measured from 3D-KDE with 1Σ in terms of the density contrast. The GSyF systems are represented by white circles with radius equal to the estimated R_{vir}. Density is indicated through a color scale displayed on the right, where denser regions are redder and less dense zones are bluer.

(see table 3.2) and 143 filaments in 40 superclusters which are summarized in table 3.7.

This table also lists the parameters used for GFiF for each supercluster: column 2 notes the segmentation parameter f, while column 3 presents the number of detected HC groups. Column 4 shows the linking length (D_{max}) used to connect HC groups. The process of filtering the connections can be followed through columns 5 to 10, which show, respectively, the number of detected edges, the number of filtered links, the number of trees detected after applying MST and the final number of filaments, N_{ske}, number of isolated bridges, N_{brid}, and tendrils, N_{tend}. Column 11 lists the minimum richness considered for GSyF systems to be accepted as bridges. Column 12 to 14 present, respectively, the fraction of these systems included in the GFiF filaments, in isolated bridges and the ones not connected by bridges. Finally, columns 15 to 17 show the number of galaxies hosted by the GFiF filaments, N_{fil}, and the filling factors calculated as V_{fil}/V and N_{fil}/N.

The list of detected filaments for each supercluster volume can be consulted in Appendix B, in the same format as the one presented in Table 3.4. In particular,

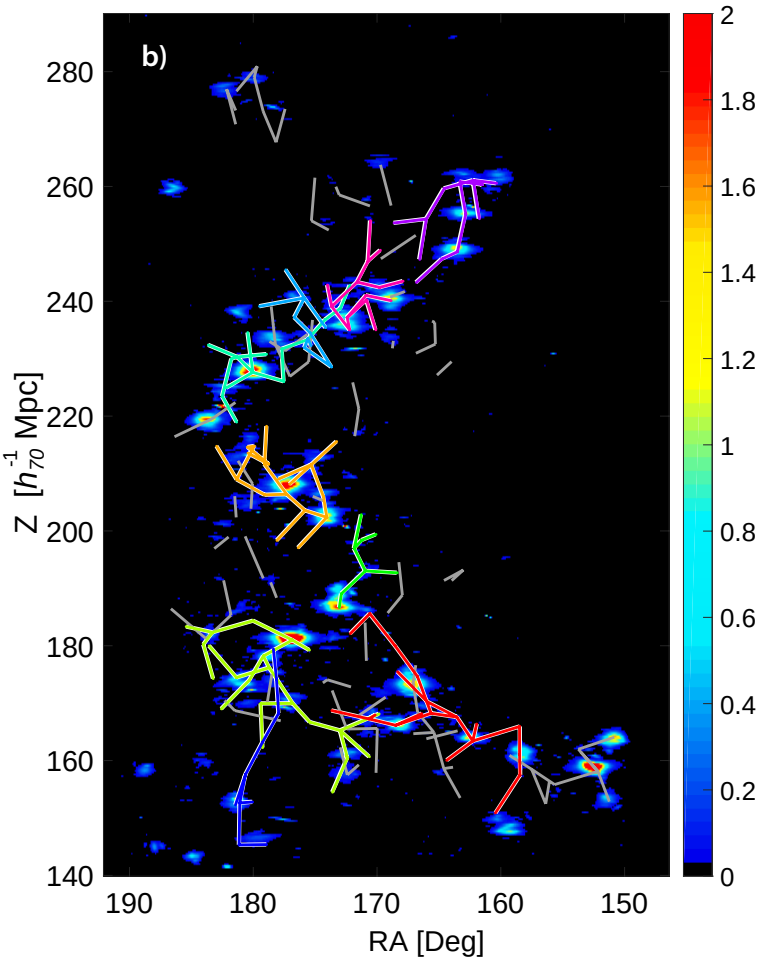

FIGURE 3.8: KDE projected results. The denser regions are represented by light red color while the less dense regions are colored as blue. The filaments are overlaid in white color. Density is indicated through a color scale displayed on the right, where denser regions are redder and less dense zones are bluer.

MSCC-55 and MSCC-579 superclusters did not have a complete coverage by the SDSS. Thus, GFiF algorithm was not able to detect filaments for these volumes.

The filaments detected by GFiF have lengths from 9 to 130 h_{70}^{-1} Mpc. Figure 3.9 depicts the length distribution for all the detected filaments. The distribution shows that the majority of the structures' lengths range from 15 up to 80 h_{70}^{-1} Mpc. There are two structures longer than 100 h_{70}^{-1} Mpc. A 130 h_{70}^{-1} Mpc long filament is located in MSCC-323, containing the Abell clusters A1449B and A1532A, the second of 105 h_{70}^{-1} Mpc in MSCC-335. Excluding these two particular cases, the mean length of the filaments is about 37 h_{70}^{-1} Mpc while the median corresponds to 29 h_{70}^{-1} Mpc.

TABLE 3.7: Summary of filaments detected by GFiF for the superclusters in table 1.

MSCC ID	f	N_{HC}	D_{max}	N_{edges}	N_{links}	N_{trees}	N_{ske}	N_{brid}	N_{tend}	N_{min}	N_{sys}			N_{fil}	$\frac{V_{fil}}{V}$	$\frac{N_{fil}}{N_{gal}}$
											filaments	bridge	isolated			
No.	#	#	#	#	#	#	#	#	#	#	%	%	%	#	%	%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
55	10	81	6	7	6	2	2	0	0	10	63.6	0.0	36.4	351	0.9	43.2
72	16	121	8	29	28	6	4	1	1	8	56.7	6.7	36.7	667	1.3	34.4
175	29	86	14	36	29	4	4	0	0	6	57.7	0.0	42.3	987	4.9	39.4
184	27	78	15	29	13	2	2	0	0	6	57.1	0.0	42.9	285	1.6	13.6
211	16	93	12	29	22	2	1	0	1	5	100.0	0.0	0.0	365	3.8	24.6
219	27	71	17	65	36	3	2	2	1	5	72.2	22.2	5.6	562	6.2	29.4
222	15	124	16	84	43	5	2	0	3	4	44.4	0.0	55.6	513	4.4	27.5
223	15	52	18	31	19	2	1	0	1	4	75.0	0.0	25.0	168	3.3	21.6
229	16	116	19	93	44	8	1	1	6	4	37.5	25.0	37.5	158	1.8	8.5
236	32	270	9	124	94	11	7	2	2	18	55.3	8.5	36.2	2547	4.5	29.5
238	26	320	18	215	115	20	6	3	11	6	39.0	10.2	50.8	1438	3.8	17.3
248	21	60	20	47	20	3	1	0	2	5	50.0	0.0	50.0	246	3.9	19.5
266	22	60	16	17	20	3	1	0	2	5	50.0	0.0	50.0	246	3.9	19.5
272	16	86	6	14	14	3	2	0	1	8	90.0	0.0	10.0	549	3.0	39.8
277	15	183	11	77	62	6	2	0	4	5	62.5	0.0	37.5	977	5.3	35.6
278	16	495	7	378	256	20	5	1	14	19	58.1	4.7	37.2	2707	4.9	34.2
283	23	101	16	32	21	5	3	0	2	4	44.4	0.0	55.6	284	1.2	12.2
295	14	1022	5	478	398	38	4	5	29	26	43.1	19.6	37.3	3294	1.7	23.0
310	16	768	8	334	273	34	9	7	18	10	52.7	12.7	34.5	3540	2.4	28.8
311	25	211	10	49	37	5	4	0	1	7	43.6	0.0	56.4	1322	2.0	25.1
314	14	40	8	13	7	2	2	0	0	7	50.0	0.0	50.0	149	2.1	26.7
317	15	56	17	56	31	3	2	1	0	5	61.5	15.4	23.1	349	7.3	41.5
323	22	151	16	46	29	5	2	2	1	4	26.3	10.5	63.2	552	3.0	16.6
333	19	104	11	39	28	6	3	1	2	7	45.8	8.3	45.8	513	4.1	26.1
335	23	135	13	108	65	8	3	0	5	8	52.9	0.0	47.1	971	5.6	31.3
343	11	244	7	51	33	5	3	1	1	7	34.5	6.9	58.6	421	1.1	15.7
360	38	58	20	54	23	4	3	0	1	6	47.6	0.0	52.4	718	7.5	32.7
386	9	362	7	165	120	20	4	2	14	9	60.6	12.1	27.3	909	2.2	27.9
407	16	362	18	2773	120	20	4	2	14	9	60.6	12.1	27.3	909	2.2	27.9
414	9	1211	6	462	386	47	15	12	20	9	42.4	16.7	41.0	2451	1.4	22.5
419	15	115	11	39	22	5	3	3	0	5	28.6	17.1	54.3	384	1.4	22.3
422	18	59	19	35	10	1	1	0	0	4	37.5	0.0	62.5	227	4.2	21.3
430	20	80	12	35	22	5	4	1	0	6	58.3	8.3	33.3	488	3.7	30.4
440	15	229	10	66	44	6	3	1	2	5	24.4	4.4	71.1	468	1.5	13.6
454	15	380	6	162	123	17	2	0	15	13	12.5	0.0	87.5	818	1.2	14.3
457	21	194	9	80	68	7	6	0	1	8	66.7	0.0	33.3	2011	7.7	49.4
460	22	159	14	102	65	6	4	1	1	5	65.9	4.5	29.5	1362	5.6	38.9
463	16	529	8	230	183	27	11	8	8	8	47.2	12.8	40.0	2605	2.1	30.8
474	15	495	5	245	209	20	7	3	10	16	48.0	12.0	40.0	2519	3.8	33.9
484	22	495	16	6732	209	20	7	3	10	16	48.0	12.0	40.0	2519	3.8	33.9
55	10	81	6	7	6	2	2	0	0	10	63.6	0.0	36.4	121	0.2	14.9
72	16	121	8	29	28	6	4	1	1	8	56.7	6.7	36.7	829	1.0	42.7
175	29	86	14	36	29	4	4	0	0	6	57.7	0.0	42.3	507	0.9	20.2
184	27	78	15	29	13	2	2	0	0	6	57.1	0.0	42.9	218	0.6	10.4
211	16	93	12	29	22	2	1	0	1	5	100.0	0.0	0.0	233	1.0	15.7
219	27	71	17	65	36	3	2	2	-1	5	72.2	22.2	5.6	279	1.1	14.6
222	15	124	16	84	43	5	2	0	3	4	44.4	0.0	55.6	97	0.2	5.2
223	15	52	18	31	19	2	1	0	1	4	75.0	0.0	25.0	11	0.1	1.4
229	16	116	19	93	44	8	1	1	6	4	37.5	25.0	37.5	32	0.1	1.7
236	32	270	9	124	94	11	7	2	2	18	55.3	8.5	36.2	1600	0.8	18.5
238	26	320	18	215	115	20	6	3	11	6	39.0	10.2	50.8	544	0.4	6.5
248	21	60	20	47	20	3	1	0	2	5	50.0	0.0	50.0	159	0.9	12.6
266	22	44	16	20	12	1	1	0	0	5	71.4	0.0	28.6	132	0.7	13.8
272	16	86	6	14	14	3	2	0	1	8	90.0	0.0	10.0	453	0.9	32.8
277	15	183	11	77	62	6	2	0	4	5	62.5	0.0	37.5	611	1.0	22.2
278	16	495	7	378	256	20	5	1	14	19	58.1	4.7	37.2	2144	1.6	27.1
283	23	101	16	32	21	5	3	0	2	4	44.4	0.0	55.6	263	0.5	11.3
295	14	1022	5	478	398	38	4	5	29	26	43.1	19.6	37.3	2992	1.0	20.9
310	16	768	8	334	273	34	9	7	18	10	52.7	12.7	34.5	2817	1.0	22.9
311	25	211	10	49	37	5	4	0	1	7	43.6	0.0	56.4	1118	0.8	21.2
314	14	40	8	13	7	2	2	0	0	7	50.0	0.0	50.0	112	0.7	20.1
317	15	56	17	56	31	3	2	1	0	5	61.5	15.4	23.1	63	0.5	7.5
323	22	151	16	46	29	5	2	2	1	4	26.3	10.5	63.2	239	0.5	7.2
333	19	104	11	39	28	6	3	1	2	7	45.8	8.3	45.8	282	0.6	14.3
335	23	135	13	108	65	8	3	0	5	8	52.9	0.0	47.1	478	1.1	15.4
343	11	244	7	51	33	5	3	1	1	7	34.5	6.9	58.6	335	0.4	12.5
360	38	58	20	54	23	4	3	0	1	6	47.6	0.0	52.4	218	0.8	9.9
386	9	362	7	165	120	20	4	2	14	9	60.6	12.1	27.3	636	0.7	19.5
407	16	70	18	49	20	3	1	0	2	4	50.0	0.0	50.0	101	0.7	9.0
414	9	1211	6	462	386	47	15	12	20	9	42.4	16.7	41.0	2232	0.8	20.5
419	15	115	11	39	22	5	3	3	-1	5	28.6	17.1	54.3	254	0.3	14.7
422	18	59	19	35	10	1	1	0	0	4	37.5	0.0	62.5	11	0.0	1.0
430	20	80	12	35	22	5	4	1	0	6	58.3	8.3	33.3	186	0.5	11.6
440	15	229	10	56	36	8	1	2	5	5	20.7	13.8	65.5	184	0.2	5.3
454	15	380	6	164	117	15	5	4	6	13	45.9	13.1	41.0	1516	1.2	26.6
457	21	194	9	80	68	7	6	0	1	8	66.7	0.0	33.3	1525	1.9	37.5
460	22	159	14	102	65	6	4	1	1	5	65.9	4.5	29.5	895	1.4	25.6

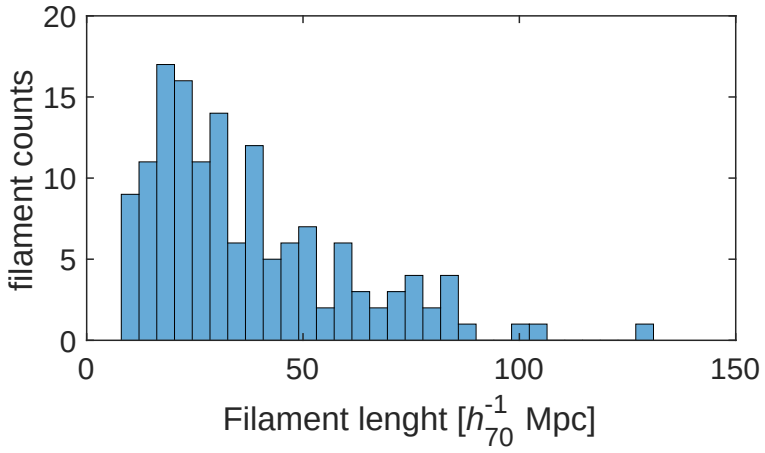

FIGURE 3.9: Distribution of filament lengths for the 143 filaments detected by GFiF. The length used corresponds to the longest path between a pair of systems. See table 3.7.

3.4.2 Distribution of galaxies along the filaments

It is expected that the environment inside filaments is denser than in the dispersed component of the supercluster or in the general field. Therefore, longitudinal density number profiles were extracted, to evaluate the environment along the filaments. Figure 3.10 depicts the longitudinal distribution of galaxies for all bridges forming filaments in the supercluster MSCC-310, along the path from one system to another. These distributions show that the density of galaxies is higher in the extremes of the bridges, as expected, and decreases through the midpoint between systems. Then, density profiles for bridges were extracted from the systems to the midpoint, by counting the galaxies that lie within a cylinder of radius 1 h_{70}^{-1} Mpc with medial axis set by the bridge skeleton. The galaxies are counted in bins of size $\Delta d = 0.5\, h_{70}^{-1}$ Mpc. We also extracted longitudinal profiles after stripping the systems (considered at 1.5 R_{vir}) from their bridges, in order to evaluate the environment of pure filaments. These profiles allow to compare the mean density of the filaments with the background density. Figure 3.11 shows the longitudinal number density profile for all filaments detected in our sample. The stacked longitudinal profile including galaxies in systems is depicted by a blue line. The dispersion about the stacked profile is represented by a blue shaded area. The pure profiles (excluding the systems' galaxies) is represented by the red line, and its corresponding dispersion by a red shaded area. As can be seen, the mean density contrast along the filament is ~ 10, that is, the filament is about 10 times denser than the background.

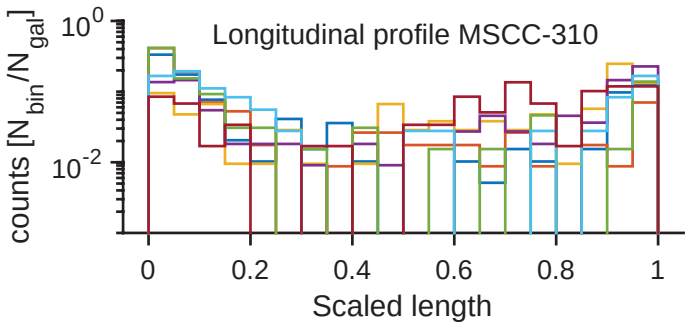

FIGURE 3.10: Galaxy density along bridges connecting pairs of systems for the nine MSCC-310 filaments. All bridges are scaled to length 1.0.

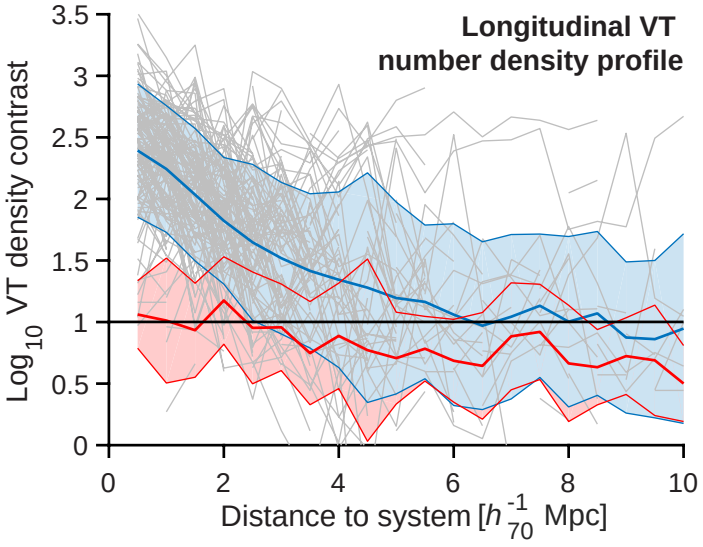

FIGURE 3.11: Longitudinal VT density distribution for galaxies for all bridges on filaments detected by GFiF. Profiles are considered from the system center to the middle of the bridge. The blue line depicts the mean longitudinal profile for bridges including galaxies in systems. The thide red line corresponds to the mean longitudinal profile for all filaments excluding galaxies belonging to systems within $1.5R_{vir}$. Blue and red shaded areas are the dispersion about the stacked profile.

3.5 Transversal profiles

The extraction of transversal profiles allows to further characterize the filament environment and their effect on the galaxies. These profiles are calculated by setting up a series of concentric cylinders with axes orientated along the filament skeletons. Then, a bin is considered to be the volume within two concentric cylinders of radius R_{cy} and $R_{cy} + \Delta R_{cy}$. The occurrence of a galaxy proxy in each bin is determined according to the galaxy distance to the filament edges D_{fil}.

The total count of galaxies per bin is weighted by the bin volume, in a similar way as making a normalized histogram. In order to compare samples of different sizes, a normalization is done by dividing the number of events in a bin by the total number of galaxies in the sample.

3.5.1 Density in transversal profiles

While the longitudinal profiles allow to characterize the mean number density along filaments, the transversal profiles allow to characterize the filaments up to and beyond their influence region.

The extraction of transversal density profiles was carried out excluding the galaxies located in systems, within a radius set as 1.5 R_{vir}. This allow to characterize pure filaments. The density profile is calculated as described in section 3.5. The cylinder radii R_{cy} were chosen from 0 up to 10 h_{70}^{-1} Mpc in steps of $\Delta R_{cy} = 0.5\, h_{70}^{-1}$ Mpc.

The transversal galaxy number density profile for filaments is calculated in two ways:

- The *local number density profile* is calculated by counting the number of galaxies within concentric cylinders and dividing them by the volume within the cylinders.
- The *VT number density profile* is calculated as the mean VT number density within concentric cylinders, d_i, as described in Sec. 2.1.1.

The local number density and VT number density profiles are expressed as a density contrast (w.r.t. d_{bas}) and stacked together.

Figure 3.12 shows the stacked profile for all filaments detected by GFiF. One can observe, in both local and VT density profiles, that the overdensity extends up to 2 - 3 h_{70}^{-1} Mpc. Beyond 5 h_{70}^{-1} Mpc, the density profiles tend to reach values below the mean density, as expected since the mean density accounts for both, the filaments and background field. However, this effect is less obvious in the VT density profile, beyond 5 h_{70}^{-1} Mpc the density profile is noisier but lies within to 2 - 3 $\times\, \bar{d}$.

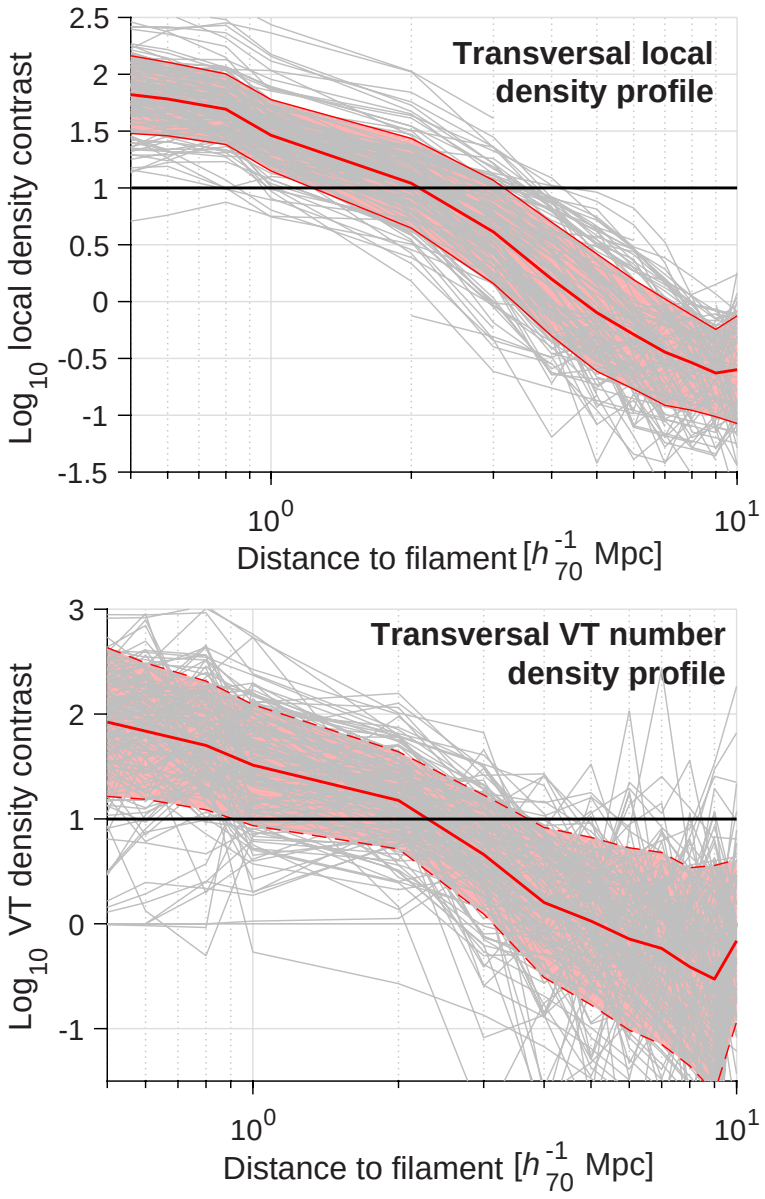

FIGURE 3.12: Stacked density profiles for the 143 filaments identified by GFiF. Individual profiles are represented by thin gray lines. Top: The red line corresponds to the mean local density (stacked) profile. Bottom: Mean VT density stacked profile. The solid line indicates the mean profile while the shaded area represents the dispersion of the profile. The solid black line depicts the density contrast 10 $\times \bar{\delta}$.

Moreover, the mean VT density number is compatible with the mean local density measured by the longitudinal profile, (see Fig. 3.11) for distances below 1 h_{70}^{-1} Mpc.

The density profiles were used to estimate the mean radius of the filaments, R_{fil}. This was achieved by considering the intersection point at which the VT density profile crosses the 10 $\times d_{bas}$ line, as indicated in figure 3.12 by the black solid

line. The mean radius as well as the mean density of each filament is noted in Table 3.7. Figure 3.13 presents the radius distribution for all the filaments. The

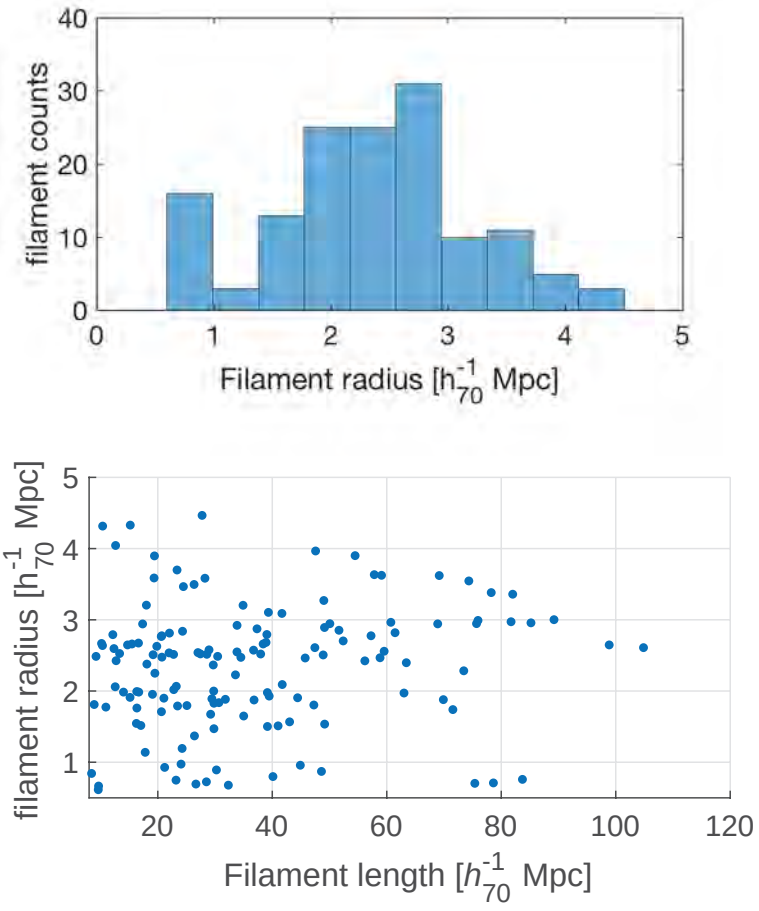

FIGURE 3.13: Comparison of filament length and radius for the 143 filaments detected by GFiF. The length used correspond to the longest path between a pair of systems. See table 3.7. Top, Distribution of radius of filaments in our sample. Bottom, filament radius as a function of the filament length.

distribution shows that radii range from 0.6 to 4.5 h_{70}^{-1} Mpc with a mean value of about 2.4 h_{70}^{-1} Mpc . The bottom panel of the figure depicts the filament radius as a function of the filament length. From these results one can conclude that the filament length does not correlate with the filament radius, as expected for a "universal" pattern for the filaments. However, it is important to note that the mean radius does not represent the radius along the whole filament path.

Afterwards, the filling factor was calculated in two ways: the first takes the ratio between the number of galaxies inside the filament radius N_{gfil} and the total number of galaxies N. The second is calculated as the rate between the filaments

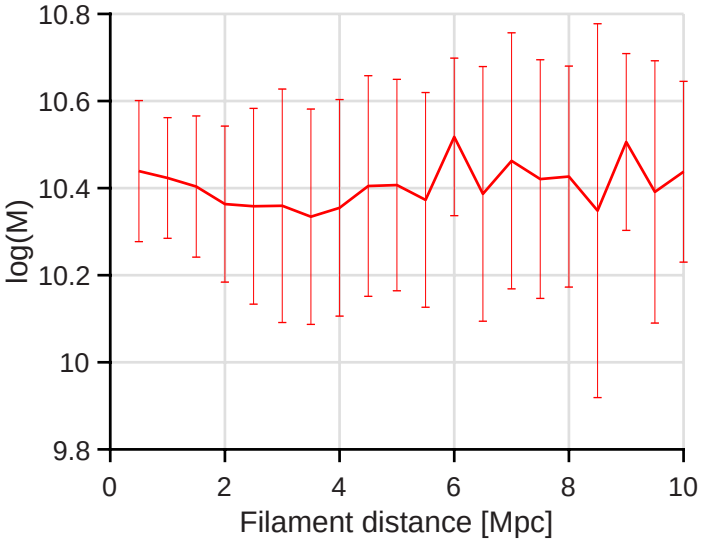

FIGURE 3.14: Stacked mass profile not weighted for 143 filaments detected by GFiF. Errors correspond to the variance of the stacked profiles.

volume, estimated from the filaments length and radius, and the total volume of the supercluster box.

3.6 Properties of galaxies in filaments

3.6.1 Stellar mass profile

The stellar mass profile for all filaments was extracted as described in section 3.5.

Before the extraction the mass was weighted by the average mass of the volume under analysis to remove the redshift dependence of the stellar mass (Chen et al., 2017b). This weighting is equivalent to a normalization of the stellar mass and allows to carry out a stacking procedure in order to increase the signal of the profiles. Without this weighting, the stacked mass profiles do not present any particular trend and the error increases, as shown in figure 3.14.

Figure 3.15 shows the stacked weighted mass profile for all filaments. The dispersion of the profile is depicted by the dashed red lines. Statistically, the fraction of massive galaxies is larger within $2\,h_{70}^{-1}$ Mpc from the filaments than outside

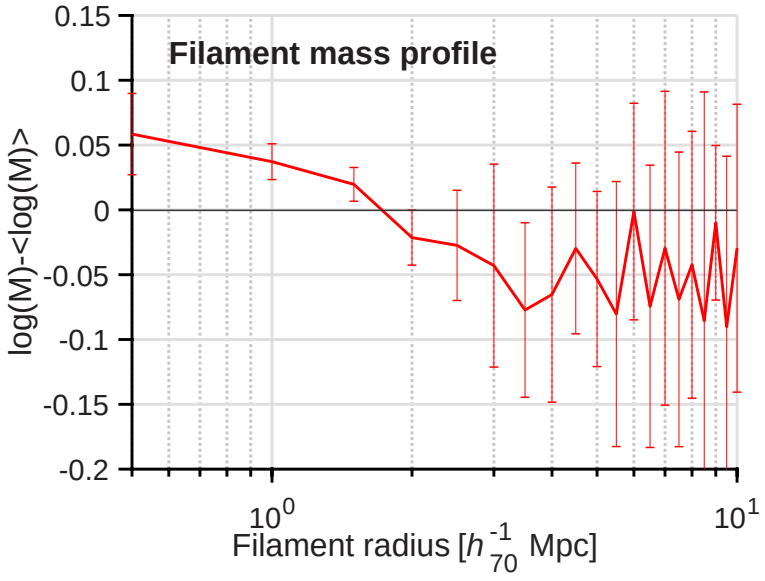

FIGURE 3.15: Stacked mass profile for 143 filaments detected by GFiF. Errors correspond to the variance of the stacked profiles.

them (beyond 3 h_{70}^{-1} Mpc), indicating that the stellar mass correlates with the distance to the filament axes. These results are in good agreement with the results presented by Chen et al. (2017b) for the MGS sample from DR7 Abazajian et al. (2009). Also they are compatible with the results presented by Kraljic et al. (2018), which present a similar trend for the filaments found for the GAMA spectroscopic survey, for a redshift range $0.03 \leq z \leq 0.25$.

3.6.2 Morphological type

Morphology profiles were calculated in order to analyze if there is some morphological trend in the population of filament galaxies (as may be expected from the morphology-density relation). For this analysis we used the Huertas-Company et al. (2011) morphological classification. They classify the galaxies in four morphological types. For this analysis we used the probability p(Early) = p(E) + p(S0) which classifies galaxies in early type p(Early)> 0.5 and late type p(Early)< 0.5.

Therefore, the distribution of both galaxy types was calculated as a function of the distance to the filament. These distributions were normalized so they can be compared and stacked for all filaments in the sample in a similar way as a profile extraction, excluding galaxies in systems, see Figure 3.16.

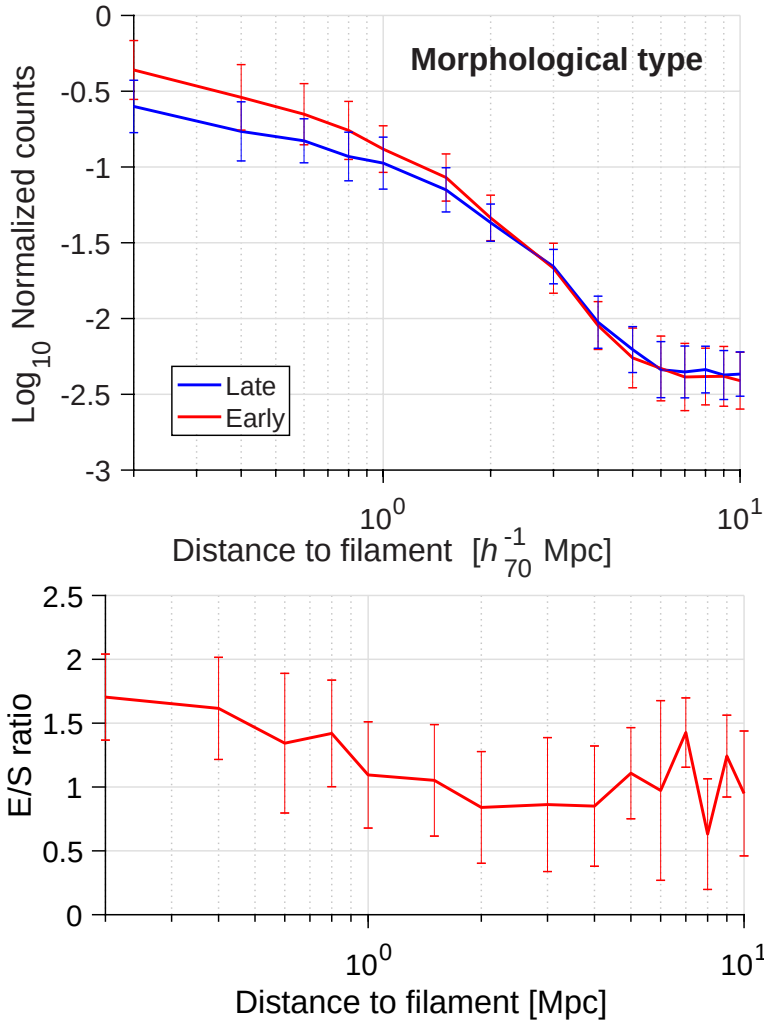

FIGURE 3.16: Top: Stacked morphological type profiles for 143 filaments detected by GFiF. Weighted error are depicted as error bars. Bottom: early to late type ratio as function of the distance to the filament axis.

The results of this analysis show that the density of early type galaxies is higher than the population of late types near the filament axis. This effect is more notorious when computed as an early to late type ratio. We can see that at distances lower than $2\,h_{70}^{-1}$ Mpc, the fraction of early types reaches almost twice the fraction of late types. At larger distances (that is, towards the dispersed supercluster population) the fractions tend to be similar. Moreover, these results are consistent with those presented by Kuutma et al. (2017) for the Huertas-Company et al. (2011) sample. They also observe that early type galaxies are more abundant near the filament axes.

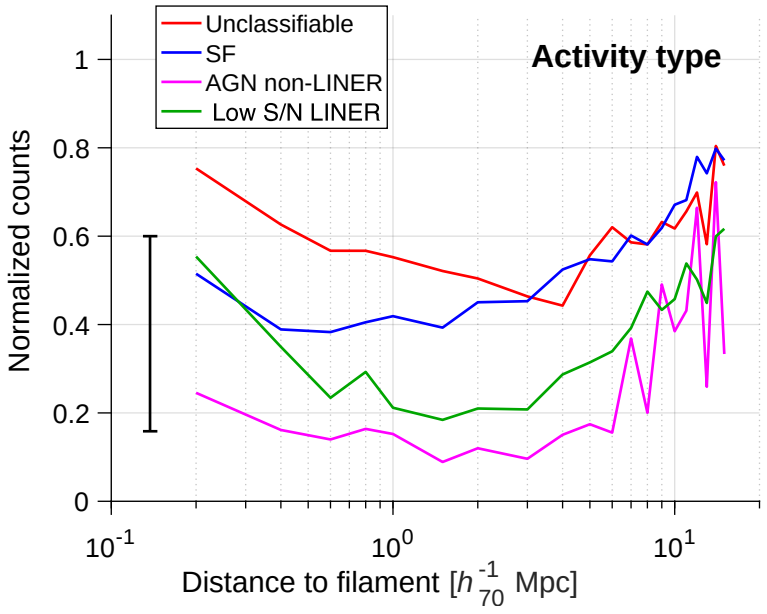

FIGURE 3.17: Stacked activity type profiles for 144 filaments detected by GFiF. The black bar on the left represents the typical errors on the stacked profiles, not overlaid for clarity.

3.6.3 Activity type

For the analysis of the activity type, we employed the activity classification from the MPA-JHU group, (Brinchmann et al., 2004; Kauffmann et al., 2003; Tremonti et al., 2004) described in Section 3.1.1. The distribution of the different galaxy activity populations as function of the filament axis distance is presented in Figure 3.17. The distributions are normalized for all filaments and then stacked together.

Although the statistical signal is poor, it seems that the three activity types (SF, AGN and LINER) have a suppression when the galaxies "approach" the filament axes; while the passive galaxies increase their number towards the filaments skeleton. This decrease in activity has its maximum at about 2 h_{70}^{-1} Mpc for AGNs and LINERS and about 0.5 h_{70}^{-1} Mpc for star forming galaxies.

3.6.4 Red sequence analysis

In order to study the color-magnitude and color-mass distributions in filaments, galaxies were separated into three environments: systems, filaments and the disperse component. Color-mass and color-magnitude diagrams are often used in galaxy clusters to separate the early type galaxy population from the late type

one. Moreover, the early type galaxies in clusters form a sequence on the color magnitude diagram which is called the “red sequence”. The best fit of the red sequence forms the so called color magnitude relation (CMR). The dispersion, color and slope of the red sequence in clusters have been used to study the evolution of their galaxies with redshifts up to z$\sim$ 1 (e.g. Bower et al., 1992; Aragon-Salamanca et al., 1993; Gladders et al., 1998; De Lucia et al., 2006; Trejo-Alonso et al., 2014, see Section 1.5.3). We used the SDSS photometry to produce color magnitude diagrams as shown in Figure 3.18. We carried out

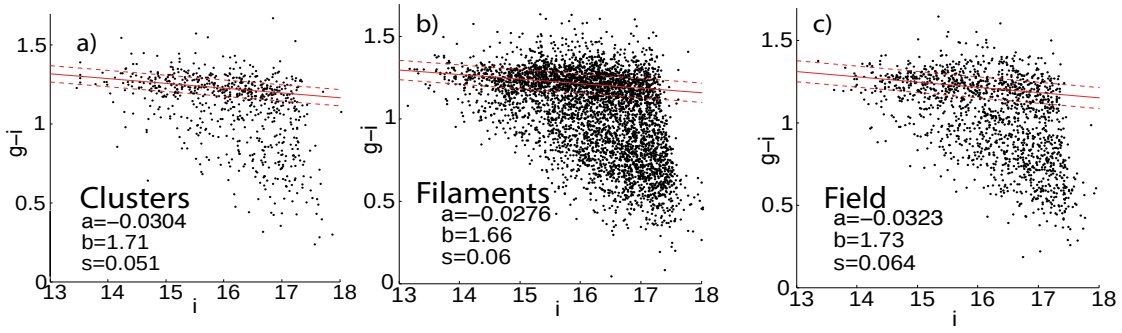

FIGURE 3.18: Color magnitude diagrams for the MSCC-310 supercluster galaxies as measured for galaxies in the clusters, filaments and field, from left to right.

this analysis using the apparent r-band magnitude m_r. Afterwards, galaxies were assigned to the red and blue clouds using a k-means classification, see Figure 3.19.

In Figure 3.18 we can observe the red sequence for system galaxies identified in our sample. Moreover, we also observe that galaxies associated to filaments, present also a red sequence while galaxies in the field and disperse component do not show the same trend. Moreover, one can observe, in Figure 3.19 clearly the red and blue clouds for clusters. One can observe a low number of galaxies in the transition region, also called green valley, located in between the red and blue clouds. An accurate classification for the blue, red and green valley requires an iterative processes. This will be calculated for future analyses in order to increase the precision in the red sequence fit. A similar trend is observed for filaments although we observe a larger number of galaxies in the transition region. As counterpart, in the field we observe a large number of galaxies in the green valley.

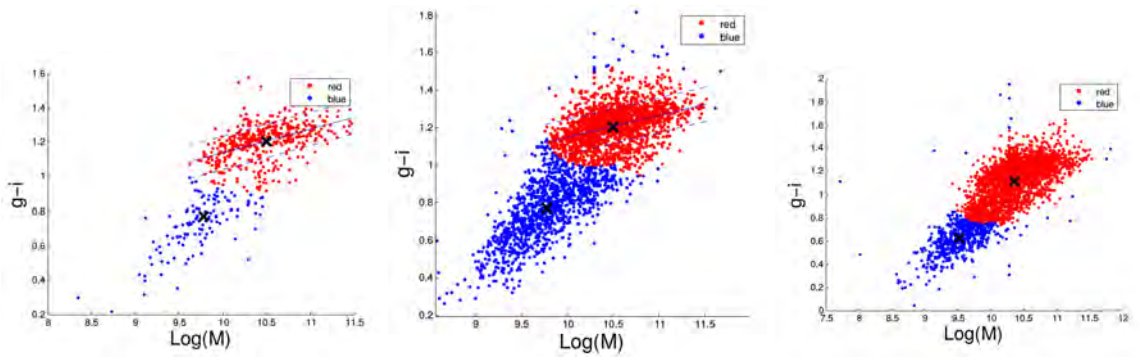

FIGURE 3.19: Color mass diagrams for the MSCC-310 supercluster for galaxies in clusters, filaments and field, from left to right. Solid and dashed lines represent the best fit for the red sequence and its dispersion respectively. Black crosses represent the centroid of the k-mean group.

3.7 Conclusions

In this Chapter we presented the analyses we carried out for the study of clusters, groups, bridges, filaments and tendrils of galaxies in the environment of superclusters of galaxies as well as the dispersed component and field galaxies. These analyses were obtained by using the implementation of the *Galaxy Systems Finding algorithm* GSyF and the *Galaxy Filament Finding algorithm* GFiF. These algorithms were applied to a sample of SDSS galaxies with spectroscopic redshifts in boxes containing 46 superclusters of galaxies selected from MSCC catalog in a redshift range from 0.02 to 0.15.

The comparison of systems detected by GSyF with previously reported system catalogs confirms the reliability of the GSyF system detection. Specifically, for rich clusters, the match rate ranges from 78% for the Abell catalog to about 100% for the BCG and C4 catalogs, while for group catalogs (T14 and MPSM) the match rate is of about 80%, as estimated for groups with richness above 5 galaxy members at redshifts Z¡0.11 and with a tolerance of 2 h_{70}^{-1} Mpc . However, we observed that for richness below $N_{gal} < 10$ the match rate increases to 90% but the difference between measured richness increases. This difference in galaxy richness is a consequence of the algorithm design, limits and search range. Moreover, the velocity dispersions and virial radii measured by GSyF algorithm are in good agreement with those measured by other system catalogs. The GSyF algorithm detected a total of 2705 systems in 45 supercluster volumes. Of

these, 159 systems with richness above 10 galaxies were not previously reported in the literature[1]. Moreover, the KDE analysis confirms that the overdensity regions correspond to clusters and rich groups of galaxies.

The GFiF algorithm detected a total of 143 filaments and 63 isolated bridges for 40 supercluster volumes. The detected filaments have skeleton lengths from 8.5 up to 130 h_{70}^{-1} Mpc while the isolated bridges have lengths between 5 and 15 h_{70}^{-1} Mpc . For most of the cases, the number density inside the filaments is measured between 5-15 times the background density. The measured radii for the filament skeleton ranges from 0.4 up to 4-5 h_{70}^{-1} Mpc .

Then, the filaments were characterized based on the properties of the galaxies that populate them. For the local number density, the conclusion is the following: i) The distribution of galaxies along the filament, within a 1 h_{70}^{-1} Mpc cylinder radius, shows that the typical density contrast of the filaments is about 10 times the background density. ii) The transversal local and VT number density profiles for pure filaments show that at distances up to 2 - 3 h_{70}^{-1} Mpc from the skeleton, the filaments are considered to have an overdensity above three times the mean box density. iii) That overdensity was used to measure the filament skeleton radius which goes from 2-3 h_{70}^{-1} Mpc. A comparison of our detected filaments with the Tempel et al. (2014) filaments has been carried out and one can observe that the filaments we define are on average four times larger, resulting in a significantly lower number of filaments. However, a comparison with isolated bridges and tendrils (a sub-product of our algorithm) shows a match of 80% with T14 filaments and a comparable filament length. A comparison of the GFiF detected filaments with those of Tempel et al. (2014) lead to the conclusion that GFiF detects as filaments the structures that present the highest probabilities to be real galaxy filaments.

The analyses regarding the galaxies stellar mass, morphology and activity type suggest a correlation with the distance of the galaxies from the filaments. These analyses lead to the following conclusions:

1. Near the filament skeleton and up to 3 h_{70}^{-1} Mpc the galaxy stellar masses are higher than the ones at larger distances from the filament. This result drives to two hypotheses: (a) that the mass of the galaxy is sensitive to the environment, or (b) the dynamical evolution brings massive galaxies into

[1] These systems are available in an electronic table.

the potential well of the filaments. This result confirms several analyses that suggest that stellar masses are sensitive to the environment (Alpaslan et al., 2015; Poudel et al., 2016; Chen et al., 2017b; Malavasi et al., 2017; Kraljic et al., 2018).

2. For the morphological types, one can observe a correlation with the filament distance. The fraction of early type galaxies has its maximum at the center of the filament and the ratio of early to late type galaxies remains above unity up to a distance of 1.5 h_{70}^{-1} Mpc . This result is compatible with a similar work carried out by Kuutma et al. (2017) for the SDSS.

3. Concerning the activity type, our analysis suggest that the number of active galaxies (AGN, SF and LINER) decreases when the galaxies are nearer than the filament skeleton and up to 2 h_{70}^{-1} Mpc , while passive galaxies appear to be more common near the filament skeleton. These analyses suggest that the filaments environment might impact the galaxies activity.

The GSyF and GFiF algorithms can be used to search for these LSS structures in different surveys, using spectroscopic or photometric redshifts. One of the future directions of this work is to incorporate to this analysis SDSS photometric redshifts in order test the improvements on the detections.

Chapter 4

Characterization of galaxy clusters using SZ

As mentioned in Chapter 1, galaxy clusters are the most massive gravity collapsed systems of the LSS. Moreover, X-ray observations had shown that the majority of their baryonic content is present in the form hot gas called intracluster medium (ICM). The inner part of these semi-spherical structures is considered to be virialized and in hydrostatic equilibrium. However, in the outskirts, the ICM is subject to shocks and turbulent motions as a consequence of LSS filament matter in-falling on the cluster potential well.

Thus, the study and characterization of the ICM in the cluster outskirts is of particular interest for the study of the evolution and dynamical state of the ICM. Recent analyses have shown that SZ observations are suitable for the characterization of the clusters ICM up to the outskirts. Following this approach we reconstructed pressure profiles for a sample of clusters using SZ observations.

In this Chapter we present and discuss the results from the recovery of pressure profiles for a sample of 31 clusters using only SZ data. Our analysis was carried out using a high resolution map, constructed from the *Planck* and ACT data. We describe the methods applied in order to reconstruct the gas pressure profiles from the SZ signal extraction as well as the profiles stacking method. We also present the validation procedure of the resulting pressure profiles.

4.1 The *Planck* and ACT sample and SZ data

4.1.1 The combined SZ map

In order to reconstruct the pressure profiles up to the outskirts for a sample of clusters, we employed the SZ maps (PACT y-map and its associated noise map) reconstructed from the linear combination of *Planck* (Planck Collaboration VIII, 2016) and ACT (Dünner et al., 2013) frequency maps. The resulting map integrates the best of the *Planck* frequency coverage and ACT spatial resolution which can be approximated by a 1.4 arcmin FWHM Gaussian (corresponding to the ACT spatial resolution). The PACT reconstruction was performed with an independent internal linear combination (ILC) method: MILCA (Modified Internal Linear Combination Algorithm; Hurier et al., 2013). This method allows to optimally combine frequency maps in a scale dependent way accounting for the intrinsic noise (instrumental and astrophysics) for the reconstruction of a targeted frequency dependent signal, the SZ effect in our case. A detailed description of the PACT map reconstruction and characterization can be found the main PACT paper (Sec. 3.2 of Aghanim et al., 2019b).

For validation purposes, we also made use of the *Planck* all-sky MILCA y-map (2015 data release Planck Collaboration et al., 2016c), which has an angular resolution of 10 arcmin FWHM (corresponding to the 100 GHz channel of the HFI instrument). We also used another (non-public) version of the MILCA y-map reconstructed at 7 arcmin FWHM. This map has been previously used for the extraction of the *Planck* SZ signal by the X-COP collaboration (Tchernin et al., 2016; Eckert et al., 2017; Ghirardini et al., 2017). Through out this Chapter, the integrated measured SZ fluxes will be expressed in arcmin^2 and the SZ luminosities in h_{70}^{-2}Mpc^2.

4.1.2 The cluster sample

The cluster samples used for this work were selected from galaxy cluster catalogs obtained from the SZ *Planck* (Planck Collaboration VIII, 2011; Planck Collaboration XXIX, 2014; Planck Collaboration XXVII, 2016, ESZ, PSZ1, PSZ2, respectively) and ACT (Hasselfield et al., 2013a; Hilton et al., 2018) surveys. For

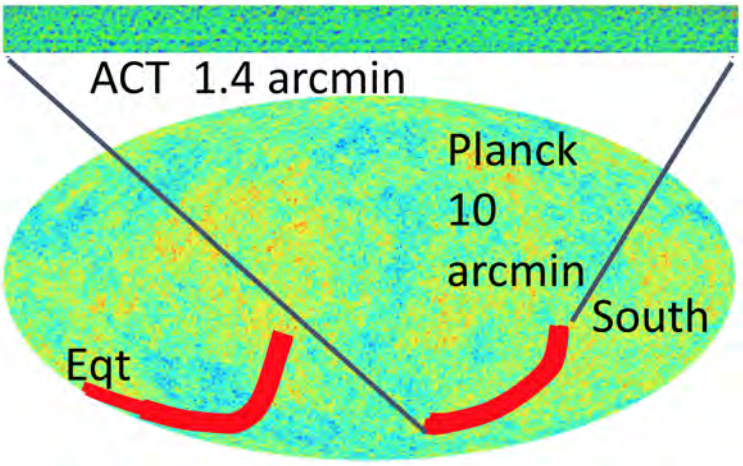

FIGURE 4.1: All-sky Comptonization y-map in Galactic coordinates (with the Galactic center in the middle, 10 arcmin resolution). The ACT equatorial and southern stripes are illustrated in red color.

instance, the PSZ2 catalog contains 1 203 confirmed clusters while a total of 119 clusters are detected within the ACT footprint. For this work we selected a sample of 34 clusters detected by both Plank and ACT. From these, 2 clusters were excluded, due to the partial coverage at the edges of ACT footprint and because its location was next to mask of the *Planck* point sources used to construct the PACT y-map. Thus, the final sample, listed in Table 4.1, comprehends 31 clusters of galaxies (hereafter referred to as PACT31) with 18 sources distributed in the equatorial strip and 13 in the southern one. By construction our sample is neither representative nor complete. The clusters range between 0.16 and 0.7 in redshift and 3.7×10^{14} to 1.3×10^{15} $M_{\odot}$ in mass. The mean signal-to-noise ratio (SNR) of the sample is 6.0 in the *Planck* catalogues and 6.26 in ACT's. The range of angular size is $2.5 < \theta_{500} < 7.9$ arcmin, with a mean value of 4.2 ± 1.1 arcmin.

Table 4.1 is organized as follows: column 1 lists the name of the cluster as we define it. Column 2, 3 and 4 list the ACT, PSZ1 and PSZ2 cluster name respectively; columns 5, 6, and 7 correspond to the cluster coordinates, right ascension, declination and redshift, respectively. Column 8 lists the cluster radius in arcmin units. Column 9 lists the cluster PSZ2 M_{500} and column 10 corresponds to the radius R_{500} in h_{70}^{-1}kpc. Columns 11,12 and 13 list the signal to noise, SNR, measured for the cluster in the different catalogs, ACT, PSZ1 and PSZ2, respectively.

Moreover, in order to assess, in Sec. 4.3, the robustness of our PACT pressure profiles (see Section 3.3.2) we derive pressure profiles of the hot intra-cluster gas for the sample of 62 massive local clusters (hereafter PLCK62) used by Planck

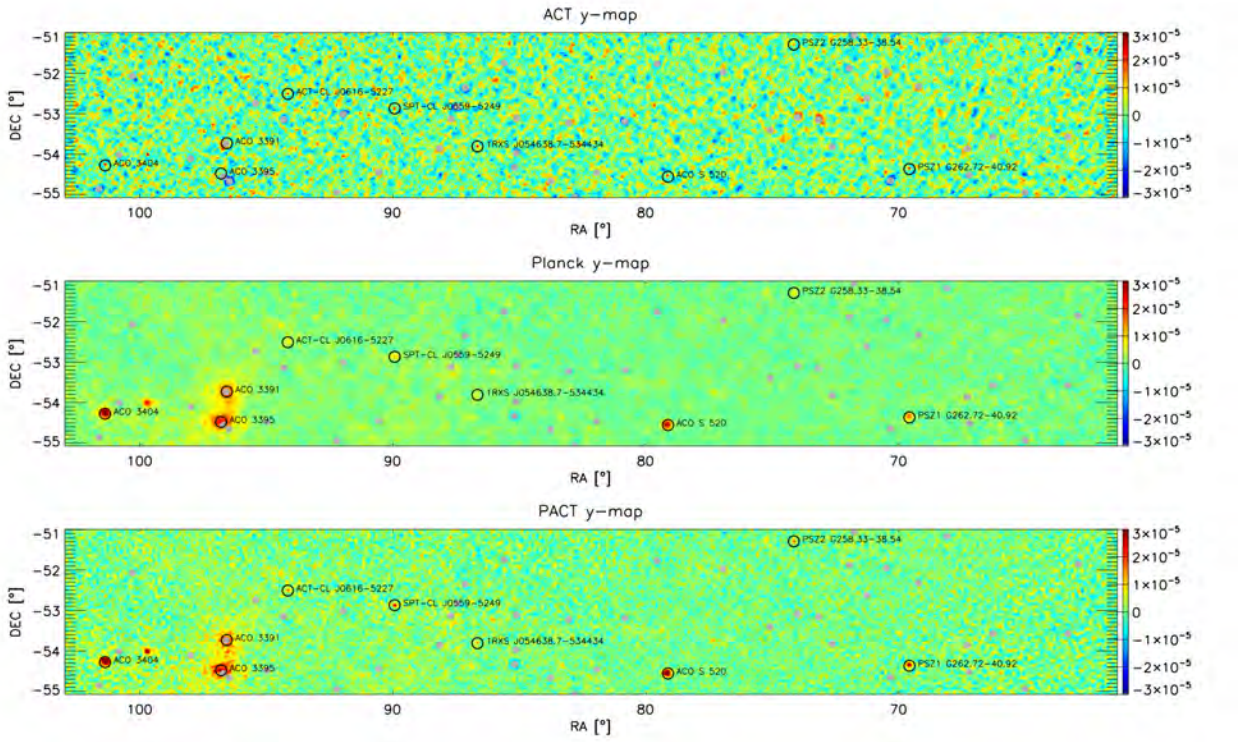

FIGURE 4.2: Maps of the Comptonization y parameter for the Southern ACT area. Top panel: 148GHz-ACT map at 1.5 arcmin resolution with scales larger than 7 arcmin filtered out. Middle panel: Planck y-map reconstructed with MILCA at 7 arcmin resolution. Lower panel: PACT y-map reconstructed with MILCA at 3 arcmin resolution. Circles indicate some known clusters in the selected area. The ACT point sources are masked in all three panels (grey areas). Figure extracted from Aghanim et al. (2019a).

Collaboration Int. V (2013). The *Planck* maps used to derive these pressure profiles are those from the full survey (i.e., 2015 data release Planck Collaboration VIII, 2016, See. Section 4.1.1). The redshifts of the PLCK62cluster range from 0.04 to 0.44. In the second *Planck* catalog, PSZ2, the PLCK62 SNR ranges from 7 to 49 (initially from 6 to 29 in the first sky survey). The covered mass interval is $2.4 \times 10^{14} < M_{500} < 1.97 \times 10^{15}$ $M_{\odot}$ for an angular size one of $3.7 < \theta_{500} < 22.8$ arcmin with a mean value of 9.8 ± 5.4 arcmin.

Figure 4.3 provides the distribution of the two samples in the $M_{500} - z$, $M_{500} - \theta_{500}$ and $\theta_{500} - z$ planes. For PACT31 the quantities are obtained from the PSZ1 and PSZ2 catalogs, while for PLCK62 we used those provided by Planck Collaboration Int. V (2013). For the last case, the masses, M_{500} are hydrostatic masses obtained from the XMM-*Newton* observations. Masses included in PSZ2 are derived from an $Y_{500} - M_{500}$ relation calibrated with *Planck* SZ integrated fluxes and XMM-*Newton* hydrostatic masses (Planck Collaboration XX, 2014; Planck

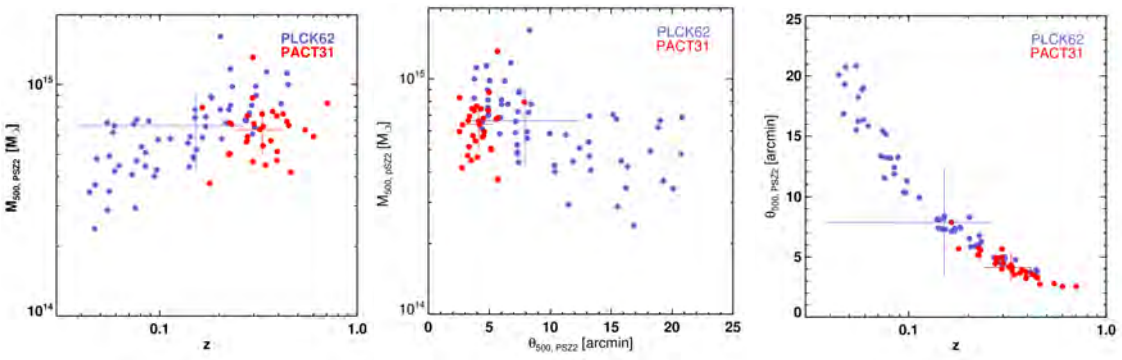

FIGURE 4.3: Distribution of the PACT31 (red dots) and PLCK62 (blue dots) in the $M_{500} - z$ *(left)*, $M_{500} - \theta_{500}$ *(middle)* and $T_{500} - z$ *(right)* planes. Crosses picture the median value and associated maximum absolute deviation of each sample. Masses, radii and temperature values are taken from the PSZ2 catalogue (Planck Collaboration et al., 2016a)

Collaboration XXIV, 2016). For PLCK62 we checked the consistency between the hydrostatic mass and scaling mass proxy estimation and derived a ratio of 0.98 ± 0.11 consistent with unity. In what follows we made use of the hydrostatic X-ray masses for PLCK62 sample, in order to fully reproduce the work by Planck Collaboration Int. V (2013).

4.2 Reconstruction of the gas pressure profile

In order to recover the pressure profiles from the SZ y-map, we strictly followed the method described and used by Planck Collaboration Int. V (2013). The general aspects of this method are described in the following paragraphs:

Let us consider a spherically symmetric galaxy cluster, in this case, the radial dependence of the pressure is assumed to follow a generalized Navarro-Frenk-White (GNFW) (Nagai et al., 2007) profile:

$$\mathbb{P}(x) = \frac{P(r)}{P_{500}} = \frac{P_0}{(c_{500}x)^\gamma \left[1 + (c_{500}x)^\alpha\right]^{(\beta-\gamma)/\alpha}}\,, \tag{4.1}$$

with $x = r/R_{500}$; P_{500} is the expected self-similar pressure at R_{500}, c_{500} is the concentration parameter defined at R_{500}; the slopes of the profile are defined by $\alpha,\ \beta$ and γ for the inner, intermediate and outer regions: ($x < 1/c_{500}$), ($x \sim 1/c_{500}$) and ($x > 1/c_{500}$) respectively.

TABLE 4.1: Properties of the PACT31 sample. It is composed of clusters which are present in the ACT (Hasselfield et al., 2013b) catalog and the Planck Catalog (Planck Collaboration et al., 2016a). The signal to noise in both detections is noted. The mass, redshift and radius values were taken from the Planck catalog.

Name		name		Ra	Dec	z	θ_{500}	M_{500}	R_{500}		SNR	
	ACT-CL	PSZ1	PSZ2	[Deg]	[Deg]		arcmin	M^{14} $M_\odot$	[kpc]	ACT	PSZ1	PSZ2
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
C00	J0008.1+020	G101.60-59.03	G101.55-59.03	2.041	2.02	0.360	3.68	5.72	1112.8	4.7	4.80	6.11
C01	J0045.2-015		G119.30-64.68	11.30	-1.88	0.545	2.80	6.37	1073.1	7.5	–	5.07
C03	J0104.8+000	G130.19-62.70	G130.21-62.60	16.21	0.04	0.277	4.53	5.70	1147.2	4.3	4.73	6.19
C05	J0152.7+010	G153.07-58.27	G153.00-58.26	28.17	1.00	0.230	5.08	5.04	1119.9	9.0	4.54	5.64
C08	J0239.8-013	G172.97-53.54	G172.98-53.55	39.97	-1.57	0.375	3.93	7.64	1218.7	8.8	6.15	7.61
C10	J0245.8-004		G173.90-51.89	41.46	-0.70	0.179	5.69	3.73	1032.1	4.1	–	4.56
C12	J0320.4+003	G181.40-44.76	G181.44-44.76	50.12	0.53	0.384	3.38	5.13	1063.7	4.9	5.08	4.63
C13	J0326.8-004	G184.23-44.26		51.70	-0.73	0.448	3.28	6.64	1130.7	9.1	4.78	–
C23	J2025.2+003	G044.57-20.46	G044.58-20.46	306.30	0.51	0.340	3.84	5.65	1116.7	6.4	5.01	6.58
C24	J2050.7+012	G048.91-25.55	G048.91-25.55	312.68	1.38	0.333	3.85	5.44	1105.6	7.4	5.00	5.56
C25	J2051.1+021	G049.83-25.22	G049.80-25.16	312.78	2.26	0.321	4.18	6.39	1171.9	5.2	6.02	7.59
C26	J2058.8+012	G050.07-27.29	G050.06-27.32	314.72	1.38	0.320	4.24	6.59	1184.8	8.3	5.51	6.78
C27	J2128.4+013	G054.94-33.37	G054.95-33.39	322.10	1.59	0.385	3.80	7.32	1196.7	7.3	5.88	7.30
C29	J2135.1-010	G053.42-36.25	G053.44-36.25	323.79	-1.03	0.330	4.31	7.45	1229.3	4.1	7.77	8.49
C30	J2135.2+012	G055.95-34.87	G055.95-34.89	323.81	1.42	0.231	5.57	6.73	1232.5	9.3	9.09	9.53
C31	J2156.1+012	G059.81-39.09	G059.81-39.09	329.04	1.38	0.224	5.17	4.99	1118.4	6.0	5.13	6.76
C32	J2327.4-020	G080.66-57.87		351.86	-2.07	0.705	2.55	8.30	1099.8	13.1	6.37	–
C33	J2337.6+001	G087.03-57.37	G087.03-57.37	354.41	0.26	0.275	4.96	7.32	1247.8	8.2	7.49	11.18
C06	J0217-5245		G276.75-59.82	34.29	-52.75	0.343	3.53	4.48	1032.6	4.1	–	5.44
C07	J0235-5121	G270.90-58.78	G270.93-58.78	38.96	-51.35	0.278	4.58	5.95	1162.8	6.2	6.02	8.95
C09	J0245-5302	G271.48-56.57	G271.53-56.57	41.38	-53.03	0.300	4.50	6.77	1204.3	9.1	7.75	10.41
C11	J0304-4921	G263.06-56.19	G263.03-56.19	46.06	-49.36	0.392	3.23	4.70	1029.8	3.9	4.64	4.65
C14	J0330-5227	G264.62-51.07	G264.60-51.07	52.72	-52.46	0.442	3.36	6.92	1149.2	6.1	7.83	10.82
C15	J0438-5419	G262.72-40.92	G262.73-40.92	69.57	-54.31	0.421	3.57	7.46	1187.6	8.0	10.8	12.70
C16	J0509-5341		G261.28-36.47	77.33	-53.70	0.461	2.75	4.17	963.72	4.8	–	5.06
C17	J0516-5430	G262.27-35.38	G262.27-35.38	79.12	-54.50	0.294	4.98	8.75	1314.9	4.6	18.6	22.87
C18	J0559-5249	G260.62-28.94	G260.63-28.94	89.92	-52.82	0.609	2.53	5.95	1023.1	5.1	4.99	7.75
C19	J0638-5358	G263.14-23.42	G263.14-23.41	99.69	-53.97	0.222	5.78	6.82	1242.3	10.0	10.8	12.75
C20	J0645-5413	G263.68-22.55	G263.68-22.55	101.37	-54.22	0.167	7.78	7.96	1333.5	7.1	17.4	21.67
C21	J0658-5557	G266.02-21.23	G266.04-21.25	104.62	-55.95	0.296	5.67	1.31	1502.6	11.5	20.4	28.38
C22	J0707-5522	G265.86-19.93	G265.86-19.93	06.80	-55.38	0.296	4.01	4.64	1063.2	3.3	4.87	5.77

Under this assumption, the Comptonization parameter y is a cylindric projection of the cluster' spherical pressure profile. Then, equation 1.59 can be expressed as:

$$y(r) = \frac{\sigma_{\rm T}}{m_{\rm e}c^2} \int_r^{R_{\rm b}} \frac{2P(r')r'dr'}{\sqrt{r'^2 - r^2}}, \tag{4.2}$$

with r is the sphere/cylinder radius. Since y is dimensionless, it can be expressed as $y(\theta)$ or $y(r)$. The integration over a frequency band accounts here for the effect of the weakly relativistic velocities of the electrons as a function of the gas temperature (e.g., Pointecouteau et al., 1998). This is however neglected in the y-map reconstruction as accounting for these effect would require a priori knowledge on the gas temperature distribution across the cluster). However, in real SZ observations, the observed Comptonization parameter profile is a geometric projection along the axis of the line of sight convolved with the instrument response (PSF), $f_{\rm PSF}$ as:

$$\tilde{y}(\theta) = f_{\rm PSF} \otimes y(\theta). \tag{4.3}$$

In order to reconstruct the pressure profile of the cluster, its y-profile needs to be

deconvolved from the effect of the PSF and deprojected. Therefore, we employ a *deconvolution + deprojection* method (Croston et al., 2006) but adapted to poorer statistics (lower number of bins per radial profile), as described in Section 4.2.3.

4.2.1 Reconstruction of the SZ profile

For the extraction of the SZ profiles we used the SZ y-maps and their associated error maps. We extracted patches of side length $20 \times \theta_{500}$ centered on the ACT cluster position from the y-map and error maps. Since the angular pixel size is different for all cluster patches, a standard pixel size was defined in constant units of θ_{500}. Then, all cluster patches have the same dimension, defined under the same scale units. However, scaling the pixel size implies a sampling of the PSF that could lead to an oversampling of the y-map pixel. We keep track and propagate this redundancy using an associated pixel index map.

The azimuthal y-profile extraction for each cluster was performed as follows:

1. Point sources (positive and negative) were masked manually from the y-map with a 2.5σ criterion clipping method.
2. The background offset was estimated for radii greater than $5 \times R_{500}$ and subtracted from the y-map. The use of the cluster surrounding area allows to take into account the astrophysical contamination in the cluster neighborhood as well as the instrumental and systematic noises.
3. Then, the y-profile is extracted on a radial grid. The y value of each radial bin is calculated as the mean y of the pixels falling on each annulus.
4. The power spectrum of the noise (background emitting sources, instrumental and systematic) was calculated in the region surrounding the cluster ($\theta > 5 \times R_{500}$).
5. The covariance matrix associated to the y-profile is calculated as:
   (a) 1 000 realizations of the noise patch are simulated, after which the same profile extraction is performed as for the y-patches.
   (b) the covariance matrix is constructed from the profiles of the noise simulations as $C = Pn^{\mathrm{T}}P_n$ where P_n is an n points $\times$ m matrix of simulated noise profiles.

4.2.2 Stacking the y profiles

Due to the expected-similarity property of the galaxy clusters, the profiles, once scaled, can be stacked together. This allows to increase the SNR by averaging out the background contaminants and noise. Stacking is a well-known method to obtain a more generalized representation of the clusters profiles. For this work we followed the stacking procedure described in Planck Collaboration Int. V (2013, Sec. 4.3.2, Eq. 14) to stack the y profiles.

Since all profiles are computed on the same grid (i.e. they are scaled in radius) the stacking consists in the calculation of the average value of all profiles at each grid radial point. Each y_i-profile and its respective covariance matrix C_i, are stacked as following:

$$\tilde{y} = \frac{1}{n}\sum_{i}^{n}\frac{y_i}{\Phi_i} \qquad \text{and} \qquad \widetilde{C} = \frac{1}{n^2}\sum_{i}^{n}\frac{C_i}{\Phi_i^2}, \tag{4.4}$$

with $\Phi = Y_{500}/R_{500}^2$, n the number of clusters in the sample, y_i the individual y profile and C_i the associated covariance matrix.

4.2.3 PSF correction and deprojection

In order to maximize the SNR, the profiles were rebinned to the largest possible radii. As mentioned before, we assumed our clusters to have a spherical symmetry. Then we applied a regularized PSF deconvolution and geometrical deprojection algorithm. Since this method was originally designed for X-ray surface brightness profiles, which have higher resolution than the SZ profiles, the methodology needed to be adapted to a lower number of bins per radial profile. The errors encoded in the y-profile covariance matrix are propagated via its randomization over $10\,000$ realizations after its Choleski decomposition (assuming correlated Gaussian noise). Each y-profile is deconvolved from the PSF and deprojected.

As noted in (Croston et al., 2006), this can be formalised as a projection matrix, $\mathcal{P}$, and a redistribution matrix, $\mathcal{B}$. The projected and PSF convolved Comptonization profile reads as:

$$y(\theta) = A_{\rm p2y}\, \mathcal{B}\, \mathcal{P}\, \mathbb{P}(x) \tag{4.5}$$

where the conversion from a universal pressure profile to an observed Comptonization parameter is expressed making use of Equation 4.1 and Equations 13 to 16 in Arnaud et al. (2010b):

$$A_{\rm p2y} = 1.65 \times 10^{-3} \left(\frac{M_{500}}{3 \times 10^{14} {\rm M}_{\odot}} \right)^{2/3+\alpha_p} \frac{\sigma_{\rm T}}{m_{\rm e} c^2} \frac{1}{D_{\rm A}} E(z)^{8/3} \tag{4.6}$$

where $\sigma_{\rm T}$ is the Thomson cross-section, $m_{\rm e}$ the mass of the electron, c the speed of light, $\alpha_{\rm P} = 0.12$. This conversion assumes a scaling relation between the pressure and the total mass, assuming self similarity.

In this process each point of the y-profile is weighted by the value of the flux dispersion in the corresponding radial bin. The covariance matrix of the pressure profile is derived from the combination of all realizations. Then the profile and covariance matrix are scaled in physical units.

4.2.4 Stacking pressure profiles

The pressure dependence with mass near a density contrasts $\delta = 500$ is almost constant with radius. Here we adopted the formulation derived by Arnaud et al. (2010b, see Equation 13) when calibrated on the REXCESS representative sample of X-ray local clusters (Böhringer et al., 2007). Then, the relation between mass and pressure can be approximated as in Arnaud et al. (2010a, Equation 9):

$$\frac{P(r)}{P_{500}} = \mathbb{P}(x) \left[\frac{M_{500}}{3 \times 10^{14} h_{70}^{-1} M_{\odot}} \right]^{0.12}. \tag{4.7}$$

From the above and Equation 4.4, the observed stacked profile can be formalised as:

$$\tilde{y}(\theta) = \frac{1}{n} \left[\sum_{i}^{n} A_{\rm p2y,i} \frac{\mathcal{B}_i\, \mathcal{P}_i}{\Phi_i} \right] \mathbb{P}(x). \tag{4.8}$$

Here we recall that r is the radius to the cluster center and $x = r/R_{500}$. Since y is dimensionless, it can be expressed as $y(\theta)$ or $y(r)$.

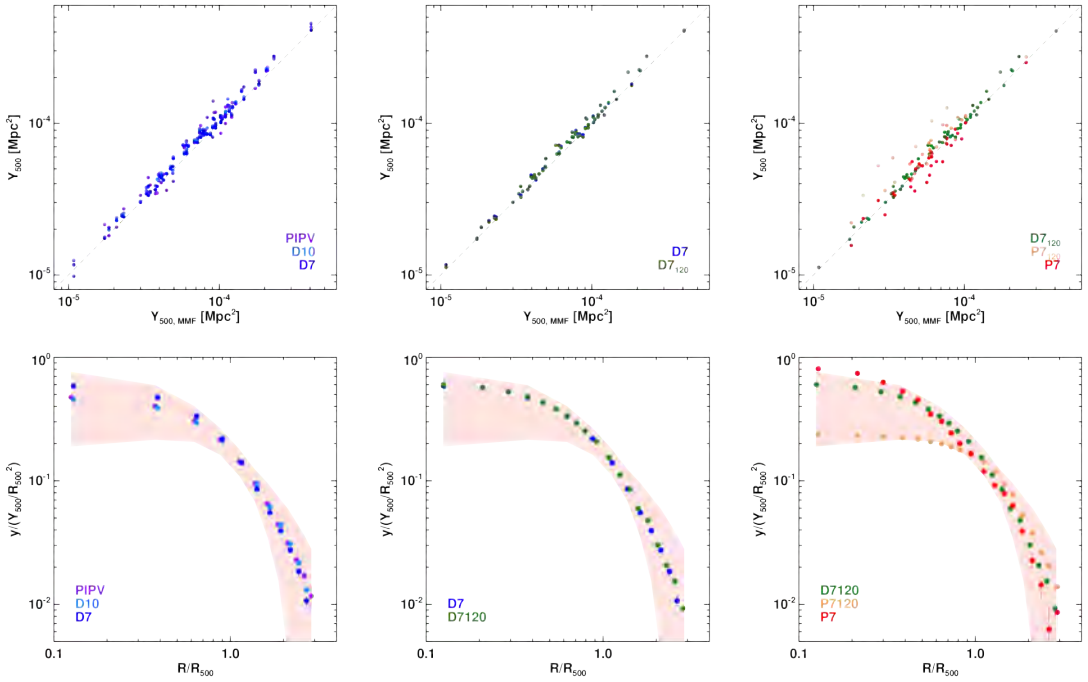

FIGURE 4.4: *(Top* One-to-one comparison of the integrated SZ flux, Y_{500}, using x-axis values computed from the PACT maps as presented in Aghanim et al. (2019b) as references. *(bottom)* Average stacked y profiles in units of Y_{500}/R_{500}^2 as function of radius in units of R_{500}. Columns present the various steps of validation: *(top left) Planck* data 2013 and 2015 and y-map resolution. *(bottom left)* Investigate of *Planck* dataset version and y map resolution. *(middle)* Check of the impact of the radial sampling factor δ_r. *(right)* Comparison of profiles derived from *Planck* and PACT maps and samples. We refer to the text (Sec. 4.3 and Table 4.2) for the definitions of the various configurations. For the profile plots the dispersion across the sample is shown as shaded area, and the error bars picture the square root of the diagonal of the covariance matrix and therefore bear a certain degree of correlation between points.

4.3 Validation of the PACT profiles

In what follows we describe the comparisons between the individual SZ y profiles and the integrated SZ flux Y_{500} over the PACT31 and PLCK62 samples for various setups. The radial profiles, $y(\theta)$, are derived from the y-map together with their correlation matrix as described in the previous section.

In the case of PACT clusters, the profile was centered at the ACT coordinates with R_{500} fixed, and was projected and convolved by the Planck PSF, while for the PACT31 clusters, the center was set at the XMM-*Newton* flux peak position. Then, the individual integrated SZ fluxes were computed for each y profile

TABLE 4.2: Definition of setups for the validation procedure

Name	Sample	y-map	FWHM	$\Delta r/R_{500}$
PIPV	PLCK62	PC-internal	10 arcmin	0.25
D10	PLCK62	DR2015	10 arcmin	0.25
D7	PLCK62	DR2015	7 arcmin	0.25
$D7_{120}$	PLCK62	DR2015	7 arcmin	0.08
$P7_{120}$	PACT31	DR2015	1.4 arcmin	0.08
P7	PACT31	PACT	1.4 arcmin	0.08

assuming a universal pressure profile (Arnaud et al., 2010b), $P(r)$. The Y_{500} optimal solution minimizing the chi-square and folding in the profile covariance matrix C reads as:

$$Y_{500} = \sigma^2_{Y_{500}} M^T C^{-1} P, \tag{4.9}$$

$$\sigma^2 = (M^T C^{-1} M)^{-1}. \tag{4.10}$$

For both the PACT31 and PLCK62 samples, we computed $y(\theta)$ and Y_{500} over the public and non-public *Planck* MILCA y-maps, with 10 and 7 arcmin FWHM, respectively.

We chose as reference fluxes the estimation of Y_{500} as derived from the multi-match filter (MMF) procedure described and used in the main PACT paper (Aghanim et al., 2019b). These fluxes were extracted with the MMF positioned at the ACTcluster coordinates and with a filter size fixed to θ_{500} for each source. For the profile comparison, we have cross-checked the stacked profile over the whole two samples, adopting as reference the y stacked profiles (and its dispersion envelope) derived for the PLCK62 sample by Planck Collaboration Int. V (2013).

4.3.1 Extraction of profiles from Planck DR2015: from 10′ to 7′

We first compared the integrated SZ fluxes and profiles derived from the first single full sky *Planck* survey (Planck Collaboration Int. V, 2013) on PLCK62 with those derived for the second *Planck* public release, which account for more than five co-added all sky surveys (hereafter DR2015) for both the all sky y-map reconstructed with a 10 arcmin (D10) and 7 arcmin (D7) resolution FWHM. We used exactly the same setup adopted by Planck Collaboration Int. V (2013) for

the radial sampling, with a bin width $\delta_r = \Delta\theta/\theta_{500} = \Delta r/R_{500} = 0.25$. The correlation between successive radial bins for each profile is encoded in the correlation matrix.

The comparison of fluxes and profiles is shown on Fig. 4.4, left-column). We can observe that all three flux estimations are consistent with each other and with the PACT MMF one. The average ratio to the PACT MMF fluxes over the sample is 1.09 ± 0.13, 1.10 ± 0.07 and 1.07 ± 0.07 for PIPV, D10 and D7, respectively. The profiles for the three cases (see Figure 4.4, *bottom left*) are also fully consistent in shape (the more peaked central D7 profiles only reflects the smaller PSF which was used to convolve the actual y profile). This agreement demonstrates that, for profiles computed with a radial sampling of $\Delta r/R_{500} = 0.25$, there is no systematic difference between the first full sky survey and the full DR2015 *Planck* survey, and neither between the flux estimations from the 10 arcmin to 7 arcmin FWHM reconstructed MILCA y-map. We thereby adopted the latest D7 setup for further comparisons.

4.3.2 Extraction of profiles from Planck DR2015 7′: improving the sampling

We recomputed the profiles over the 7 arcmin FWHM y-map, with a sampling of $\Delta r/R_{500} = 0.08$ and calculated the resulting associated fluxes (i.e., D7$_{120}$ setup). The comparison of the D7 setup from the previous section with the D7$_{120}$ setup is depicted in Figure 4.4, middle column.

The value $\delta_r = 0.08$ was chosen in order to properly sample the PACT PSF accounting for the angular extension of our clusters (see Table 4.1). As we extracted our profiles over a regular grid in units of R_{500} out to a value of 10, the number of radial bin is fixed by the cluster in our sample with the largest angular extension, i.e., C32 with $\theta_{500} = 8.30$ arcmin. This led to a minimum of 112 bins with $10 \times \theta_{500}$, that was rounded up to 120 points. The PSF sampling is thus compliant with the Nyquist-Shannon criteria for all our objects, with the PSF oversampling rate increasing towards the cluster with smaller angular extension, and ranging from 2.1 to 6.6 with an average value of 4.2.

The profiles derived with this setup as well as their integrated fluxes are fully consistent with the MMF fluxes having an average ratio of 1.06 ± 0.07 D7$_{120}$. This

leads to the conclusion that no bias is introduced by further oversampling the PSF, as long as the increased bin-to-bin correlation is being properly encoded in the correlation matrix.

4.3.3 Samples and y-maps: from PLCK to PACT

Finally, we studied the PACT31 sample and computed the profiles and fluxes from the *Planck* 7 arcmin FWHM DR2015 map with a sampling factor of $\Delta r/R_{500} = 0.08$, i.e., P7$_{120}$ setup, and over the PACT maps using the same sampling factor, i.e., the P7 setup (see Table 4.2). The latter is the nominal setup for the results on the PACT map and sample presented hereafter in this paper. The comparison of fluxes and profiles (see Fig 4.4, right-column) allows us to assess that no bias is introduced due to the change from the PLCK62to the PACT31sample.

The differences between the D7$_{120}$ and P7$_{120}$ stacked profiles exhibits the intrinsic difference between the PACT31 and PLCK62 samples. As shown in Fig. 4.5, they are sampling different regions of the mass and redshift plane. While the range of mass is similar, though the PLCK62 sample covers a broader range in mass, the main difference lie in the redshift coverage of the two samples. As a consequence, the angular sizes of the PACT31 clusters are smaller, hence the increased dilution in the *Planck* beam explains the flatter y profile for the P7$_{120}$ setup. As a further consequence, the profile convolution by the beam re-distributes the power towards larger scales and explains the slightly shallower shape at larger radii.

Unsurprisingly the stacked y-profiles for the PACT sample as measured from the PACT map are more peaked at $r < R_{500}$ than the one obtained from the *Planck* data only, which in turn are shallower at radii larger than R$_{500}$. This is due to the difference in smoothing by the PSF, i.e., 1.4 arcmin versus 10 arcmin for PACT and *Planck* respectively. However the match in fluxes demonstrates that no bias is introduced in the total integrated SZ flux when switching to the PACT maps. The average ratios to the reference MMF values are 1.24 ± 0.27 and 0.95 ± 0.12 for P7$_{120}$ and P7 respectively. The PACT results are further discussed in the next section. Table 4.2 gathers the definition of the various setups used in the validation procedure.

4.4 PACT profiles

Following the validation procedure presented in the previous section, in what follows only PACT maps with the adopted configuration P7 as defined in Table 4.2 for the PACT31 sample are considered. For the following analyses, we employed the cluster mass value M_{500} as provided by the PSZ2 and PSZ1 (for clusters not detected by PSZ2) catalogs (Planck Collaboration XXVII, 2016; Planck Collaboration XXIX, 2014).

4.4.1 PACT31 y-profiles

Figure 4.5, *left*, presents the individual y-profiles for the whole PACT31 sample, along with their corresponding stacked average and median profiles. The two stacked profiles do not present any significant differences, since, we chose to use their average for further analyses. The *right* panel shows the individual and associated stacked profiles for both equatorial and southern strip subsamples, i.e. 18 and 13 clusters, respectively. Within the limit of our statistics, we did not find any significant difference between the two subsamples, and hereafter the PACT31 sample is considered as a whole.

To further consolidate our result, we followed the procedure of Planck Collaboration Int. V (2013) and stacked the individual y-maps across the sample. Each individual map, m_i, was rescaled by the factor Φ_i (see Equation. 4.4) and randomly rotated by 0, 90°, 180° or 270° before averaging. The stacked map was finally back normalized in units of the Comptonisation parameter by $\langle\Phi_i\rangle$. A null test map is built from $\sum(-1)^i m_i$. The RMS values of this null test map and of the stacked SZ map, outside $5 \times R_{500}$, are 8.9×10^{-6} and 4.7×10^{-6}, respectively, and are consistent. They are also compatible with the average value of the stacked error y map, 8.8×10^{-6}. The the stacked map is shown on Fig. 4.6.

4.4.2 Stacked pressure profiles

After verifying the reliability of the PACT31 y-profiles obtained from the PACT maps, we reconstructed the pressure profiles following the methodology presented in Planck Collaboration Int. V (2013) and recalled in Sec. 4.2.4). Then,

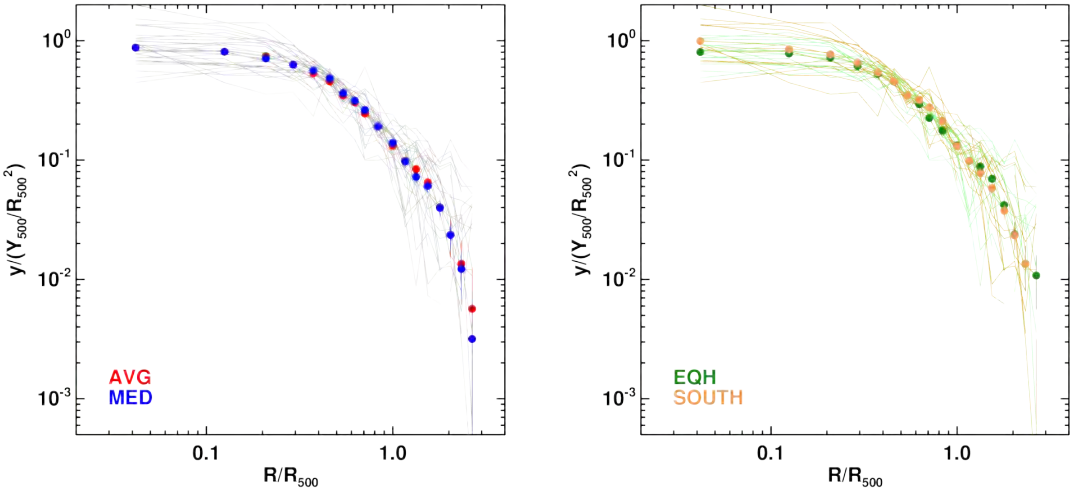

FIGURE 4.5: y profiles for the PACT31 sample. *(left)* Stacked average (red) and median(blue) profiles. *(right)* Average stacked profiles for the clusters in the Equatorial (green) and Southern (yellow) strips. The reported errors are correlated and correspond to the square root of the diagonal value of the covariance matrix for the y profile. Individual profiles are shown as solid lines.

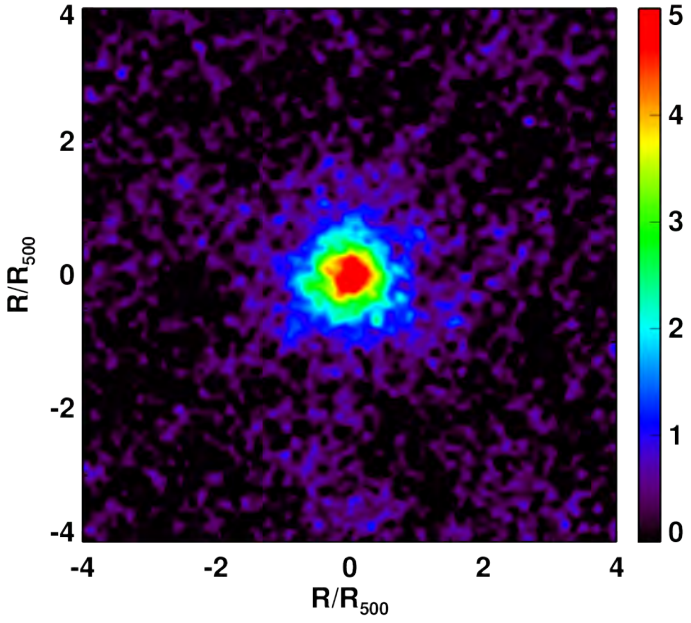

FIGURE 4.6: Stacked scaled y map over the PACT31 sample (side size corresponds to $8 \times R_{500}$). Individual maps were rescaled by Φ_i (see Sec. 4.2.2) before averaging, and then multiplied by $\langle \Phi_i \rangle$.

in a similar way to the SZ profiles, we stacked the individual PACT31 sample pressure profiles. The profiles were scaled in units of R_{500} and P_{500} according to

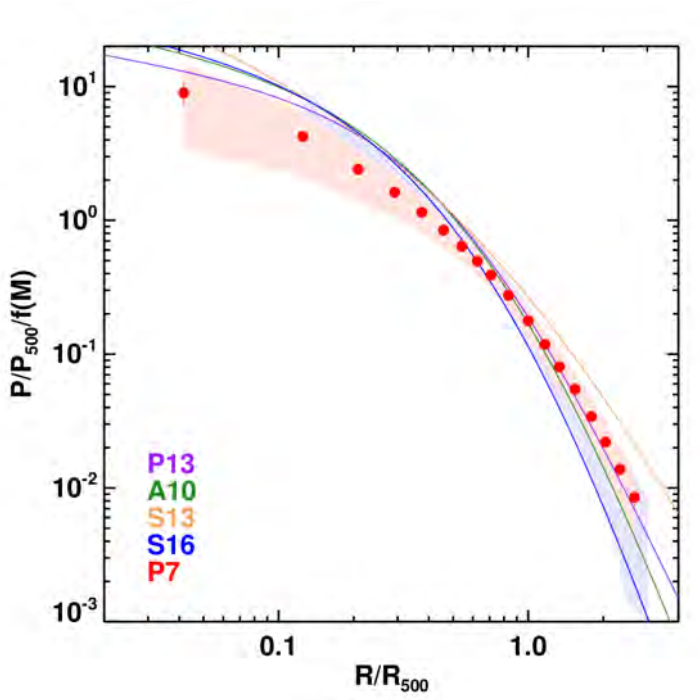

FIGURE 4.7: Stacked average pressure profiles over the PACT31 sample. The red solid line, labelled P7, shows the best gNFW pressure profile fitted over the data. The profiles by Arnaud et al. (2010a); Planck Collaboration Int. V (2013); Sayers et al. (2013, 2016) are over-plotted as green, purple, yellow and blue lines and labelled as A10, P13, S13 and S16, respectively. The red shaded area pictures the dispersion of the stacked profiles in each radial bin for the PACT31 (red) and the PLCK62 (purple) samples (as published by Planck Collaboration Int. V, 2013), respectively. The reported errors on the data points are correlated and correspond to the square root of the diagonal valued of the covariance matrix for the pressure profile.

equations 4.6 and 4.7. One can observe in Figure 4.7 that our PACT31 average pressure profile is fully compatible with the pressure profile for the PLCK62 sample as derived by the *Planck* Collaboration (Planck Collaboration Int. V, 2013).

In order to further quantify the compatibility with previous works we adopted a Monte Carlo Markov Chain to fit our pressure profile, see Table 4.3. Accounting for the difficulties to efficiently apply our algorithm of PSF deconvolution and deprojection (Croston et al., 2006) to our sample, Monte Carlo Markov Chain is used to directly adjust a gNFW pressure profile (see Equation 4.1) to our stacked y profile.

TABLE 4.3: Best fit parameter for the gNFW pressure profile.

	P_0	c_{500}	γ	α	β	$\tilde{\chi}^2_{\rm stat}$	$N_{\rm dof}$
A10	6.32	1.02	0.31	1.05	5.49	3.8	15
P13	6.41	1.81	0.31	1.33	4.13	0.9	13
S13	6.41	1.81	0.31	1.33	4.13	0.9	13
S16	6.41	1.81	0.31	1.33	4.13	0.9	13
P7	**6.41**	**1.81**	0.31	**1.33**	**4.13**	0.9	13

Note: A10, P13, S13 and S16 are the parameterizations provided by Arnaud et al. (2010b), Planck Collaboration et al. (2013a), Sayers et al. (2013, 2016), respectively, and tested against our average stacked profile. P7 corresponds to the best fit parametrization for this work. Values of best fit parameters are printed in boldface.

4.5 Conclusions and perspectives

We extracted y-profiles for 31 low SNR galaxy clusters using the new PACT SZ map. This map, product of the combination of the *Planck* and ACT data, takes the best of the ACT higher spatial resolution and the *Planck* large scale coverage.

From our stacked average y-profile, we detect SZ signal out to a radius of $\sim 2.5R_{500}$ and reach $r = 0.04 \times R_{500}$ towards the cluster center, i.e. the PACTmap allows us to have more detail near the cluster center. This can be compared to the SZ standalone and joint X-ray plus SZ derived pressure profiles for the PLCK62 sample by Planck Collaboration Int. V (2013), which reach 0.125 and $0.02 \times R_{500}$, respectively (see Figure 4.7).

Our stacked average y-profile is slightly less extended than the *Planck* profiles (Planck Collaboration Int. V, 2013), and we attribute this difference to the difference in sample and to their coverage of the $M - z$ plane.

Even though the PACT data does not reach the resolution of X-ray observations (i.e. does not arrive as close to the cluster center, we had demonstrated that the PACT data is suitable to investigate the ICM gas pressure from standalone SZ observation, allowing a better sampling towards the cluster centers. Therefore SZ observations are a good alternative for the characterization of clusters when X-ray observations are not available.

We carried out a validation using different sampling setups for the ACT 1.4 arcmin FWHM and *Planck* 10 and 7 arcmin FWHM DR2015 map, using the PLCK62 and PACT31 cluster samples. We demonstrated that the integrated SZ flux and

derived y-profiles for the D10 and D7 setups are compatible and consistent with the reference fluxes Y_{500}, as derived from the multi-match filter (MMF). Then, we showed that the oversampling of the PSF does not have effects in the extracted SZ signal. Finally, using the improved sampling of the PSF, we extracted y-profiles from the PACT map for the PACT31 sample. The comparisons of the integrated SZ flux Y_{500} measured from the derived y-profiles, are consistent with the reference fluxes Y_{500}, as derived from the MMF. Moreover, the stacked y-profile as well as the derived pressure profile, as obtained from the PACT map, are compatible (within the dispersion) with the previously published profiles by Planck Collaboration Int. V (2013).

As shown in Figure 4.3, the PLCK62 and PACT31 samples describe two different areas of the mass-redshift plane. As a consequence, both samples cannot be directly compared. Therefore, considering only clusters in the common redshift and mass interval of $0.15 < z < 0.45$ and $5 \times 10^{14} < M_{500} < 10 \times 10^{14}$ $M_{\odot}$, we obtained averaged values for the SZ flux (i.e., the integrated Comptonisation parameter Y_{500}) of 2.4×10^{-3} and 1.3×10^{-3} arcmin^2, from 25 and 20 clusters for the PLCK62 and PACT31 samples, respectively. We observe that the PACT31 clusters are, on average, fainter in their SZ signal than the PLCK62 clusters by a factor of two. As a consequence, the detection of the SZ signal for the PACT31 clusters is more difficult, especially out to larger radii.

Recalling that the PACT31 sample was built from detected clusters in both ACT and *Planck* catalogs, it still has a reasonable statistical size. However it is not representative of the cluster population. A proper assessment of the pressure distribution from SZ-only observation in clusters of galaxies would need to be built from a representative sample spanning the mass range sufficiently large to characterize possible variation with the halo mass. As done for X-ray studies with, e.g., the REXCESS sample (Böhringer et al., 2007), it would provide a solid reference in the local Universe, a mandatory step to provide the necessary data for evolution studies (that requires also, to sample the redshift range in a representative way).

Conclusions and perspectives

The objective of this thesis was to study the different structures that conform the cosmic web. In this context, we used the galaxy positions to study filaments, clusters and groups, and the gas component to extract pressure profiles of galaxy clusters. Since clusters and superclusters are structures that have formed under the effects of gravity, they are of particular interest to study the evolution of their galaxies and gas components.

Concerning the analysis using the optical galaxies, we selected SDSS galaxies with spectroscopic redshifts. This selection was based on the need of an accurate measurement on the line of sight velocity for the galaxy position, because our methodology is sensitive to this coordinate, as described in Chapter 2. We implemented a methodology that identifies various structures embedded in superclusters by using the optical galaxy positions.

For the detection and characterization of clusters and groups, we developed the Galaxy System Finding algorithm (GSyF). As described in Chapters 2 and 3, one of the main objectives of this algorithm is the correction of the Finger of God effects. However, in order to achieve this correction, an accurate detection of clusters and groups was necessary. Then, we created a new catalog of systems as a sub-product of this work. From the application of GSyF methods, we were able to detect a total of 2 705 groups and galaxy clusters, of which 159 were not previously reported. We also performed a comparison with other cluster catalogs, which leads us to conclude that GSyF is in good agreement with both, cluster (Abell, C4 and BCG) and group catalogs (MSPM and T11) available in the literature.

For the identification of the galaxy bridges and filaments we developed a new methodology implementing machine learning techniques, the Galaxy Filament Finding Algorithm (GFiF). Regarding this subject, the main result of our work

consists in a new catalog of filaments inside superclusters and the identification of galaxies as filament members. GFiF identified a total of 143 filaments which were characterized by their membership, radius and galaxy density.

This catalog provides the properties of bridges of galaxies linking three or more galaxy clusters. Also, as a sub-product, smaller bridged structures were also catalogued as isolated bridges and tendrils. The nature of these smaller structures can be of interest for future analyses of less dense LSS structures. Once identified, we characterized the detected structures by their length, radius and member galaxies. We also searched for correlations with the properties of the galaxies embedded in them. Our results show that the galaxy morphology and mass correlates with the distance to the nearest filament. This result confirms previous results by Chen et al. (2017b) and Kraljic et al. (2018).

Regarding the gas component of galaxy clusters, our analyses were focused on the gas component, studied through the Sundayev Zeldovich effect. We carried out an extraction of the gas pressure profiles for a sample of 31 galaxy clusters using a high resolution y-map. The PACT maps combine the large frequency range from the Planck satellite observation with the high resolution of the ACT map. The analysis of these pressure profiles using the PACT map allowed us to extract y-profiles for the inner part of the clusters, from 0.03 R_{500} up to 2 R_{500}.

We have shown that the resulted stacked profile is in good agreement with the previously reported pressure profiles from Planck Collaboration Int. V (2013). The differences we observed in the pressure profiles can be related with the fact that these clusters occupy a different range in redshift and have different angular sizes as compared to the PIP-V clusters. We have shown that the use of the PACT map allows the characterization of clusters with angular sizes up to 4 arcmin. Since the SZ signal does not depend on the redshift, only on the angular size, the use of high angular resolution maps open the possibility to carry out further analyses of clusters at higher redshifts using the SZ effect.

Further analyses of galaxy properties can be carried out in order to better constrain the possible environmental effects on galaxies. For instance, an analysis of galaxy metallicity and color index can help to further investigate the effect of environment in filament galaxies.

Other analyses are orientated towards the SDSS spectra using different diagnostic diagrams, in order to further investigate the effects of environment on the

galaxy activity. Moreover, another follow up to be done is to measure the position angle of galaxies belonging to filaments, with respect to the filament axis, e.g. Zhang et al. (2015) only studied the spin of spiral galaxies but not the ellipticals. This analysis can bring also information regarding the environment in which a galaxy resides. Following this approach, statistical analyses regarding the preferential orientation of radio galaxies with respect to the environment they reside can also be carried out.

Moreover, our new filament catalog aims to identify the best sky regions to carry out X-ray and millimeter observations. Therefore, future analyses orientated to the gas component of filaments can shed light on the evolution and dynamical state of this structures. One direction of this work is to better constrain the most concentrated regions in superclusters, i.e. their filament core. In this region galaxy systems are connected by bridges of high galaxy density, therefore, they are subject to interaction effects with the IGM in the filaments. This makes the study of filament cores of particular interest in order to understand the dynamical processes involved in such interactions. Then, a possible follow up is the characterization of the gas component for the systems belonging to the superclusters' cores. An analysis of this nature can allow to better understand some mechanisms in galaxy clusters such as their gas metallicity enrichment. Another direction of this work is to compare with the current analyses of gravitational lensing in the direction of superclusters' cores. Since gravitational lensing is an indirect tracer of the matter distribution, it can allow to better constrain the filaments path in order to carry out further observations of the gas component.

Conclusiones y perspectivas

El objetivo de esta tesis de Doctorado fue estudiar las diferentes estructuras que conforman la estructura a gran escala del Universo. Bajo este contexto, hemos utilizado las posiciones de las galaxias para estudiar filamentos, cúmulos y grupos así como la componente gaseosa de una muestra de cúmulos para extraer sus perfiles de presión. Los cúmulos y supercúmulos son estructuras de gran interés para el estudio de la evolución de las galaxias y gas que habitan en ellos dado que fueron formados en conjunto bajo el efecto de la gravedad.

Para el estudio de las galaxias, seleccionamos galaxias con redshift espectroscópico a partir de la base de datos del SDSS. Esta selección fue basada en la necesidad de una medida precisa de la velocidad de las galaxias en la línea de visión dado que nuestro método es sensible a esta coordenada, como hemos descrito en el Capítulo 2. La metodología implementada propone la identificación de estructuras dentro de supercúmulos utilizando la posición de las galaxias. Para la deteccion y caracterizacion de cúmulos y grupos desarrollamos el algoritmo “Galaxy System Finding (GSyF)”. Como hemos descrito en los Capítulos 2 y 3, uno de los objetivos principales de este algoritmo es la corrección del efecto de dedos de Dios. Sin embargo, para lograr dicha corrección, los cúmulos y grupos deben ser detectados con precisión. Como subproducto de dicha detección obtuvimos un nuevo catálogo de sistemas.

A partir de la aplicación de GSyF detectamos un total de 2 705 grupos y cúmulos de galaxias, de los cuales 159 no han sido reportados previamente. Hicimos una comparación con otros catalogos de cúmulos a partir de la cual podemos concluir que GSyF es compatible con catálogos de cúmulos (Abell, C4 and BCG) y de grupos (MSPM and T11) disponibles en la literatura.

Para la identificación de puentes de galaxias y filamentos, desarrollamos una nueva metodología que implementa técnicas de machine learning en nuestro

algoritmo Galaxy Filament Finding (GFiF). Como resultado de la aplicación de GFiF en nuestra muestra de galaxias obtuvimos un nuevo catálogo de filamentos dentro de superculos, así como la identificación de sus galaxias miembro. GFiF identificó un total de 143 filamentos caracterizados por sus miembros, radio y densidad de galaxias. Este catálogo provee las propiedades de puentes de galaxias conectando tres o más cúmulos de galaxias. Como subproducto de la aplicación de GFiF catalogamos puentes cortos y aislados llamados por nosotros tendrils. La naturaleza de estas estructuras aisladas puede ser de interés en futuros análisis de la componente menos densa de la estructura a gran escala del Universo.

Una vez identificadas, las estructuras fueron carcterizadas por su radio, longitud y galaxias miembro. Además, analizamos las propiedades de las galaxias miembro de estructuras para buscar correlaciones. A partir de este análisis observamos un relación entre la morfología de las galaxias y la distancia a la estructura más cercana. Nuestros resultados confirman el análisis previo de Chen et al. (2017b) y Kraljic et al. (2018).

Respecto a la componente gaseosa de los cúmulos de galaxias, nuestro estudio se enfocó en caracterizar el gas utilizando observaciones del efecto Sunyaev Zeldovich Para esto llevamos a cabo un estraccion del perfil de presión de gas para una muestra de 31 cúmulos utilizando un mapa de Comptonizacion de alta resolución, PACT. El mapa PACT combina el amplio rango de frecuencia del satélite Planck con la alta resolución del para ACT. Nuestro análisis de los perfiles de presión utilizando el mapa PACT nos permitio extraer perfiles y desde la parte interna de los cúmulos, 0.03 R_{500}, hasta radios de 2 R_{500}.

Como parte de nuestros resultados mostramos que el perfil promedio de presión es compatible con los resultados previamente obtenidos para perfiles de presión por Planck Collaboration Int. V (2013). Las pequeñas diferencias que observamos en los perfiles de presión pueden estar relacionadas con el hecho de que estos cúmulos cubren un rango diferente en tamaño angular y masa comparado con los cúmulos de la muestra de cúmulos PIP-V, utilizada por Planck Collaboration Int. V (2013). Hemos mostrado que el uso del mapa PACT permite la caracterización de los cúmulos con tamaños angulares de hasta 4 arcmin. Dado que la señal SZ no depende del redshift pero el tamaño angular si, la utilización de mapas de alta resolución nos da la posibilidad de caracterizar cúmulos a redshifts mayores a través del efecto SZ.

A partir de los resultados presentados en esta Tesis se pueden llevar a cabo análisis dedicados a la caracterización de los efectos ambientales sobre las galaxias. Por ejemplo, un análisis de la metalicidad e índice de color puede ser utilizado par estudiar el efecto del ambiente de los filamentos sobre la evolución de las galaxias.

Adicionalmente, se pueden llevar a cabo otros análisis utilizando diagramas de diagnóstico sobre espectros del SDSS con la finalidad de investigar más profundamente dichos efectos ambientales sobre la actividad de las galaxias. Otra propuesta es medir las orientaciones del ángulo de posición de las galaxias con respecto al eje del filamento al cual corresponden, por ejemplo, Zhang et al. (2015) estudió el spin de galaxias espirales pero no la orientación de galaxias elípticas. Análisis de esta naturaleza pueden dar información respecto al ambiente en el cual las galaxias se encuentran. Bajo este contexto, se puede llevar a cabo un análisis estadístico de la orientación del eje preferencial de radio galaxias con respecto al ambiente en el cual ellas residen.

Dado que nuestro nuevo catálogo de filamentos tiene como objetivo ayudar a identificar regiones del Universo con alta densidad de materia, puede ser utilizado como referencia para llevar a cabo observaciones en rayos X y en longitudes de onda milimétricas, favoreciendo de este modo estudios de la componente gaseosa en filamentos. Esto permitirá caracterizar la evolución y estado dinámico de estas estructuras. Otra propuesta a futuro es delimitar las regiones más concentradas en supercúmulos, es decir su núcleo. Es esta región las galaxias están conectadas por puentes de alta densidad de gas. Esto puede provocar efectos de interacción entre el medio en filamentos y las galaxias que residen en el. Por esta razón, estudiar estas regiones es de un interés particular para comprender los procesos dinámicos de estas interacciones. Otra propuesta es estudiar la componente gaseosa de los cumulos y sistemas que forma los puentes del centro del supercúmulo para comprender diversos mecanismos como por ejemplo el enriquecimiento del gas.

Finalmente, como propuesta de trabajo a futuro, proponemos comparar nuestro resultados con los resultados obtenidos por lentes gravitacionales en la dirección de los núcleos de supercúmulos. Dado que las lentes gravitacionales ponen en evidencia la distribución de materia oscura, mapas de lentes gravitacionales pueden ayudar a delimitar la distribución de los filamentos para llevar a cabo futuras observaciones de la componente gaseosa.

Conclusions et perspectives

L'objectif de cette thèse de Doctorat a été d'étudier les différentes structures qui composent la structure à grande échelle de l'Univers. Dans ce contexte, nous avons utilisé la position des galaxies d'une part pour l'étude des structures filamentaires et d'autre part pour l'étude de la composante gazeuse, afin de faire l'extraction des profils de pression pour les groupes et amas de galaxies. Les amas et superamas sont des structures de grande importance pour l'étude de l'évolution des galaxies et des gaz qui s'y trouvent, en raison de leur formation liée à des effets gravitationnelles.

Pour l'étude des galaxies, nous avons sélectionné un échantillon de galaxies avec redshift spectroscopique provenant de la base de données SDSS. Cette sélection a permis d'avoir une mesure précise de la troisième coordonnée de la vitesse des galaxies dans la ligne de vision. Cette sélection est nécessaire, car notre méthode est dépendante de cette troisième coordonnée, comme nous l'avons décrit dans le Chapitre 2 de cette thèse.

La méthodologie implémentée propose l'identification des structures à l'intérieur des superamas en utilisant la position des galaxies. Pour la détection et la caractérisation des amas et des groupe, nous avons développé l'algorithme Galaxy System Finding (GSyF). Comme décrit dans les Chapitre 2 et 3, l'un des objectifs de cet algorithme est la correction de l'effet fingers of God. Cependant, pour avoir cette correction, les amas et les groupes doivent être identifiés avec précision. A partir de cette détection, nous avons obtenu un nouveau catalogue de systèmes. Après l'application de l'algorithme GSyF nous avons détecté au total 2 705 groupes et amas de galaxies dont 159 n'ont pas été référencés dans la littérature. Ensuite, nous avons fait une comparaison avec d'autres catalogues

à partir desquels nous pouvons conclure que notre catalogue GSyF est en accord avec les catalogues d'amas (Abell, C4 and BCG) et des groupes (MSPM and T11) existant dans la littérature.

Pour l'identification des ponts des galaxies et des structures filamentaires, on a développé une nouvelle méthodologie en implémentant des techniques de machine learning dans notre algorithme, Galaxy Filament Finding (GFiF).

Le résultat de l'implémentation de l'algorithme GFiF à notre échantillon de galaxies, a permis d'obtenir un nouveau catalogue des structures filamentaires à l'intérieur des superamas, permettant ainsi l'identification des galaxies liées à ces filaments.

GFiF a identifié un total de 143 structures filamentaires qui ont été caractérisées par leur radius, leurs galaxies membres et leur densité de galaxies.

Ce catalogue nous donne les propriétés des ponts de galaxies qui font la jonction d'au moins trois amas de galaxies. En plus, nous avons produit un catalogue des ponts de moins de trois amas isolés que nous appelons ici tendrils. La nature de ces structures isolées peut avoir un intérêt pour les analyses futures de la composante moins dense de la structure à grande échelle de l'Univers.

Nous avons analysé les propriétés des galaxies membres des structures avec pour objectif de trouver des relations entre elles. A partir de cet analyse nous avons observé qu'il existe une relation entre la morphologie de la galaxie et sa distance à la structure à laquelle elle appartient. Nos résultats ont confirmé les analyses du Chen et al. (2017b) et Kraljic et al. (2018).

D'autre part, nous avons analysé la composante gazeuse des amas de galaxies en utilisant des observations de l'effet Sunyaev Zeldovich. Cet étude a été faite par l'extraction des profils de pression de gaz et de densité pour un échantillon de 31 amas à partir d'une carte de Comptonization de haute résolution (PACT). Les résultats que nous avons obtenus de cet extraction nous ont permis d'analyser les profils des amas de la partie centrale 0.03 R_{500} jusqu'à 2 R_{500}.

En plus, nous avons montré que le profil moyen de pression est en accord avec ceux obtenus par Planck Collaboration Int. V (2013). Les différences observées peuvent être liées à la différence entre la taille angulaire et la masse moyenne de notre échantillon (PACT) par rapport à l'échantillon d'amas utilisé par (Planck Collaboration Int. V, 2013, PIP-V). D'autre part, nos résultats ont

montré que l'utilisation de la carte de haute résolution PACT permettait de faire la caractérisation des amas qui ont un taille angulaire jusqu'à 4 arcmin. Vue que l'effet SZ ne garde pas la relation avec redshift mais la taille angulaire elle oui, l'utilisation des cartes de haute résolution SZ peuvent donner la possibilité de caractériser les amas à haute redshifts.

Les résultats obtenus pendant cette thèse peuvent être poursuivies par des étude dédiés à la caractérisation des effets environnementaux du milieu intergalactique chaud sur les galaxies. Par exemple, une analyse de la métallicité et l'indice de couleur des galaxies qui habitent dans structures filamentaires peut donner des informations sur l'environnement et son influence sur l'évolution des galaxies.

De plus, une analyse de l'effet environnemental sur l'activité des galaxies peut être appliquée en utilisant le spectre SDSS des galaxies. Aussi, une analyse de l'orientation de l'angle de position des galaxies par rapport à l'axe du filament à laquelle elles appartiennent. Par exemple Zhang et al. (2015) ont étudié l'orientation de la rotation des galaxies spirales. Un étude statistique de l'orientation des galaxies peut donner des informations sur l'environnement dans lequel la galaxie se trouve. Vue que notre nouveau catalogue de structure filamentaires a pour objectif d'identifier des zones à haute densité de matière dans l'Univers, il peut être utilisé comme référence pour effectuer des observations futures au rayons X et micro ondes. Ce type d'étude donne la possibilité d'étudier en profondeur la composant gazeuse des structures filamentaires et mieux comprendre leur dynamique.

Une autre possibilité d'étude est d'identifier les filaments les plus denses au centre du noyau des superamas. Dans cette région les galaxies sont liées par des pont de haute densité gazeuse. Cela peut entraîner des effets d'interaction avec les galaxies qui y sont localisées. Dans le même contexte, l'étude de l'interaction entre les groupes et amas avec les filaments qui les lient peut donner des informations sur divers effets comme l'enrichissement du gaz dans ce milieu. Finalement, une comparaison entre les superamas que nous avons étudié et les résultats de lentilles gravitationnelles peut nous informer sur la distribution de matière dans le milieu.

Appendix A

Algorithms

Algorithm 1: Mock maps generator

Input: Supercluster volume and number of galaxies.
Output: Simulated distribution of galaxies in clusters, groups and field of the supercluster.

1. The simulated volume is filled with N_{synth} synthetic systems of galaxies with $N_{elem} = 10 - 200$. The number of systems in the volume is set using the power function $\log_{10}[N_{synth}(N_{elem})][h_{70}^3\ Mpc^{-3}] = m \log_{10}(N_{elem})$ (multiplicity function) with a slope m set according to:

$$m = \begin{cases} -2.48 & \text{if } z < 0.08 \\ -2.72 & \text{otherwise} \end{cases}$$

2. Set mock systems center position randomly in the volume.

3. Calculate synthetic system proxies (Mass M_{500}, radius R_{vir} and velocity dispersion σ_v).

$$\begin{cases} \log_{10} M_{500} = 1.03 \log_{10}(N_{elem} - 2.63) + 0.34 \\ \log_{10} R_{vir} = 1.05\ \log_{10}(M_{500} - 0.35) + 8.48 \\ \log_{10} \sigma_v^3 = 2.33\ \log_{10}(M_{500} - 0.21) + 3.04 \end{cases}$$

4. Fill the systems with galaxies (elements) following a normal distribution $N_{elem}(\mu_i, R_{vir})$ with $\mu_i = \{\alpha_i, \delta_i\}$.

5. Add FoG effects to galaxy systems by adding velocity dispersion $N_{elem}(z_i, \sigma_v)$.

6. Add random galaxies to the box volume following the ratio: 60% galaxies are located in field, and 40% in systems.

Algorithm 2: Galaxy systems finding algorithm (GSyF)

Input: Supercluster's galaxy positions $\alpha,\ \delta$ and z.
Output: Galaxy systems, membership and member galaxy positions corrected for FoG effect.

(i) Compute local surface density of galaxies using VT.

(ii) Construct 1,000 randomizations of galaxy positions to calculate the baseline surface density, d_{bas}.

(iii) Calculate the density contrast of galaxies, δ_i, with respect to d_{bas}.

(iv) Select galaxies with density contrast above the reference value $(1+g)\ d_{bas}$.

(v) Group galaxies by their position ($\alpha,\ \delta$ and $1,000 \times z$), using the HC algorithm.

(vi) Filter resulting groups by number of members $N_i \geq 3$.
Virial refinement and FoG correction

(a) *Select galaxies in a cylinder of radius $R_a = 1\ h_{70}^{-1}$ Mpc projected in the sky, centered on the brightest galaxy close to the group's centroid, and within $\Delta cz \pm$ 3000 km s^{-1} along the line-of-sight, centered on the group mean z.*

(b) *Calculate virial radius using Eq. 2.9 using the bi-weighted velocity dispersion, harmonic radius and redshift.*

(c) *Update R_a by R_{vir}, mean z by v_{LOS} and Δz by $3 \times \sigma_v$.*

(d) Compute iteratively virial radius for each group until $R_a \rightarrow R_{vir}$.

(e) Calculate N_{mem} and M_{vir}.

(f) Correct co-moving distance of the member galaxies for FoG effect by re-scaling the cylinder length to the R_{vir} size.

(g) Calculate galaxy corrected rectangular coordinates.

Appendix B

Filament properties

TABLE B.1: Main properties of the filaments extracted through GFiF.

Fil. Nr. (1)	N$_{sys}$ systems (2)	N_{gfil} gals. (3)	redshift [mean, (4)	min, (5)	max] (6)	mean density [h_{70}^{3} Mpc^{-3}] (7)	R_{fil} [h_{70}^{-1} Mpc] (8)	N$_{nod}$ filament (9)	skeleton (10)	L_{fil} [h_{70}^{-1} Mpc] (11)
MSCC-55-F1	4	297	0.0619	0.0529	0.0706	0.1570	4.31	3	2	10.5
MSCC-55-F2	3	211	0.0585	0.0527	0.0680	0.1911	0.60	2	2	9.7
MSCC-72-F1	5	370	0.0790	0.0722	0.0868	0.5636	3.57	6	6	28.4
MSCC-72-F2	4	354	0.0801	0.0718	0.0866	0.3688	3.46	5	5	24.6
MSCC-72-F3	5	299	0.0777	0.0717	0.0839	0.3043	3.89	6	5	19.5
MSCC-72-F4	3	238	0.0850	0.0787	0.0917	0.2012	3.49	4	4	26.5
MSCC-175-F1	6	429	0.0912	0.0841	0.0996	0.1541	3.08	8	6	41.8
MSCC-175-F2	3	298	0.0978	0.0885	0.1032	0.2260	3.19	3	3	35.0
MSCC-175-F3	3	221	0.0937	0.0883	0.1001	0.1714	1.79	5	4	47.4
MSCC-175-F4	3	195	0.0923	0.0867	0.0995	0.2184	2.86	3	3	37.5
MSCC-184-F1	5	768	0.1071	0.0994	0.1171	0.0488	2.88	8	5	49.3
MSCC-184-F2	3	332	0.0980	0.0911	0.1029	0.0838	3.09	3	3	39.5
MSCC-211-F1	3	555	0.1205	0.1100	0.1297	0.0465	4.46	13	8	27.9
MSCC-219-F1	10	608	0.1125	0.1063	0.1207	0.1169	2.96	17	6	81.8
MSCC-219-F2	3	170	0.1235	0.1183	0.1287	0.0347	0.79	3	3	40.2
MSCC-222-F1	3	408	0.1422	0.1311	0.1522	0.0418	1.87	14	8	70.0
MSCC-222-F2	5	377	0.1349	0.1238	0.1460	0.0461	0.70	11	8	78.8
MSCC-223-F1	3	342	0.1367	0.1286	0.1472	0.0245	0.86	12	8	48.7
MSCC-229-F1	3	204	0.1445	0.1361	0.1514	0.0389	1.96	9	6	63.1
MSCC-236-F1	5	545	0.0324	0.0228	0.0415	0.3382	1.49	16	9	39.3
MSCC-236-F2	4	492	0.0411	0.0361	0.0461	0.2409	2.65	7	6	38.5
MSCC-236-F3	5	452	0.0331	0.0274	0.0399	0.4889	2.62	8	5	20.0
MSCC-236-F4	3	358	0.0296	0.0255	0.0364	0.3971	2.24	5	4	19.6
MSCC-236-F5	3	337	0.0354	0.0286	0.0422	0.4448	2.47	5	5	20.8
MSCC-236-F6	3	247	0.0330	0.0276	0.0373	2.3032	2.66	2	2	16.8
MSCC-236-F7	3	177	0.0333	0.0297	0.0391	0.4359	1.53[1]	4	3	16.4
MSCC-238-F1	5	673	0.1190	0.1075	0.1309	0.0558	2.95	11	7	85.3
MSCC-238-F2	6	645	0.0915	0.0802	0.1010	0.0928	2.98	14	10	76.1
MSCC-238-F3	3	553	0.1054	0.0979	0.1115	0.0782	2.64	11	8	99.0
MSCC-238-F4	3	176	0.0957	0.0896	0.1024	0.1467	0.71	3	3	28.6
MSCC-238-F5	3	162	0.1016	0.0934	0.1099	0.0285	0.67	3	2	32.5
MSCC-238-F6	3	128	0.1096	0.1040	0.1156	0.0398	0.88	4	3	30.4
MSCC-248-F1	3	476	0.1256	0.1161	0.1355	0.0895	3.35	10	7	82.1
MSCC-266-F1	5	452	0.1282	0.1188	0.1344	0.0366	2.94	9	8	75.8
MSCC-272-F1	6	432	0.0752	0.0694	0.0808	0.8405	2.93	6	4	17.5
MSCC-272-F2	3	200	0.0757	0.0714	0.0814	0.1245	2.66	4	3	10.3
MSCC-277-F1	10	798	0.1124	0.1031	0.1208	0.1268	2.99	21	11	89.3
MSCC-277-F2	5	650	0.1053	0.0956	0.1135	0.0805	2.93	16	10	69.0
MSCC-278-F1	7	1262	0.0328	0.0251	0.0398	0.9871	2.41	36	14	56.3
MSCC-278-F2	3	443	0.0322	0.0273	0.0361	0.4963	1.70	24	13	20.8
MSCC-278-F3	4	352	0.0319	0.0266	0.0386	0.3810	1.18	14	7	24.4
MSCC-278-F4	8	301	0.0254	0.0221	0.0298	1.0335	1.86	27	10	36.9
MSCC-278-F5	3	185	0.0348	0.0297	0.0387	0.4401	1.76	7	6	11.1
MSCC-283-F1	4	532	0.1339	0.1245	0.1470	0.0701	3.96	7	4	47.7
MSCC-283-F2	5	347	0.1364	0.1284	0.1465	0.0655	3.26	6	5	49.1
MSCC-283-F3	3	304	0.1357	0.1296	0.1496	0.0397	0.96	2	2	24.2
MSCC-295-F1	7	1233	0.0230	0.0160	0.0284	2.1243	1.89	37	28	44.6
MSCC-295-F2	7	1060	0.0228	0.0158	0.0299	2.2694	2.37	34	12	18.2

Continued on next page

Table B.1 – continued from previous page

Fil.	N_{sys}	N_{gfil}		redshift		mean density	R_{fil}	N_{nod}		L_{fil}
Nr.	systems	gals.	[mean,	min,	max]	[h_{70}^3 Mpc^{-3}]	[h_{70}^{-1} Mpc]	filament	skeleton	[h_{70}^{-1} Mpc]
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
MSCC-295-F3	4	626	0.0232	0.0189	0.0283	0.5824	1.36	44	17	26.5
MSCC-295-F4	4	209	0.0218	0.0188	0.0255	0.6254	1.83[1]	10	9	30.8
MSCC-310-F1	10	714	0.0609	0.0518	0.0689	0.4180	2.81	18	11	61.6
MSCC-310-F2	8	696	0.0502	0.0443	0.0588	0.5210	2.84	22	11	51.8
MSCC-310-F3	7	506	0.0481	0.0427	0.0528	0.3486	2.50	19	10	49.0
MSCC-310-F4	8	490	0.0656	0.0585	0.0710	0.4263	2.46	14	10	59.0
MSCC-310-F5	7	461	0.0700	0.0642	0.0774	0.5303	2.60	13	7	47.6
MSCC-310-F6	6	409	0.0725	0.0651	0.0791	0.2986	1.97	12	9	39.3
MSCC-310-F7	4	385	0.0551	0.0479	0.0619	0.4062	2.76	7	5	20.7
MSCC-310-F8	4	254	0.0546	0.0485	0.0617	0.2118	2.22	9	7	33.7
MSCC-310-F9	4	155	0.0464	0.0437	0.0528	0.1464	1.13	9	6	17.9
MSCC-311-F1	13	1140	0.0813	0.0753	0.0892	0.2812	3.54	14	9	74.5
MSCC-311-F2	5	372	0.0847	0.0769	0.0937	0.3235	3.69	4	3	23.5
MSCC-311-F3	3	322	0.0888	0.0821	0.0937	0.1900	4.32	2	2	15.3
MSCC-311-F4	3	302	0.0826	0.0746	0.0901	0.2447	3.58	2	2	19.5
MSCC-314-F1	4	100	0.0819	0.0769	0.0893	0.1201	1.88	4	4	29.6
MSCC-314-F2	3	83	0.0777	0.0715	0.0844	0.1610	2.78	2	2	12.3
MSCC-317-F1	4	187	0.1331	0.1224	0.1417	0.0282	1.73[1]	15	10	71.7
MSCC-317-F2	4	176	0.1187	0.1102	0.1264	0.0528	0.69	10	7	75.5
MSCC-323-F1	7	802	0.1380	0.1262	0.1540	0.0436	3.73	14	11	129.3
MSCC-323-F2	3	441	0.1383	0.1309	0.1535	0.0261	1.66	5	4	29.4
MSCC-333-F1	3	219	0.0760	0.0723	0.0833	0.1205	2.36	5	4	29.8
MSCC-333-F2	4	173	0.0803	0.0752	0.0880	0.1519	2.83	5	4	24.5
MSCC-333-F3	4	126	0.0794	0.0729	0.0835	0.2052	2.46	5	5	34.6
MSCC-335-F1	10	745	0.0779	0.0705	0.0843	0.1260	2.60	21	11	105.0
MSCC-335-F2	4	219	0.0665	0.0613	0.0731	0.0650	0.75	9	8	83.8
MSCC-335-F3	4	195	0.0762	0.0712	0.0814	0.1113	2.45	6	4	45.9
MSCC-343-F1	4	291	0.0822	0.0786	0.0874	0.1394	2.51	7	6	28.7
MSCC-343-F2	3	202	0.0798	0.0733	0.0876	0.2638	2.50	3	3	19.3
MSCC-343-F3	3	194	0.0824	0.0738	0.0876	0.1991	2.65	6	5	15.6
MSCC-360-F1	4	633	0.1050	0.0967	0.1159	0.1367	2.95	9	6	60.8
MSCC-360-F2	3	348	0.1074	0.1004	0.1152	0.0870	1.52	6	4	49.3
MSCC-360-F3	3	203	0.1033	0.0973	0.1093	0.0416	1.46	3	2	29.9
MSCC-386-F1	7	422	0.0734	0.0657	0.0805	0.4163	2.91	18	10	34.0
MSCC-386-F2	4	157	0.0708	0.0636	0.0765	0.1969	1.92	10	8	39.6
MSCC-386-F3	5	115	0.0627	0.0591	0.0705	0.1760	1.56	13	10	43.2
MSCC-386-F4	4	44	0.0617	0.0594	0.0664	0.1703	0.92[1]	7	6	21.3
MSCC-407-F1	5	327	0.1388	0.1254	0.1468	0.0177	3.61	10	6	59.2
MSCC-414-F1	4	556	0.0626	0.0537	0.0691	0.6269	2.51	19	9	38.1
MSCC-414-F2	6	294	0.0628	0.0548	0.0667	0.3793	2.54	12	7	34.0
MSCC-414-F3	6	262	0.0750	0.0688	0.0809	0.5198	2.53	12	7	27.1
MSCC-414-F4	5	227	0.0641	0.0585	0.0675	0.3705	2.52	10	5	13.5
MSCC-414-F5	4	216	0.0748	0.0684	0.0810	0.4251	2.53	10	8	22.1
MSCC-414-F6	3	173	0.0657	0.0613	0.0701	0.3368	2.01	7	6	22.9
MSCC-414-F7	3	159	0.0616	0.0576	0.0647	0.2329	1.97	10	5	14.1
MSCC-414-F8	3	157	0.0657	0.0603	0.0722	0.1197	1.82	7	6	29.9
MSCC-414-F9	4	143	0.0663	0.0594	0.0726	0.3001	2.50	5	5	22.9
MSCC-414-F10	4	133	0.0546	0.0518	0.0590	0.4163	1.99	6	6	29.9
MSCC-414-F11	5	127	0.0616	0.0589	0.0657	0.1262	1.64	14	8	35.1
MSCC-414-F12	3	101	0.0762	0.0696	0.0801	0.1126	0.66	4	4	9.8

Continued on next page

Table B.1 – continued from previous page

Fil.	N$_{sys}$	N_{gfil}		redshift		mean density	R_{fil}	N$_{nod}$		L_{fil}
Nr.	systems	gals.	[mean,	min,	max]	[h_{70}^3 Mpc^{-3}]	[h_{70}^{-1} Mpc]	filament	skeleton	[h_{70}^{-1} Mpc]
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
MSCC-414-F13	4	89	0.0640	0.0614	0.0674	0.1692	1.94	4	4	19.2
MSCC-414-F14	4	82	0.0727	0.0686	0.0758	0.4275	1.78	7	6	23.6
MSCC-414-F15	3	50	0.0607	0.0563	0.0640	0.1070	1.80	3	3	9.0
MSCC-419-F1	3	226	0.1139	0.1084	0.1210	0.1849	4.03	2	2	12.8
MSCC-419-F2	3	137	0.1099	0.1044	0.1182	0.1816	3.20	2	2	18.1
MSCC-419-F3	4	86	0.1127	0.1080	0.1191	0.0693	1.79	5	4	25.2
MSCC-422-F1	3	398	0.1424	0.1321	0.1528	0.0291	0.95	8	6	45.0
MSCC-430-F1	4	361	0.0975	0.0880	0.1066	0.1391	2.78	7	4	39.2
MSCC-430-F2	4	215	0.0937	0.0867	0.0999	0.1684	1.50	5	4	41.1
MSCC-430-F3	3	126	0.0942	0.0891	0.1015	0.0691	1.89	2	2	21.2
MSCC-430-F4	3	13	0.1024	0.0973	0.1041	0.0225	0.00[2]	2	2	18.3
MSCC-440-F1	6	697	0.1173	0.1074	0.1284	0.1550	2.93	8	6	50.2
MSCC-454-F1	13	689	0.0389	0.0334	0.0466	2.7338	2.39	23	13	63.5
MSCC-454-F2	6	587	0.0446	0.0377	0.0500	2.3032	2.57	20	9	29.1
MSCC-454-F3	3	148	0.0520	0.0464	0.0557	4.0214	2.48	3	3	9.3
MSCC-454-F4	3	140	0.0490	0.0425	0.0551	1.5093	2.42	3	3	12.8
MSCC-454-F5	3	41	0.0438	0.0424	0.0460	1.1194	0.83[1]	3	3	8.6
MSCC-457-F1	19	1287	0.0786	0.0709	0.0877	0.4021	3.37	23	11	78.4
MSCC-457-F2	8	508	0.0758	0.0673	0.0830	0.1594	2.69	12	8	52.5
MSCC-457-F3	6	209	0.0846	0.0768	0.0896	0.2454	2.48	7	5	30.6
MSCC-457-F4	3	197	0.0829	0.0757	0.0881	0.1691	2.80	3	3	22.2
MSCC-457-F5	3	101	0.0756	0.0676	0.0785	0.0982	1.97[1]	3	3	16.8
MSCC-457-F6	3	39	0.0816	0.0779	0.0872	0.0649	0.74	3	3	23.3
MSCC-460-F1	12	885	0.1142	0.1067	0.1242	0.1862	3.61	17	9	69.3
MSCC-460-F2	6	485	0.1139	0.1043	0.1237	0.1195	3.62	9	6	57.9
MSCC-460-F3	6	310	0.1080	0.1003	0.1159	0.0695	2.55	6	5	59.7
MSCC-460-F4	5	253	0.1218	0.1155	0.1304	0.0522	2.51	7	5	27.6
MSCC-463-F1	9	575	0.0722	0.0652	0.0788	0.3732	2.77	16	10	57.4
MSCC-463-F2	6	530	0.0765	0.0699	0.0839	0.3129	2.67	15	10	39.0
MSCC-463-F3	11	388	0.0768	0.0696	0.0842	0.2787	2.27	13	11	73.6
MSCC-463-F4	3	361	0.0657	0.0585	0.0743	0.4589	2.77	5	4	20.8
MSCC-463-F5	5	358	0.0655	0.0600	0.0730	0.4039	2.08	10	8	41.9
MSCC-463-F6	7	328	0.0837	0.0775	0.0894	0.2868	2.56	7	6	36.9
MSCC-463-F7	4	191	0.0688	0.0654	0.0739	0.0878	0.68[1]	8	7	26.8
MSCC-463-F8	4	159	0.0663	0.0628	0.0698	0.4765	2.59	6	4	12.5
MSCC-463-F9	4	147	0.0706	0.0660	0.0764	0.2416	1.87	6	5	31.9
MSCC-463-F10	3	133	0.0701	0.0651	0.0754	0.3595	2.05	4	3	12.7
MSCC-463-F11	3	130	0.0802	0.0748	0.0861	0.1864	2.64	4	4	14.8
MSCC-474-F1	3	765	0.0369	0.0300	0.0444	2.7548	2.63	13	5	10.5
MSCC-474-F2	5	404	0.0340	0.0296	0.0387	1.3043	2.06	13	7	23.4
MSCC-474-F3	3	277	0.0368	0.0318	0.0412	0.2423	1.98	9	8	16.5
MSCC-474-F4	3	260	0.0321	0.0284	0.0388	0.7372	1.75	7	5	16.5
MSCC-474-F5	3	229	0.0337	0.0292	0.0372	0.3332	1.51	10	6	17.2
MSCC-474-F6	4	205	0.0365	0.0324	0.0423	0.5697	1.90	6	6	15.3
MSCC-474-F7	3	23	0.0379	0.0328	0.0406	0.1777	0.00[1]	10	7	28.3
MSCC-484-F1	4	681	0.1361	0.1243	0.1501	0.0130	3.89	7	6	54.6

Notes. [1] The radius was calculated using $3\times\bar{\delta}$. [2] The filament local density is below $3\times\bar{\delta}$.

Bibliography

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, *ApJ Suppl.*, 182, 543

Abell, G. O. 1961, *Astron. J.*, 66, 607

Abell, G. O., Corwin, Jr., H. G., & Olowin, R. P. 1989, *ApJ Suppl.*, 70, 1

Aghanim, N., Douspis, M., Hurier, G., et al. 2019a, arXiv e-prints, arXiv:1902.00350

Aghanim, N., Douspis, M., Hurier, G., et al. 2019b, arXiv e-prints, arXiv:1902.00350

Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *ApJ Suppl.*, 219, 12

Albareti, F. D., Allende Prieto, C., Almeida, A., et al. 2017, The Astrophysical Journal Supplement Series, 233, 25

Alpaslan, M., Driver, S., Robotham, A. S. G., et al. 2015, *MNRAS*, 451, 3249

Alpaslan, M., Robotham, A. S. G., Obreschkow, D., et al. 2014, *MNRAS*, 440, L106

Altay, G., Colberg, J. M., & Croft, R. A. C. 2006, *MNRAS*, 370, 1422

Aragón-Calvo, M. A., Platen, E., van de Weygaert, R., & Szalay, A. S. 2010, *ApJ*, 723, 364

Aragon-Salamanca, A., Ellis, R. S., Couch, W. J., & Carter, D. 1993, *MNRAS*, 262, 764

Arnaud, M., Neumann, D. M., Aghanim, N., et al. 2001, *A&A*, 365, L80

Arnaud, M., Pointecouteau, E., & Pratt, G. W. 2005, *A&A*, 441, 893

Arnaud, M., Pratt, G. W., Piffaretti, R., et al. 2010a, *A&A*, 517, A92

Arnaud, M., Pratt, G. W., Piffaretti, R., et al. 2010b, *A&A*, 517, A92

Bahcall, J. N. & Casertano, S. 1985, *ApJ (Lett.)*, 293, L7

Basilakos, S., Plionis, M., & Rowan-Robinson, M. 2001, *MNRAS*, 323, 47

Battaglia, N., Bond, J. R., Pfrommer, C., Sievers, J. L., & Sijacki, D. 2010, *ApJ*, 725, 91

Battaner, E. & Florido, E. 2000, Fund. Cosmic Phys., 21, 1

Beers, T. C., Flynn, K., & Gebhardt, K. 1990, *Astron. J.*, 100, 32

Bender, A. N., Kennedy, J., Ade, P. A. R., et al. 2016, *MNRAS*, 460, 3432

Berlind, A. A., Frieman, J., Weinberg, D. H., et al. 2006, *ApJ Suppl.*, 167, 1

Bertschinger, E. 1998, *Ann. Rev. Astron. Ap.*, 36, 599

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

Binney, J. & Tremaine, S. 1987, Galactic dynamics

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

Blumenthal, G. R., Pagels, H., & Primack, J. R. 1982, *Nature*, 299, 37

Böhringer, H., Schuecker, P., Pratt, G. W., et al. 2007, *A&A*, 469, 363

Bolton, A. S., Schlegel, D. J., Aubourg, É., et al. 2012, *Astron. J.*, 144, 144

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

Bond, J. R. & Szalay, A. S. 1983, *ApJ*, 274, 443

Bond, J. R., Szalay, A. S., & Turner, M. S. 1982, Physical Review Letters, 48, 1636

Bond, N. A., Strauss, M. A., & Cen, R. 2010, *MNRAS*, 409, 156

Bourdin, H., Mazzotta, P., Kozmanyan, A., Jones, C., & Vikhlinin, A. 2017, *ApJ*, 843, 72

Bower, R. G., Lucey, J. R., & Ellis, R. S. 1992, *MNRAS*, 254, 589

Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, *MNRAS*, 351, 1151

Carter, D. & Metcalfe, N. 1980, *MNRAS*, 191, 325

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

Cavaliere, A. & Fusco-Femiano, R. 1976, *A&A*, 49, 137

Cen, R. & Ostriker, J. P. 1999, *ApJ*, 514, 1

Chambers, S. W., Melott, A. L., & Miller, C. J. 2002, *ApJ*, 565, 849

Chen, Y.-C., Ho, S., Brinkmann, J., et al. 2016, *MNRAS*, 461, 3896

Chen, Y.-C., Ho, S., Mandelbaum, R., et al. 2017a, *MNRAS*, 466, 1880

Chen, Y.-C., Ho, S., Mandelbaum, R., et al. 2017b, *MNRAS*, 466, 1880

Chernin, A. D. 1981, AZh, 58, 25

Chow-Martinez, M., Andernach, H., Caretta, C. A., & Trejo-Alonso, J. J. 2014, Monthly Notices of the Royal Astronomical Society, 445, 4073

Ciotti, L. & Bertin, G. 1999, *A&A*, 352, 447

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

Conroy, C., Gunn, J. E., & White, M. 2009, *ApJ*, 699, 486

Costa-Duarte, M. V., Sodré, L., J., & Durret, F. 2011, *MNRAS*, 411, 1716

Cowsik, R. & McClelland, J. 1973, *ApJ*, 180, 7

Croston, J. H., Arnaud, M., Pointecouteau, E., & Pratt, G. W. 2006, *A&A*, 459, 1007

Croston, J. H., Pratt, G. W., Böhringer, H., et al. 2008, *A&A*, 487, 431

Cuceu, A., Farr, J., Lemos, P., & Font-Ribera, A. 2019, arXiv e-prints, arXiv:1906.11628

Cybulski, R., Yun, M. S., Fazio, G. G., & Gutermuth, R. A. 2014, *MNRAS*, 439, 3564

Czakon, N. G., Sayers, J., Mantz, A., et al. 2015, *ApJ*, 806, 18

Darvish, B., Mobasher, B., Sobral, D., Scoville, N., & Aragon-Calvo, M. 2015, The Astrophysical Journal, 805, 121

Davis, M., Huchra, J., Latham, D. W., & Tonry, J. 1982, *ApJ*, 253, 423

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

Dietrich, J. P., Bocquet, S., Schrabback, T., et al. 2019, *MNRAS*, 483, 2871

Dietrich, J. P., Werner, N., Clowe, D., et al. 2012, *Nature*, 487, 202

Dijkstra, E. W. 1959, Numer. Math., 1, 269

Dolag, K., Komatsu, E., & Sunyaev, R. 2016, *MNRAS*, 463, 1797

Doroshkevich, A. G. & Khlopov, M. I. 1984, *MNRAS*, 211, 277

Doroshkevich, A. G., Zeldovich, Y. B., Syunyaev, R. A., & Khlopov, M. Y. 1980, Pisma v Astronomicheskii Zhurnal, 6, 457

Dressler, A. 1980, *ApJ*, 236, 351

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

Dünner, R., Hasselfield, M., Marriage, T. A., et al. 2013, *ApJ*, 762, 10

Eckert, D., Ettori, S., Pointecouteau, E., et al. 2017, Astronomische Nachrichten, 338, 293

Eckert, D., Molendi, S., Vazza, F., Ettori, S., & Paltani, S. 2013, *A&A*, 551, A22

Einasto, J. 2009, ArXiv e-prints

Einasto, J. 2014, arXiv e-prints, arXiv:1410.6932

Einasto, M., Einasto, J., Tago, E., Müller, V., & Andernach, H. 2001, *Astron. J.*, 122, 2222

Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, *ApJ*, 633, 560

Ellis, R. S., Smail, I., Dressler, A., et al. 1997, The Astrophysical Journal, 483, 582

Ettori, S., Tozzi, P., Borgani, S., & Rosati, P. 2004, *A&A*, 417, 13

Fixsen, D. J. 2009, *ApJ*, 707, 916

Ghirardini, V., Ettori, S., Eckert, D., et al. 2017, ArXiv e-prints

Giodini, S., Lovisari, L., Pointecouteau, E., et al. 2013a, *Space Sci. Rev.*, 177, 247

Giodini, S., Lovisari, L., Pointecouteau, E., et al. 2013b, *Space Sci. Rev.*, 177, 247

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

Gladders, M. D., López-Cruz, O., Yee, H. K. C., & Kodama, T. 1998, *ApJ*, 501, 571

Gladders, M. D. & Yee, H. K. C. 2000, *Astron. J.*, 120, 2148

Godłowski, W. & Flin, P. 2010, *ApJ*, 708, 920

Graham, R. L. & Hell, P. 1985, Annals of the History of Computing, 7, 43

Guennou, L., Adami, C., Durret, F., et al. 2014, *A&A*, 561, A112

Guth, A. H. & Pi, S.-Y. 1982, Physical Review Letters, 49, 1110

Hao, J., Kubo, J. M., Feldmann, R., et al. 2011, *ApJ*, 740, 39

Harrison, E. R. 1970, Phys. Rev. D, 1, 2726

Hasselfield, M., Hilton, M., Marriage, T. A., et al. 2013a, J. Cosmology Astropart. Phys., 7, 008

Hasselfield, M., Hilton, M., Marriage, T. A., et al. 2013b, J. Cosmology Astropart. Phys., 7, 008

Hilton, M., Hasselfield, M., Sifón, C., et al. 2018, *ApJ Suppl.*, 235, 20

Hotokezaka, K., Nakar, E., Gottlieb, O., et al. 2019, Nature Astronomy, 385

Huchra, J. P., Macri, L. M., Masters, K. L., et al. 2012, *ApJ Suppl.*, 199, 26

Huertas-Company, M., Aguerri, J. A. L., Bernardi, M., Mei, S., & Sánchez Almeida, J. 2011, *A&A*, 525, A157

Hurier, G., Macías-Pérez, J. F., & Hildebrandt, S. 2013, *A&A*, 558, A118

Jeans, J. H. 1902, Philosophical Transactions of the Royal Society of London Series A, 199, 1

Joeveer, M. & Einasto, J. 1978, in IAU Symposium, Vol. 79, Large Scale Structures in the Universe, ed. M. S. Longair & J. Einasto, 241–250

Kaiser, N. 1986, *MNRAS*, 222, 323

Kaiser, N., Squires, G., & Broadhurst, T. 1995, *ApJ*, 449, 460

Kauffmann, G. & Charlot, S. 1998, Monthly Notices of the Royal Astronomical Society, 294, 705

Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, *MNRAS*, 341, 33

Kewley, L. J., Groves, B., Kauffmann, G., & Heckman, T. 2006, *MNRAS*, 372, 961

King, I. R. 1972, *ApJ (Lett.)*, 174, L123

Klypin, A. A., Trujillo-Gomez, S., & Primack, J. 2011, *ApJ*, 740, 102

Kodama, T. & Arimoto, N. 1997, *A&A*, 320, 41

Kotov, O. & Vikhlinin, A. 2005, *ApJ*, 633, 781

Kraljic, K., Arnouts, S., Pichon, C., et al. 2018, *MNRAS*, 474, 547

Kuutma, T., Tamm, A., & Tempel, E. 2017, *A&A*, 600, L6

Lambas, D. G., Groth, E. J., & Peebles, P. J. E. 1988, *Astron. J.*, 95, 975

Lauer, T. R. & Postman, M. 1992, *ApJ (Lett.)*, 400, L47

Lauer, T. R., Postman, M., Strauss, M. A., Graves, G. J., & Chisari, N. E. 2014, 82

Lee, J. & Evrard, A. E. 2007, *ApJ*, 657, 30

Liivamägi, L. J., Tempel, E., & Saar, E. 2012, *A&A*, 539, A80

Linde, A. D. 1982, Physics Letters B, 108, 389

Lintott, C., Schawinski, K., Bamford, S., et al. 2011, *MNRAS*, 410, 166

Lintott, C. J., Schawinski, K., Slosar, A., et al. 2008, *MNRAS*, 389, 1179

Luparello, H., Lares, M., Lambas, D. G., & Padilla, N. 2011, *MNRAS*, 415, 964

Malavasi, N., Arnouts, S., Vibert, D., et al. 2017, *MNRAS*, 465, 3817

Matthews, T. A., Morgan, W. W., & Schmidt, M. 1964, *ApJ*, 140, 35

Miller, C. J., Nichol, R. C., Reichart, D., et al. 2005, The Astronomical Journal, 130, 968

Mroczkowski, T., Nagai, D., Basu, K., et al. 2019, *Space Sci. Rev.*, 215, 17

Murtagh, F. & Contreras, P. 2011, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2, 86

Murtagh, F. & Legendre, P. 2014, Journal of Classification, 31, 274

Nagai, D., Vikhlinin, A., & Kravtsov, A. V. 2007, *ApJ*, 655, 98

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *ApJ*, 490, 493

Oort, J. H. 1940, *ApJ*, 91, 273

Ostriker, J. P. & Hausman, M. A. 1977, *ApJ (Lett.)*, 217, L125

Ostriker, J. P. & Peebles, P. J. E. 1973, *ApJ*, 186, 467

Padmanabhan, T. 2002, Theoretical Astrophysics - Volume 3, Galaxies and Cosmology, 638

Padmanabhan, T. 2003, *Physics Reports*, 380, 235

Pagels, H. & Primack, J. R. 1982, Physical Review Letters, 48, 223

Paz, D. J., Sgró, M. A., Merchán, M., & Padilla, N. 2011, *MNRAS*, 414, 2029

Pearson, R. J., Ponman, T. J., Norberg, P., Robotham, A. S. G., & Farr, W. M. 2015, *MNRAS*, 449, 3082

Peebles, P. J. E. 1980, The large-scale structure of the universe (Princeton University Press)

Peebles, P. J. E. 1982, *ApJ*, 258, 415

Peebles, P. J. E. 1993, Principles of Physical Cosmology

Penzias, A. A. & Wilson, R. W. 1965, *ApJ*, 142, 419

Percival, W. J. 2001, *MNRAS*, 327, 1313

Plagge, T., Benson, B. A., Ade, P. A. R., et al. 2010, *ApJ*, 716, 1118

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2013a, *A&A*, 550, A131

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2013b, *A&A*, 550, A134

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2011, *A&A*, 536, A11

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016a, *A&A*, 594, A27

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016b, *A&A*, 594, A13

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2018, arXiv e-prints, arXiv:1807.06210

Planck Collaboration, Aghanim, N., Arnaud, M., et al. 2016c, *A&A*, 594, A22

Planck Collaboration VIII. 2011, *A&A*, 536, A8

Planck Collaboration X. 2011, *A&A*, 536, A10

Planck Collaboration XX. 2014, *A&A*, 571, A20

Planck Collaboration XXIX. 2014, *A&A*, 571, A29

Planck Collaboration VIII. 2016, *A&A*, 594, A8

Planck Collaboration XXIV. 2016, *A&A*, 594, A24

Planck Collaboration XXVII. 2016, *A&A*, 594, A27

Planck Collaboration Int. III. 2013, *A&A*, 550, A129

Planck Collaboration Int. V. 2013, *A&A*, 550, A131

Platen, E., van de Weygaert, R., Jones, B. J. T., Vegter, G., & Calvo, M. A. A. 2011, *MNRAS*, 416, 2494

Plionis, M. & Basilakos, S. 2002, *MNRAS*, 329, L47

Pointecouteau, E., Arnaud, M., & Pratt, G. W. 2005, *A&A*, 435, 1

Pointecouteau, E., Giard, M., & Barret, D. 1998, *A&A*, 336, 44

Poudel, A., Heinämäki, P., Nurmi, P., et al. 2016, *A&A*, 590, A29

Pratt, G. W., Arnaud, M., Biviano, A., et al. 2019, *Space Sci. Rev.*, 215, 25

Pratt, G. W., Croston, J. H., Arnaud, M., & Böhringer, H. 2009, *A&A*, 498, 361

Press, W. H. & Schechter, P. 1974, *ApJ*, 187, 425

Rees, M. J. & Sciama, D. W. 1968, *Nature*, 217, 511

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, *Astron. J.*, 116, 1009

Riess, A. G., Macri, L. M., Hoffmann, S. L., et al. 2016, *ApJ*, 826, 56

Rines, K., Diaferio, A., & Natarajan, P. 2007, *ApJ*, 657, 183

Rines, K., Geller, M. J., Kurtz, M. J., & Diaferio, A. 2003, *Astron. J.*, 126, 2152

Roberts, M. S. 1966, *ApJ*, 144, 639

Romero, C. E., Mason, B. S., Sayers, J., et al. 2017, *ApJ*, 838, 86

Roncarelli, M., Pointecouteau, E., Giard, M., Montier, L., & Pello, R. 2010, *A&A*, 512, A20

Rubin, V. C. & Ford, Jr., W. K. 1970, *ApJ*, 159, 379

Ruppin, F., Mayet, F., Pratt, G. W., et al. 2018, *A&A*, 615, A112

Sachs, R. K. & Wolfe, A. M. 1967, *ApJ*, 147, 73

Sakelliou, I. & Ponman, T. J. 2004, *MNRAS*, 351, 1439

Santanu, R. S. 2014, Graph Theory with Algorithms and Its Applications: In Applied Science and Technology (Springer Publishing Company, Incorporated)

Santiago-Bautista, I., Caretta, C., Bravo-Alfaro, H., & Pointecouteau, E. 2019b, in prep., n/a

Sastry, G. N. 1968, *P. A. S. P.*, 80, 252

Sayers, J., Golwala, S. R., Mantz, A. B., et al. 2016, *ApJ*, 832, 26

Sayers, J., Mroczkowski, T., Zemcov, M., et al. 2013, *ApJ*, 778, 52

Scoville, N., Arnouts, S., Aussel, H., et al. 2013, The Astrophysical Journal Supplement Series, 206, 3

Serna, A. & Gerbal, D. 1996, *A&A*, 309, 65

Sérsic, J. L. 1968, Atlas de Galaxias Australes

Shectman, S. A., Landy, S. D., Oemler, A., et al. 1996, *ApJ*, 470, 172

Smith, A. G., Hopkins, A. M., Hunstead, R. W., & Pimbblet, K. A. 2012, *MNRAS*, 422, 25

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Nature*, 435, 629

Starobinsky, A. A. 1980, Physics Letters B, 91, 99

Strauss, M. A., Weinberg, D. H., Lupton, R. H., et al. 2002, *Astron. J.*, 124, 1810

Sunyaev, R. A. & Zeldovich, Y. B. 1972, Comments on Astrophysics and Space Physics, 4, 173

Suyu, S. H., Bonvin, V., Courbin, F., et al. 2017, *MNRAS*, 468, 2590

Szalay, A. S. & Marx, G. 1976, *A&A*, 49, 437

Tanaka, M., Hoshi, T., Kodama, T., & Kashikawa, N. 2007, *MNRAS*, 379, 1546

Tanimura, H., Hinshaw, G., McCarthy, I. G., et al. 2017, ArXiv e-prints

Tchernin, C., Eckert, D., Ettori, S., et al. 2016, *A&A*, 595, A42

Tempel, E., Stoica, R. S., Martínez, V. J., et al. 2014, *MNRAS*, 438, 3465

Tempel, E., Tago, E., & Liivamägi, L. J. 2011, 106, 1

Theodoridis, S. & Koutroumbas, K. 2009, in Pattern Recognition (Fourth Edition), fourth edition edn., ed. T. Sergios & K. Konstantinos (Boston: Academic Press), 595 – 625

Theodoridis, S., Pikrakis, A., Koutroumbas, K., & Cavouras, D. 2010, Introduction to Pattern Recognition: A Matlab Approach (Academic Press)

Tittley, E. R. & Henriksen, M. 2001, *ApJ*, 563, 673

Trejo-Alonso, J. J., Caretta, C. A., Laganá, T. F., et al. 2014, *MNRAS*, 441, 776

Tremaine, S. & Gunn, J. E. 1979, Phys. Rev. Lett., 42, 407

Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, *ApJ*, 613, 898

Vogelsberger, M., Genel, S., Springel, V., et al. 2014, *Nature*, 509, 177

Voronoi, G. 1908, Journal für die reine und angewandte Mathematik, 134, 198

Wang, Y., Park, C., Yang, X., Choi, Y.-Y., & Chen, X. 2009, *ApJ*, 703, 951

Werner, N., Finoguenov, A., Kaastra, J. S., et al. 2008, *A&A*, 482, L29

West, M. J. 1994, *MNRAS*, 268, 79

Zel'dovich, Y. B. 1970, *A&A*, 5, 84

Zhang, Y., Yang, X., Wang, H., et al. 2015, *ApJ*, 798, 17

Zucca, E., Zamorani, G., Scaramella, R., & Vettolani, G. 1993, *ApJ*, 407, 470

Author : Iris Santiago-Bautista

Title: The environmental effects of the LSS: Characterization of the baryonic components

Abstract:
The baryonic component of the Large Scale Structure (LSS) of the Universe is composed by concentrations of gas and galaxies forming groups, clusters, elongated filaments and widely spread sheets which probably underline the distribution of dark matter. Nevertheless, according to the current cosmological models, most of the baryonic material in the Universe has not yet been directly observed.
Numerical simulations suggest that from one-half to two-thirds of all baryons may be located out of clusters of galaxies, pervading the structures between them. The most concentrated structures, which we call systems of galaxies (i.e., groups and clusters) usually contain high density hot gas (1 - 10 keV) that cools radiatively, emits at X-rays wavelengths and interacts with the cosmic microwave background at millimeter wavelengths (Sunyaev Zel'dovich effect, SZ). For the less dense structures, filaments and sheets, the baryons are probably in moderately hot gas phase (0.01 - 1 keV), commonly named as warm hot intergalactic medium (WHIM).
In this PhD Thesis, we study the environmental effects associated to the different components of the LSS. For the galaxy systems, we aim to characterize the intra cluster medium (ICM) through the analysis of the S-Z effect. We employ the ACT and Planck data to analyze the gas pressure profiles of a sample of low mass galaxy clusters. For the least dense structures, we assembled a sample of filament candidates composed by chains of clusters that are located inside superclusters of galaxies. We aim to probe the filament structure skeletons and characterize their components (galaxies, groups/clusters and gas).

Key words: galaxies, galaxy groups, clusters, large scale structure of the Universe, cosmology, galaxies evolution, data analisis