Universidade Federal do Rio de Janeiro
Centro de Ciências Matemáticas e da Natureza
Observatório do Valongo

Galaxy Distribution Systems as Fractals

Sharon Teles de Azevedo Chaves

Rio de Janeiro
Setembro de 2023

Galaxy Distribution Systems as Fractals

Sharon Teles de Azevedo Chaves

Dissertação de Graduação submetida ao Observatório do Valongo, Universidade Federal do Rio de Janeiro, como requisito necessário para a obtenção do título de Bacharel.

Orientador: Dr. Marcelo Byrro Ribeiro
Coorientadora: Dra. Amanda Reis Lopes

Rio de Janeiro
Setembro de 2023

Universidade Federal do Rio de Janeiro

Departamento de Astronomia

Aos treze dias do mês de setembro de dois mil e vinte e três, às quatorze horas, via Zoom Meeting (https://us02web.zoom.us/j/83064086278?pwd=R2UxVENKVXM4aDNtam5DL0pHQ2ExQT09), reuniu-se a banca examinadora do trabalho de conclusão de curso em Astronomia da aluna SHARON TELES DE AZEVEDO CHAVES, DRE 116055276, intitulado **Galaxy Ditribution Systems as Fractals**. Compuseram a banca examinadora os professores: Prof. Paulo Afrânio Augusto Lopes – Presidente (OV/UFRJ), Prof. Marcelo Byrro Ribeiro – Orientador (IF/UFRJ), Profª. Amanda Reis Lopes - Coorientadora (Inst. de Astrof. De La Plata/Argentina), Prof. Sérgio E. C. E. Jorás – (IF/UFRJ) e Profª Beatriz B. Siffert (Caxias/UFRJ). Após a exposição oral, a candidata foi arguida pelos membros da banca que se reuniram reservadamente, e decidiram, **por unanimidade pela aprovação**. Para constar, redigi à presente Ata, que aprovada por todos os presentes, vai assinada por mim, secretário acadêmico de graduação, e pelos demais membros da banca e a aluna.

Rio de Janeiro, 13 de setembro de 2023.

Documento assinado digitalmente
gov.br PAULO AFRANIO AUGUSTO LOPES
Data: 13/09/2023 17:40:41-0300
Verifique em https://validar.iti.gov.br

Prof. Paulo Afrânio Augusto Lopes – **Presidente da Banca**

Documento assinado digitalmente
gov.br MARCELO BYRRO RIBEIRO
Data: 13/09/2023 17:54:16-0300
Verifique em https://validar.iti.gov.br

Prof. Marcelo Byrro Ribeiro – **Orientador**

Documento assinado digitalmente
gov.br AMANDA REIS LOPES
Data: 13/09/2023 18:08:58-0300
Verifique em https://validar.iti.gov.br

Profª. Amanda Reis Lopes – **Coorientadora**

Documento assinado digitalmente
gov.br SERGIO EDUARDO DE CARVALHO EYER JORAS
Data: 14/09/2023 12:33:36-0300
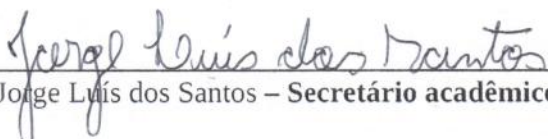Verifique em https://validar.iti.gov.br

---

Prof. Sérgio E. C. E. Jorás – **Membro da banca**

Documento assinado digitalmente
gov.br BEATRIZ BLANCO SIFFERT
Data: 14/09/2023 17:11:10-0300
Verifique em https://validar.iti.gov.br

Profª. Beatriz B. Siffert – **Membro da banca**

Documento assinado digitalmente
gov.br SHARON TELES DE AZEVEDO CHAVES
Data: 24/09/2023 01:27:56-0300
Verifique em https://validar.iti.gov.br

Sharon Teles de Azevedo Chaves – **Aluna**

Jorge Luís dos Santos – **Secretário acadêmico**

Declaration of Originality

I hereby certify that I am the sole author of this dissertation. All used materials, references to the literature and the work of others have been referred to. The results presented here are original and were published in Refs. [1] and [2]. This dissertation has not been presented for examination anywhere else.

Dedication

Dedico esse trabalho às mulheres da história da minha família que não tiveram oportunidade de ser cientistas, e a todas que um dia serão.

I dedicate this work to the women in my family's history who did not have the opportunity to be scientists, and to all who will one day be.

Acknowledgments

Agradecimentos

Abstract

Galaxy Distribution Systems as Fractals

Author: Sharon Teles de Azevedo Chaves

1st Adviser: Marcelo Byrro Ribeiro

2nd Adviser: Amanda Reis Lopes

This dissertation seeks to test if the large-scale galaxy distribution can be characterized as a fractal system. The standard ΛCDM cosmology having $H_0 = (70 \pm 5)$ km/s/Mpc is adopted on the study of almost a million objects from the UltraVISTA DR1, COSMOS2015 and SPLASH surveys, alongside the number density equations required for the description of these galaxy distribution systems as fractals with dimension D. The luminosity distance d_L, redshift distance d_Z and galaxy area distance (transverse comoving distance) d_G are the relativistic distance definitions used to estimate the galaxy number densities in the redshift interval of $0.1 \leq z \leq 4$ at volume limited subsamples defined by absolute magnitudes in the J-band for the UltraVISTA DR1 data and in the K-band for the COSMOS2015 and SPLASH data. Applying the appropriate relations for the description of galaxy fractal structures with single dimension D in the relativistic settings to these surveys datasets it is possible to state that for $z < 1$ the UltraVISTA DR1 galaxies presented an average of $D = (1.58 \pm 0.20)$, the COSMOS2015 galaxies produced $D = (1.39 \pm 0.19)$ and the SPLASH galaxies generated $D = (1.00 \pm 0.12)$. For $1 \leq z \leq 4$ the dimensions respectively decreased to $D = (0.59 \pm 0.28)$, $D = 0.54^{+0.27}_{-0.26}$ and $D = 0.83^{+0.36}_{-0.37}$. These results are robust under the Hubble constant uncertainty assumed here. Analysis of blue and red galaxies subsamples in the COSMOS2015 and SPLASH surveys show that the fractal dimensions of blue galaxies present essentially no alteration from the values above, although the ones for the red galaxies changed mostly to smaller values, meaning that D may be assumed as a more intrinsic property of the distribution of objects in the Universe, thus allowing for the fractal dimension to be used as a tool to study different populations of galaxies. All results confirm the decades old theoretical prediction of a decrease in the fractal dimension for $z > 1$ suggesting that either there are yet unclear observational biases causing such decrease in the fractal dimension, or the galaxy clustering was possibly more sparse and the universe void dominated in a not too distant past.

Keywords: cosmology, galaxy distribution, fractals

Resumo

Sistemas de Distribuição de Galáxias como Fractais

Autora: Sharon Teles de Azevedo Chaves

Orientador: Marcelo Byrro Ribeiro

Coorientadora: Amanda Reis Lopes

Esta dissertação busca testar se a distribuição de galáxias em grande escala pode ser caracterizada como um sistema fractal. A cosmologia padrão ΛCDM com constante de Hubble $H_0 = (70 \pm 5)$ km/s/Mpc é adotada neste estudo de quase um milhão de objetos dos levantamentos UltraVISTA DR1, COSMOS2015 e SPLASH, juntamente com equações de densidade numérica necessárias para a descrição desses sistemas de distribuição de galáxias como fractais com dimensão D. A distância de luminosidade d_L, distância de desvio para o vermelho d_Z e distância por área da galáxia (distância comóvel transversal) d_G são usadas como definições de distância relativística para estimar as densidades numéricas de galáxias no intervalo de desvio para o vermelho de $0.1 \leq z \leq 4$ em subamostras limitadas por volume definidas por magnitudes absolutas na banda J para os dados do UltraVISTA DR1 e na banda K para os dados do COSMOS2015 e SPLASH. Aplicando aos conjuntos de dados desses levantamentos as relações apropriadas para a descrição de estruturas fractais de galáxias com dimensão única D nas configurações relativísticas é possível afirmar que para $z < 1$ as galáxias do UltraVISTA DR1 apresentaram em média $D = (1.58 \pm 0.20)$, as galáxias do COSMOS2015 produziram $D = (1.39 \pm 0.19)$ e as galáxias do SPLASH geraram $D = (1.00 \pm 0.12)$. No intervalo $1 \leq z \leq 4$ as dimensões diminuíram respectivamente para $D = (0.59 \pm 0.28)$, $D = 0.54^{+0.27}_{-0.26}$ e $D = 0.83^{+0.36}_{-0.37}$. Esses resultados são robustos sob a incerteza na constante de Hubble assumida aqui. Análises de subamostras de galáxias azuis e vermelhas nos levantamentos COSMOS2015 e SPLASH surveys mostram que as dimensões fractais das galáxias azuis praticamente não apresentam alterações às mencionadas acima, embora as dimensões fractais das galáxias vermelhas tenham mudado principalmente para valores menores de desvio para o vermelho, o que significa que D pode ser assumido como uma propriedade intrínseca da distribuição dos objetos no Universo, permitindo assim que a dimensão fractal seja utilizada como ferramenta para estudar diferentes populações de galáxias. Todos os resultados confirmam a previsão teórica de décadas de uma diminuição na dimensão fractal para $z > 1$ sugerindo que ainda existem vieses observacionais causando tal diminuição na dimensão fractal, ou o agrupamento de galáxias era possivelmente mais esparso e o universo dominado por vazios em um passado não muito distante.

Palavras-chave: cosmologia, distribuição de galáxias, fractais

List of Figures

List of Tables

Table of Contents

Introduction

Several recent redshift galactic surveys display the large scale structure of the universe showing voids that shatter the luminous matter in an apparent irregular pattern, which combined with the cosmic microwave background subtle inhomogeneity can lead to ponder about some underlying assumptions of the standard ΛCDM cosmological model. Thus, it was brought to light the importance of defining the model scale limits and of testing the assumed homogeneity through observations to understand the evolution of the structure and get a glimpse of the universe's initial conditions. Analysis of the galaxies in these redshift surveys suggest that the distribution is irregular, possibly self similar and, therefore, fractal, which means that they can be represented by a power law. Self similarity is when the structure does not change with the scaling factor. Hence the distribution looks the same from afar or zoomed in. Additionally, the structure's irregularity can be characterized by the concept of fractal dimension which is a fractional value in the exponent of a power law. The mathematician who studied irregular objects with this approach was Benoît Mandelbrot, who from the 1950s to 1970s established adequate techniques to represent irregular shapes and forms by means of the fractal concept.

The fractal idea applied to galaxy distribution was originally known as the *hierarchical cosmology* a century ago. The first attempt to mathematically describe the possible star hierarchy was published by E. Fournier d'Albe who proposed that the stars were infinitely distributed in a hierarchy of spheres in space where the mass inside a sphere was directly proportional to its radius $M(R) = R$ and the average matter density was zero [3]. In 1929, M. Amoroso Costa admitted the general relativistic implications on the universe spatial finitude. In his publication, Amoroso Costa investigated the possibilities of average density, potential and gravitational force for an one-dimensional star system both in a finite and infinite universe [4]. Succeeding that there was a publication by Carpenter who detected that the distribution of the most luminous nebulae was not random but followed a specific regularity [5]. However, the concept of hierarchy was lacking adequate mathematical representation and as a consequence Carpenter's results and the hierarchical cosmology were abandoned until 1970 when G. de Vaucouleurs and James Wertz reinterpreted them. Their suggestion pointed to an universal size-density power law existence that later was found to be in agreement with Mandelbrot's approach in his theory of fractals [6]. In 1980 the first redshift galaxy catalogs emerged and using the CfA redshift survey data de Lapparent et al. published evidence that the large scale structure of the universe did not appear as a smooth homogeneous distribution [7]. In the following years other authors confirmed the power law behaviour in the galaxy distribution to even deeper distances and some authors concluded that the universe could be described by a simple fractal, a succession of single fractal

dimensions or even by a multifractal spectrum of dimensions [8, 9, 10, 11, 12].

Up to this point we discussed the cosmology only in the Newtonian context, however, a complete cosmology study needs general relativity. In this theory space and time are connected forming the space-time which is curved by the gravitational effects of the matter distribution. The object's trajectory in the curved space-time is described by a line called geodesic, thus, the light that was emitted by a galaxy in the past travels through the geodesic with constant speed c forming the past light cone. There is also the future light cone formed by the light that objects are emitting now, however we can not observe it. Another important general relativistic aspect is that distance is not a unique concept because it depends on how it is measured. G.R.F. Ellis [13] observational relativistic distance definitions along with the ideas for density-size distribution considered by Wertz and Pietronero [9, 14] were reviewed by Ribeiro [8] who presented a relativistic model for the universe large scale structure as a self similar fractal system in a Lemaître-Tolman-Bondi (LTB) space-time. The analysis in Ribeiro [15] recognized that the spatial homogeneity of the Einstein-de Sitter cosmology (EdS) was observed only for nearby redshift values of $z \sim 0.01$, which suggests that the perceived fractal structure could be extended in the past light cone to higher z values. In this way, the fractal distribution was not conflicting with the Cosmological Principle with regard to spatial homogeneity because general relativity allows us to distinguish between two types of homogeneity: observational homogeneity, represented in the past light cone where astronomical measurements are made, and spatial homogeneity, defined on the spatial hyper-surface where the density-energy term of Einstein's equation is constant in the standard cosmological model of Friedmann-Lemaître-Robertson-Walker (FLRW). Therefore, fractal galaxy distribution regards observational inhomogeneity.

Rangel Lemos & Ribeiro [16] analyzed the consequences of distinguishing spatial and observational homogeneity simulating the galaxy number counts behaviour taking into consideration the four cosmological distances: luminosity distance d_L, observer area distance d_A, galactic area distance d_G and redshift distance d_z. Furthermore, they found the dependence between the radial densities and the distance definitions, and also the tendency of the fractal dimension to approach 3 as $z \to 0$, for all distances, and the asymptotic tendency of it going to zero as $z \to \infty$. This analyses was carried out using the spatially homogeneous EdS model and led to the curious result that spatially homogeneous cosmologies do not necessarily present observational homogeneity and, therefore, there is not contradiction with spatially homogeneous cosmological models like the FLRW which presents observational inhomogeneities at high redshift ranges [16]. With these results the fractal dimension calculated from the observable quantities show that the relativistic fractal model can be a valid tool to characterize the galaxy distribution irregularity without conflicting with the standard model postulates. Additionally, obtaining more than one fractal dimension value in the same scaling range suggests that the distribution could have various values for the single fractal dimensions D at different scaling ranges.

Of notable interest to this dissertation is the study presented by Conde-Saavedra et al., who tested the fractal galaxy distribution hypothesis using the FORS Deep Field (FDF) dataset consisting of 5558 galaxies in the range $0.45 \leq z \leq 5.0$ [17]. They concluded that at $z \lesssim 1.3 - 1.9$ the sample produced an average single fractal dimension of $D = 1.4^{+0.7}_{-0.6}$ while beyond this threshold they obtained $D = 0.5^{+1.2}_{-0.4}$. Although the volume-limited samples are acquired from an indirect luminosity function method that produced high data uncertainties, this work contributed with further observational support for the theoretically predicted decreasing in fractal dimension at larger scales mentioned here. In this dissertation we followed this line of investigation by using volume-limited samples directly obtained from UltraVISTA DR1 [18], COSMOS2015 [19] and SPLASH [20] data, assuming a FLWR cosmological model. The determination of D occurs by taking theses galaxy datasets with volume limited samples and plotting the number density vs. distance. Comparing with previous results, ours included better defined thresholds for moderate and high scaling ranges, smaller uncertainties and final results more in line with each other considering all cosmological distance definitions.

Our calculations showed that in all three surveys the galaxy distribution can be well characterized as a fractal system possessing two consecutive scaling ranges. The UltraVISTA DR1 dataset containing 219,300 objects presented the following median fractal dimensions and uncertainties: $D = (1.58 \pm 0.20)$ for $z < 1$, and $D = (0.59 \pm 0.28)$ for $1 \leq z \leq 4$ [1] which is also in agreement with the early theoretical prediction of a decrease in the fractal dimension at larger scales [15]. Extending the study, the COSMOS2015 dataset including 578,379 objects generated $D = 1.39 \pm 0.19$ for $z < 1$ and $D = 0.54^{+0.27}_{-0.26}$ for $1 \leq z \leq 4$ and the SPLASH dataset consisting of 390,362 objects presented $D = 1.00 \pm 0.12$ for $z < 1$ and $D = 0.83^{+0.36}_{-0.37}$ for $1 \leq z \leq 4$. These numbers turned out to be robust under the adopted Hubble constant uncertainty of $H_0 = (70 \pm 5)$ km/s/Mpc and provided further empirical support that the galaxy distribution can be characterized by two subsequent fractal scaling ranges with no detectable transition to homogeneity up to the redshift limits of these surveys [2].

Additional subsamples were generated by selecting blue, star forming galaxies and red, quiescent, ones using color-color diagrams or specific star formation rates. The fractal dimension of blue galaxies turned out either unaltered or only marginally altered as compared to the unselected samples. However, the red galaxies had their fractal dimensions becoming evidently smaller in most cases, apart from the red COSMOS2015 data whose D values increased noticeably. Such results indicate that single fractal dimensions may be used as descriptors of galaxy distributions as well as tools to trace galaxy types and/or their evolutionary stages at different redshift ranges. Supplementarily, all results collected here provide extra empirical confirmation of the theoretical decrease prediction in the fractal dimension at redshift values higher than the unity [15] and that the galaxy distribution can be described by two fractal scaling ranges without any transition to homogeneity up to the redshift limits of the three

surveys.

The plan of the dissertation is as follows. Chapter 1 goes through the basic concepts and tools needed to understand relativistic fractal geometry. Chapter 2 specifies the observational details of the UltraVISTA DR1, COSMOS2015 and SPLASH redshift surveys used in this work. Chapter 3 features the results of the fractal analysis of the galaxy distribution surveys. Conclusion is used to discuss the results. Appendix A presents the python code used for the galaxy data processing. Appendix B exhibits mathematical details for the number-distance relation, essential to the discussions presented in Chapter 1.

1.   Basic Concepts

This chapter presents a discussion about fractals and their application to the relativistic domain. Section 1.1 will present fractals' fundamental concepts and Sec. 1.2 their application to galaxy distribution in the context of relativistic cosmology.

1.1   Fractals

Benoit Mandelbrot (1924-2010) developed the concept of fractals under the viewpoint that nature is too complex to be described by classically idealized geometrical shapes like lines, planes and geometrically smooth solids. The original concept of fractals had the aim of calculating the structure's irregularity by using a parameter called fractal dimension D that represents how rough the structure is. This is useful because in our world there are sea waves, mountains, coastlines, clouds; everything surrounding us is irregular if we look close enough.

Perhaps the first natural phenomenon studied by Mandelbrot was the length of coastlines, which is a useful illustration for our fractal analyses of the galaxy distribution [21]. The idea is to be able to measure the length of a coastline using the concept of fractals, since it is not possible to find a unique number representation applying Euclidean geometry. This is because using our intuitive notion of length, a coastline would have many different extension values depending on how much detail it is being measured, and it would get to a point where if every ondulation is considered, the length of the coast would be infinite. In Fig. 1 there is a drawing showing the different lengths the British coastline would have depending on how much detail is being measured using Euclidean geometry. On the left of Fig. 1 there are 169 straight lines each being 25 kilometers long. The middle one has 72 straight lines 50 kilometers long, and the right one has 29 straight lines 100 kilometers long. From this we get the results that the length of the British coastline would be 4,225 kilometers, 3,600 kilometers and 2,900 kilometers from left to right in Fig. 1. This means we could decrease the size of the reference line as much as we wanted and only get to a relative value for the coastal length, since every time we look closer at an edge that is not smooth we will notice new curves. The fractal analyses is more practical due to the lack of this inconsistency problem as fractals have the important property of being self similar, meaning that they have the same irregularity level independent of measurement scales [22]. So, the British coastline would have an absolute value representation of its length, which would be proportional to its roughness.

One of the methods that can be used to calculate the fractal dimension D is called
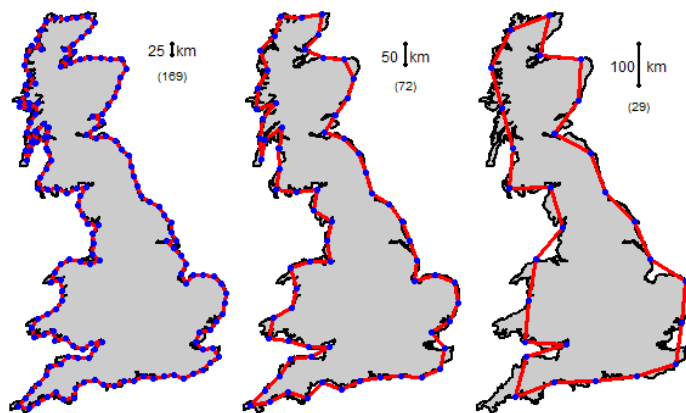
Figure 1. Different British coastline lengths depending on the size of the measured reference line. On the left there are 169 straight lines 25 kilometers long. On the middle there are 72 straight lines 50 kilometers long. On the right there are 29 straight lines 100 kilometers long. The length of the British coastline would be 4,225 kilometers, 3,600 kilometers and 2,900 kilometers from left to right depending on the chosen yardstick. Figure from Ref. [23].

box-counting. A grid is put over the map of Britain and the boxes touching the coastline are scaled up and down, as shown in Fig. 2. From that, a plot is made using the number of boxes touching the coastline in the Y axis and the size of the box in the X axis generating the log log plot shown in Fig. 3. And finally, the slope of the adjusted line can lead directly to the fractal dimension D. It is important to mention that, as a comparison, if the box-counting method were to be applied to a smooth straight line the dimension found would be 1, and if applied to a smooth plane the dimension found would be 2.
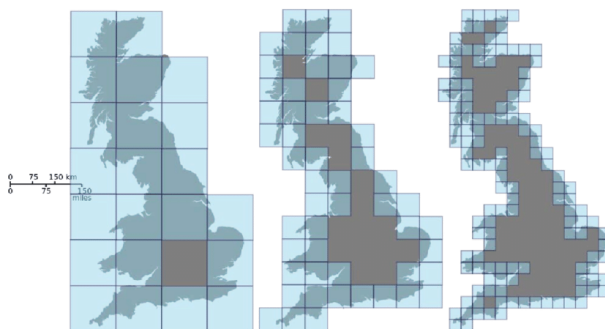
Figure 2. Box-counting method for the length of the British coastline. The boxes are scaled down from left to right. Figure from Ref. [23].
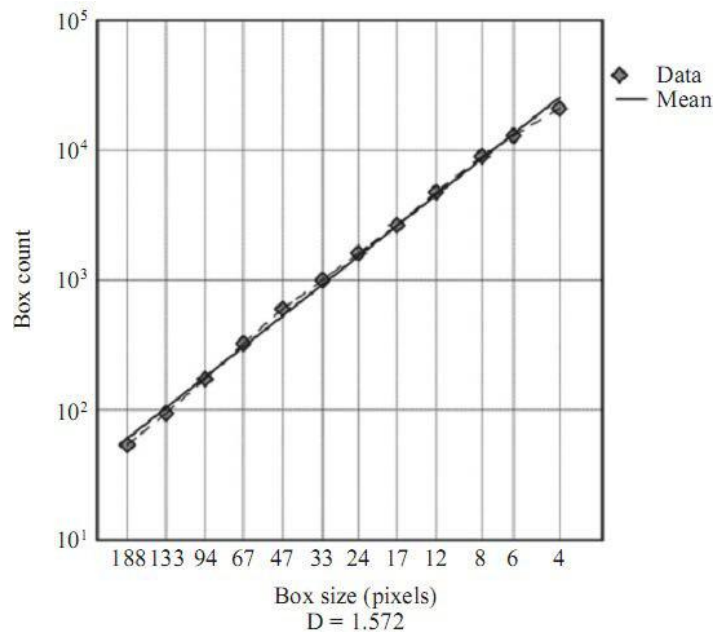
Figure 3. Fractal dimension D of the British coastline from the log log plot of the number of boxes vs. the size of the boxes in the box-counting method. The fractal dimension of the British coastline is D = 1.572. Figure from Ref. [23].

1.2 Fractal Application to Galaxy Distributions

The concept of what we modernly call fractal galaxy distribution actually started in the eighteenth century with Emanuel Swedenbord, who suggested in his book *The Principia* that stellar systems form a non-uniform infinite hierarchical structure [24]. In the book entitled *Two New Worlds* [3], Edmund Fournier d'Albe (1907) approached mathematically the description of stars as a hierarchy of spheres distributed in an infinite space, as shown in Fig. 4, where the mass inside a given sphere is proportional to the radius of this sphere, $M(R) \propto R$, unlike the usual uniform approach where the mass is proportional do the radius cubed.
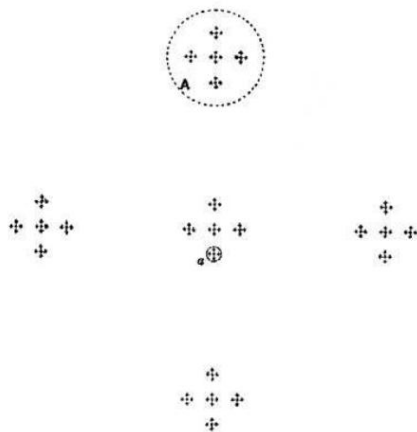
Figure 4. Star distribution by Fournier (1907) [25].

19

Gérard de Vaucouleurs worked on a density-size law using multiple data and found that:

$$log\rho_M^* \simeq -1.9(logR - 10.7), \tag{1.1}$$

where ρ_M^* is the average density of a galaxy cluster of radius R. In view of this, he advanced the following proposition without the logarithm [26]:

$$M(R) \propto R^{3-D}, \tag{1.2}$$

at which the fractal dimension D is associated with the galaxy distribution's inhomogeneity degree.

At this point, the hierarchical concept was lacking mathematical formalism until the books *Fractals: form, chance and dimension* and *The fractal geometry of nature* [6, 27], where Benoit Mandelbrot presented mathematical fractal properties for galaxy distribution, generalizing Einstein's cosmological principle, where D = 3 (spherically homogeneous distribution), to obtain a non-uniform distribution of galaxies with D < 3. Mandelbrot also proposed the "Conditional Cosmological Principle" stating that observers would only perceive the same structure around them if they are observing from a material point like a galaxy but not in vacuum, so the detection of decrease in density would have to be scaled larger than the size of the vacuum [28].

Nevertheless, the galaxy distribution inhomogeneity had its most important observational support with the release of the CfA Redshift Survey in 1988 by means of a famous picture of the galaxy distribution as shown in Fig. 5.
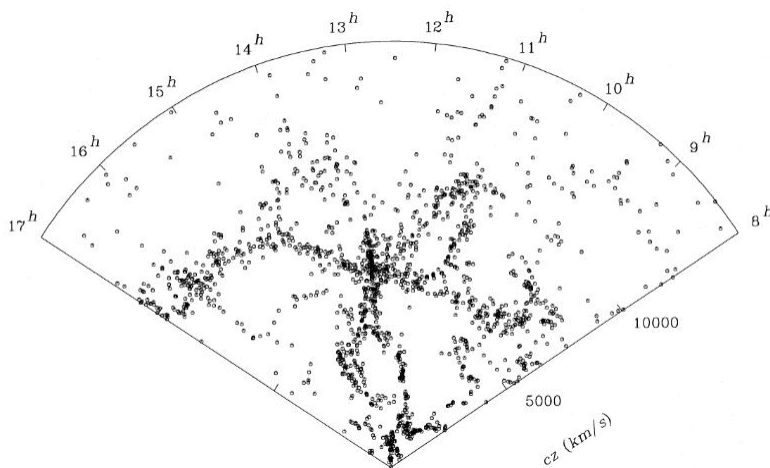
Figure 5. CfA Redshift Survey: Map of galaxy distribution containing 1761 galaxies [29].

Almost simultaneously, Luciano Pietronero proposed the conditional density Γ as a

scaled mean density since in a fractal system the density is dependent on position and volume observed [9]. Such density was defined as

$$\Gamma(r) = \frac{D}{4\pi} B r^{D-3}, \tag{1.3}$$

where B is a scale parameter and r is the distance from the observer.

Relativistic fractal cosmology considers the fact that the expanding universe is curved by the presence of matter, and so the distances are not absolute as in the Newtonian approach. Additionally, our universe can be represented by a light cone, which is the path a light beam would take through spacetime when moving in all directions, as shown in Fig. 6. The FLRW cosmology describes a homogeneous and isotropic universe, which is not in disagreement with a fractal model since the homogeneity of the FLRW cosmological model is spatial and the inhomogeneity of the fractal model is observational, that is, along the spacetime lightcone.
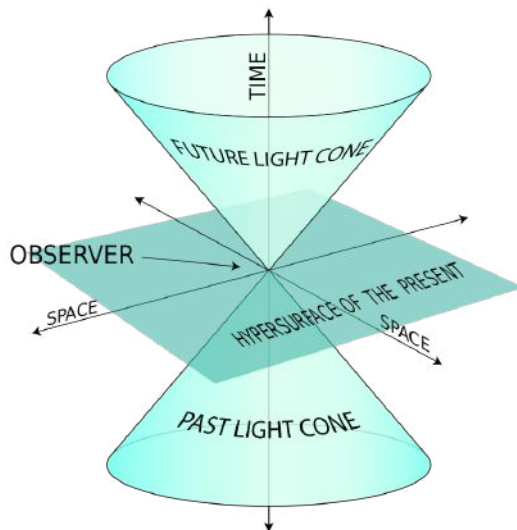
Figure 6. Observer's light cone. The observer is in the origin of the cones as the lower one is the past light cone, which is the light observed from the past, while the upper one is the future light cone, representing the light that is emanating from the observer.

1.2.1 Newtonian Cosmology

In the Newtonian hierarchical fractal cosmology, V_{obs} is defined as the observational volume, d_{obs} is an observational distance, γ^*_{obs} is the observed volume number density and N_{obs} is the observed cumulative number counts of galaxies. From these terms we can write the following

expressions,

$$V_{obs} = \frac{4}{3}\pi(d_{obs})^3, \tag{1.4}$$

$$\gamma^*_{\mathrm{obs}} = \frac{N_{\mathrm{obs}}}{V_{\mathrm{obs}}}. \tag{1.5}$$

The *Pietronero-Wertz hierarchical (fractal) cosmology* [22, 30] starts by assuming a critical phenomenological expression for the galaxy observed cumulative number counts N_{obs} that depends on the distance d_{obs} [1] via the power law called the *number-distance relation* given by

$$N_{\mathrm{obs}} = B\,(d_{\mathrm{obs}})^D, \tag{1.6}$$

where B is a scale parameter and D is the single fractal dimension.[2] Such relation gives the number of sources per unit of observational volume out to a distance d_{obs}. Additionally, N_{obs} is a radial quantity, so, it is not understood in statistical sense considering that it does not average all points against all points. From Eqs. (1.4), (1.5) and (1.6) we define the *De Vaucouleurs density power-law* γ^*_{obs} for a fractal structure as

$$\gamma^*_{\mathrm{obs}} = \frac{N_{\mathrm{obs}}}{V_{\mathrm{obs}}} = \frac{3B}{4\pi}(d_{\mathrm{obs}})^{D-3}, \tag{1.7}$$

where the term $3B/4\pi$ is a positive constant [9, 30].

If in the previous expression $D = 3$, N_{obs} increases with V_{obs} and galaxies would be evenly distributed along all regions of the observed space. However, if $D < 3$, as d_{obs} increases N_{obs} grows at a smaller rate than V_{obs}, forming then gaps in the galactic distribution, in other words, regions devoid of galaxies. Close to these galactic gaps there would be regions where galaxies clump. Hence, voids and galactic clumpiness would be a consequence of a fractal galaxy structure whose fractal dimension is smaller than the topological dimension in which the galaxy structure is embedded. In this case the fractal dimension becomes a descriptor of galactic clumpiness or of void dominance in the galactic structure.

Since N_{obs} is a cumulative quantity, it does not increase with d_{obs} if beyond a certain distance there are no longer galaxies. However, if objects are still detected and counted, even

---

[1]Since this is a galaxy distribution scale, the minimum observational distance is the order of the local galaxy group and the maximum distance will depend on the observational samples.
[2]See the Appendix B for the heuristic derivation of Eq. (1.6)

at irregular rate, then N_{obs} continues to increase. This amount of growth can be influenced by observational biases, possibly leading to an irregular behavior, however N_{obs} is required to grow or remain constant and, so, the exponent in Eq. (1.6) must be positive or zero. Therefore, if the galaxy distribution fractal dimension is $D < 3$, then the observational number density decays as a power-law. If $D = 3$ galaxies are uniformly distributed and the galactic structure is called *observationally homogeneous*. In this instance the number density is constant and distance independent. Lower values of D imply a more abrupt power-law decay and, as a consequence, more gaps in the galaxy distribution.

1.2.2   Relativistic Cosmology

In relativistic cosmology the observer's past light cone is the geometrical locus of observations, which means that even spatially homogeneous cosmologies like the FLRW do not generate observationally constant number densities at moderate and high redshift ranges, since observational and spatial homogeneities are different relativistic concepts in cosmology. So, $\gamma_{\mathrm{obs}}^{*}$ is an average relativistic density. In addition, in general relativity distances may vary, especially for $z > 1$, a situation which indicates that at high redshift d_{obs} would have different values for each distance definition at the same redshift z. This means that as distance in cosmology is not uniquely defined d_{obs} must be replaced by d_i in the previous equations of this chapter [13, 31, 32, 33, 34].

The distance definitions applicable to relativistic fractal cosmologies are the luminosity distance d_L, the redshift distance d_Z and the galaxy area distance d_G, also known as transverse comoving distance. The luminosity distance is defined according to,

$$F = \frac{L}{4\pi d_L^2} \, , \tag{1.8}$$

where F is the flux density of the object and L is its total amount of electromagnetic energy emitted per unit of time. From an operational and numerical viewpoint, d_L comes from the python parameter ΛCDM from the package astropy.cosmology with the following settings: Hubble constant $H_0 = 70 km \cdot s^{-1} \cdot Mpc^{-1}$, Matter density parameter $\Omega_{m0} = 0.3$ and Dark energy density parameter $\Omega_{\Lambda 0} = 0.7$, for the FLRW cosmology assumed.

Using d_L we defined d_G considering the Etherington reciprocity law,

$$d_G = \frac{d_L}{1 + z}. \tag{1.9}$$

23

The redshift distance is defined as follows,

$$d_z = \frac{c\,z}{H_0},\tag{1.10}$$

where c is the speed of light in vacuum and H_0 is the Hubble constant. The previous expressions must be rewritten as below in order to become suitable for relativistic cosmologies,

$$d_{\text{obs}} = d_i,\tag{1.11}$$

$$V_{\text{obs}} = V_i = \frac{4}{3}\pi(d_i)^3,\tag{1.12}$$

$$N_{\text{obs}} = N_i = B_i\,(d_i)^{D_i},\tag{1.13}$$

$$\gamma^*_{\text{obs}} = \gamma^*_i = \frac{N_i}{V_i} = \frac{3B_i}{4\pi}(d_i)^{D_i-3},\tag{1.14}$$

where $i = (L,\, z,\, G)$ corresponding to the chosen distance definition, which is attached to each specific constant B_i and fractal dimension D_i. This is because N_i is counted considering the limits given by every particular distance definition, so for a given z each d_i will generate its respective V_i, N_i, B_i and D_i. Therefore, every quantity is connected to a certain distance definition.

Subsequently, to calculate the fractal dimension D_i it was necessary the linearization of the graph γ^*_i vs. d_i so we plotted the log-log results for γ^*_i vs. d_i to obtain a value for the slope of the points and used the relation in the following equation:

$$D_i = 3 + \text{ slope },\tag{1.15}$$

where D_i $(i = L,Z,G)$.

2.   Galaxy Surveys Description

The three catalogues chosen for the study are presented in this chapter. There are two important features that were taken into account for their choice: high values of redshift, which is needed for our relativistic analysis, and a large number of galaxies, due to the statistical nature of the fractal analysis. In Fig. 7 there is a histogram showing the galaxy distribution numbers in terms of redshift for the UltraVISTA DR1 [18], the COSMOS2015 [19] and the SPLASH [20] surveys.
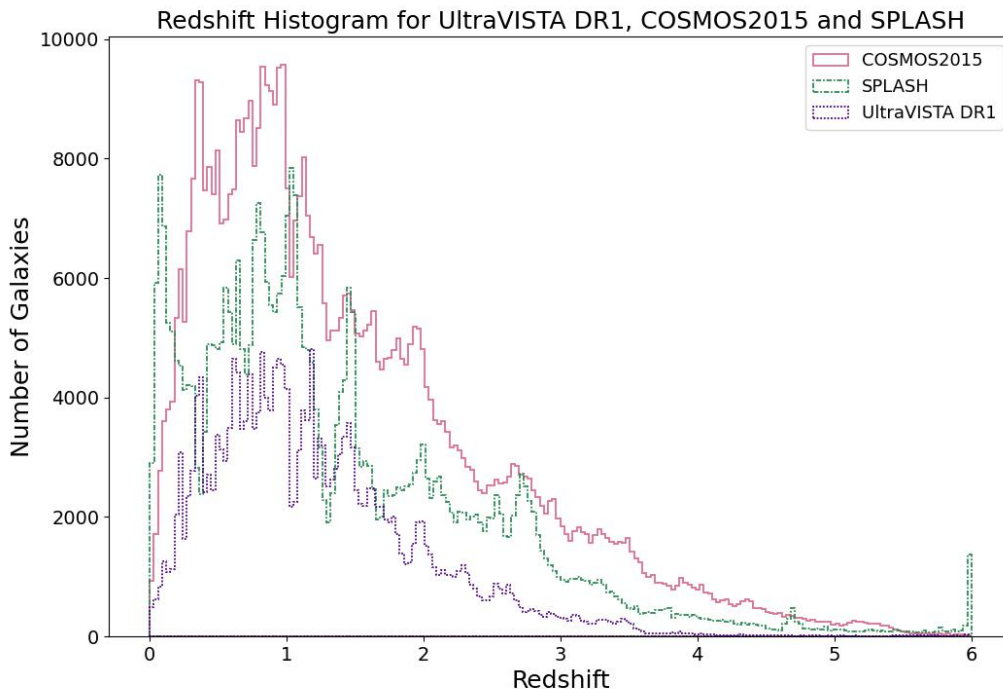
Figure 7. Histogram showing the galaxy distribution numbers in terms of redshift for the UltraVISTA DR1, COSMOS2015 and SPLASH surveys datasets studied here. This graph has $\Delta z \approx 0.03$ as the redshift bins' size.

2.1   UltraVISTA DR1

The first data release of the UltraVISTA galaxy survey is centered on the COSMOS field [35] with an effective area of $1.5 \deg^2$. The observations were made in four near infrared filters, Y, J, H and K_S, described in Ref. [18]. After applying the SED fitting technique to 29 bands, the photometric redshifts were calculated by Ref. [36] using the ones from UltraVISTA and another set of broad and narrow bands from other surveys covering the mid infrared, near

infrared, optical and ultraviolet regimes. The original dataset consisted in a sample of 219300 galaxies in the redshift range of $0.1 \leq z \leq 6$. In Fig. 8 there is a histogram of the galaxy distribution numbers in terms of right ascension and declination for the UltraVISTA DR1 survey, which are $1.61 \leq \text{DEC (deg)} \leq 2.81$ and $149.30 \leq \text{RA (deg)} \leq 150.78$, with limiting apparent magnitude of $K_S = 24.0$.

Figure 8. On the left is the histogram showing the galaxy distribution numbers in terms of right ascension (deg) and on the right is the histogram showing the galaxy distribution numbers in terms of declination (deg), both for the UltraVISTA DR1 survey dataset studied here.

2.2 COSMOS2015

The COSMOS2015 catalog includes infrared data from the Spitzer Large Area survey over 2 deg^2 in the COSMOS field, Y-band images from Subaru/Hyper-Suprime-Cam and YJHK$_S$ images from the UltraVISTA DR2 [35]. The UltraVISTA DR2 region contains 578379 galaxies where the object detection is performed by the χ^2 sum of the YJHK$_S$ and z^{++} images. This means that the sample has 359079 new objects as compared to the UltraVISTA DR1. The observed area has the range of coordinates as ilustrated in Fig. 9, which are: $1.43 \leq \text{DEC (deg)} \leq 3.05$ and $149.28 \leq \text{RA (deg)} \leq 151.05$, with hundreds of thousands objects at redshift $0.1 \leq z \leq 6$ and limiting apparent magnitude of $K_S = 24.0$ at 3σ in a 3" diameter aperture and parts of the field covered by the "ultra-deep stripes" (0.62 deg^2) have limiting apparent magnitude $K_S = 24.7$ at 3σ in a 3" diameter. The photometric redshifts were computed using LePHARE [37, 38] following Ref. [36].

Figure 9. On the left is the histogram showing the galaxy distribution numbers in terms of right ascension (deg) and on the right is the histogram showing the galaxy distribution numbers in terms of declination (deg), both for the COSMOS2015 survey dataset studied here.

2.3 SPLASH

The SPLASH survey is a multi-wavelenght photometry catalog with coverage area of $2.4\,\mathrm{deg}^2$. The sources were identified using a detection image defined as a χ^2 combination of ugri images from CFHT Legacy Survey (CFHTLS), u image from Megacam Ultra-deep Survey: U-Band Imaging (MUSUBI), YJHKs images from VISTA Deep Extragalactic Observations (VIDEO), JHK images from Ultra Deep Survey (UDS) DR11 and grzy images from Hyper-Suprime-Cam (HSC) DR1. This survey follows a different strategy for the detection of the objects, as it includes the u-band, in order to recover the bluest objects. This catalog contains 390362 galaxies at $0 < z < 6$, where the redshifts were measured using LePhare with a similar method to that used for the COSMOS field [36, 19]. In Fig. 10 there is a histogram of the galaxy distribution numbers in terms of right ascension and declination for the SPLASH survey, which are $-6.04 \leq \mathrm{DEC}\,(\mathrm{deg}) \leq -3.96$ and $33.45 \leq \mathrm{RA}\,(\mathrm{deg}) \leq 35.55$, with limiting apparent magnitude of $K = 24.7$. The SPLASH galaxies were mapped in a different region of the sky and it follows a different strategy for object detection as compared to the other two surveys, which validates its inclusion in the fractal analyses, along with the sizable portion of galaxies in the redshift interval $1 < z < 4$ as shown in Fig. 7.
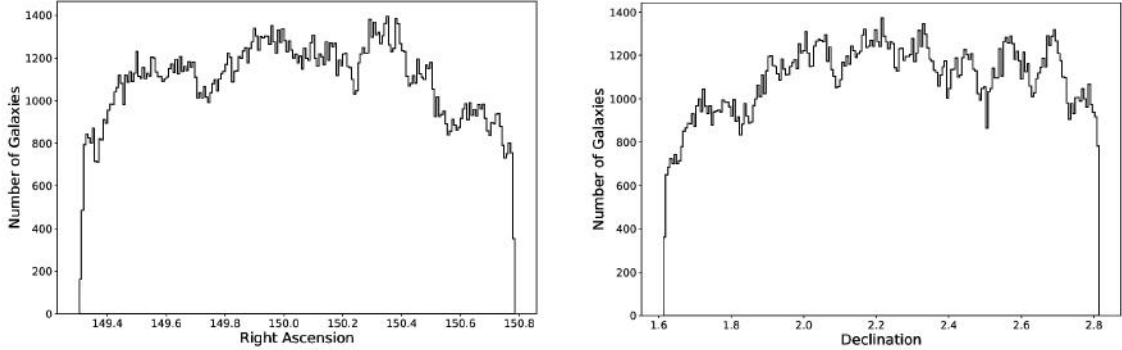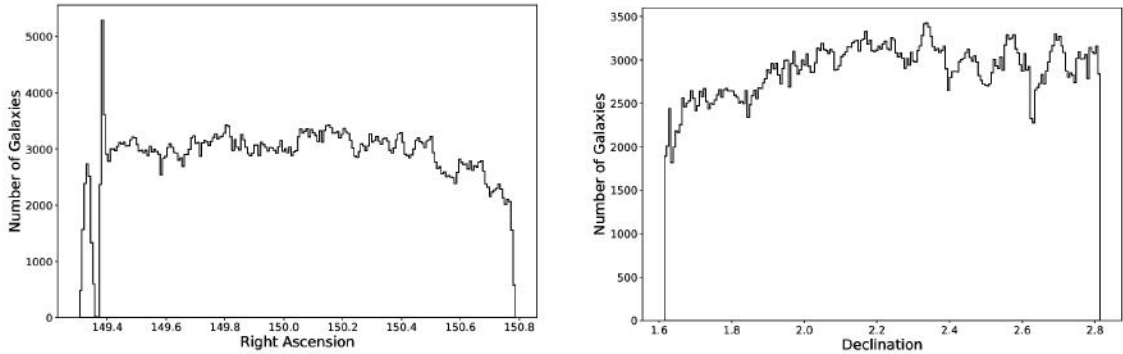
27
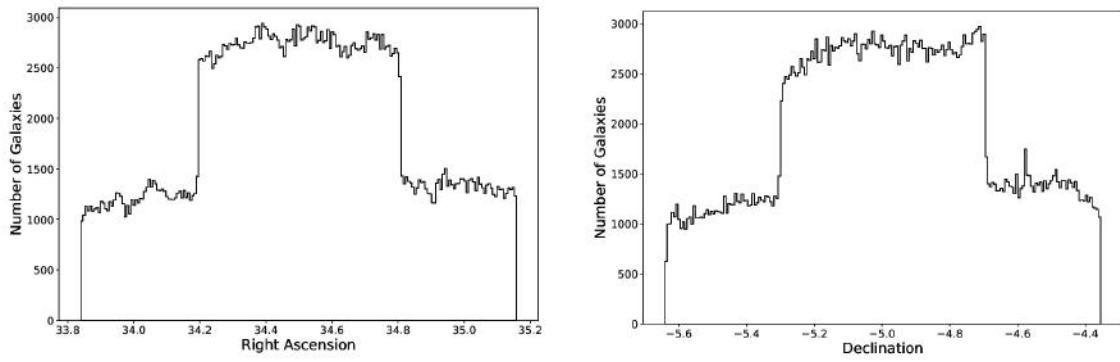
Figure 10. On the left is the histogram showing the galaxy distribution numbers in terms of right ascension (deg) and on the right is the histogram showing the galaxy distribution numbers in terms of declination (deg), both for the SPLASH survey dataset studied here.

3. Fractal Analysis

As previously discussed, this study aims at testing whether or not the fractal galaxy distribution hypothesis holds at very large scales of the observed Universe, hence we chose to perform a fractal analysis with the data provided by the UltraVISTA DR1, COSMOS2015 and SPLASH galaxy surveys exactly because each of them contains hundreds of thousands of galaxies with measured high redshift values.

3.1 UltraVISTA DR1

3.1.1 Data selection

The establishment of volume limited samples is needed for the fractal analytical approach detailed in the first chapter. Knowing that galaxy surveys are limited by apparent magnitude, we tested if reduced subsamples would follow increasing redshift bins for the sake of rendering the reduced galaxy data distributed over limited volume bins. The data reduction strategy was to select galaxies below a certain absolute magnitude threshold determined by the limiting apparent magnitude on the survey. This was made on the plot of absolute magnitude in terms of their measured redshifts using the following expression:

$$M = m - 5 \log d_L(z) - 25, \tag{3.1}$$

where M is the absolute magnitude, m is the apparent magnitude and d_L is the luminosity distance given in Mpc. Afterwards we chose the apparent magnitude threshold m, its passband and verified if the resulting data is distributed along the measured redshift bins to establish a redshift window range for the final subsample distribution.
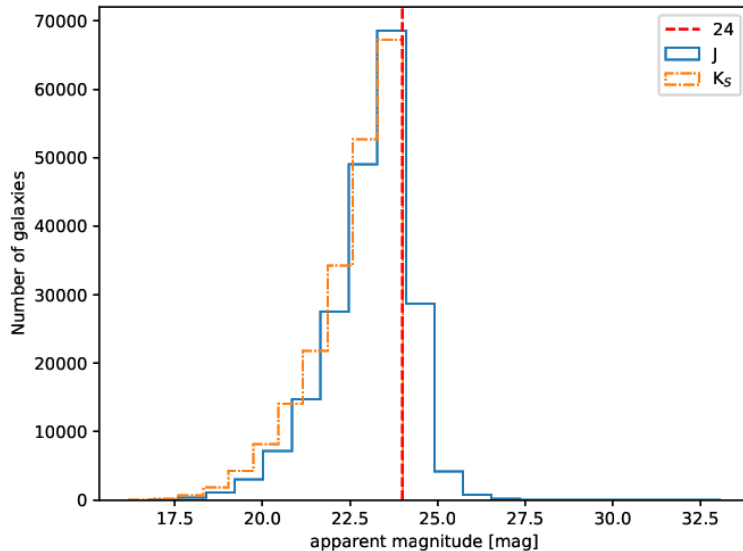
Figure 11. Histogram of galaxy distribution *vs* apparent magnitude (mag) for the UltraVISTA DR1 survey dataset studied here.

The UltraVISTA DR1 survey provided absolute magnitudes calculated in the NUV, B, R and J passbands. Since the near infrared filter J is less affected by dust extinction it is the best choice for our objectives here. In Fig. 11 it is shown an histogram of the UltraVISTA DR1 galaxy numbers in terms of apparent magnitudes in the K_S and J passbands where we can see that the number distribution peaks at the apparent magnitude value 24 for both wavebands, an information that led us to adopt $J = 24$ as our apparent magnitude threshold. So, Eq. (3.1) can be rewritten according to the following expression, which provides the dividing line between the selected and unselected data.

$$M_J = 24 - 5 \log d_L(z) - 25. \tag{3.2}$$

Fig. 12 displays a plot of the 219300 UltraVISTA galaxies' absolute magnitudes in the J-band against their respective measured redshifts. Green filled circles are those whose absolute magnitudes supplied by the survey are smaller than M_J in Eq. (3.2), and are then a part of the reduced subsample, while the open blue circles are outside this threshold.

30
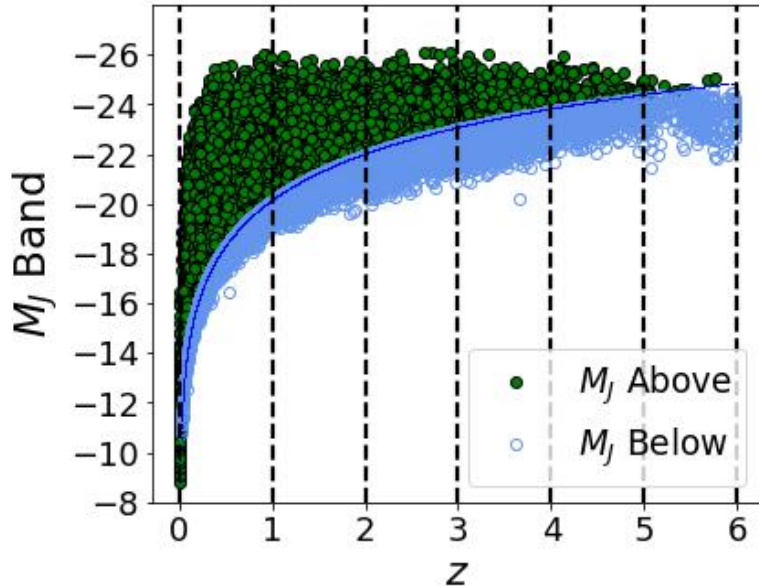
Figure 12. Plot of the absolute magnitudes for the UltraVISTA DR1 galaxies in terms of their photometrically measured redshift values. The dividing line corresponds to apparent magnitude $J = 24$ and all galaxies having M_J above this cutoff and $z \leq 4$ were included in the subsample.

Since the data generally follow the redshift bins of increasing values we noticed that the absolute magnitude cut based on the J-band is consistent with a volume-limited subsample. Additionally, we noticed that there are fewer galaxies at the tail of the distribution, so our subsample went through another cut at $z = 4$. Hence, we came to a *reduced UltraVISTA DR1 subsample* of 166566 galaxies cut by J-band absolute magnitudes M_J and restricted up to $z = 4$. The remaining 52734 galaxies outside this subsample, that is, below the dividing line in Fig. 12 or having $z > 4$, were not considered for further analysis.

3.1.2 Data Analysis

In the python algorithm we chose to separate the data using a fixed distance interval $\Delta d_i = 200$Mpc, and from that, the volume ΔV_i was determined and the number of galaxies N_i in this volume was calculated. We also defined the volume density γ_i^* as in Eq. (1.14). The algorithm in the Appendix A was designed to calculate the cumulative value of Δd_i, ΔV_i, N_i and γ_i^*, so the first distance interval is up to 200Mpc away from earth, the second interval is up to 400Mpc and so on. We tested different fixed distance intervals in the data analyses to check if it would alter the distribution of the points in the graph γ_i^* vs. d_i and it did not, as seen in Fig. 13.
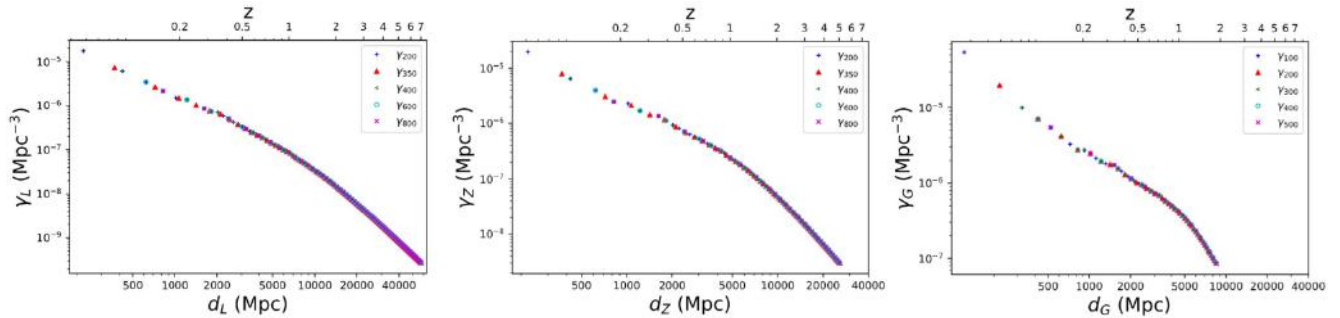
Figure 13. Graph showing the results for γ_L^* vs. d_L, γ_Z^* vs. d_Z and γ_G^* vs. d_G, from left to right, obtained with the UltraVISTA DR1 galaxy redshift survey dataset for distance bin values varying from 200Mpc to 800Mpc.

Now one needs to choose a cosmological model to acquire the observational distances d_i ($i = G$, L, z) from the calculated photometric redshift values. We assumed the FLRW cosmology with $\Omega_{m_0} = 0.3$, $\Omega_{\Lambda_0} = 0.7$ and $H_0 = 70$ km s^{-1} Mpc^{-1}. The next steps were to establish the minimum redshift value z_0 from where the analysis starts, the respective minimum distances $d_{i_0} = d_{i_0}(z_0)$, and the incremental distance interval Δd_i. The process of sorting out the data was initiated by counting the number of observed galaxies N_{i_1} in the first interval between d_{i_0} and $d_{i_1} = d_{i_0} + \Delta d_i$ and calculating the respective volume density $\gamma_{i_1}^*$. The first bin was defined by this first interval. Next, the size of the bin was increased by Δd_i and cumulative values for N_{i_2} and $\gamma_{i_2}^*$ were acquired at the distance $d_{i_2} = d_{i_0} + 2\Delta d_i$. This algorithm was reproduced n times until the farthest group of galaxies were included and all related quantities were also counted and calculated. In order to find out whether or not different bin size increments Δd_i affect the results we tested many values for each distance definition. This test turned out negative, meaning that the obtained results are independent of bin size increment. Then, we chose the value $\Delta d_i = 200$ Mpc which was implemented in all calculations and presented a very reasonable amount of data points for each quantity from where simple linear regression analyses were ready to be executed. The determination of the fractal dimension itself was the next and final step. According to Eq. (1.14) if the galaxy distribution really formed a fractal system, the graphs of γ_i^* vs. d_i would behave as decaying power-law curves and the fractal dimension D_i of the distribution would be determined by the linear fit slopes in the log-log plots.

3.1.3 Results

Reduced subsample

The reduced UltraVISTA DR1 galaxy survey subsample characterized according to the criteria set at Sec. 3.1.1 agreed, to a good approximation, with the predictions that the galaxy distribution does form a fractal system as seen in graphs for log-log values of γ_i^* vs. d_i. In addition, two power-law decaying regions were noted in the data, for $z < 1$ and $z > 1$,

generating different single fractal dimensions. The results for all distance definitions adopted here are in Fig. 14, from where it can be determined that for $z < 1$ the fractal dimension is in the range $1.38 - 1.78$, while for $1 \leq z \leq 4$ the resulting range is fairly smaller, $0.31 - 0.87$. Table 1 assemble these results.

Table 1. Results for the UltraVISTA DR1 galaxy survey fractal analysis in two redshift scales in the *reduced* subsample. The single fractal dimensions D_L, D_z and D_G were acquired from this galaxy distribution respectively using the luminosity distance d_L, redshift distance d_z and galaxy area distance (transverse comoving distance) d_G.

	D_L	D_z	D_G
$z < 1$	1.40 ± 0.02	1.61 ± 0.02	1.75 ± 0.03
$1 \leq z \leq 4$	0.32 ± 0.01	0.38 ± 0.02	0.81 ± 0.06
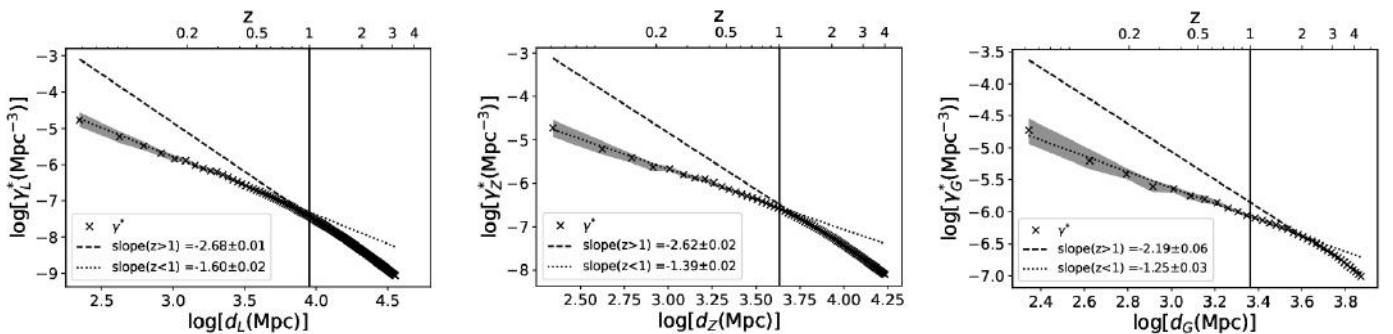
Figure 14. Log-log plots of γ_L^* vs. d_L, γ_Z^* vs. d_Z and γ_G^* vs. d_G, from left to right, obtained with the UltraVISTA DR1 galaxy redshift survey dataset. The dotted line is the linear fit for galaxies at $z < 1$, while the dashed line represents those with $1 \leq z \leq 4$. The grey area illustrates the 1 sigma error of the bins.

These results indicate, in summary, that the UltraVISTA DR1 galaxy survey presents a galaxy distribution dataset that can be interpreted as a fractal system with the following approximate single fractal dimensions values: $D\,(z < 1) = (1.58 \pm 0.20)$ and $D\,(1 \leq z \leq 4) = (0.59 \pm 0.28)$. The potential reasons to why the fractal dimension is so much reduced at the deepest range $z > 1$ will be discussed later.

Complete (unselected) survey data

The fractal analysis of the UltraVISTA DR1 galaxy subsample was established by a plot of absolute magnitudes in the *J*-band, as defined in Sec. 3.1.1, in terms of galaxies measured redshifts, as shown in Fig. 12. However, this plot suggests that the whole sample may be volume-limited, since even galaxies outside the absolute magnitude cut are distributed along increasing redshift bins. So, it is pertinent to apply the same fractal methodology developed here to the whole survey data, not only to the subsample, in order to compare the results.

Table 2. Results for the UltraVISTA DR1 galaxy survey fractal analysis in two redshift scales in the *complete* subsample. The single fractal dimensions D_L, D_z and D_G were acquired from this galaxy distribution respectively using the luminosity distance d_L, redshift distance d_z and galaxy area distance (transverse comoving distance) d_G.

	D_L	D_z	D_G
$z < 1$	1.44 ± 0.02	1.65 ± 0.02	1.80 ± 0.03
$1 \leq z \leq 6$	0.24 ± 0.01	0.30 ± 0.02	0.75 ± 0.06
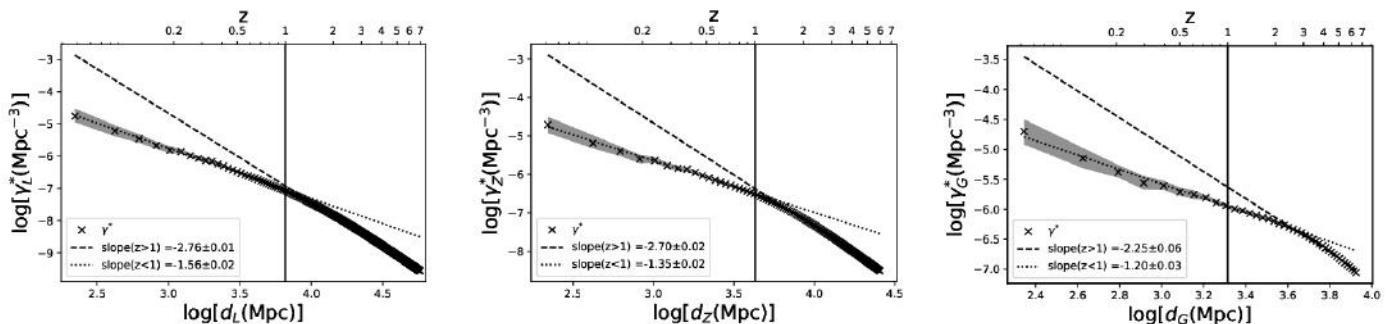
Figure 15. Log-log plots of γ_L^* vs. d_L, γ_Z^* vs. d_Z and γ_G^* vs. d_G, from left to right, obtained with the UltraVISTA DR1 galaxy redshift survey dataset. The dotted line is the linear fit for galaxies at $z < 1$, while the dashed line represents those with $1 \leq z \leq 7$. The grey area illustrates the 1 sigma error of the bins.

The results for the fractal analysis of the complete UltraVISTA DR1 survey data are in Fig. 15. This dataset of galaxy distribution clearly presents fractal features in both regions, above and below the redshift value $z = 1$. The respective single fractal dimensions for each distance definitions are in the range $1.42 - 1.83$ for $z < 1$, while for $1 \leq z \leq 6$ the dimension is significantly smaller, in the range $0.23 - 0.81$. Table 2 assembles these results and show that they are very much alike with the values from Table 1 for the reduced subsample, although they do not overlap even considering their respective uncertainties. So, it is evident that the whole UltraVISTA DR1 galaxy survey can also be described as a fractal system having two consecutive single fractal dimensions: $D(z < 1) = (1.63 \pm 0.20)$ and $D(1 \leq z \leq 6) = (0.52 \pm 0.29)$.

3.2 COSMOS2015 & SPLASH

3.2.1 Data Selection

As stated above, fractal analysis of galaxy surveys demands arranging data along volume-limited distributions. Due to the fact that galaxy surveys are limited by apparent magnitude, we need to continue by reducing the data into subsamples so that they follow

increasing redshift bins. This is done the same way as in Sec. 3.1.1 using Eq.(3.1) and assuming the FLRW cosmology with $\Omega_{m_0} = 0.3$, $\Omega_{\Lambda_0} = 0.7$ and $H_0 = (70 \pm 5)$ km s^{-1} Mpc^{-1}. A homogeneous volume-limited subsample could be created by assuming the apparent magnitude threshold in the K band to be $K = 24.7$, which is a mean limiting value acceptable to both surveys [1], and by changing Eq. (3.1) to the expression below,

$$M_K = 24.7 - 5 \log d_L(z) - 25. \tag{3.3}$$

From this expression we get the cutoff line between the selected and unselected galaxies. Finally, we took into consideration the uncertainty in the Hubble constant in order to test whether our results would be affected by its error margin. Fig. 16 shows the data selection of the COSMOS2015 survey obtained according to the method described previously. Only galaxies with absolute magnitudes M_K above the cutoff line were selected as part of the subsample for further analysis. In addition, since galaxies having $z > 4$ are in an amount too small to be considered representative, as we can see in Fig. 7, we disregarded them. The outcome of this process was the creation of three subsamples containing 230705 galaxies in the redshift range $0 < z \leq 4$ from the original 578379 objects for three values of the Hubble constant within its uncertainty.
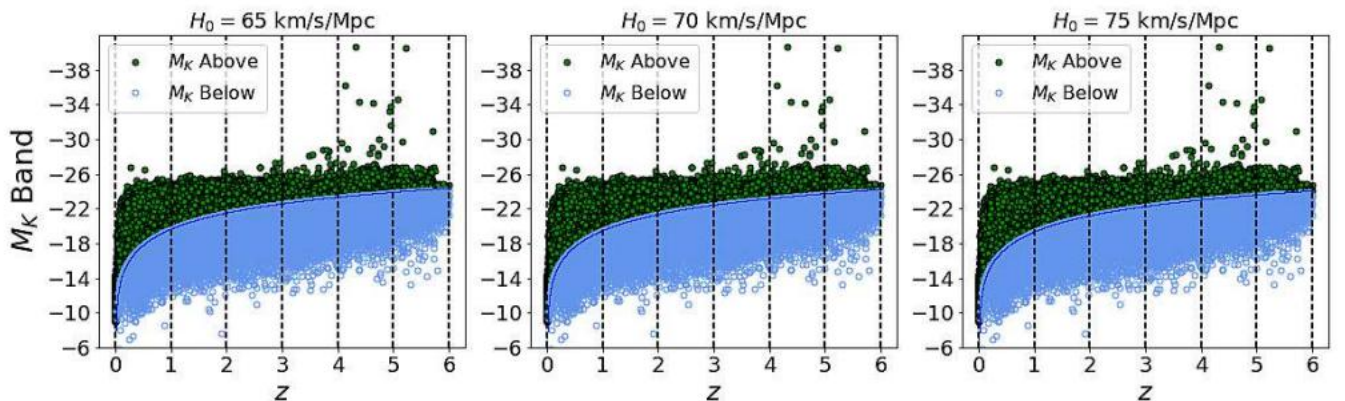
Figure 16. Plot of the absolute magnitudes for the COSMOS2015 galaxies in terms of their photometrically measured redshift values. The dividing line corresponds to apparent magnitude $K_S = 24.7$ and all galaxies having M_K above this cutoff and $z \leq 4$ were included in the subsample that assumed three different values of the Hubble constant within its uncertainty.

Fig. 17 shows the same selection procedure executed with the SPLASH survey dataset also generating three other subsamples. The original 390362 objects were then reduced to 171548 galaxies.
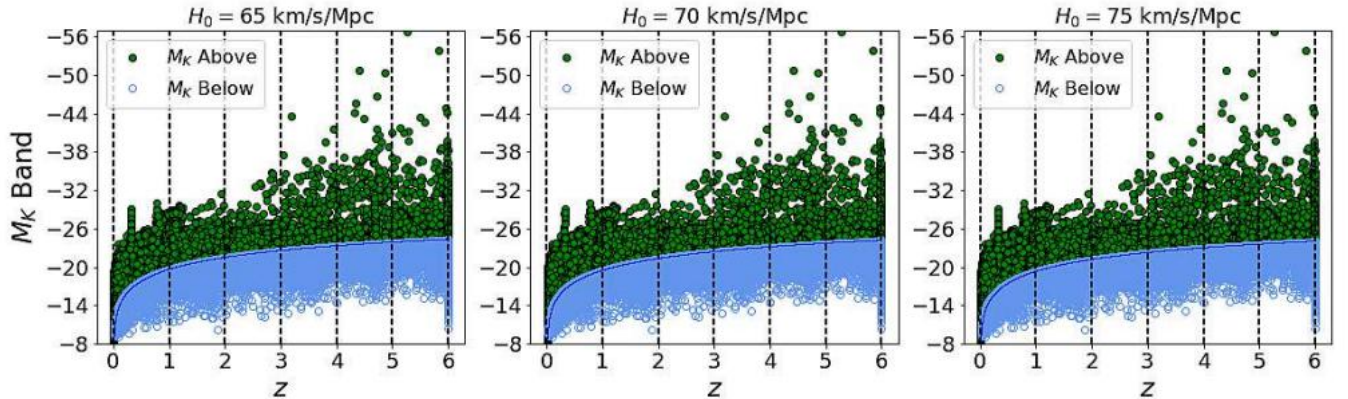
Figure 17. Plot of the absolute magnitudes for the SPLASH galaxies in terms of their photometrically measured redshift values. The dividing line corresponds to apparent magnitude $K = 24.7$ and all galaxies having M_K above this cutoff and $z \leq 4$ were included in the subsample that assumed three different values of the Hubble constant within its uncertainty.

3.2.2 Data Analysis

The FLRW cosmology parameters allowed the calculation of the observational distances d_i ($i = G, L, z$) using theoretical expressions relating distance to redshift given by this cosmological model. Together with the photometric redshift values supplied by the galaxy surveys these observational distances can be calculated. The data sorting algorithm necessary to implement a fractal analysis under the theoretical model discussed in Sec. 1.2.2 is the same as described in Sec. 3.1.2. For each subsample, we start by setting the minimum redshift z_0, its respective minimum distance $d_{i_0} = d_{i_0}(z_0)$, and the incremental distance interval Δd_i. The value $\Delta d_i = 200$ Mpc was kept in all cases, as provides a large amount of data elements for all quantities involved, allowing enough points to perform an adequate regression analysis. Then, the algorithm establishes the first bin by counting the number of galaxies N_{i_1} within d_{i_0} and $d_{i_1} = d_{i_0} + \Delta d_i$, and calculating its respective volume density $\gamma_{i_1}^*$. The next step is to increase the bin size by Δd_i, leading to a distance of $d_{i_2} = d_{i_0} + 2\Delta d_i$. At a interval between d_{i_0} and d_{i_2}, we count the number of galaxies N_{i_2} and estimate $\gamma_{i_2}^*$. These steps are repeated n times until all galaxies are included. Finally, if the galaxy distribution truly formed a fractal system, the behavior of γ_i^* vs. d_i plots would be as decaying power-law curves according to Eq. (1.14). So the linear fit slopes in log-log plots would make it possible to directly determine the fractal dimensions D_i of the distribution.

3.2.3 Results

Figs. 18 and 19 show a grid of the γ_i^* vs. d_i log-log graphs with their respective Hubble constant values choices for the COSMOS2015 and SPLASH surveys. The results reveal that the galaxy distribution in both surveys display power-law decays in two scale ranges: for $z < 1$

and $1 \leq z \leq 4$, which is coherent with a fractal system presenting two single fractal dimensions at different distance ranges. Using Eq. (1.15) we can simply calculate the fractal dimension in both scaling ranges from the slopes of the fitted straight lines. These results obtained with this procedure are shown in Tables 3 and 4, where one finds two single fractal systems in sequence at different data ranges with decreasing values for D as the redshift increases. Hence, the results confirm the theoretically predicted decay in the fractal dimension the deeper one goes in the surveys (see above in the Introduction). In addition, the fractal dimensions are unaffected by variations in the Hubble constant as demonstrate by the estimated values being within 1 sigma error in both COSMOS2015 (Table 3) and SPLASH (Table 4). Summarizing the results of Tables 3 and 4 by rounding off the main values and uncertainties around their medians, we can determine that for $z < 1$ the COSMOS2015 survey presented $D = 1.4 \pm 0.2$, while the SPLASH galaxies generated $D = 1.0 \pm 0.1$. For $1 \leq z \leq 4$ we respectively found $D = 0.5 \pm 0.3$ and $D = 0.8 \pm 0.4$. Evidently the SPLASH galaxies obtained fractal dimensions somewhat smaller than the COSMOS2015 ones for $z < 1$, but the opposite was true for $z > 1$, even though with overlapping uncertainties.

Next, there will be a discussion for the potential reasons for such differences and a comparison with previous studies. Finally, Table 5 contain the comparison of the fractal dimensions D calculated with the UltraVISTA DR1, COSMO2015, SPLASH and FDF (see Ref. [17]) surveys when adopting $H_0 = 70$ km/s/Mpc.
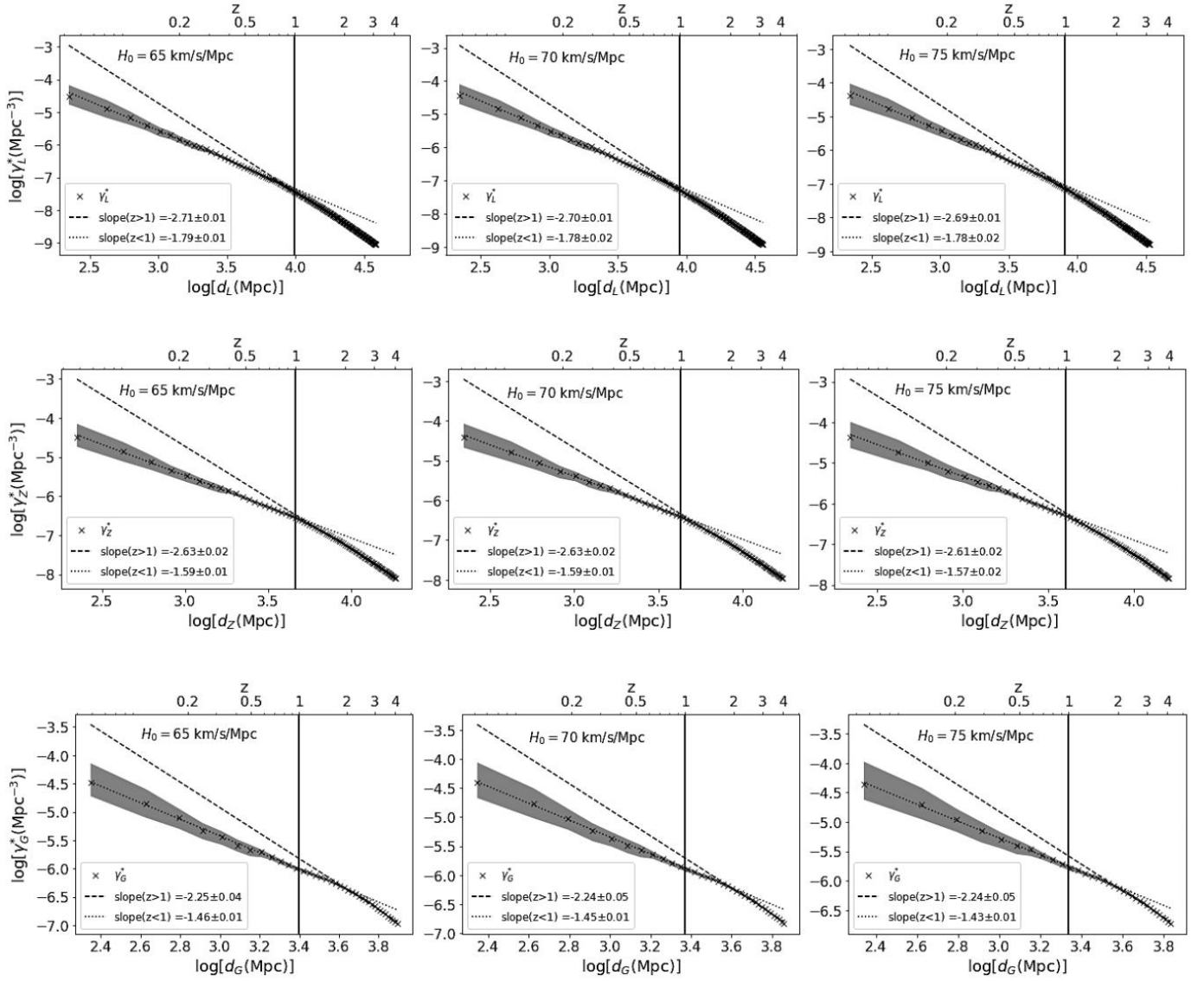
Figure 18. Log-log plots of γ_L^* vs. d_L (top), γ_Z^* vs. d_Z (middle) and γ_G^* vs. d_G (bottom), with three different values of the Hubble constant, obtained with the COSMOS2015 galaxy redshift survey dataset. The dotted line is the linear fit for galaxies at $z < 1$, while the dashed line represents those with $1 \leq z \leq 4$. The grey area illustrates the 1 sigma error of the bins.
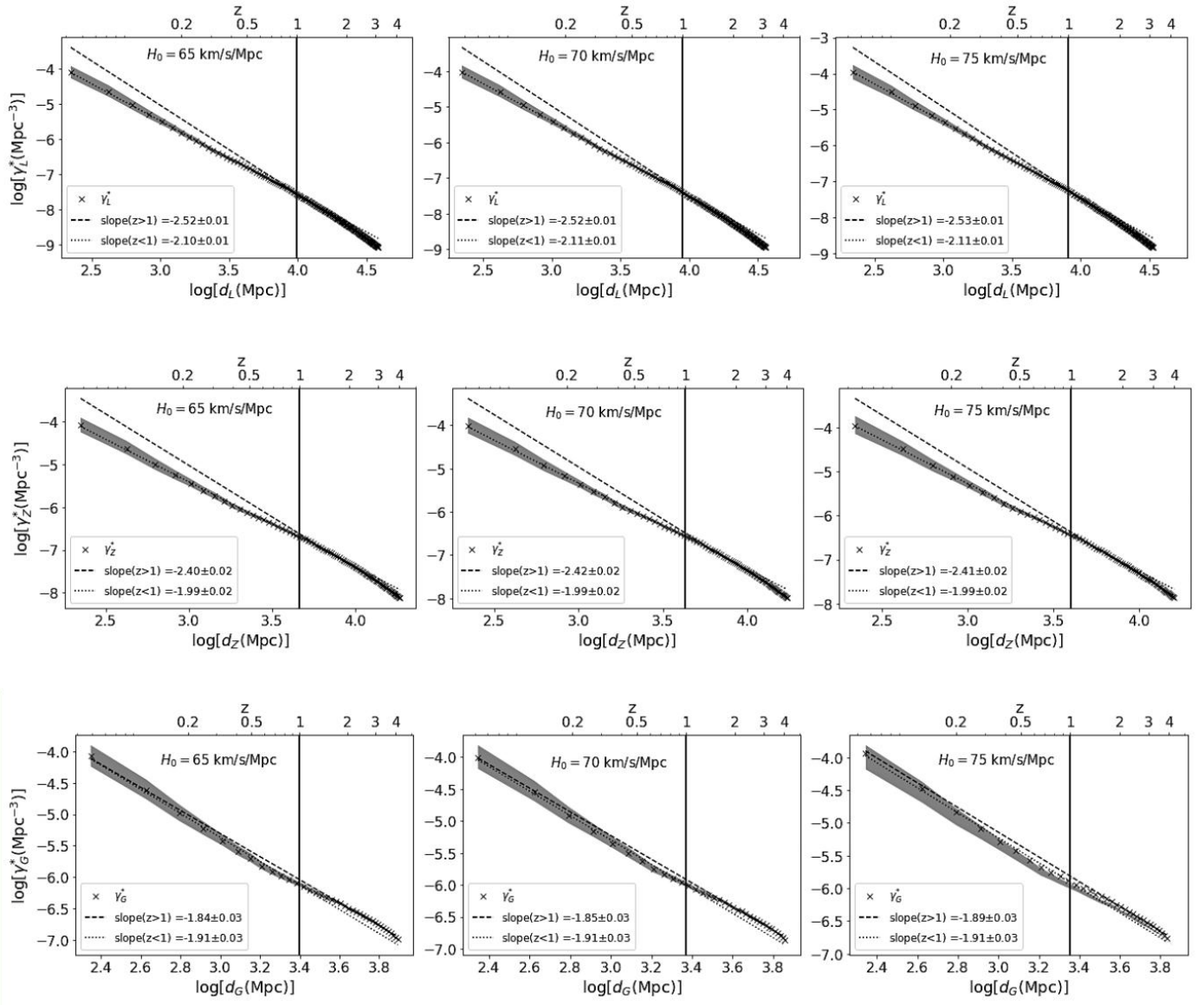
Figure 19. Log-log plots of γ_L^* vs. d_L (top), γ_Z^* vs. d_Z (middle) and γ_G^* vs. d_G (bottom), with three different values of the Hubble constant, obtained with the SPLASH galaxy redshift survey dataset. The dotted line is the linear fit for galaxies at $z < 1$, while the dashed line represents those with $1 \leq z \leq 4$. The grey area illustrates the 1 sigma error of the bins.

Table 3. Fractal dimensions calculated in the selected subsamples of the COSMOS2015 galaxy redshift survey in the ranges $z < 1$ and $1 \leq z \leq 4$. The single fractal dimensions D_L, D_Z and D_G were obtained from the galaxy distributions respectively using the luminosity distance d_L, redshift distance d_Z and galaxy area distance (transverse comoving distance) d_G.

H_0	$D_L(z<1)$	$D_Z(z<1)$	$D_G(z<1)$	$D_L(z>1)$	$D_Z(z>1)$	$D_G(z>1)$
65 km/s/Mpc	1.21 ± 0.01	1.41 ± 0.01	1.54 ± 0.01	0.29 ± 0.01	0.37 ± 0.02	0.75 ± 0.04
70 km/s/Mpc	1.22 ± 0.02	1.41 ± 0.01	1.55 ± 0.01	0.30 ± 0.01	0.37 ± 0.02	0.76 ± 0.05
75 km/s/Mpc	1.22 ± 0.02	1.43 ± 0.02	1.57 ± 0.01	0.31 ± 0.01	0.39 ± 0.02	0.76 ± 0.05

Table 4. Fractal dimensions calculated in the selected subsamples of the SPLASH galaxy redshift survey in the ranges $z < 1$ and $1 \leq z \leq 4$. The single fractal dimensions D_L, D_Z and D_G were obtained from the galaxy distributions respectively using the luminosity distance d_L, redshift distance d_Z and galaxy area distance (transverse comoving distance) d_G.

H_0	$D_L(z < 1)$	$D_Z(z < 1)$	$D_G(z < 1)$	$D_L(z > 1)$	$D_Z(z > 1)$	$D_G(z > 1)$
65 km/s/Mpc	0.90 ± 0.01	1.01 ± 0.02	1.09 ± 0.03	0.48 ± 0.01	0.60 ± 0.02	1.16 ± 0.03
70 km/s/Mpc	0.89 ± 0.01	1.01 ± 0.02	1.09 ± 0.03	0.48 ± 0.01	0.58 ± 0.02	1.15 ± 0.03
75 km/s/Mpc	0.89 ± 0.01	1.01 ± 0.02	1.09 ± 0.03	0.47 ± 0.01	0.59 ± 0.02	1.11 ± 0.03

Table 5. The comparison displayed of all recently calculated single fractal dimensions applying similar analytical tools as presented here to various galaxy distribution surveys. These results were obtained with the UltraVISTA DR1 [18], COSMOS2015 [19] and SPLASH [20], and FDF [39, 40] catalogs, all considering $H_0 = 70$ km/s/Mpc. There is a clear tendency for decreasing values of D at $z > 1$ in virtually all results, as theoretically predicted in the Introduction. Such a decrease is, nonetheless, less pronounced in the SPLASH data, which is the only galaxy distribution shown here to have been surveyed in the southern hemisphere.

	UVista DR1	$(0.2 < z < 4)$	COSMO2015	$(0.1 < z < 4)$	SPLASH	$(0.1 < z < 4)$	FDF	$(0.45 < z < 5)$
	$z < 1.0$	$z > 1.0$	$z < 1.0$	$z > 1.0$	$z < 1.0$	$z > 1.0$	$z \lesssim 1.2$	$z \gtrsim 1.2$
D_L	1.40 ± 0.02	0.32 ± 0.01	1.22 ± 0.02	0.30 ± 0.01	0.89 ± 0.01	0.48 ± 0.01	1.2 ± 0.3	0.5 ± 0.2
D_z	1.61 ± 0.02	0.38 ± 0.02	1.41 ± 0.01	0.37 ± 0.02	1.01 ± 0.02	0.58 ± 0.02	1.5 ± 0.4	0.6 ± 0.2
D_G	1.75 ± 0.03	0.81 ± 0.06	1.55 ± 0.01	0.76 ± 0.05	1.09 ± 0.03	1.15 ± 0.03	1.8 ± 0.3	1.0 ± 0.7

3.2.4 Blue and red subsamples

A question that can be raised regarding this fractal analysis is the possibility that the single fractal dimension D would depend on galactic type, features or potential evolutionary stages, and if there are dependencies on some, or all, of these characteristics. In such a case then the fractal dimension could, perhaps, be used as a tracer of galactic properties or their evolutionary stages. In this subsection we introduce a simple test of this possible dependency using the COSMOS2015 and SPLASH surveys studied here.

These two surveys allow us to compute D in two galaxy subsamples: the red, or quiescent, and the blue, or star forming galaxies. Therefore, this grants us with a selection that provides an initial and direct way for testing the concept of using D as a tracer of galactic characteristics. The criteria for creating these subsamples uses color-color diagrams or star formation rates as supplied in both surveys databases. Regarding the COSMOS2015 dataset, the classification introduced in Ref. [19] is obtained from the location of galaxies in the NUV-r vs. r-J color-color diagram [41]. In addition to using these colors, this selection has the absolute magnitudes' estimation at rest-frame based on the apparent magnitude at $\lambda_{\text{rest-frame}}(1 + z_{\text{gal}})$,

which reduces the k-correction dependency [42]. Such approach prevents the mixing between red dusty galaxies and quiescent ones. In implementation, selected quiescent galaxies have $M_{\mathrm{NUV}} - M_r > 3(M_r - M_J) + 1$ and $M_{\mathrm{NUV}} - M_r > 3.1$, while the other galaxies were considered star-forming [18]. The redshift dependent evolution of this distribution is in Fig. 20, where there is an accumulation of quiescent galaxies at low redshift inside the box and a relative decrease in bright, star-forming galaxies outside the box.
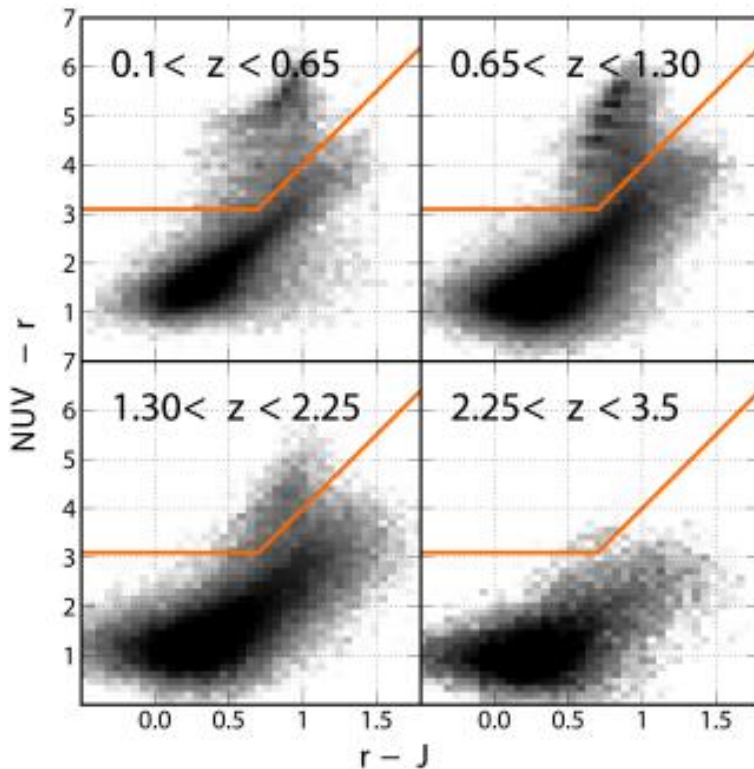
Figure 20. NUV - r/r - J COSMOS2015 galaxy distributions. Quiescent galaxies lie in the top-left corner. The objects fainter than limiting magnitudes are not considered [19].

For the SPLASH survey, following Refs. [43] and [44] we divided the galaxy sample using the *specific star formation rate* (SSFR). Galaxies with $\log \mathrm{SSFR} > -11$ were classified as star-forming, whereas the ones with $\log \mathrm{SSFR} < -11$ were set as quiescent. SSFR is the ratio between star forming rate and stellar mass, being a simple approach for quantifying in a uniform way the star formation degree. Such separation by stellar mass indicates how strong is the star formation in a certain galaxy, which is particularly important when comparing galaxies having different masses and sizes. In Ref. [36] this classification was compared to the color-color selection applied to COSMOS2015, and the conclusion was that both methods provide similar results at $z < 1$, but towards high redshift the one based on SSFR tends to be more conservative.

Two subsamples were generated for each survey taking into consideration the criteria explained above. Fig. 21 displays histograms of redshift distribution of both subsamples up

to $z = 4$, where it is evident that the number of blue star forming galaxies is greater than red quiescent ones. This selection methodology led the COSMOS2015 survey to turn out as 527899 star forming galaxies and 31424 quiescent ones. The SPLASH was respectively divided into 359021 and 20045 galaxies. Then, the same volume-limited filtering process of obsolute magnitude cutoff was applied to these four subsamples as defined by Eq. (3.3). In Fig. 22 is presented the outcome of this filtering procedure, which filtered the COSMOS2015 subsamples to 208005 blue galaxies and 22824 red ones, while reducing the SPLASH data to 205012 and 13491 galaxies respectively.
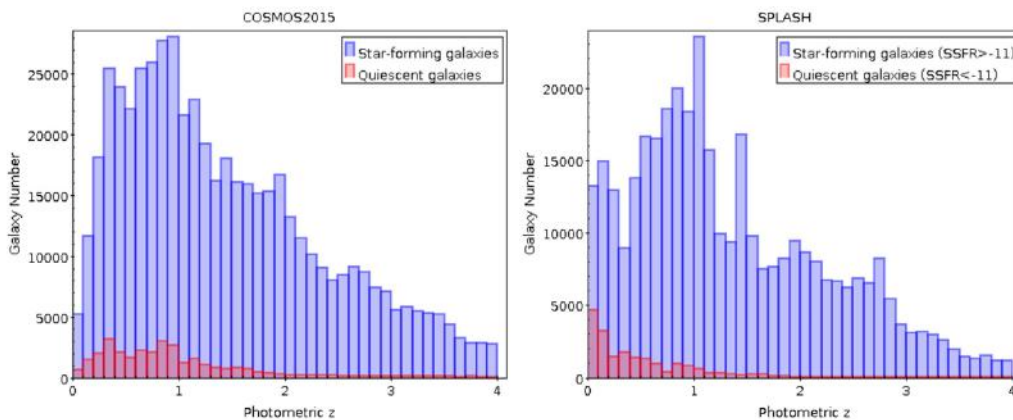
Figure 21. Histograms displaying the galaxy distribution numbers in terms of the redshift for the COSMOS2015 (left) and SPLASH (right) subsamples of blue star forming galaxies and red quiescent ones. Labels are as in the legends. COSMOS2015 galaxies were classified by color and SPLASH galaxies were separated considering the specific stellar formation rate (SSFR) using -11 as the cutoff value (see main text). Clearly the number of blue galaxies is much higher than the red ones in both surveys.
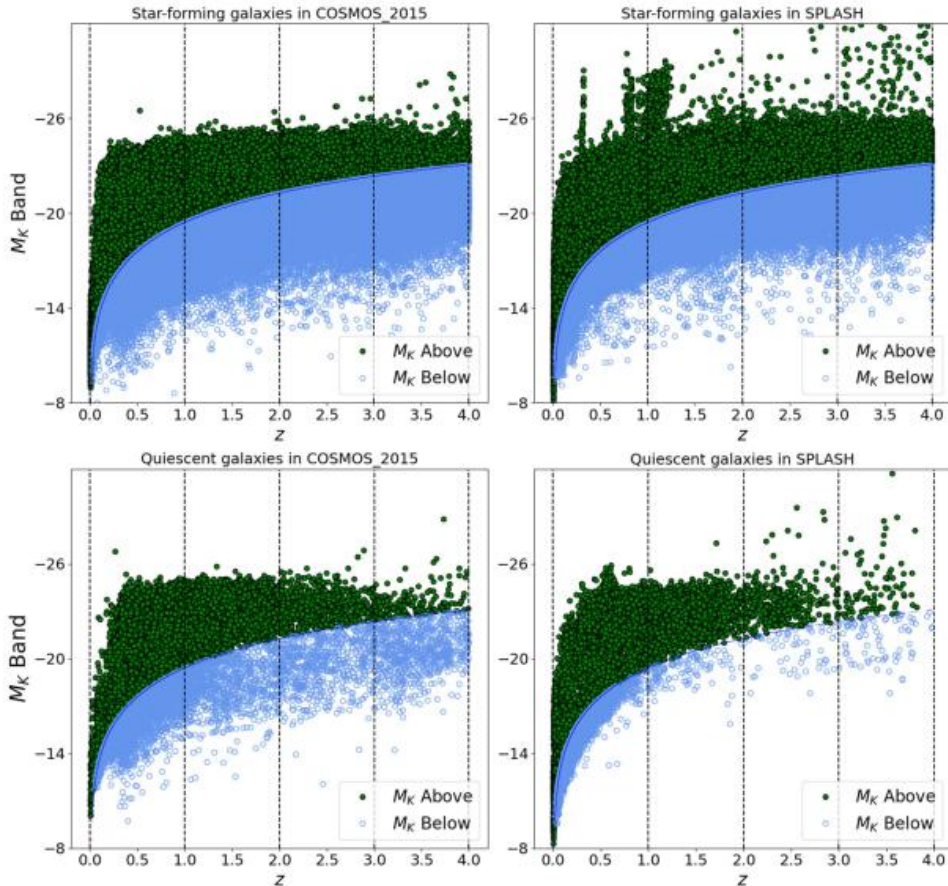
Figure 22. Plots of the absolute magnitudes for the blue star forming galaxies (top), and red quiescent ones (bottom) for the COSMOS2015 (left) and SPLASH (right) surveys in terms of their photometrically measured redshift values. Specifications are as in the title of each plot. As in Figs. 16 and 17, the dividing line corresponds to apparent magnitude $K = 24.7$, so only blue and red galaxies having M_K above this cutoff and $z \leq 4$ were included in the blue and red COSMOS2015 and SPLASH subsamples. Since previous results indicate that the fractal dimension is not affected by altering the Hubble constant within its currently accepted uncertainty, here only $H_0 = 70$ km/s/Mpc was assumed.

The calculation method for the γ_i^* number densities from Sec. 3.2.2 was adopted for these four selected and filtered galaxies subsamples, then the resulting data points were linear fitted against their respective distance measurement values d_i ($i = G, L, z$). Figs. 23 and 24 show the decaying power-law curves and data fits in the ranges $z < 1$ and $1 \leq z \leq 4$. Considering the robust results for the fractal dimension against changes in the Hubble constant as presented in Tables 3 and 4, in this subsection we calculated all results using only $H_0 = 70$ km/s/Mpc.
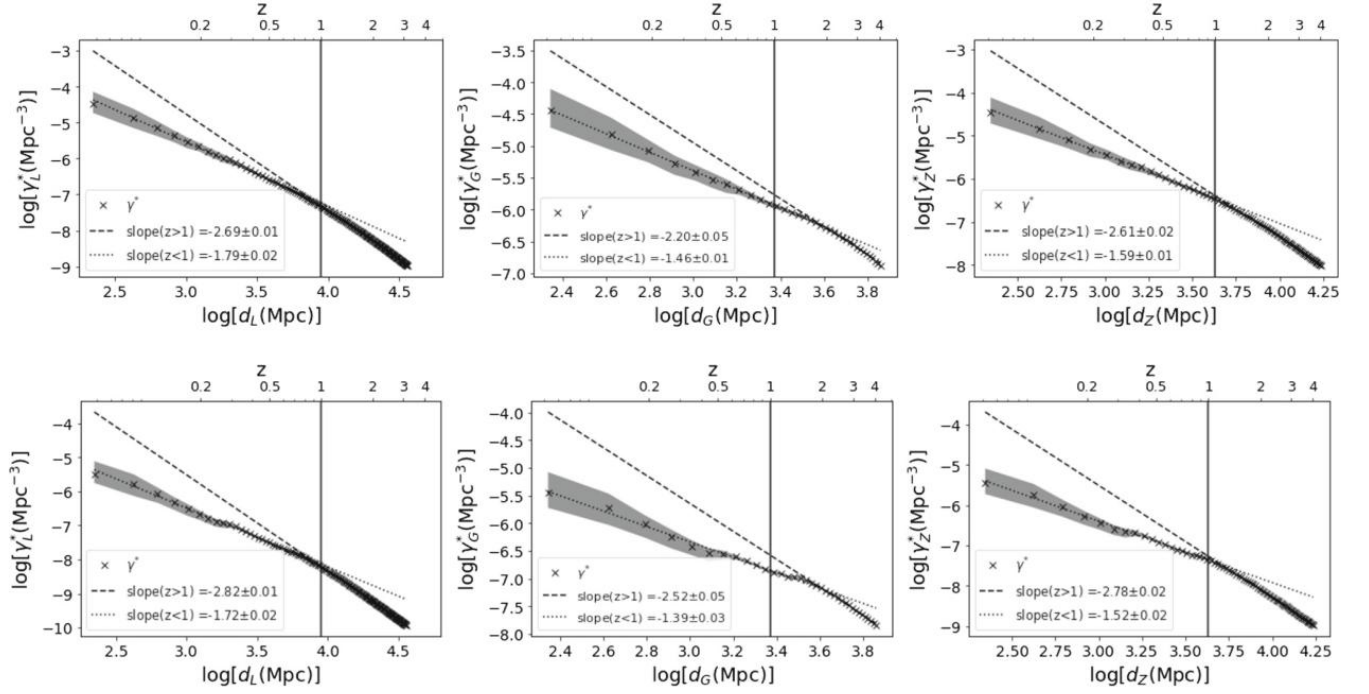
Figure 23. Log-log graph of γ_L^*, γ_G^* and γ_z^* respectively vs. d_L, d_G and d_z obtained with the blue COSMOS2015 star forming galaxy subsample (top) and with the red COSMOS2015 quiescent galaxy subsample (bottom). The dotted line is the linear fit for galaxies at $z < 1$, while the dashed line represents those with and $1 \leq z \leq 4$. The grey area illustrates the 1 sigma error of the bins. All results were obtained considering $H_0 = 70$ km/s/Mpc.
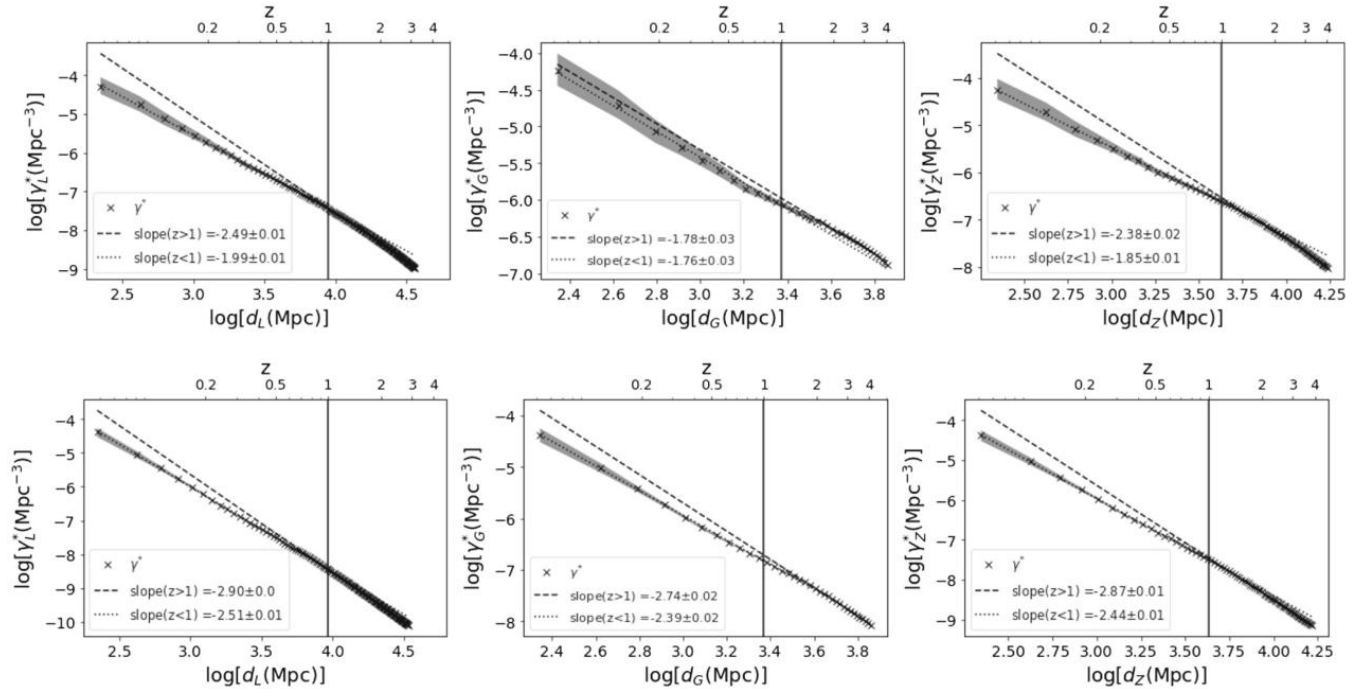
Figure 24. Log-log graph of γ_L^*, γ_G^* and γ_z^* respectively vs. d_L, d_G and d_z obtained with the blue SPLASH star forming galaxy subsample (top) and with the red SPLASH quiescent galaxy subsample (bottom). The dotted line is the linear fit for galaxies at $z < 1$, while the dashed line represents those with and $1 \leq z \leq 4$. The grey area illustrates the 1 sigma error of the bins. All results were obtained considering $H_0 = 70$ km/s/Mpc.

The fractal dimensions obtained from the graphs in Figs. 23 and 24 are assembled in Table 6, which presents results that can now be compared with the 3rd to 6th columns of Table 5. The blue COSMOS2015 galaxies had their fractal dimensions essentially unchanged as compared to the unselected samples, while the blue SPLASH ones had D moderately increased for $z < 1$, but remained basically unchanged for $z > 1$ within the uncertainties. However, the fractal dimensions of the quiescent galaxies had more evident changes. For $z < 1$ the red COSMOS2015 subsample had higher values for D in all distance measures, but experienced a fractal dimension reduction at $z > 1$. The same substantial decrease in D happened for the red SPLASH galaxies in both ranges, $z < 1$ and $z > 1$. Finally, in all cases for every distance measures the fractal dimension reduction in the range $z > 1$ is observed, as the theory predicted.

Table 6. Fractal dimensions determined in the selected blue-star-forming and red-quiescent, and then volume-limit filtered, galaxy subsamples of the COSMOS2015 and SPLASH redshift surveys in the range $z \leq 4$. The single fractal dimensions D_L, D_z and D_G were acquired from the galaxy distributions respectively using the luminosity distance d_L, redshift distance d_z and galaxy area distance (transverse comoving distance) d_G. The results were calculated considering only $H_0 = 70$ km/s/Mpc (see main text). A comparison of these numbers with the ones in the 3rd to 6th columns of Table 5 shows that the fractal dimensions do vary according to the blue-red selection used here, this being especially the case for the quiescent galaxies. Similarly, as in the results shown in Table 5, all values of D decrease in the range $z > 1$, in some cases quite substantially.

	Blue COSMOS2015	$(0.1<z<4)$	Blue SPLASH	$(0.1<z<4)$	Red COSMOS2015	$(0.1 < z < 4)$	Red SPLASH	$(0.1<z<4)$
	$z < 1.0$	$z > 1.0$	$z < 1.0$	$z > 1.0$	$z < 1.0$	$z > 1.0$	$z < 1.0$	$z > 1.0$
D_L	1.21 ± 0.02	0.31 ± 0.01	1.01 ± 0.01	0.51 ± 0.01	1.28 ± 0.02	0.18 ± 0.01	0.49 ± 0.01	0.10 ± 0.01
D_z	1.41 ± 0.01	0.39 ± 0.02	1.15 ± 0.01	0.62 ± 0.02	1.48 ± 0.02	0.22 ± 0.02	0.56 ± 0.01	0.13 ± 0.01
D_G	1.54 ± 0.01	0.80 ± 0.05	1.24 ± 0.03	1.22 ± 0.03	1.61 ± 0.03	0.48 ± 0.05	0.61 ± 0.02	0.26 ± 0.02

Conclusion

This dissertation pursued to empirically test if the large-scale galaxy distribution could be represented as a fractal system and if galaxy types could possibly be dependent on the single fractal dimension D. Some of the Newtonian hierarchical cosmology relations were extended and applied to relativistic cosmological models in order to describe possible galaxy fractal structures by means of D at deep redshift scales. The application of these tools to the UltraVISTA DR1, COSMOS2015 and SPLASH datasets comprised almost one million objects spanning the redshift interval of $0.1 \leq z \leq 6$.

Absolute magnitudes were calculated using the measured redshifts of the surveys in the interest of acquiring volume-limited subsamples. These absolute magnitudes were adopted to obtain the luminosity distances d_L by assuming the ΛCDM relativistic cosmological model and the apparent magnitude limit of 24 in the J-band for the UltraVISTA DR1 and 24.7 in the K-band for the COSMOS2015 and SPLASH, as well as data redshift cutoff at z=4. Since these subsamples showed that their galaxies followed increasing redshift bins, they were considered as effectively forming a volume-limited distribution. This procedure provided three subsamples with about 568k objects, the first containing 166566 UltraVISTA DR1 galaxies, the second containing 230705 COSMOS2015 galaxies, and the third containing 171548 SPLASH ones. Fractal analysis was then performed in these three subsamples.

As relativistic cosmologies have several definitions of observed distance [31], only three were used here, which are the luminosity distance d_L, the redshift distance d_Z and the galaxy area distance (transverse comoving distance) d_G. The adoption of several cosmological distance measures originates from the fact that relativistic effects become strong enough for redshift ranges larger than $z \gtrsim 0.3$, so these distance definitions generate different results for the same redshift value at those ranges. A data sorting algorithm was executed to plot ghaphs for number densities vs. relativistic distances. Straight lines were then fitted to the data in two scales, $z < 1$ and $1 \leq z \leq 4$, whose slopes directly produced the single fractal dimensions.

The results indicate that the subsamples have two consecutive redshift ranges behaving as single fractal structures. Rounding them off and their respective uncertainties we found that for $z < 1$ the UltraVISTA DR1 galaxies presented $D = (1.6 \pm 0.2)$, the COSMOS2015 galaxies produced $D = (1.4 \pm 0.2)$ and the SPLASH galaxies generated $D = (1.0 \pm 0.1)$. For $1 \leq z \leq 4$ the dimensions respectively decreased to $D = (0.6 \pm 0.3)$, $D = (0.5 \pm 0.3)$ and $D = (0.8 \pm 0.4)$. These results were not affected by changes in the Hubble constant within the

assumed uncertainties. Additionally, no transition to observational homogeneity was found in the data.

Subsamples were then created by selecting blue star forming galaxies and red quiescent ones from the COSMOS2015 and SPLASH datasets. After filtering the data by the same absolute magnitude limits that were applied to the unselected samples, the results showed that up to two decimal digits the fractal dimension of blue galaxies remained essentially unchanged, whereas some red galaxies showed evident reduction at the same precision. This demonstrates that the fractal dimensions of both surveys are dominated by blue galaxies. These results also suggest that the fractal dimensions can be valuable tools as tracers of galaxy types and evolution in addition to being descriptors of galaxy distributions. So, in this context the fractal dimension would be seen as an intrinsic property of the objects' distribution in the Universe that could be adjusted depending on the types of galaxy formed, which means that D can be applied as a tool to study different galaxy populations.

Finally, the results introduced here raise the question of why there is such considerable decrease in the fractal dimension for redshift values larger than the unity. The reason could be an observational effect caused by data bias, because many galaxies are not detected beyond $z > 1$ reducing then the observed galaxy clustering and consequently reducing the associated fractal dimension. Another potential reason is bias associated to the small angular areas of these surveys so that they would not provide representative measurements of the entire sky distribution. Apart from observational biases, it is possible that the decrease in the fractal dimension is due to real physical effects. Considering galaxy evolution and large-scale structure dynamics of the Universe, it is plausible that there might be much less galaxies at high z. These interpretation is in agreement with the theoretical prediction that states a sharp decrease in observed number density at $z > 1$ as shown in Ref. [15] and Fig. 6 in Ref. [45]. So, the shift in D to smaller values could be the empirical verification of this theoretical prediction, meaning that the Universe was void dominated at those epochs as a result of galaxies being much more sparsely distributed and in reduced numbers.

A. Python code used in this study

```
1
  # Analysis of the COSMOS2015 survey for z<4
3
  # Code to perform fractal analysis in large astronomical surveys
5 import pandas as pd
  import matplotlib.pyplot as plt
7 import numpy as np
  import astropy.units as u
9 from scipy import constants
  from scipy import stats
11 from scipy.stats import linregress
  from scipy import odr
13 import matplotlib.ticker as ticker
  from astropy.cosmology import LambdaCDM, z_at_value
15 paramet = LambdaCDM(H0=70, Om0=0.3, Ode0=0.7)


17 # Importing tables
  tabela_qst = pd.read_csv("quiescent_cosmos2015.csv")
19 tabela_sf = pd.read_csv("sf_cosmos2015.csv")


21 # Selecting sample
  tabela=[]
23 sel_uvista = (tabela_sf['FLAG_HJMCC']==0)&(tabela_sf['TYPE']==0)&(tabela_sf['
    ZPDF']>0.0)
  tabela = tabela_sf[sel_uvista]
25 dl = paramet.luminosity_distance(tabela.ZPDF).values.value
  selacima = (tabela['MK'] <= (-5*np.log10(dl) - 0.3)) & (tabela['ZPDF'] < 4.0) &
      (tabela['MK'] > -90)
27 MK_ABOVE = tabela[selacima]


29 dl = paramet.luminosity_distance(MK_ABOVE.ZPDF).values.value


31 # Unity of H0 is km/Mpc.s
  H0 = paramet.H(0)
33 c = 299792 #km/s
  z = MK_ABOVE.ZPDF
35 dz = (c*z) / H0
  dg = dl/(1+z)
37
  # Inferior uncertainty
39 dl_inf = paramet.luminosity_distance(MK_ABOVE.ZPDF_L68).values.value
  H0 = paramet.H(0)
41 c = 299792 #km/s
```

```python
   z_inf = MK_ABOVE.ZPDF_L68
43 dz_inf = (c*z_inf) / H0
   dg_inf = dl_inf/(1+z_inf)

45
   # Upper uncertainty
47 dl_sup= paramet.luminosity_distance(MK_ABOVE.ZPDF_H68).values.value
   H0 = paramet.H(0)
49 c = 299792 #km/s
   z_sup = MK_ABOVE.ZPDF_H68
51 dz_sup = (c*z_sup) / H0
   dg_sup = dl_sup/(1+z_sup)

53
   # Functions to assist in our analysis
55 # 1) Find redshift from dl, in order to find dg(z) relation
   def find_z(dist):
57     z_test=np.arange(0.001,7.05,0.0001)
       dg_test = paramet.luminosity_distance(z_test).value/(1+z_test)
59     i=0
       test= dg_test[i]
61     while round(test,2)<round(dist,2):
           i=i+1
63         test=dg_test[i]
           z_out=z_test[i]
65         dist_out = test
       return z_out

67
   # 2) Function to generate gama density in respect to a given distance
69 def gama_dist(dbin, dist, dist_sup, dist_inf, dist_type):
   list_count = []
71 list_count_sup = []
   list_count_inf = []
73 list_gama = []
   list_gama_sup = []
75 list_gama_inf = []
   list_dist = []
77 list_z = []
   count = 0
79 dmin = min(dist)

81 while count < len(dist):
     d = dmin + dbin
83   # Center values:
     sel = dist<(d)
85   dist_sel=dist[sel]
     count = len(dist_sel)
87   list_count.append(count)
```

```python
89      # Upper values:
        sel_sup = dist_sup<(d)
91      dist_sel_sup=dist_sup[sel_sup]
        count_sup = len(dist_sel_sup)
93      list_count_sup.append(count_sup)

95      # Inferior values:
        sel_inf = dist_inf<(d)
97      dist_sel_inf=dist_inf[sel_inf]
        count_inf= len(dist_sel_inf)
99      list_count_inf.append(count_inf)

101     if dist_type=='dL':
          z_lim_dist = z_at_value(paramet.luminosity_distance, d*u.Mpc)
103     if dist_type=='dZ':
          z_lim_dist = (d*H0.value)/c
105     if dist_type=='dG':
          z_lim_dist = find_z(d)
107
        volume = (4/3)*np.pi*((d)**3)
109     gama = count/volume
        gama_sup = count_sup/volume
111     gama_inf = count_inf/volume

113     list_gama.append(gama)
        list_gama_sup.append(gama_sup)
115     list_gama_inf.append(gama_inf)
        list_dist.append(d)
117     list_z.append(z_lim_dist)
        dmin = d
119
    return np.array(list_gama), np.array(list_gama_sup), np.array(list_gama_inf),
        np.array(list_dist), np.array(list_z)
121
  # 3) Linear regression with error in both axis
123 def func(p,x):
        a,b=p
125     return a*x+b

127 # 4) Axis label format
  def label_form(x, pos):
129     if x<1:
          xla=float(x)
131       x=xla
        if x>=1:
133       xla=int(x)
          x=xla
```

```python
135         return str(x)


137 # 5) Function to find slopes of gama vs dist (linear regression)
    def slopes(dist_type, dist, gama, gama_lim_sup):
139
        # Cut related to z=1
141     listateste = []
        if dist_type == 'dL':
143         for i in listaz1:
                if i > 1:
145                 listateste.append(i)
            cut_dist = len(listaz1) - len(listateste)
147     if dist_type == 'dZ':
            for i in listazz1:
149           if i > 1:
                  listateste.append(i)
151         cut_dist = len(listazz1) - len(listateste)
        if dist_type == 'dG':
153         for i in listazg2:
                if i > 1:
155                 listateste.append(i)
            cut_dist = len(listazg2) - len(listateste)
157
        # Linear regression for z > 1
159     data_x = np.log10(dist[cut_dist:-1:])
        data_y = np.log10(gama[cut_dist:-1:])
161     data_y_err = np.log10(gama_lim_sup[cut_dist:-1:]) - np.log10(gama[cut_dist
            :-1:])
        # Linear regression with error in both axis:
163     def func(p, x):
            a, b=p
165         return a*x+b
        from scipy import odr
167     linear_model = odr.Model(func)
        data = odr.RealData(data_x, data_y, sy=data_y_err)
169     odr = odr.ODR(data, linear_model, beta0=[1., 1.])
        out = odr.run()
171     slope_g = out.beta[0]
        err_slope_g = out.sd_beta[0]
173     interp_g = out.beta[1]
        err_interp_g = out.sd_beta[1]
175     # Linear regression for z < 1
        from scipy import odr
177     data_x = np.log10(dist[:cut_dist:])
        data_y = np.log10(gama[:cut_dist:])
179     data_y_err = np.log10(gama[:cut_dist:]) - np.log10(gama_lim_sup[:cut_dist:])
        # Linear regression with error in both axis:
```

```python
181    def func(p,x):
           a,b=p
183        return a*x+b
       linear_model = odr.Model(func)
185    data = odr.RealData(data_x,data_y,sy=data_y_err)
       odr = odr.ODR(data,linear_model,beta0=[1.,1.])
187    out = odr.run()
       slope_l = out.beta[0]
189    err_slope_l = out.sd_beta[0]
       interp_l = out.beta[1]
191    err_interp_l = out.sd_beta[1]
       return slope_g, err_slope_g, interp_g,err_interp_g, slope_l, err_slope_l,
           interp_l,err_interp_l

193


195 # 6) Function to create gama plots
    def gama_plot(dist,gama,z,type_dist,gama_lim_sup,gama_lim_inf,slope_l,
        err_slope_l,interp_l,slope_g,err_slope_g,interp_g):
197    plt.figure(figsize=(15,15))
       # Graph of gama_di vs d_i where i=l,z,g with bin of 200mpc and z plotted above
199    fig = plt.figure()
       # Create first subplot:
201    ax1 = fig.add_subplot(111)


203    if type_dist=='L':
         # For the first plot:
205      ax1.plot(np.log10(dist),np.log10(gama),'kx',label='$\gamma^{*}$')
         ax1.plot(np.log10(dist),interp_g_dl + slope_g_dl*np.log10(dist),'k—',label=
             'slope(z>1) ='+str("%.2f" % round(slope_g_dl,2))+'$\pm$'+str(round(
             err_slope_g_dl,2)))
207      ax1.plot(np.log10(dist),interp_l_dl + slope_l_dl*np.log10(dist),'k:',label='
             slope(z<1) ='+str("%.2f" % round(slope_l_dl,2))+'$\pm$'+str(round(
             err_slope_l_dl,2)))
         # Hatched area with error only on y axis
209      ax1.fill_between(np.log10(dist),np.log10(gama_lim_inf),np.log10(gama_lim_sup
             ),facecolor='grey')
         ax1.set_xlabel('$\log[d_{L}$(Mpc)]', fontsize=18)
211      ax1.set_ylabel('$\log[\gamma_{L}^{*}$(Mpc$^{-3}$)]', fontsize=18)

213    if type_dist=='Z':
         ax1.plot(np.log10(dist),np.log10(gama),'kx',label='$\gamma^{*}$')
215      ax1.plot(np.log10(dist),interp_g_dz + slope_g_dz*np.log10(dist),'k—',label=
             'slope(z>1) ='+str("%.2f" % round(slope_g_dz,2))+'$\pm$'+str(round(
             err_slope_g_dz,2)))
         ax1.plot(np.log10(dist),interp_l_dz + slope_l_dz*np.log10(dist),'k:',label='
             slope(z<1) ='+str("%.2f" % round(slope_l_dz,2))+'$\pm$'+str(round(
             err_slope_l_dz,2)))
```

```python
217     ax1.fill_between(np.log10(dist),np.log10(gama_lim_inf),np.log10(gama_lim_sup
            ),facecolor='grey')
        ax1.set_xlabel('$\log[d_{Z}$(Mpc)]', fontsize=18)
219     ax1.set_ylabel('$\log[\gamma_{Z}^{*}$(Mpc$^{-3}$)]', fontsize=18)


221   if type_dist=='G':
        ax1.plot(np.log10(dist),np.log10(gama),'kx',label='$\gamma^{*}$')
223     ax1.plot(np.log10(dist),interp_g_dg + slope_g_dg*np.log10(dist),'k--',label=
            'slope(z>1) ='+str("%.2f" % round(slope_g_dg,2))+'$\pm$'+str(round(
            err_slope_g_dg,2)))
        ax1.plot(np.log10(dist),interp_l_dg + slope_l_dg*np.log10(dist),'k:',label='
            slope(z<1) ='+str("%.2f" % round(slope_l_dg,2))+'$\pm$'+str(round(
            err_slope_l_dg,2)))
225     ax1.fill_between(np.log10(dist),np.log10(gama_lim_inf),np.log10(gama_lim_sup
            ),facecolor='grey')
        ax1.set_xlabel('$\log[d_{G}$(Mpc)]', fontsize=18)
227     ax1.set_ylabel('$\log[\gamma_{G}^{*}$(Mpc$^{-3}$)]', fontsize=18)


229   ax1.legend(loc=3, prop={'size': 11.5})
      plt.xticks(fontsize=14)
231   plt.yticks(fontsize=14)
      # Create second subplot keeping the y axis of first plot
233   ax2 = ax1.twiny()
      # For the second plot
235   ax2.plot(z,np.log10(gama),alpha=0)
      ax2.semilogx()
237   ax2.set_xticks([0.2,0.5,1,2,3,4])
      ax2.xaxis.set_major_formatter(ticker.FuncFormatter(label_form)) # Formatting
          number values of second axis
239   ax2.set_xlabel('z',fontsize=18)
      plt.xticks(fontsize=13)
241   # Vertical line
      plt.axvline(x=1, color='k', linestyle='-')
243   plt.show()
      return

245

  # Applying function 2
247 listagamadl1, listagamadl1_sup, listagamadl1_inf, listadl1, listaz1 = gama_dist(
      dbin=200, dist=dl, dist_sup=dl_sup, dist_inf=dl_inf, dist_type='dL')
  listagamadz1, listagamadz1_sup, listagamadz1_inf, listadz1, listazz1 = gama_dist
      (dbin=200, dist=dz, dist_sup=dz_sup, dist_inf=dz_inf, dist_type='dZ')
249 listagamadg2, listagamadg2_sup, listagamadg2_inf, listadg2, listazg2 = gama_dist
      (dbin=200, dist=dg, dist_sup=dg_sup, dist_inf=dg_inf, dist_type='dG')


251 # Applying function 5
  slope_g_dl, err_slope_g_dl, interp_g_dl, err_interp_g_dl, slope_l_dl,
      err_slope_l_dl, interp_l_dl, err_interp_l_dl = slopes('dL',listadl1,
```

```
        listagamadl1 , listagamadl1_sup )
253 slope_g_dz ,  err_slope_g_dz ,  interp_g_dz ,  err_interp_g_dz ,  slope_l_dz ,
        err_slope_l_dz ,  interp_l_dz ,  err_interp_l_dz = slopes ('dZ',listadz1 ,
        listagamadz1 , listagamadz1_sup )
    slope_g_dg ,  err_slope_g_dg ,  interp_g_dg ,  err_interp_g_dg ,  slope_l_dg ,
        err_slope_l_dg ,  interp_l_dg ,  err_interp_l_dg = slopes ('dG',listadg2 ,
        listagamadg2 , listagamadg2_sup )
255
    # Applying  function  6  to  create  the  graphs
257 gama_plot ( listadl1 , listagamadl1 , listaz1 , 'L',listagamadl1_sup , listagamadl1_inf ,
        slope_l_dl , err_slope_l_dl , interp_l_dl , slope_g_dl , err_slope_g_dl , interp_g_dl )
    gama_plot ( listadz1 , listagamadz1 , listazz1 , 'Z',listagamadz1_sup , listagamadz1_inf ,
        slope_l_dz , err_slope_l_dz , interp_l_dz , slope_g_dz , err_slope_g_dz , interp_g_dz )
259 gama_plot ( listadg2 , listagamadg2 , listazg2 , 'G',listagamadg2_sup , listagamadg2_inf ,
        slope_l_dg , err_slope_l_dg , interp_l_dg , slope_g_dg , err_slope_g_dg , interp_g_dg )
```

B. Heuristic derivation of the fractal number-distance relation

Suppose a hierarchy as in Fig. 4, which is a self-similar structure, discrete at order n. Let us suppose now that at each radius there is a certain number of objects:

$$\text{radius } d_0 \rightarrow N_0 \text{ objects.} \tag{B.1}$$

$$\text{radius } d_1 \rightarrow N_1 \text{ objects.} \tag{B.2}$$

Proportionally,

$$d_1 = kd_0 \,, \tag{B.3}$$

$$N_1 = k_* N_0 \,, \tag{B.4}$$

where k and k_* are proportionality constants. So, in general

$$d_n = k^n d_0 \,, \tag{B.5}$$

$$N_n = k_*^n N_0 \,, \tag{B.6}$$

and by writing Eq. (B.5) we get:

$$\log(d_n) = \log(k^n d_0) \,, \tag{B.7}$$

$$\log\left(\frac{d_n}{d_0}\right) = n \log(k) \,, \tag{B.8}$$

as well as doing the same in Eq. (B.6) yields,

$$\log\left(\frac{N_n}{N_0}\right) = n \log(k_*). \tag{B.9}$$

Dividing Eq. (B.9) by Eq. (B.8) we have:

$$\frac{\log(N_n/N_0)}{\log(d_n/d_0)} = \frac{n \log(k_*)}{n \log(k)} \,, \tag{B.10}$$

$$\log\left(\frac{N_n}{N_0}\right) = \frac{\log(k_*)}{\log(k)} \log\left(\frac{d_n}{d_0}\right) \,, \tag{B.11}$$

and then we have that, where

$$D = \frac{\log(k_*)}{\log(k)} \,, \tag{B.12}$$

$$\log\left(\frac{N_n}{N_0}\right) = \log\left(\frac{d_n}{d_0}\right)^D . \tag{B.13}$$

Removing the logarithms we get:

$$\frac{N_n}{N_0} = \left(\frac{d_n}{d_0}\right)^D \,, \tag{B.14}$$

$$N_n = \frac{N_0 d_n^D}{d_0^D} \,, \tag{B.15}$$

and defining B as:

$$B = \frac{N_0}{d_0^D} \,, \tag{B.16}$$

we subsequently attain

$$N_n = B d_n^D . \tag{B.17}$$

Thus, in general, for a smooth proportionality we reach to the number distance relation for a self-similar fractal structure.

$$N = B d^D . \tag{B.18}$$

Bibliography

[1] **Sharon Teles**, Amanda R. Lopes, and Marcelo B. Ribeiro. Fractal analysis of the UltraVISTA galaxy survey. *Physics Letters B*, 813:136034, 2021.

[2] **Sharon Teles**, Amanda R. Lopes, and Marcelo B. Ribeiro. Galaxy distributions as fractal systems. *The European Physical Journal C*, 82:896, 2022.

[3] E. Fournier d'Albe. *Two New Worlds*. Longmans, Green, and Co., 1907.

[4] M. Amoroso Costa. *Annals of Brazilian Acad. Sci. (in Portuguese)*, 1:51, 1929.

[5] E. F. Carpenter. Some characteristics of associated galaxies. I. a. density restriction in the metagalaxy. *Astrophys.J*, 88:344, 1938.

[6] Benoît B. Mandelbrot. *Fractals: form, chance, and dimension*. Freeman, 1977.

[7] V. de Lapparent, M. J. Geller, and J. P. Huchra. A Slice of the Universe. *Astrophys.J.*, 302:L1, 1986.

[8] Marcelo B. Ribeiro. On Modeling a Relativistic Hierarchical (Fractal) Cosmology by Tolman's Spacetime. I. Theory. *Astrophys.J*, 388:1, 1992.

[9] L. Pietronero. The fractal structure of the universe: Correlations of galaxies and clusters and the average mass density. *Physica A*, 144:257–284, 1987.

[10] Bernard J. T. Jones et al. Multifractal description of the large-scale structure of the universe. *Astrophys.J. Lett.*, 332:L1, 1988.

[11] Roger Balian and Richard Schaeffer. Galaxies - fractal dimensions, counts in cells, and correlations. *Astrophys.J*, 335:L43, 1988.

[12] V. J. Martínez. Fractals and multifractals in the description of the cosmic structure. *Vistas in Astronomy*, 33:337–356, 1990.

[13] G. F. R. Ellis. Relativistic cosmology. In R. K. Sachs, editor, *General Relativity and Cosmology: Proceedings of the International School of Physics "Enrico Fermi", Course 47*, pages 104–182. Academic Press, 1971. Reprinted in *Gen. Rel. Grav.*, 41:581-660, 2009.

[14] James R. Wertz. *Newtonian Hierarchical Cosmology*. PhD thesis, University of Texas at Austin, 1970.

[15] Marcelo B. Ribeiro. Observations in the Einstein–de Sitter cosmology: Dust statistics and limits of apparent homogeneity. *Astrophys.J.*, 441:477–487, 1995.

[16] L. J. Rangel Lemos and Marcelo B. Ribeiro. Spatial and observational homogeneities of the galaxy distribution in standard cosmologies. *A&A*, 488:55, 2008.

[17] G. Conde-Saavedra, A. Iribarrem, and Marcelo B. Ribeiro. Fractal analysis of the galaxy distribution in the redshift range $0.45 < z < 5.0$. *Physica A*, 417:332–344, 2015.

[18] H. J. McCracken et al. Ultravista: A new ultra-deep near-infrared survey in COSMOS. *A&A*, 544:A156, 2012.

[19] C. Laigle et al. The COSMOS2015 catalog: Exploring the $1 < z < 6$ universe with half a million galaxies. *Astrophys.J Supp.*, 224(2):24, 2016.

[20] Vihang Mehta et al. SPLASH-SXDF multi-wavelength photometric catalog. *Astrophys.J Supp.*, 235(2):36, 2018.

[21] Benoit Mandelbrot. How long is the coast of britain? Statistical self-similarity and fractional dimension. *Science*, 156(3775):636–638, 1967.

[22] Marcelo B. Ribeiro and A. Y. Miguelote. Fractals and the distribution of galaxies. *Brazilian J. Phys.*, 28(2):132–160, 1998.

[23] Kris Gurung. *Fractal Dimension in Architecture: An Exploration of Spatial Dimension.* PhD thesis, 2017.

[24] E. Swedenborg and A. Clissold. *The Principia: Or, The First Principles of Natural Things, Being New Attempts Toward a Philosophical Explanation of the Elementary World.* Swedenborg Scientific Association, 1976.

[25] I. Baryshev and P. Teerikorpi. *The Discovery of Cosmic Fractals.* World Scientific, Singapore, 2002.

[26] G. De Vaucouleurs. The apparent density of matter in groups and clusters of galaxies. *Astrophys.J*, 131:585, 1960.

[27] Benoit B Mandelbrot. *The fractal geometry of nature.* Freeman, San Francisco, CA, 1982.

[28] Benoit B. Mandelbrot, Boaz Kol, and Amnon Aharony. Angular Gaps in Radial Diffusion-Limited Aggregation: Two Fractal Dimensions and Nontransient Deviations from Linear Self-Similarity. *Physical Review Letters*, 88(5):055501, 2002.

[29] J. P. Huchra et al. The CFA Redshift Survey. *International Astronomical Union. Symposium no. 130, Kluwer Academic Publishers*, 130:105, 1988.

[30] Marcelo B. Ribeiro. Relativistic fractal cosmologies. In D. W. Hobill, A. Burd, and A. Coley, editors, *Deterministic Chaos in General Relativity*, volume 332 of *B*, page 269. NATO ANSI, Plenum, 1994.

[31] G. F. R. Ellis. On the definition of distance in general relativity: I. M. H. Etherington (Philosophical Magazine ser. 7, vol. 15, 761, 1933). *Gen. Rel. Grav.*, 39:1047–1052, 2007.

[32] R. F. L. Holanda, J. A. S. Lima, and Marcelo B. Ribeiro. Testing the distance-duality relation with galaxy clusters and supernovae ia. *Astrophys.J. Lett.*, 722:L233–L237, 2010.

[33] R. F. L. Holanda, J. A. S. Lima, and Marcelo B. Ribeiro. Cosmic distance duality relation and the shape of galaxy clusters. *A&A*, 528:L14, 2011.

[34] R. F. L. Holanda, J. A. S. Lima, and Marcelo B. Ribeiro. Probing the cosmic distance-duality relation with the Sunyaev-Zel´dovich effect, x-ray observations and supernovae Ia. *A&A*, 538:A131, 2012.

[35] N. Scoville et al. The cosmic evolution survey (COSMOS): Overview. *Astrophys.J Supp.*, 172(1):1–8, 2007.

[36] O. Ilbert et al. Mass assembly in quiescent and star-forming galaxies since z from UltraVISTA. *A&A*, 556:A55, 2013.

[37] S. Arnouts et al. Measuring the redshift evolution of clustering: the Hubble Deep Field South. *MNRAS*, 329(2):355–366, 2002.

[38] O. Ilbert et al. Accurate photometric redshifts for the cfht legacy survey calibrated using the vimos vlt deep survey. *A&A*, 457(3):841–856, 2006.

[39] A. Gabasch et al. The evolution of the luminosity functions in the FORS deep field from low to high redshift. *A&A*, 421(1):41–58, 2004.

[40] A. Gabasch et al. The evolution of the luminosity functions in the FORS deep field from low to high redshift. *A&A*, 448(1):101–121, 2006.

[41] Rik J. Williams et al. Detection of quiescent galaxies in a bicolor sequence from z = 0-2. *Astrophys.J*, 691(2):1879–1895, 2009.

[42] O. Ilbert et al. The VIMOS-VLT deep survey. Evolution of the galaxy luminosity function up to z = 2 in first epoch data. *A&A*, 439(3):863–876, 2005.

[43] O. Ilbert et al. Galaxy Stellar Mass Assembly Between $0.2 < z < 2$ from the S-COSMOS Survey. *Astrophys.J.*, 709(2):644–663, 2010.

[44] H. Domínguez Sánchez et al. The evolution of quiescent galaxies at high redshifts (z≥ 1.4). *MNRAS*, 417(2):900–915, 2011.

[45] V. V. L. Albani et al. Differential density statistics of the galaxy distribution and the luminosity function. *Astrophys.J.*, 657:760–772, 2007.